# NATURE OF BIOLOGY

VCE UNITS 3 AND 4 FIFTH EDITION

# 2

# NATURE OF BIOLOGY

VCE UNITS 3 AND 4 FIFTH EDITION

2

Judith KINNEAR
Marjory MARTIN

jacaranda
A Wiley Brand

Fifth edition published 2017 by
John Wiley & Sons Australia, Ltd
42 McDougall Street, Milton, Qld 4064

First edition published 1992
Second edition published 2000
Third edition published 2006
Fourth edition published 2013

Typeset in 10.5/12 pt Utopia Std



National Library of Australia
Cataloguing-in-Publication entry

| | |
|---|---|
| Creator: | Kinnear, Judith, author. |
| Title: | Nature of biology. Book 2: VCE units 3 and 4 / Judith Kinnear, Marjory Martin. |
| Edition: | Fifth edition. |
| ISBN: | 978-0-7303-2878-0 (set) |
| | 978-0-7303-2882-7 (ebook) |
| | 978-0-7303-2879-7 (paperback) |
| | 978-0-7303-3110-0 (studyON) |
| Notes: | Includes index. |
| Target Audience: | For secondary school age. |
| Subjects: | Biology—Study and teaching (Secondary) |
| | Biology—Textbooks. |
| | Biology—Problems, exercises, etc. |
| | Victorian Certificate of Education examination—Study guides. |
| Other Creators/ Contributors: | Martin, Marjory, author. |
| Dewey Number: | 570.712 |



Cover images: © Pics by Nick/Shutterstock; Phillip Dyhr Hobbs/Shutterstock

Cartography by Spatial Vision, Melbourne

Typeset in India by Aptara

Illustrated by various artists, Aptara and Wiley Composition Services

Printed in Singapore by
Markono Print Media Pte Ltd

10 9 8 7 6 5 4 3 2

All activities have been written with the safety of both teacher and student in mind. Some, however, involve physical activity or the use of equipment or tools. **All due care should be taken when performing such activities**. Neither the publisher nor the authors can accept responsibility for any injury that may be sustained when completing activities described in this textbook.


*This book is dedicated to friends and colleagues who generously shared their stories.*

# Contents

# Preface

This fifth edition of *Nature of Biology Book 2* builds on previous editions that were positively received by teachers and students of biology. It has been thoroughly revised and updated, and reflects current curriculum decisions regarding the key knowledge and skills expected of biology students.

This book seeks to convey a multifaceted sense of biology: as a rigorous scientific discipline with explanatory models that organise the living world for us in a meaningful way; as a dynamic science whose explanations are subject to testing and change, rather than as a fixed and unchanging body of knowledge; as a science that impacts on everyday life, both at the level of the individual and at a societal level.

We continue to emphasise recent developments in biotechnology as exemplified by the CRISPR gene-editing technique, monoclonal antibodies and an update on genetically modified organisms (GMOs), both plant and animal. We have placed emphasis on examples relevant to Australia, including the establishment of the Australian Criminal Intelligence Commission (ACIC) in 2016, trends in organ and tissue transplantation, as well as examples of global interest such as the spread of the Zika virus. New material in the fifth edition includes updates relating to continuing curriculum topics, such as hominin evolution, and new material that reflects curriculum changes, including the immune system's components, mode of operation and the defects that can arise. Importantly, to provide the context for the study of the immune system, we have provided coverage of the types of microbes and agents that cause the myriad diseases against which our immune system provides defence.

Included in each unit are examples to assist students to understand how biological knowledge and skills are applied in a variety of settings. Throughout the book, the history of various significant discoveries, such as the hormone insulin and the Taung fossil, is outlined so that students may appreciate how advances in biology have been achieved. Likewise, aspects of the development of key concepts, such as the germ theory and the theory of evolution by natural selection, are described to illustrate that theories must be based on evidence and be subject to testing. The 'Biologist at work' profiles are intended to increase student awareness of vocational opportunities. Updated or new profiles introduce a range of people working in diverse roles. In particular, the updated box on the Walter and Eliza Hall Research Institute in Melbourne introduces students to some of the outstanding researchers who are involved in cutting-edge research programs. It is our hope that the range of profiles may inspire some readers to explore the domain of biology further through their tertiary studies and become the researchers and the practitioners of the future, or the inspiring teachers of future students of biology.

We have enjoyed writing this book and we hope that our readers will also enjoy reading the text and exploring the visual images, and gain confidence as they grapple with and master the questions associated with each chapter.

This project was greatly enhanced by the generous cooperation of many academic colleagues and friends. In particular, we owe a special debt of gratitude to the following:

Matthew Ames and family, Ivy and Van Steel and family, Dr Robyn Pickering (University of Cape Town), Dr Christophe Lasseur (European Space Agency), Christine Tomlin (Zoos Victoria, Healesville), Brian Versteeg (Canada), Professor Dietrich Stout (Emory University), Dr Olivier Schwartz (Insitut Pasteur, Paris), Professor Gillian Griffiths (Cambridge Instutute for Medical Research), Dr Vanessa Solomon (WEHI), Penny Fannin (WEHI), Dr Song Tan (Penn State University), Professor Colin Brooks (University of Newcastle, UK), Professor Nina van Sorge (University Medical Centre, Utrecht), Dr Heinrich Feldmann (NIH/NIAID, USA), Elizabeth Fischer (NIH/NIAID, USA), Dr Jessie Maisano (University of Texas), Dave Conley (AquaBounty Technologies), Dr Kristin Friedrich (La Brea Tar Pits and Museum), Dr Linzi Wilson-Wilde (ANZPAA NIFS), Dr Jessica Mazerik (NCI, Office of Cancer Genomics), Dr Kate Charlton-Robb (Australian Marine Mammal Conservation Foundation), Steven Mileto (Monash University), Dr Alex Johnson (University of Melbourne), Catherine Cavallo (Monash University), Sonia Sanchez (Monash University), Dr Dadna Hartman (Victorian Institute of Forensic Medicine), Professor John Paterson (University of New England), Professor Leigh Ackland (Deakin University), Drew Berry (WEHI), Professor Marie-Liesse Asselin-Labat (WEHI), Susie Moreton (Epworth Hospital), Shaun McKenna (Creative Photography), Professor Peter Timms (University of the Sunshine Coast), Frank Oncken (Stone Hill Construction Inc. VA, USA), Jonathan Franks and Dr Donna Stolz (University of Pittsburg).

# About eBookPLUS and studyON

Access your online Jacaranda resources anywhere, anytime, from any device in three easy steps:

**STEP 1** Go to **www.jacplus.com.au** and create a user account.
**STEP 2** Enter your registration code.
**STEP 3** Instant access!

## eBook*plus* + study*on*

**eBookPLUS is an electronic version of the textbook, together with a targeted range of supporting multimedia resources.**

**eBookPLUS features:**

- **eBook** — the entire textbook in electronic format
- **Digital documents** designed for easy customisation and editing
- **Interactivities** to reinforce and enhance students' learning
- **eLessons** — engaging video clips and supporting material
- **Weblinks** to relevant support material on the internet

**eGuidePLUS features assessment and curriculum material to support teachers.**

**studyON is an interactive and highly visual online study, revision and exam practice tool designed to help students and teachers maximise exam results.**

**studyON features:**

- **Concept summary screens** provide concise explanations of key concepts, with relevant examples.
- **Access 1000+ past VCAA questions** or custom-authored practice questions at a concept, topic or entire course level, and receive immediate feedback.
- **Sit past VCAA exams** (Units 3 & 4) or **topic tests** (Units 1 & 2) in exam-like situations.
- **Video animations and interactivities** demonstrate concepts to provide a deep understanding (Units 3 & 4 only).
- **All results and performance in practice and sit questions** are tracked to a concept level to pinpoint strengths and weaknesses.

---

**NEED HELP? Go to www.jacplus.com.au and select the Help link.**

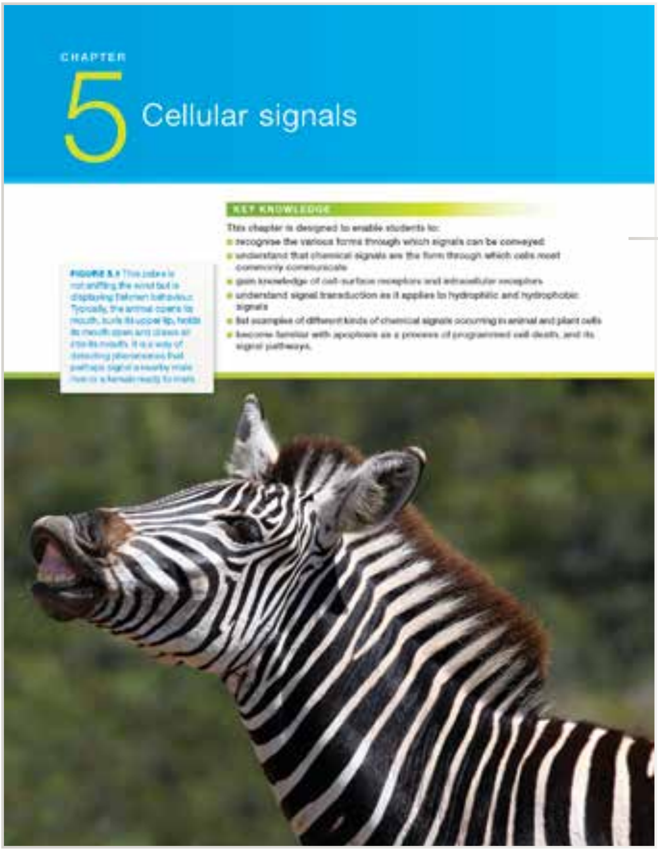
- Visit the JacarandaPLUS Support Centre at **http://jacplus.desk.com** to access a range of step-by-step user guides, ask questions or search for information.
- **Contact** John Wiley & Sons Australia, Ltd.
  **Email:** support@jacplus.com.au
  **Phone:** 1800 JAC PLUS (1800 522 7587)

**Minimum requirements**

JacarandaPLUS requires you to use a supported internet browser and version, otherwise you will not be able to access your resources or view all features and upgrades. Please view the complete list of JacPLUS minimum system requirements at **http://jacplus.desk.com**.

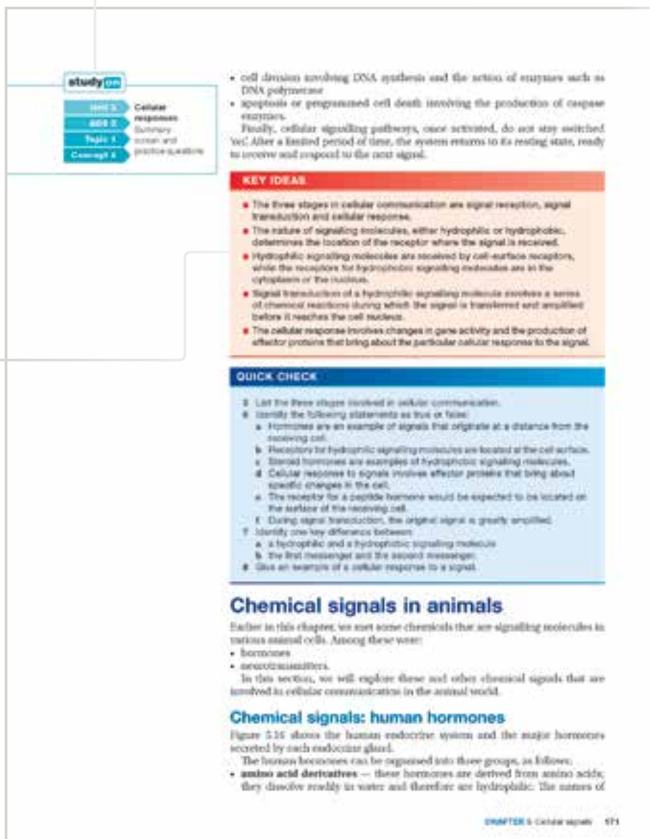
# About this book

*Nature of Biology 2 VCE Units 3 and 4* has the following features:



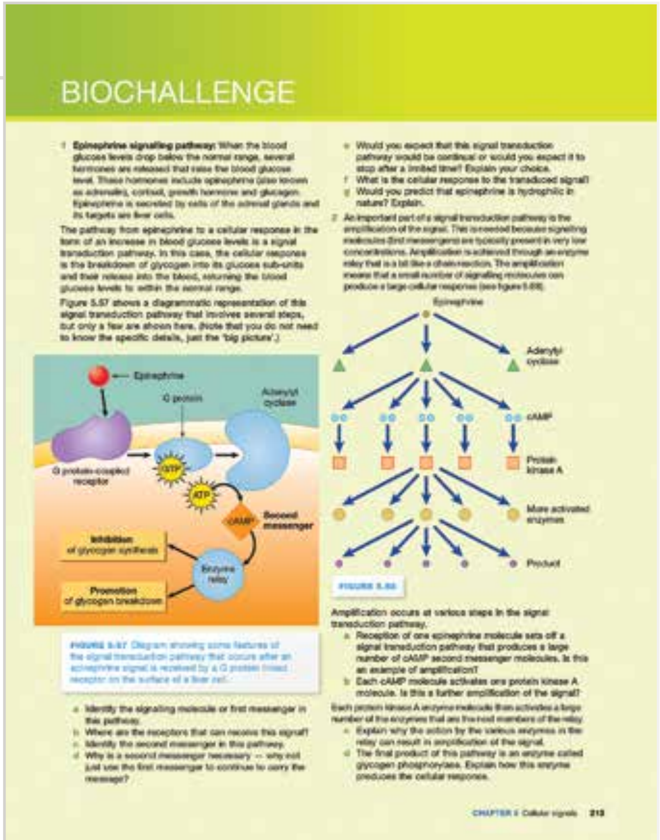

An overview of the Key knowledge and outcomes addressed in this chapter

References to studyON summary screens encapsulate key knowledge and provide questions with answers and worked solutions, to reinforce learning.




Key ideas panels provide a summary of the main points covered in this section. These are followed by Quick check revision questions.

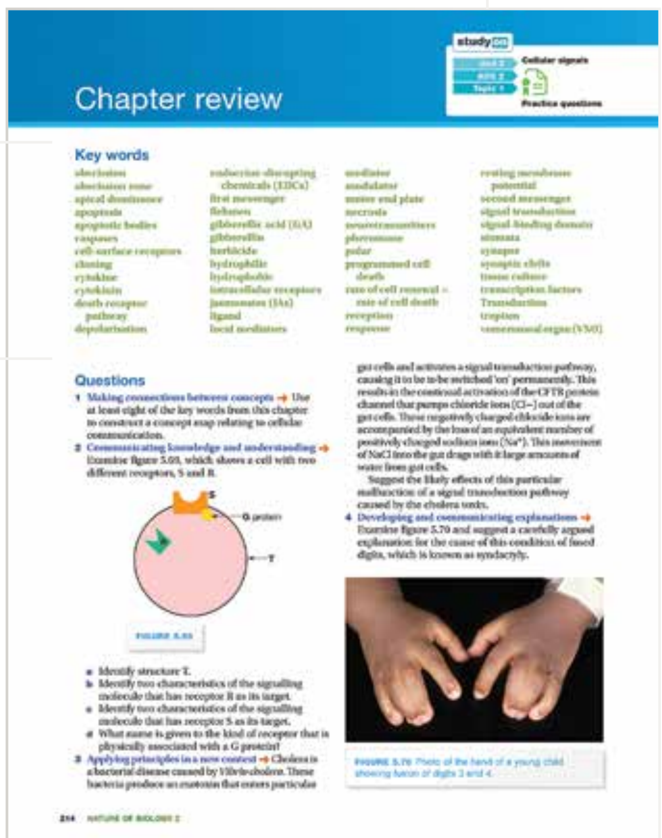
Biochallenge sections focus on applying knowledge in response to visual stimuli and data.


BIOCHALLENGE

The studyON topic review has additional multiple choice, short answer and extended response questions; these are different to the end-of-chapter review questions.

A list of the key words used in the chapter to enhance the vocabulary of the student is provided.

Chapter review questions check and challenge students' understanding.


Chapter review

Key words

Questions

# Acknowledgements

The authors and publisher would like to thank the following copyright holders, organisations and individuals for their assistance and for permission to reproduce copyright material in this book.

**Images**

• Dr. Alan R. Hildebrand • AAP Newswire: **750**/PR Handout Image • AIDS Info: **374**/AIDSinfo, Public Domain. • AIHW: **272**/Source: Australian Institute of Health and Welfare; **375**(a), **375**(b)/Based on Australian Institute of Health and Welfare material • Alamy Australia Pty Ltd: **7**/© Science Photo Library; **73**/© The Science Picture Company; **130** (bottom)/© Enviro Images; **207** (top), **214** (right)/Mediscan; **218**/© FALKENSTEINFOTO; **223**(b), **503**(b)/© Mary Evans Picture Library; **247**/© Stocktrek Images, Inc.; **267** (top right)/ © Agencja Fotograficzna Caro; **267** (bottom)/© Image Source; **353**/© amomentintime; **358**/© Papa Kay; **401**/ © KrystynaSzulecka; **402**/Trinity Mirror / Mirrorpix; **410**/© Classic Image; **411**/© George S de Blonsky; **412**, **432**; **431**, **704**/© Nigel Cattlin; **444**(a), **472**(b), **564**, **566**/© The Natural History Museum; **444**(b)/© GL Archive; **446** (left), **447**/Pictorial Press Ltd; **446**/World History Archive; **468**(e)/© Lourens Smak; **471**/© PjrStudio; **473**/ZUMA Press Inc; **490**(c)/© markku murto/art; **490**(e)/© Auscape International Pty Ltd; **493**(b)/© blickwinkel; **506**, **606**, **610**/© epa european pressphoto agency b.v.; **541**/© All Canada Photos; **584**, **608**/© Sabena Jane Blackbird; **593**/Granger, NYC.; **598**/© liszt collection; **601**/© dpa picture alliance; **604**/© ClassicStock; **626**/Cultura Creative RF; **642**/The Art Archive; **644**/© Eric Carr; **659**/AF archive; **670**/© TravelStockCollection — Homer Sykes; **695**/© Cultura Creative RF; **742**/© GRANT ROONEY PREMIUM; **746**/ guy harrop • Alex Johnson: **717** • Applied Biosystems: **45**(a), **676**/Applied Biosystems • Aqua Bounty: **718** • Arabella Smith, Dr: **624** • Australia Post: **39**(a)/Designer: Rod Oliver, Fragile Design; © Australian Postal Corporation 2003 • Australian Bone Marrow Donor Registry: **269** • Australian Criminal Intelligence Commission: **679** • BlueStar Forensic: **54** (bottom a), **54** (bottom b)/ Photo©Bluestar® • Braford: **425** • Brian McStay, Dr: **66**(b) • Bridgeman Art Library : **140**/Werner Forman Archive / Bridgeman Art Library • Canadian Biotechnology Action Network: **708** • CanStockPhoto: **234**(bii)/ rob3000 / CanStock Photo • Center for Biologic Imaging: **295**/Center for Biologic Imaging, University of Pittsburgh, USA; **317**/Jonathan Franks M.S. & Donna Stolz PhD. • Centers for Disease Control and Prevention: **730**/James Gathany / CDC/Douglas Jordan • Centers for Disease Control and Prevention: **3**/"Legionella pneumophila multiplying inside a cultured human lung fibroblast." CDC/Dr. Edwin P. Ewing, Jr; **733**, **736**, **757**, **760**/Centers for Disease Control and Prevention; **748**, **749**, **749** • Centers for Disease Control and Prevention: **217**/Centers for Disease Control and Prevention; **234**(aii) • Christine Tomlin: **496**/© Christine Tomlin • Copyright Clearance Center: **512**, **673**, **674–675** • Copyright Clearance Center: **689**; **745**/Bentham Science Publishers Ltd; **746**/Figure 6a, Antiviral adhesion molecular mechanisms for influenza: W. G. Laver's lifetime obsession, Elspeth F. Garman, 22.12.2014. • CSIRO Publishing: **400**/CSIRO Publishing • Department of Health: **345**/National Immunisation Program Schedule From 20 April 2015, Department of Health, 2016.; **349**/ Figure 1, Annual Report of the Australian Meningococcal Surveillance Programme, 2014, Monica M Lahra, Rodney P Enriquez, Department of Health.; **371**/ Figure 6.1.1, Appendix 6.1, Communicable Diseases Intelligence Volume 34, Department of Health, 2010.; **372**/Table 3, Australia's notifiable disease status, 2013: Annual Report, CDI Vol 39, No 3, 2015, Department of Health. • Dr Dadna Hartman: **681** • Dr Dietrich Stout: **585**, **585**/Photos by Emory University • Dr Emmanuelle Charpentier: **612**/Hallbauer&Fioretti • Dr Jennifer Doudna: **611** • Dr Linzi Wilson-Wilde: **685** • Dr Robyn Pickering: **567**/Photo credit: Warrick Joe • Dr. Fred Cohen: **255** • Dr. Kate Charlton-Robb: **407**(a), **407**(b) • Elizabeth Fischer: **234**(ai)/Elizabeth R. Fischer/NIH/ NIAID/Rocky Mountain Laboratories • ESA: **105**(b), **105** (bottom)/ESA • FDA: **720** • Genomic Diagnostics: **689**, **690**, **690** • Getty Images Australia: **9**(a), **31**(a), **43**, **44**, **54**(a), **120** (bottom), **277**, **325**, **415**, **589**, **619**, **653**/ Science Photo Library; **30**/P Motta & T Naguro/SPL; **34**/ Don W Fawcett; **42**/DR ELENA KISELEVA; **45**/© Tek Image/Science Photo Library; **54**(b)/SUSUMU NISHINAGA; **54**(c)/DR KLAUS BOLLER; **69** (bottom)/Mark Harmel; **105**(a), **620**/PASCAL GOETGHELUCK; **114**(a), **280**(c), **280**(d)/Ed Reschke; **117**/Dr. Terry Beveridge; **118**(b)/Paul Zahl; **127**/Ben Radford; **144**/BILL LONGCORE; **181** (top), **192**(b)/© David Scharf / Science

Photo Library; **182**(b)/© Astrid & Hanns-Frieder Michler / Science Photo Library; **193**/© SPL/ Geoff Kidd; **209**/© Photo Researchers; **222**(b)/Science & Society Picture Library; **243**/BSIP/UIG; **251**, **276**, **335**/ Bettmann; **255**/Hubert Raguet/Eurelios/SPL; **261**/ BIOPHOTO ASSOCIATES; **280**(a)/Ed Reschke; **280**(b); **114**(a), **280**(c), **280**(d)/Ed Reschke; **293**/STEVE GSCHMEISSNER; **318**/M I WALKER; **320**/SECCHI-LECAQUE/ROUSSEL-UCLAF/CNRI; **336**/DR UCT; **337**(a)/Bloomberg; **443**/Xavier ROSSI; **504**/Dan Kitwood; **530**/UniversalImagesGroup; **551**, **551**, **557**, **559**/John Reader/SPL; **568**/ALEXANDER JOE; **586**/ Kenneth Garrett / National Geographic; **602**/Cyril Ruoso/Minden Pictures; **623**/Kevork Djansezian; **653**/ Volker Steger/SPL; **654**/VOLKER STEGER; **668**/DENIS CHARLET / AFP; **676**/© Image Tek; **683**/Patrick AVENTURIER • Global Crop Diversity Trust: **436**(a), **436**(b) • Glofish: **699** • Golden Rice Humanitarian Board: **716**/ Golden Rice Humanitarian Board. www.goldenrice.org • Institute of Human Origins: **561** • ISAAA: **707**, **707**, **724**/ISAAA, 2015 • John Paterson: **480**/Katrina Kenny, David Elkins. • Judith Kinnear: **9**(b), **112**(a), **141** (left), **384**(aii), **384**(bi), **384**(bii), **384**(biii), **384**(ciii), **385**(ei), **385**(eii), **385**(eiii), **385**(f), **390**, **413** (bottom a), **413** (bottom b), **426**(a), **426**(b), **426**(c), **427** (top b), **427** (top c), **427** (top d), **434**, **434**(a), **434**(b), **445**, **448**, **449**, **450**(a), **450**(b), **453**, **459**(a), **463**, **463**, **464**, **465**, **468**(a), **468**(b), **468**(c), **468**(d), **469**, **472**(a), **472**, **544**, **583**, **583**, **669**, **682**, **686**, **691**, **709**, **715**, **715**, **715**, **715**, **715**; **192**(a), **197** (bottom), **603**/Judith Kinnear • Lawrence Berkeley Laboratory: **56**/Lawrence Berkeley Laboratory • Leigh Ackland: **97** (left), **97** (right) • Macfarlane Burnet Institute: **361**/Image was taken by the laboratory of Prof Johnson Mak Deakin University while he was Head of the Virology Program at Burnet Institute. • MAPgraphics: **441**, **441**(a), **441**(b), **451**, **598**/MAPgraphics Pty Ltd, Brisbane • Maret Vesk, Dr: **61** • Mariano Rocchi, Dr: **515** • Marjory Martin, Prof: **26**, **341**, **421**(a), **421**(b), **421**(c), **641**, **754**, **758**, **758**, **758** • Matthew Ames: **241**/ Kate Ames • Melissa Lutz Blovin: **575**/Melissa Lutz Blouin, University of Arkansas • Molecular Science Student Workbench: **539** • NASA: **109**/NASA • National Human Genome: **64**/Courtesy: National Human Genome Research Institute • National Human Genome Research Institute : **516**/www.genome.gov • National Human Genome Research Institute NHGRI, National Institutes of Health: **658**/Courtesy: National Human Genome Research Institute, https://www.genome.gov/ • National Toxicology Program: **264**/National Toxicology Program • NDNAD: **680**, **680**/NATIONAL DNA DATABASE STRATEGY BOARD ANNUAL REPORT 2012-13, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/252885/NDNAD_Annual_Report_2012-13.pdf • New England Biolabs, Inc.: **618**/Print below image: "Reprinted from www.neb.com 2016 with permission from New England Biolabs, Inc.". • Newspix: **267** (top left)/David Caird; **337**(b)/Patrick Hamilton; **667**/Michael Marschall • NIAID: **234**(bi)/NIAID • Novozymes Biologicals, Inc.: **118**(a) • NSW Department of Primary Industries : **424**/ NSW DPI Prime Facts 579, March 2007 • NSW Health: **729**, **761**/Reproduced by permission, NSW Ministry of Health © 2016 • Out of Copyright: **26**/Le Petit Journal • Pfizer Animal Genetics: **691**/Photo courtesy of Zoetis • Phillip Island Nature Parks: **437**(a), **437**(b) • Photodisc: **379**, **545**, **545**, **545**, **545**, **547**, **548**, **584**, **605**/© PhotoDisc, Inc • PlasmidFactory: **631** • PLOS: **693**/Source: Coble M.D. et al., 2009. Mystery solved: the identification of the two missing Romanov children using DNA analysis, PLoS ONE Vol. 4, No. 3: e4838.doi:10.1371/journal.pone.0004838 • PNAS: **260**/Figure 3, Adaptation of the bovine spongiform encephalopathy agent to primates and comparison with Creutzfeldt–Jakob disease: Implications for human health, PNAS, Vol. 98, 27.03.2001. Copyright 2001 National Academy of Sciences, U.S.A. • Professor Colin Brooks: **296**(a)/Colin Brooks, Professor of Cellular Immunology, University of Newcastle, UK. • Professor Gillian Griffiths: **324** • Promega Corporation: **677**/©Promega Corporation 2016. All rights reserved, Promega Corporation, 2800 Woods Hollow Rd, Madison, WI 53711 USA • Public Domain: **703**; **727**/CDC / Rebecca Myers; **729**/http://collections.museumvictoria.com.au/items/1602148 • Public Library of Science: **532**/Gareth J Fraser, C Darrin Hulsey, Ryan F Bloomquist, J Todd Streelman, An Ancient Gene Network Is Co-opted for Teeth on Old and New Jaws, Mar 2009 · PLoS Biology • RCSB PDB: **81**/Tsujino, S., Tomizaki, T., www.rcsb.org; **82**/Kloos, M., Straeter, N., www.rcsb.org; **83**(a)/Feng, J., Zhao, S., Liu, L., www.rcsb.org • Royal Mint: **39**(b)// Royal Mint • Science Photo Library: **181**(a)/Sinclair Stammers; **235**(ei)/STEVE GSCHMEISSNER; **264**/HERVE CONGE, ISM; **399**/EYE OF SCIENCE • Shutterstock: **4**(a)/ martynowi.cz; **4**(b)/www.royaltystockphoto.com; **4**(c)/ MichaelTaylor3d; **5**(a)/Juan Gaertner; **7**/LCleland; **52**(d)/molekuul_be; **77**/Helen Sushitskaya; **79**/effe45; **111**/jeka84; **125**(a), **125**(b), **135**(b), **207** (bottom), **212**, **238**, **279**(a), **314**/Designua; **130** (top)/Martin Lisner; **131**/Rudmer Zwerver; **133**/Radu Bercan; **135**(a)/Tewan Banditrukkanka; **139**/Rena Schild; **141** (right)/kaband; **149**/Heiti Paves; **159**/Four Oaks; **160**/Blaj Gabriel; **178** (top)/joshya; **178** (bottom), **341**/Everett Historical; **180** (top)/eastern light photography; **181**(b)/alexsvirid; **182**(a)/Alf Ribeiro; **194**/Tricia Daniel; **200**/logoboom; **203** (bottom left)/azure; **203** (bottom right)/ Eduard Kyslynskyy; **225**/Satirus; **226**/alexmak7; **231**/ Smith1972; **235**(ci)/Kateryna Kon; **235**(cii)/Mukhina Viktoriia; **235**(eii), **278**, **331**, **351**(b), **359**, **360**/Alila Medical Media; **239**(b)/toeytoey; **244**/villorejo; **256**/ thelefty; **279**(b)/Sebastian Kaulitzki; **284**/Blamb; **289** (top)/Jose Luis Calvo; **289**(a)/Jose Luis Calvo; **290**, **305**(b)/Kondor83; **303**/FCG; **307**(a)/Dirima;

**328**/Tatiana Shepeleva; **339**/Dmitry Lobanov; **343**(a)/jomphong; **344**/pruit phatsrivong; **351**(a)/BlueRing-Media; **352**/Jarun Ontakrai; **354**/Africa Studio; **357**/Lichtmeister; **357**/Anton Gvozdikov; **377**(a)/Catzatsea; **377**(b)/photomaster; **378**, **381**/ChameleonsEye; **384**(ai)/Malota; **384**(ci)/Raywoo; **384**(cii)/2shrimpS; **385**(di)/Martin Fowler; **385**(dii)/Steve McWilliam; **386**(a)/Ivan Kuzmin; **387**/muratart; **405**/© Leksele; **413** (top a)/Xseon; **413** (top b)/Deamles for Sale; **413** (top c), **440**/Eric Isselee; **423**/Artush; **425**(a)/eltoro69; **425**(b)/Paul Looyen; **427** (bottom a)/© Halina Yakushevich; **427** (bottom b)/© Margo Harrison; **427** (bottom c)/© Sari ONeal; **427** (bottom d)/© chris2766; **427** (bottom e)/Ewan Chesser; **429**(a)/AsyaPozniak; **429**(b)/© Eric Isselée; **429**(c)/© Eric Isselée; **431**(a)/© Viorel Sima; **431**(b)/© dcwcreations; **464**(b)/© Shutterstock.com / Inc; **482**/stihii; **484**(a)/N Mrtgh; **484**(b)/© Steve Bower; **486**(a)/mark higgins; **486**(b)/Kelsey Green; **486**(c)/2630ben; **486**(d)/© Karel Gallas; **487**(a)/© nialat; **487**(b)/© Martha Marks; **490**(a)/Bildagentur Zoonar GmbH; **490**(b)/worldswildlifewonders; **490**(d)/Kristian Bell; **491**(a)/Susan Flashman; **491**(b)/© Timothy Craig Lubcke; **491**(c)/© dirkr; **492**/Brooke Whatnall; **497**(a), **608**/Bardocz Peter; **497**(b)/Sergei Drozd; **524**/annalisa e marina durante; **527**/dimair; **531**/AndreJakubik; **542**/GUDKOV ANDREY; **543**/jo Crebbin; **544**/Volodymyr Burdiak; **544**/J. Norman Reid; **545**/© Sergey Uryadnikov; **550**, **578**, **673**/Rawpixel.com; **570**/Kjersti Joergensen; **593**/rnl; **593**/MaluStudio; **593**/nikolae; **593**/Stephen Marques; **593**/Andrey Burmakin; **602**/Oksana Kuzmina; **604**/Monkey Business Images; **604**/3Dsculptor; **644**/ittipon; **651**/anyaivanova; **661**/Doro Guzenda; **687**/John Carnemolla; **687**/fivespots; **696**/goodluz; **708**/Jerry Horbert; **710**/Fotokostic; **714**/Mykola Komarovskyy; **734**/Veremeya; **742**/adriaticfoto • Song Tan: **42**(b)/FOR PRINT: Song Tan, Penn State University. FOR ELECTRONIC VERSIONS: www.personal.psu.edu/sxt30.; **313**/Song Tan, Penn State University • SpaceHabs: **103** • State Government of Victoria: **374** • Stephen Vasquez: **197** (top)/UC Regents & photographers Stephen Vasquez & Matthew Fidelibus, UC Cooperative, Viticulture Farm Advisor and UC Cooperative Viticulture Specialist. • Steven Mileto: **634** • Susie Moreton: **535** • Suzana Herculano-Houzel: **547** • Tariq M. Rana, Ph.D: **713**/Tariq M. Rana, Ph.D., University of California, San Diego. • The Genetics Society of Japan: **524**/Takashi Kitano, Norimasa Matsuoka, Naruya Saitou, Phylogenetic relationship of the genus Oncorhynchus species inferred from nuclear and mitochondrial markers, Issue 72, 30, 1997. Genetics Society of Japan. • The Kirby Institute: **359**/The Kirby Institute • The University of Texas at Austin: **470**/The University of Texas High-Resolution X-ray CT Facility and the Texas Memorial Museum. • Thermo Fisher Scientific: **627**, **627**/Photo courtesy of Thermo Fisher Scientific • Torsten Wittmann: **1** • United States Geological Survey: **498**(a) • University Autònoma of Barcelona, Spain: **108** • University of Pennsylvania: **702**/Image by R.L Brinster, School of Veterinary Medicine, University of Pennsylvania • USDA Economic Research: **725**/U. S. Department of Agriculture. • Walter & Eliza Hall Institute: **33**, **269**, **321**, **321**, **330**, **367**, **368** (top), **368** (bottom), **369**, **370**, **744** • Washington University In St. Louis: **420**/AM Bauer and SCR Elgin • Wellcome Trust, The: **220**/The Wellcome Library • Wikimedia Commons: **5**(c)/Dartmouth Electron Microscope Facility, Dartmouth College; **29**(a)/© Public Domain / Louisa Howard; **219**, **222**(a), **223**(a), **229**, **360**, **379**, **411**, **492**, **493**(a), **503**(a), **521**, **533**, **644**, **694**, **704**, **736**/© Wikimedia Commons; **235**(di)/Rocky Mountain Laboratories, NIAID, NIH http://www3.niaid.nih.gov/biodefense/Public/Images.htm; **235**(dii)/Centers for Disease Control and Prevention; **258**/Public Health Image Library; **464**(a)// Wikimedia Commons • Wiley Owned: **427** (top a), **593** • Willem Rens, Dr: **517**/Willem Rens • World Health Organization: **733**/World Health Organization

**Text**

• Selected extracts from the VCE **Biology Units 3 and 4 (2017-2021)** are copyright Victorian Curriculum and Assessment Authority (VCAA), reproduced by permission. VCE® is a registered trademark of the VCAA. The VCAA does not endorse this product and makes no warranties regarding the correctness or accuracy of its content. To the extent permitted by law, the VCAA excludes all liability for any loss or damage suffered or incurred as a result of accessing, using or relying on the content. Current VCE Study Designs and related content can be accessed directly at www.vcaa.vic.edu.au. Teachers are advised to check the VCAA Bulletin for updates.

Every effort has been made to trace the ownership of copyright material. Information that will enable the publisher to rectify any error or omission in subsequent reprints will be welcome. In such cases, please contact the Permissions Section of John Wiley & Sons Australia, Ltd.

CHAPTER

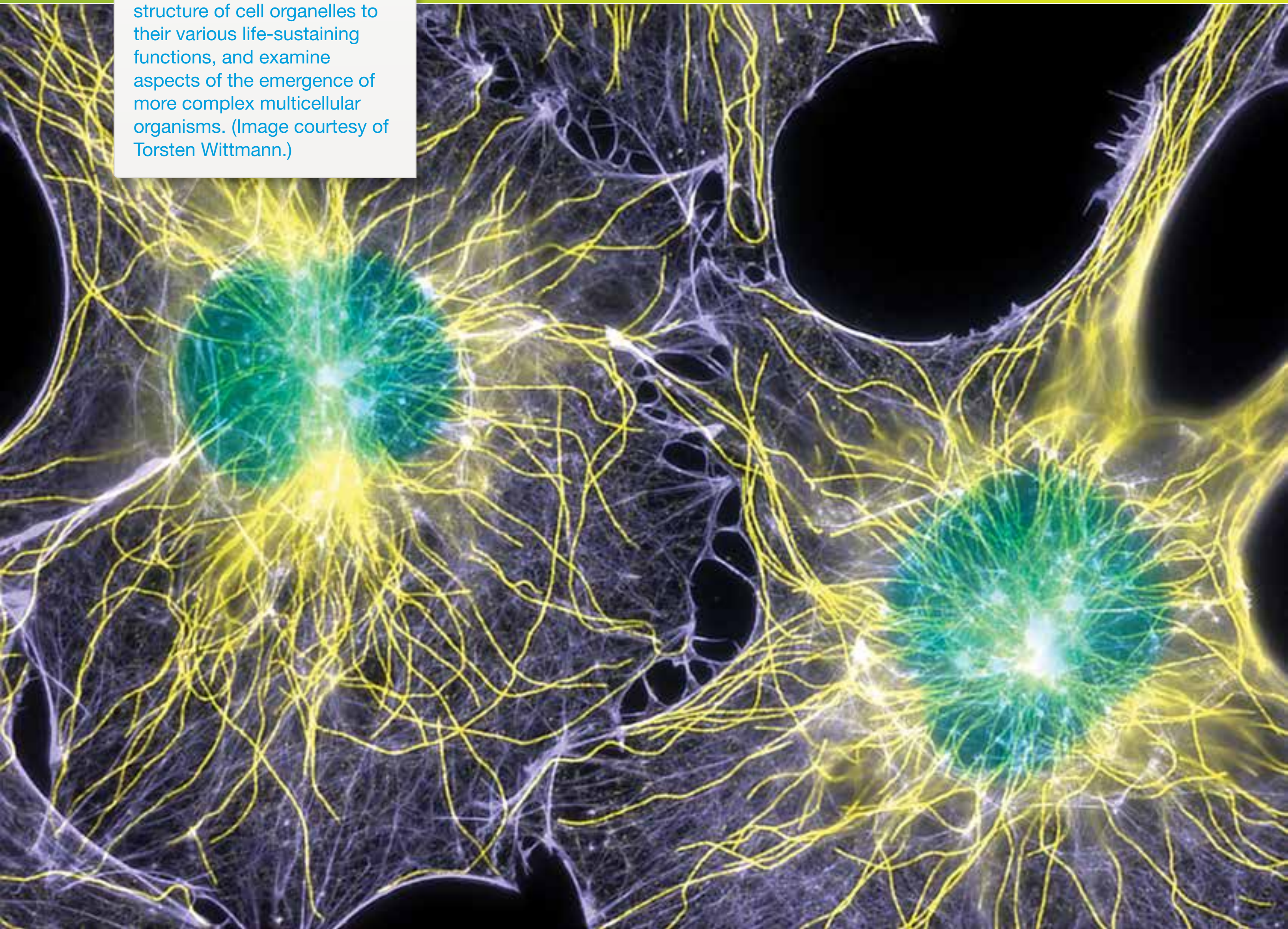
# 1 Cells: a review

**FIGURE 1.1** This prize-winning image shows eukaryotic cells stained with fluorescent probes. The actin filaments (purple) and the microtubules (yellow) are part of the cytoskeleton of these cells. The nucleus is stained green. Note that the cell on the left appears to be in the process of dividing. In this chapter, we will explore the infrastructure of eukaryotic cells (both plant and animal), relate the structure of cell organelles to their various life-sustaining functions, and examine aspects of the emergence of more complex multicellular organisms. (Image courtesy of Torsten Wittmann.)

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- review and expand on key knowledge from Units 1 and 2
- understand that cells are the basic units of structure and function of living organisms
- understand and apply the concept of surface-area-to-volume ratio
- list the defining characteristics of prokaryotic and eukaryotic cells
- recognise the plasma membrane as the boundary separating the cell from its external environment
- describe the various modes of transport across the plasma membrane
- identify cell organelles in eukaryotic cells and recognise their various functions.



# Cells: the basic units of life

**Cells are the basic structural and functional units of life, and all living organisms are built of one or more cells.** Cells, with only a very few exceptions, are too small to be seen with an unaided eye. Their existence was not recognised until after the development of the first simple microscopes. This enabled the first observations of cells to be made, in the 1660s. However, the recognition of cells as the basic unit of life did not occur until almost 200 years later.

1 millimetre (mm) = 1000 micrometres (μm)

1 micrometre (μm) = 1000 nanometres (nm)

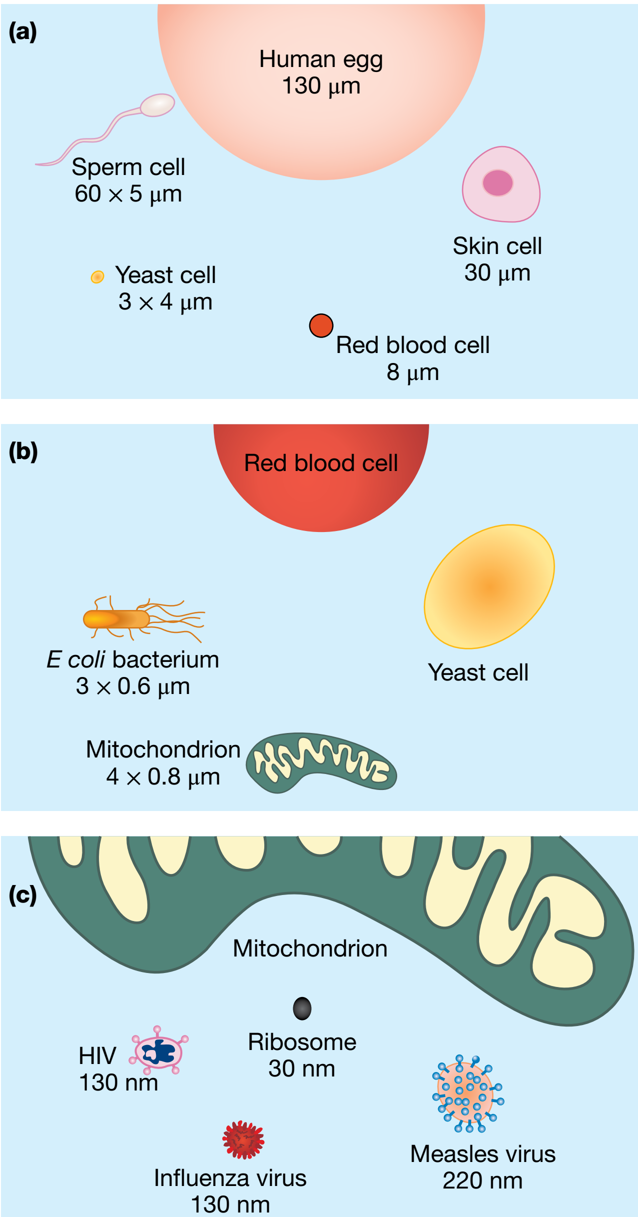



**FIGURE 1.2** Diagrams, at increasing levels of magnification, showing cells, cell organelles and viruses. Note the extreme differences in size. **(a)** Some human cells showing variation in cell size **(b)** A bacterial cell with a mitochondrion, a cell organelle and other small cells shown for comparison **(c)** A mitochondrion compared with three viruses and a ribosome

## Cells: how big?

**Cells are typically microscopic (not visible with an unaided eye).** Only a few single cells are large enough to be seen with an unaided human eye, for example, human egg cells with diameters about 0.1 mm and the common amoeba (*Amoeba proteus*), a unicellular organism with an average size ranging from 0.25 to 0.75 mm. (You would see an amoeba as about the size of a full stop on this page.) Contrast this with one of the smallest bacteria, *Pelagibacter ubique*, consisting of a cell just 0.2 μm diameter. How many of these bacteria could fit across an amoeba that is 0.5 mm wide?

- Most animal cells fall within the size range of 10 to 40 μm. Among the smallest human cells are red blood cells with diameters for normal cells in the range of 6 to 8 μm.
- Plant cells typically fall in the range of 10 to 100 μm.
- Microbial cells, both bacterial and archaeal, are much smaller than plant and animal cells. Most bacterial cells have diameters in the range of 0.4 to 2.0 μm and 0.5 to 5 μm in length. **On average, microbial cells are about 10 times smaller than plant and animal cells, with sizes typically in the few micrometres range.**

A *non-living* microworld exists beyond that of microbes. This is occupied by viruses, which are non-cellular particles that are generally regarded as belonging to the grey area between living and non-living. Why? Because viruses do not have a cellular structure, they cannot carry out metabolic activities in isolation and they cannot self-replicate. (Viruses can replicate only inside and with the assistance of living cells.) Viruses range in diameter from 20 to 300 nanometres (nm). (Refer to chapter 6, page 245, for more detail about viruses, in particular their role in human diseases.) Figure 1.2 shows a sample of the range of sizes seen in selected cells. (Other non-cellular structures are included for size comparison.)

Microbial cells are relatively much smaller than the cells of animals and plants, so some animal bacterial infections can involve the invasion of bacterial cells into the cells of the host, where they multiply. Examine figure 1.3 of a human lung fibroblast and note the presence of numerous bacterial cells in a single cell. This image highlights the size difference between microbial cells and the cells of animals (and plants).

**FIGURE 1.3** Transmission electron microscope (TEM) image of a lung fibroblast infected with many bacterial cells (shown as small dark circular and ovoid shapes). The bacteria are *Legionella pneumophila*, the cause of several infections in people, including Legionnaires' disease.

**studyon**

Unit 3 | AOS 1 | Topic 1 | Concept 1

**Key skills for Units 3 and 4**
Summary screen and practice questions

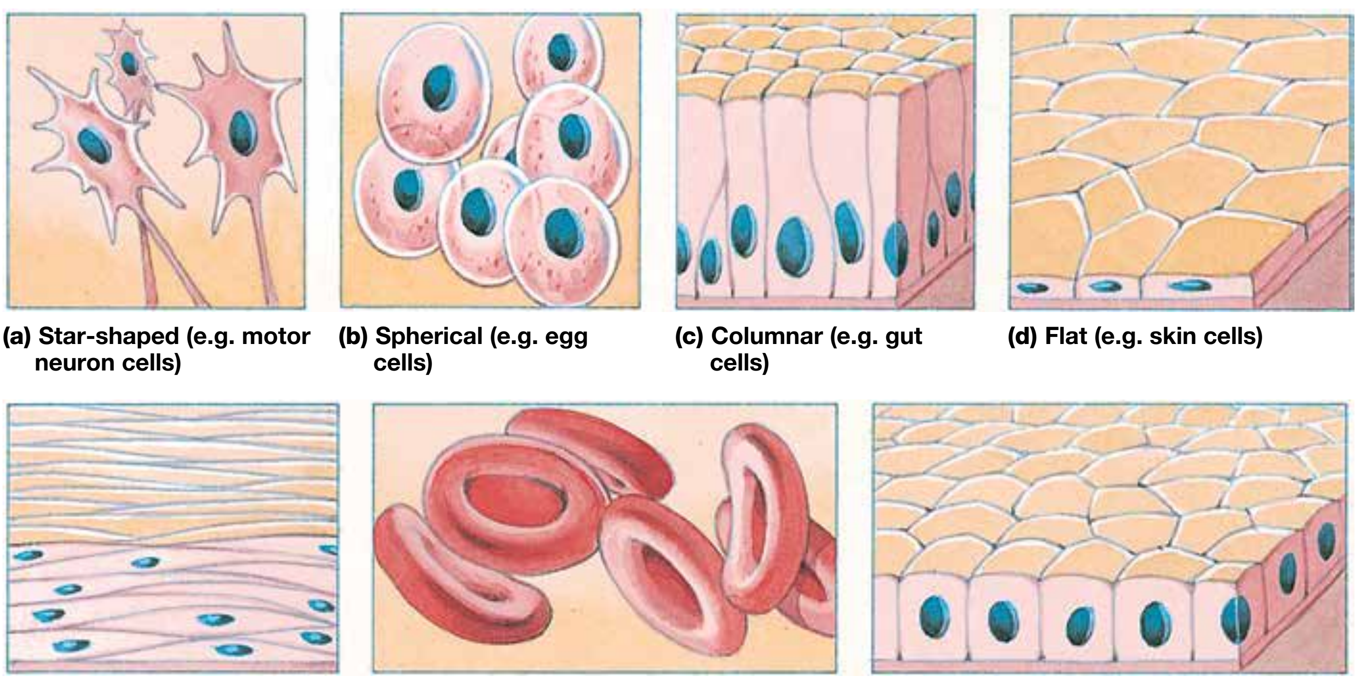
**studyon**

Unit 3 | AOS 1 | Topic 1 | Concept 1

**Do more**
Catalase and temperature

## Cells: all sorts of shapes

There is no fixed shape for cells. Cells vary in shape and their shapes often reflect their functions. Figure 1.4 shows some examples of cell shapes. Scan this figure and note that some cells are thin and flattened, others are column-shaped, yet others are spherical.



**(a) Star-shaped (e.g. motor neuron cells)**

**(b) Spherical (e.g. egg cells)**

**(c) Columnar (e.g. gut cells)**

**(d) Flat (e.g. skin cells)**

**(e) Elongated (e.g. human smooth muscle cells)**

**(f) Disc-shaped (e.g. human red blood cells)**

**(g) Cuboidal (e.g. human kidney cells)**

**FIGURE 1.4** Examples of variations in cell shape: **(a)** star-shaped **(b)** spherical **(c)** columnar **(d)** flat **(e)** elongated **(f)** disc-shaped **(g)** cuboidal

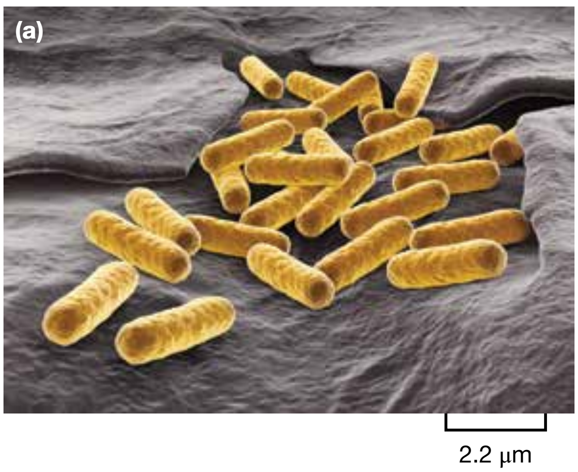
Look at figure 1.4a. Note the long axon, which is a distinctive feature of motor neuron cells. These cells transmit nerve impulses from a person's spinal cord to voluntary muscles throughout the body. In this case, the shape of the nerve cell is fitted to its conductive function. Can you estimate the approximate length of a motor neuron that has its cell body in the lower spinal cord with its axon reaching to your big toe?

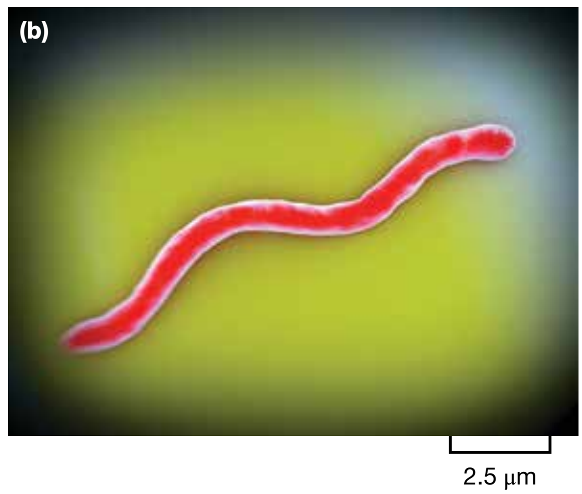
Look at figure 1.4e. Note the spindle-shaped smooth muscle cells. Smooth muscle cells contain special proteins that criss-cross the cell, and when these proteins contract the smooth muscle fibres shorten. The spindle shape of these cells is suited to their contractile function. Bundles of smooth muscle cells are found in the gut wall, in the walls of blood vessels, in ducts of secretory glands and in the wall of the uterus. These bundles of smooth muscle cells can generate sustained involuntary contractions in these organs.

Microbial cells also vary in shape (see figure 1.5). Note that some bacteria are rod-shaped, such as the gut-dwelling bacterium *Escherichia coli*; some are corkscrew-shaped, such as *Borrelia burgdorferi,* the causative agent of Lyme disease; while others are more or less spherical, such as *Streptococcus pneumonia*, the cause of many infections, including pneumonia.

**ODD FACT**

Motor neurons in animals, such as the giant squid (*Architeuthis* sp.), may be as long as 12 metres.



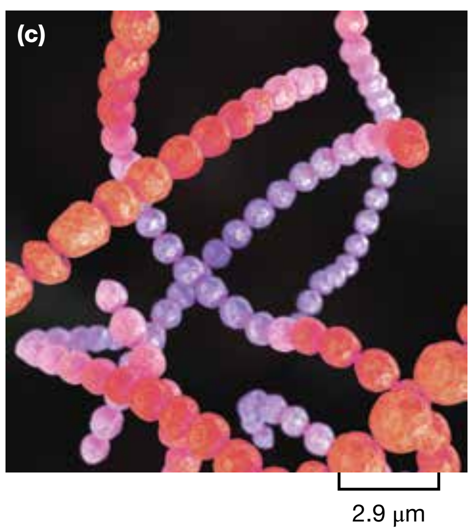




**FIGURE 1.5** Bacterial cells come in many shapes. Some are **(a)** rod-shaped bacilli (singular: bacillus) **(b)** spiral-shaped and **(c)** spherical cocci (singular: coccus).

**studyon**

Unit 3 | AOS 1 | Topic 2 | Concept 1

**Cells of living organisms**
Summary screen and practice questions

**studyon**

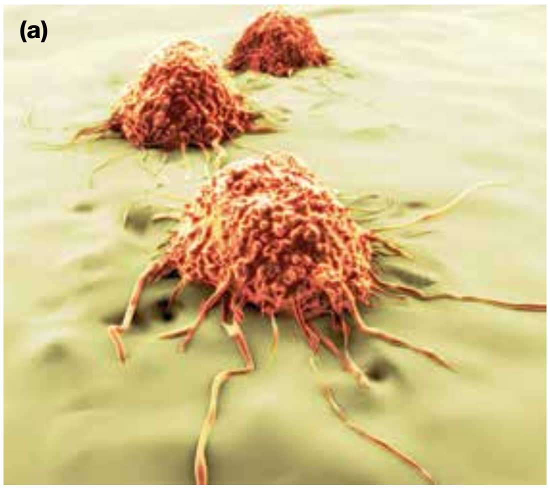
Unit 3 | AOS 1 | Topic 2 | Concept 1

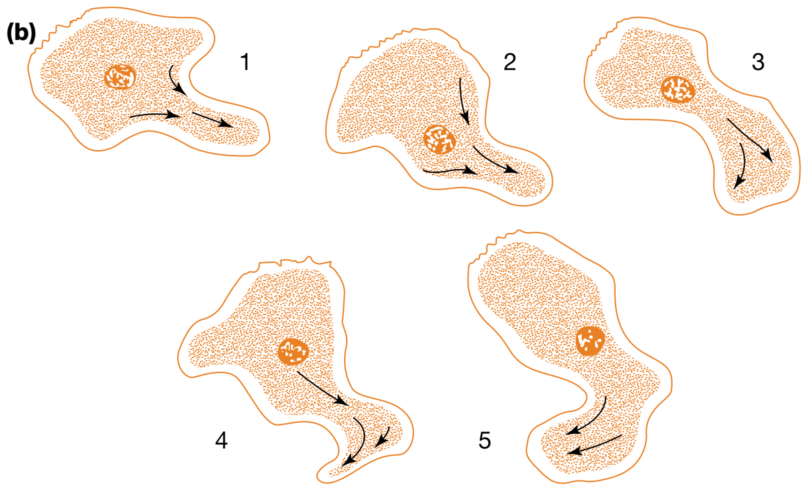
**See more**
Living organisms are made of cells

Not all cells have a fixed shape. For example, some cells are able to move actively, and these self-propelled cells do not have fixed shapes because their outer boundary is their flexible plasma membrane. So, as these cells move, their shapes change. Examples of cells capable of active self-propelled movement include:

- cancer cells that migrate into capillaries and move around the body when a malignant tumour undergoes metastasis (see figure 1.6a). The thread-like protrusions (known as filopodia) that fold out from the plasma membrane of cancer cells make a cancer cell self-mobile and able to migrate from a primary tumour and invade other tissues.
- white blood cells that can squeeze from capillaries into the surrounding tissues where they travel to attack infectious microbes (refer to figure 1.12, p. 13)
- amoebas as they move across surfaces (see figure 1.6b).

Some other cells that have a fixed shape because of the presence of a rigid cell wall outside their plasma membranes can self-propel. However, this ability depends on the presence of cilia or flagella to power their movement. For example, the green alga, *Chlamydomonas* sp., moves due to the beating of its two flagella (see figure 1.6c).



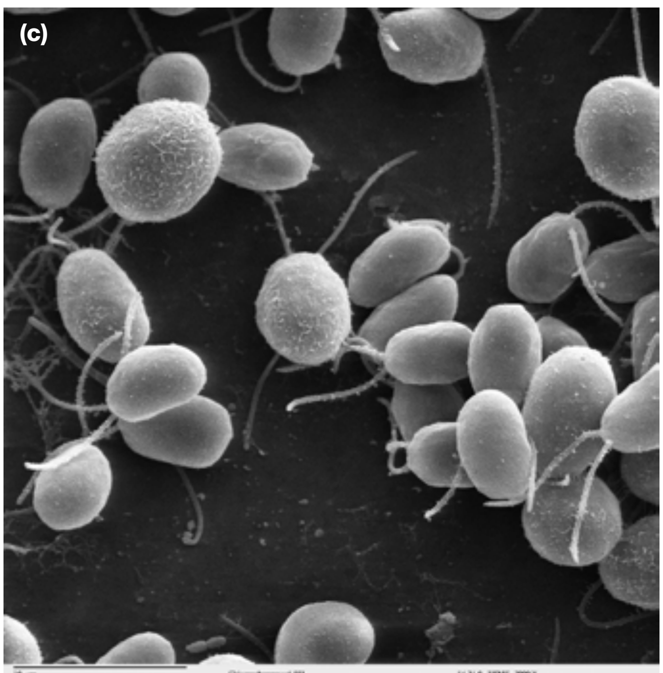




**FIGURE 1.6 (a)** Cancer cells. Note the many thread-like projections (filopodia) that enable these cancer cells to be mobile or self-propelling. The ability to move is an important factor in the spread of a malignant cancer. **(b)** Outlines showing the changing shape of an amoeba as it moves **(c)** Scanning EM image of *Chlamydomonas reinhardtii*, a single-celled green alga. Note the presence of two flagella that make this organism able to self-propel. What is the reason for the fixed shape of this organism? Like all algae, this organism has a rigid cell wall that defines its shape.

## Cells: why so small?

Why are cells microscopically small? Would it be more efficient to have a larger macroscopic unit to carry out cellular processes rather than many smaller units occupying the same space? To answer these questions we need to look at the concept of surface-area-to-volume ratio.

### Surface-area-to-volume ratio

Every living cell must maintain its internal environment within a narrow range of conditions, such as pH and the concentrations of ions and chemical compounds. At the same time, a cell must carry out a variety of functions that are essential for life. These functions include trapping a source of energy, obtaining the chemical building blocks needed for cellular repair, growth and reproduction, taking up water and nutrients, and removing wastes.

- These essential functions require a constant exchange of material between the cell and its external environment.
- The site of exchange where materials are moved into or out of a cell is the plasma membrane, also termed the cell membrane. The plasma membrane must enable enough exchange between the external and internal environments to support these life functions.
- The exchange of materials must occur at rates sufficient to ensure that substances are delivered fast enough into cells to meet their nutrient needs and that wastes are removed fast enough from the cells to avoid their accumulation.

A critical issue in keeping a cell alive is the surface area of plasma membrane available to supply material to or remove wastes from the metabolically active cytoplasm of the cell. This can be quantified by a measure termed the **surface-area-to-volume ratio**, abbreviated SA:V ratio. This ratio provides a key clue to the answer to the question: why are cells so small?

Let us look at the SA:V ratio for some identical shapes of different sizes.

Consider some cubes.

The surface area of a cube is given by the equation:

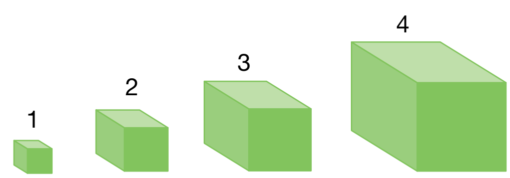
$$\mathbf{SA = 6L^2}$$

where L = the length of one side of the cube

The volume of a cube is given by the equation:

$$\mathbf{V = L^3}$$

Examine figure 1.7. Note that as the cubes increase in size, their volumes enlarge faster than their surface areas expand. When the side length doubles from 1 to 2, the volume increases by 8, but the surface area only increases by 4. This is reflected in a decrease in the SA:V ratio as the cube grows bigger.




| | | | | |
|---|---|---|---|---|
| Length of side | 1 | 2 | 3 | 4 |
| Surface area | 6 | 24 | 54 | 96 |
| Volume | 1 | 8 | 27 | 64 |
| SA:V | 6:1 | 3:1 | 2:1 | 3:2 |

**FIGURE 1.7** The surface area (SA) and volume (V) of cubes with increasing side lengths (L). With each increase in the length of a side, an increase occurs in both the surface area and the volume of the cube. Do these two measures increase at the same rate? If not, which parameter — surface area or volume — increases more rapidly?

This generalisation applies to other shapes; that is, the SA:V ratio of a smaller object is higher than that of a larger object with the same shape. The higher the SA:V ratio, the greater efficiency of two-way exchange of materials across the plasma membrane; that is, efficient uptake and output of dissolved material is favoured by a high SA:V ratio.

The same principle applies to cells. **As cells increase in size through an increase in cytoplasm, both their surface areas and volumes increase, but *not* at the same rate. The internal volumes of cells expand at a greater rate than the areas of their plasma membrane.** This means that the growth of an individual cell is accompanied by a relative decrease in the area of its plasma membrane.

The metabolic needs of a cell increase in proportion to the volume of metabolically active cytoplasm. But, the inputs/outputs of materials to meet these needs increase only in proportion to the cell surface area. So, as a cell increases in cytoplasmic volume, its metabolic needs increase faster than the cell's ability to transport the materials into and out of the cell to meet those needs. The continued decrease in SA:V ratio as metabolically active cells increase in size places an upper limit on cell size. This is the one clue to why metabolically active cells are so small.

In general, the rate at which nutrients enter and wastes leave a cell is inversely proportional to the cell size, as measured in metabolically active cytoplasm; in other words, the larger the cell, the slower the rate of movement of nutrients into, and wastes out of, a cell. **Beyond a given cell size, the two-way exchange of materials across the plasma membrane cannot occur fast enough to sustain the volume of the cell contents.** If that cell is to carry out the functions necessary for living, it must divide into smaller cells or die.

Some cells show features that compensate for the decrease in SA:V ratio with increasing size. This occurs, for example, in the cells that function in the absorption of digested nutrients from your small intestine. These cells greatly increase their surface area with only a minimal increase in cell volume. How is this achieved? This is achieved by means of extensive folding of the plasma membrane on the cell surface that faces into the gut lumen (see figure 1.8). These folds are termed microvilli (singular: microvillus). Surfaces of other cells with either a major absorptive or secretory function also show microvilli. Another compensatory strategy seen in some cells involves their overall shape; SA:V ratio is higher in a long thin cell than in a spherical cell.

**ODD FACT**

Surface-area-to-volume ratio considerations apply not only to individual cells but also to entire organisms; for example, the sea anemone has many thin tentacles, each armed with stinging cells — these provide a greatly increased surface area for gaining nutrients for the whole animal (as compared with a single flat sheet of cells).

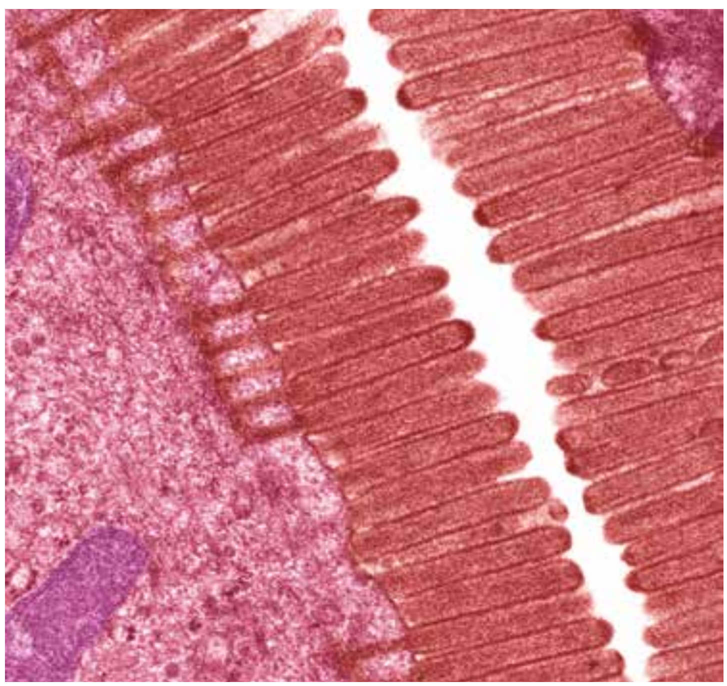


**FIGURE 1.8** TEM image showing a section through part of two cells from the lining of the small intestine. Note the multiple folds of the plasma membranes on these cells (the multiple folds face into the intestinal space, or lumen). These folds, known as microvilli, produce a great increase in the surface area for absorption of digested nutrients. Does the folding also produce a great increase in cell volume?

**ODD FACT**

A single cell lining the small intestine may have up to 10 000 microvilli on its apical surface facing into the gut lumen. How would this affect the surface area available for absorption of digested nutrients, compared with a cell with no microvilli?

**KEY IDEAS**

- Cells are the basic structural and functional units of life.
- Cells are typically too small to be seen by an unaided eye.
- The unit of measurement used for cell size is the micrometre (μm), one millionth of a metre.
- Microbial cells are much smaller than plant and animal cells.
- The metabolic needs of a cell are determined by its metabolically active cytoplasmic volume.
- The ability of a cell to meet its metabolic needs is determined by the surface area of the cell.
- As a cell increases in size, its internal volume expands at a greater rate than the area of its plasma membrane.
- The surface-area-to-volume ratio (SA:V ratio) of a smaller object is higher than that of a larger object with the same shape.
- The continued decrease in SA:V ratio as metabolically active cells increase in size places an upper limit on cell size.

**QUICK CHECK**

1 Identify whether each of the following statements is true or false:
   a Cells are typically too small to be seen with an unaided eye.
   b Bacterial cells are typically larger than animal cells.
   c Viral particles are smaller than microbial cells.
   d As a given shape increases in size, its surface-area-to-volume ratio increases.
   e Beyond a given cell size, the two-way exchange of materials across the cell surface cannot occur at a rate sufficient to meet the needs of a cell.
2 Two spheres (A and B) have different diameters, with A being larger than B. Which has the higher SA:V ratio?

## Prokaryotes: no nuclear envelope!

The microscopically tiny creatures that we call 'microbes' are a diverse group of organisms. In fact, the microbes comprise two different classification groups, namely bacteria and archaea. The cells of all these microbes can be readily distinguished from the cells of the other major groups of living organisms: fungi, plants and animals. **The key distinguishing feature of archaea and bacteria is that their cells lack a membrane-bound nucleus** (see figure 1.9a). Cells with this characteristic are described as **prokaryotic** cells and organisms displaying this feature are called **prokaryotes**. Prokaryotes are generally assumed to be the oldest existing form of life on planet Earth. The absence of a distinct nucleus does *not* mean that prokaryotes, such as archaea and bacteria, lack genetic material. Like all other kinds of organism, archaea and bacteria have DNA in their cells, but the DNA in prokaryotic cells is dispersed, not enclosed within a separate membrane-bound compartment.

In contrast, the cells of all other organisms — protists, fungi, plants and animals — have a definite nucleus (see figure 1.9b). The nucleus is enclosed by a double membrane, called the **nuclear envelope**. Organisms with this feature are termed **eukaryotes** and their cells are described as being eukaryotic. The nucleus of a **eukaryotic** cell contains DNA, the genetic material of cells. In addition, eukaryotic cells contain many membrane-bound cell organelles that are not present in prokaryotic cells (see table 1.1).

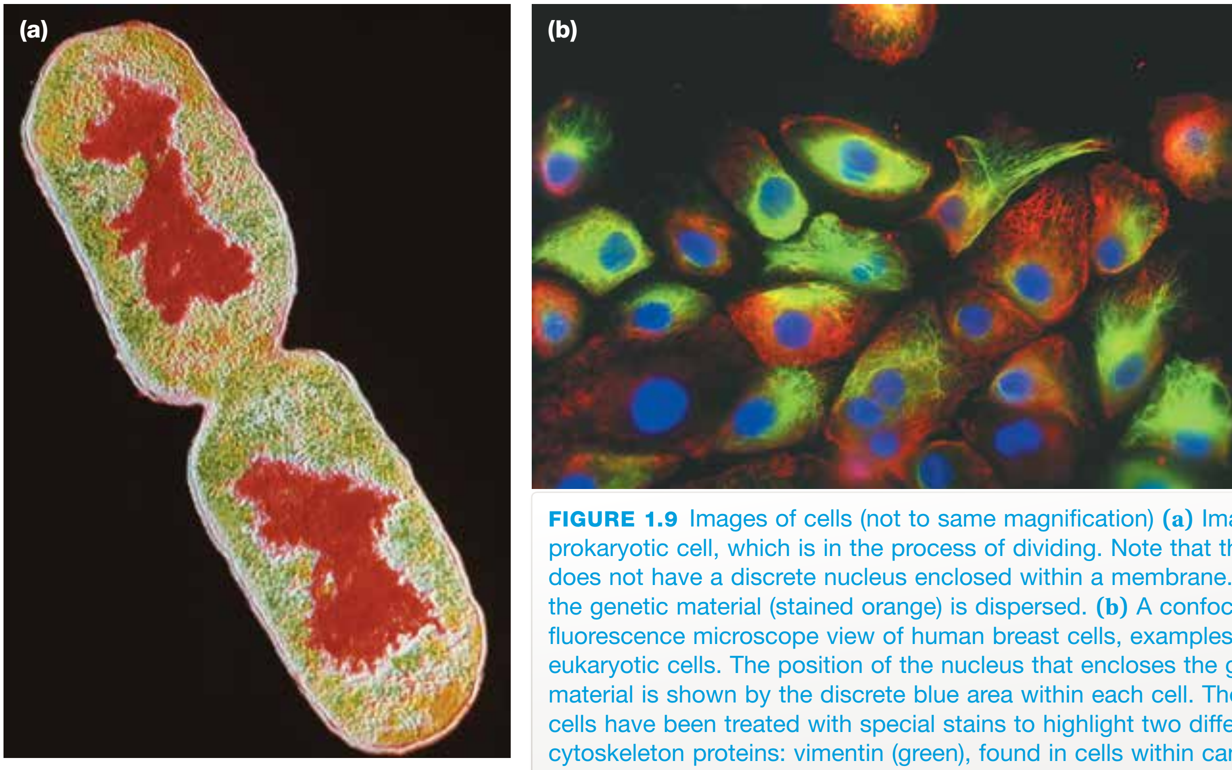



**FIGURE 1.9** Images of cells (not to same magnification) **(a)** Image of a prokaryotic cell, which is in the process of dividing. Note that the cell does not have a discrete nucleus enclosed within a membrane. Instead, the genetic material (stained orange) is dispersed. **(b)** A confocal fluorescence microscope view of human breast cells, examples of eukaryotic cells. The position of the nucleus that encloses the genetic material is shown by the discrete blue area within each cell. The cells have been treated with special stains to highlight two different cytoskeleton proteins: vimentin (green), found in cells within cancerous tissue, and keratin (red). (Image (b) courtesy of Judith Kinnear)

*pro* = before + *karyon* = kernel, nucleus

*eu* = well, good + *karyon* = kernel, nucleus

## Comparing prokaryotes with eukaryotes

Figure 1.10 shows the structure of a typical prokaryotic cell in comparison with a eukaryotic cell. Note that a prokaryotic cell has a simple architecture in contrast to a eukaryotic cell, which has a more elaborate structure owing to the presence of many membrane-enclosed compartments within the cell.

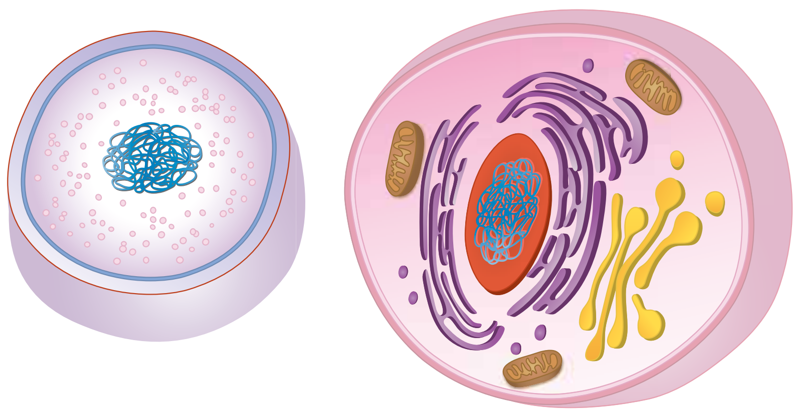


**FIGURE 1.10** Diagram showing a basic comparison of a prokaryotic cell (left) and a eukaryotic cell (right). The key distinction between these cells is the presence, only in eukaryotic cells, of membrane-bound organelles, in particular the nucleus that contains the genetic material, DNA. (Cells are not drawn to scale.)

Table 1.1 outlines a comparison between the structures of prokaryotic and eukaryotic cells. The critical difference is the absence of membrane-enclosed organelles in prokaryotes, in contrast to eukaryotic cells in which the nucleus and many other cell organelles are membrane-enclosed.

In general, prokaryotic cells are about 10 times smaller than eukaryotic cells. However, size is *not* an absolute distinction; there are some rare exceptions:

- Large prokaryotic cells exist, such as the giant bacterium *Thiomargarita namibiensis* (0.1 × 0.3 mm in diameter), which lives in the muddy sea floor off the coast of Namibia.
- Relatively small eukaryotic cells exist, such as the single-celled green alga *Ostreococcus tauri*, which is just 0.8 μm in diameter.

**TABLE 1.1** Comparison of prokaryotic and eukaryotic cells.

| Feature | Prokaryote | Eukaryote |
|---|---|---|
| size | small: typically ~1–2 μm diameter | larger: typically in range 10–100 μm |
| chromosomes | present as single circular DNA molecule | present as multiple linear DNA molecules |
| ribosomes | present: small size (70s) | present: large size (80s) |
| plasma membrane | present | present |
| cell wall | present and chemically complex | present in plants, fungi, and some protists, but chemically simple; absent in animal cells |
| membrane-bound nucleus | absent | present |
| membrane-bound cell organelles | absent | present; e.g. lysosomes, mitochondria |
| cytoskeleton | absent | present |

In general, prokaryotes are unicellular and the great majority of eukaryotes are multicellular. However, it cannot be assumed for certain that a unicellular organism is a prokaryote (either a bacterium or an archaean), because a number of eukaryotes are unicellular. Unicellular eukaryotes are protists, such as *Amoeba*, *Paramecium* and *Euglena*, some algae such as *Chlorella* and the diatoms, and fungi such as yeasts.

While there are some differences in aspects of the structure of eukaryotic and prokaryotic cells, there are many similarities in their structures and functioning. These common features reflect the inter-connectedness of life forms.

For example, both prokaryotic and eukaryotic cells:

- have DNA as their genetic material; however, in comparison with eukaryotes, prokaryotes have only about 0.001 times the amount of DNA
- have plasma membranes that selectively control the entry and exit of dissolved materials into and out of the cell
- use the same chemical building blocks including carbon, nitrogen, oxygen, hydrogen and phosphorus, to build the organic molecules that form their structure and enable their function
- produce proteins through the same mechanism (transcription of DNA and translation of mRNA on ribosomes)
- use ATP as their source of energy to drive the energy-requiring activities of their cells.

## One or more compartments?

What is strikingly different is that every prokaryotic cell is a single compartment, with no further subdivisions of the cell. This is in contrast to eukaryotic cells, which are organised internally into various compartments, each enclosed by a membrane. Because of the multi-compartmental structure of eukaryotic cells, their ultrastructure is more complex than that of prokaryotic cells.

Table 1.2 shows the relative volumes of the different compartments in a eukaryotic cell, a liver cell.

**TABLE 1.2** Relative volumes of the major compartments within a liver cell

| Intracellular compartment | Percentage of total cell volume* |
|---|---|
| cytosol | 54 |
| mitochondria | 22 |
| rough endoplasmic reticulum | 9 |
| smooth endoplasmic reticulum | 6 |
| nucleus | 6 |
| lysosomes, peroxisomes, endosomes | 3 |

*More than half of the cell volume is occupied by the cytosol.

The multicompartment structure of a eukaryotic cell enables it to maintain different conditions within each membrane-enclosed compartment that are suitable for the particular function of that compartment. Think about a house that is subdivided into rooms with different functions: you shower in the bathroom, not in the kitchen; the stove is in the kitchen, not in the bedroom. A eukaryotic cell can be likened to a house — its many compartments are like different rooms where different tasks are carried out. (Later in this chapter, we will explore some of the various compartments in eukaryotic cells.)

**KEY IDEAS**

- Prokaryotic cells lack a membrane-bound nucleus, and organisms lacking a nuclear envelope are termed prokaryotes — bacteria and archaea.
- Prokaryotes differ from eukaryotes in the absence of membrane-bound cell organelles of any kind.
- Eukaryotic cells are typically about ten times larger than prokaryotic cells.
- Eukaryotic cells have a membrane-bound nucleus in addition to other membrane-bound organelles.
- Organisms built of cells that have a nucleus enclosed within a nuclear envelope are termed eukaryotes — protists, fungi, plants and animals.

**QUICK CHECK**

**3** Identify whether each of the following statements is true or false:
  **a** Prokaryotes are unicellular organisms, comprising bacteria and archaea.
  **b** The presence of a membrane-bound nucleus in its cells provides evidence that an organism is a eukaryote.
  **c** All eukaryotes are multicellular organisms.
  **d** The various compartments within eukaryotic cells would be expected to have identical conditions.

**4** List two similarities between prokaryotic and eukaryotic cells.

**5** A unicellular organism was found in a sample of pond water. Is it reasonable to conclude that this organism must be either a bacterium or an archaeon? Briefly explain.

# Plasma membrane: the gatekeeper

The cells of all living organisms have a boundary that separates their internal environment from their surroundings. From single-celled organisms, such as amoebae or bacteria, to multicellular organisms, such as mushrooms, palm trees and human beings, each of their cells has an active boundary called the **plasma membrane**, also known as the **cell membrane**.

studyon

Unit 3 | AOS 1 | Topic 2 | Concept 2

**Structure of the plasma membrane**
Summary screen and practice questions

The plasma membrane forms the outer boundary of the living compartment of every cell. Within this compartment, conditions can be established that differ from those in the external environment and that support the living state. The plasma membrane can exclude some substances from entering the cell, while permitting entry of other substances and elimination of yet other substances. Without such a boundary, life could not exist, and indeed could not have evolved.

The plasma membrane boundary can be thought of as a busy gatekeeper selectively controlling the entry and exit of materials into and out of cells. As such, the plasma membrane is said to be **semipermeable** or **selectively permeable**, meaning that it allows only some substances to cross it — in or out.

This gatekeeper function ensures that materials required by the cell are supplied and that excesses and wastes are removed, both entry and exit occurring at rates sufficient to maintain the internal environment of the cell within narrow limits. This is quite a cellular balancing act! In addition to its role in transportation of materials into and out of the cell, the plasma membrane plays other important cellular roles (see p. 16).

The cells of fungi, plants and many bacteria and archaea have rigid cell walls outside their plasma membranes. The cells of animals do not have a cell wall. Cell walls do *not* control which materials enter or leave cells; instead, cell walls provide strength and give a fixed shape to those cells that possess them (see chapter 2). A cell wall is fully permeable so that gases and dissolved solutes can pass freely across it. It is only when these substances reach the plasma membrane that their passage may be blocked.

**studyon**

| | |
|---|---|
| Unit 3 | **A selectively permeable membrane** |
| AOS 1 | Summary screen and practice questions |
| Topic 2 | |
| Concept 3 | |

## Structure of the plasma membrane

Small, but vitally important, the plasma membrane is just 8 nanometres (nm) wide and so is only visible using a transmission electron microscope (TEM). A TEM image of the plasma membrane has a 'train track' appearance with two dark lines separated by a more lightly stained region. These images were important clues in elucidating the structure of the plasma membrane (see Biochallenge, p. 34).

The plasma membrane has two major components:

1. **phospholipids**. These are the main structural component of the plasma membrane (see below).
2. **proteins**. Most proteins are embedded in the plasma membrane, while others are attached at the membrane surface (see next page).

Let's consider each of these components in turn.

**ODD FACT**

In addition to phospholipids, the plasma membrane of animal cells also contains cholesterol as part of its structure.

### Phospholipids

The plasma membrane consists of a double layer (bilayer) of phospholipids. Each phospholipid molecule consists of two fatty acid chains joined to a phosphate-containing group. The phosphate-containing group of a phospholipid molecule constitutes its water-loving (**hydrophilic**) head. The fatty acid chains constitute the water-fearing (**hydrophobic**) tail of each phospholipid molecule.

Hydrophilic (water-loving) molecules dissolve readily in water.

Lipophilic substances dissolve readily in organic solvents such as benzene.

Hydrophobic (water-fearing) molecules are usually lipophilic (lipid-loving).

Examine figure 1.11. Notice that the two layers of phospholipids are arranged so that the hydrophilic heads are exposed at both the external environment of the cell and at the cytosol (the internal environment of the cell). In contrast, the two layers of hydrophobic tails face each other in the central region of the plasma membrane. Water and lipids do not mix.

At human body temperature, the fatty acid chains in the inner portion of the plasma membrane are not solid. Instead, they are viscous fluids — think about thick oil or very soft butter — this makes the plasma membrane flexible, soft and able to move freely. This property of the plasma membrane is very important as it enables cells to change shape (provided they do not have a cell wall outside the plasma membrane). For example, red blood cells are about 8 micrometres (μm) in diameter. When circulating red blood cells reach a capillary bed, they must deform themselves by bending and stretching in order to squeeze through capillaries, some of which have diameters as narrow as 5 micrometres. Likewise, when white blood cells reach sites of infection, they must squeeze out of small gaps between the single layer of cells that forms the capillary walls (see figure 1.12). Shape changes by animal cells are only possible because of the flexible nature of the lipids in the plasma membrane. Flexibility and shape changes are not possible for cells with cell walls.

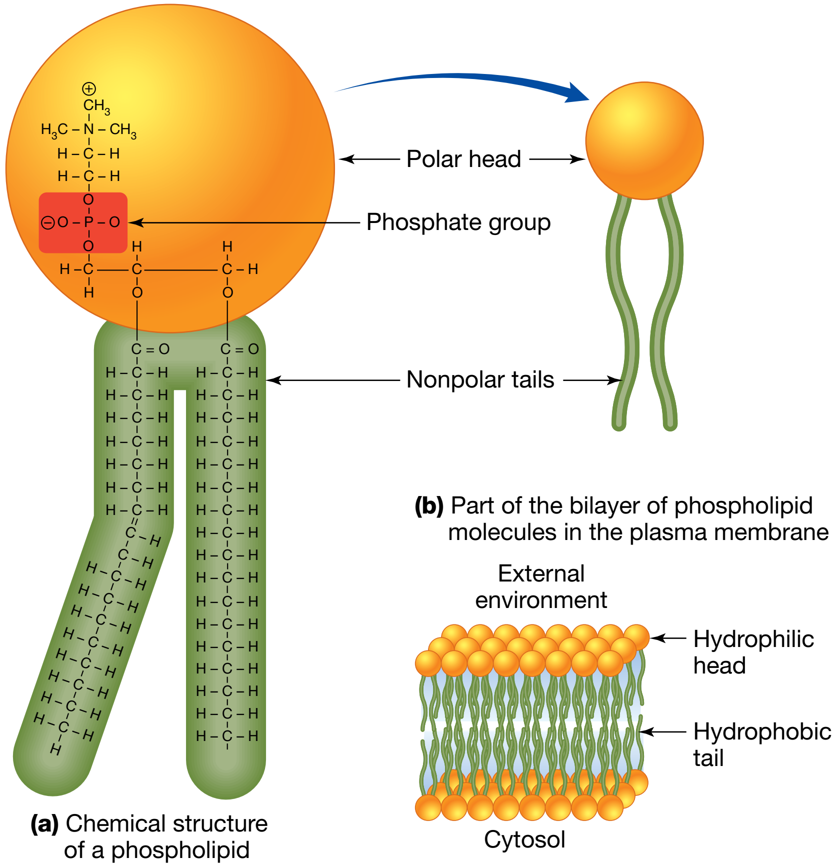



**FIGURE 1.11** **(a)** Chemical structure of a phospholipid (left) and a stylised representation (right) showing the hydrophilic head and the two fatty acid chains that make up its hydrophobic tail **(b)** Diagram showing part of the bilayer of phospholipid molecules in the plasma membrane. Notice that the tails face each other and are enclosed in the central region of the membrane, while the heads face outwards to the cell's external environment and inwards to its cytoplasm.

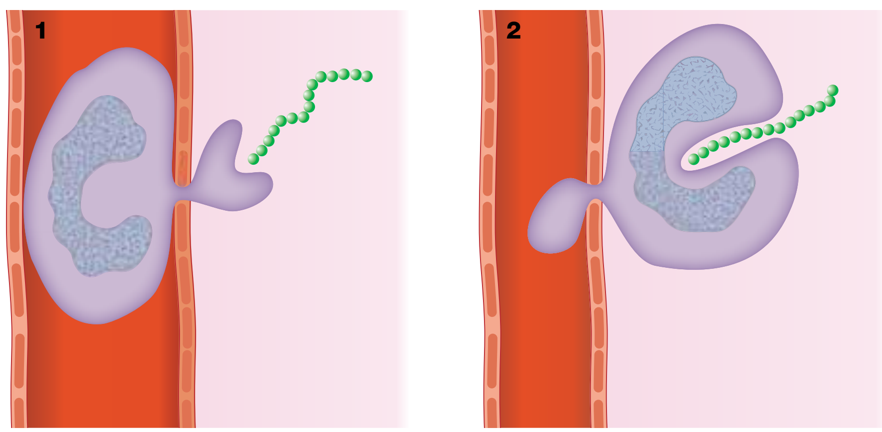



**FIGURE 1.12** Diagram showing a white blood cell squeezing between the cells of the one-cell-thick wall (endothelium) of a capillary to engulf bacteria. What feature of the plasma membrane enables the white blood cell to do this?

## Proteins

Proteins form the second essential part of the structure of the plasma membrane. Many different kinds of protein comprise part of the plasma membrane. They can be broadly grouped into:

- integral proteins
- peripheral proteins.

**Integral proteins**, as their name implies, are fundamental components of the plasma membrane. These proteins are embedded in the phospholipid bilayer. Typically, they span the width of the plasma membrane with part of the protein being exposed on both sides of the membrane (see figure 1.13). Proteins like this are described as being **trans-membrane**. In some cases carbohydrate groups, such as sugars, are attached to the exposed part of these proteins on the outer side of the membrane, creating a combination called a **glycoprotein**. Integral proteins can be separated from the plasma membrane only by harsh treatments that disrupt the phospholipid bilayer, such as treatment with strong detergents.

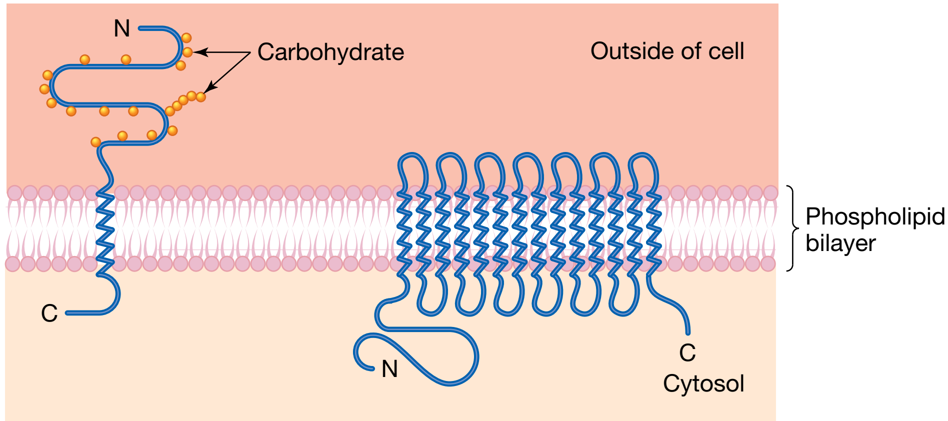



**FIGURE 1.13** Diagram showing two integral proteins embedded in and spanning the plasma membrane. Part of each protein is exposed on each side of the membrane. Note that carbohydrate groups are attached to the exposed part of one protein on the outer side of the membrane. What name could be given to this kind of protein?

The prefix 'glyco' means sugar. sugars attached to a protein = glycoprotein

sugars attached to a lipid = glycolipid

**Peripheral proteins** are either anchored to the exterior of the plasma membrane through bonding with lipids or are indirectly associated with the plasma membrane through interactions with integral proteins in the membrane. Peripheral proteins can be more easily separated from the plasma membrane than integral proteins.

The various roles of the proteins in the plasma membrane are outlined on pages 16–17 of this chapter.

## Fluid mosaic model of plasma membrane

Both phospholipids and proteins are key components of the structure of the plasma membrane. But how are they organised?

An early view was that the proteins present in the plasma membrane were concealed within the phospholipid bilayer. However, in 1972, Singer and Nicolson proposed the **fluid mosaic model** of membrane structure. This is now generally accepted as the structure for the plasma membrane. The fluid mosaic model also applies to the membranes that form the outer boundary of cell organelles, such as the membranes that surround the cell nucleus and other cell organelles.

The fluid mosaic model proposes that the plasma membrane and other intracellular membranes should be considered as two-dimensional fluids in which proteins are embedded.

The term 'fluid' comes from the fact that the fatty chains of the phospholipids are like a thick oily fluid, and the term 'mosaic' comes from the fact that the external surface (when viewed from above) has the appearance of a mosaic because of the various embedded proteins set in a uniform background. Figure 1.14 shows a diagrammatic representation of the fluid mosaic model.

A good working definition of the plasma membrane is that it is the active boundary around all living cells that consists of a phospholipid bilayer and associated proteins and which separates the cell contents from their external environment.

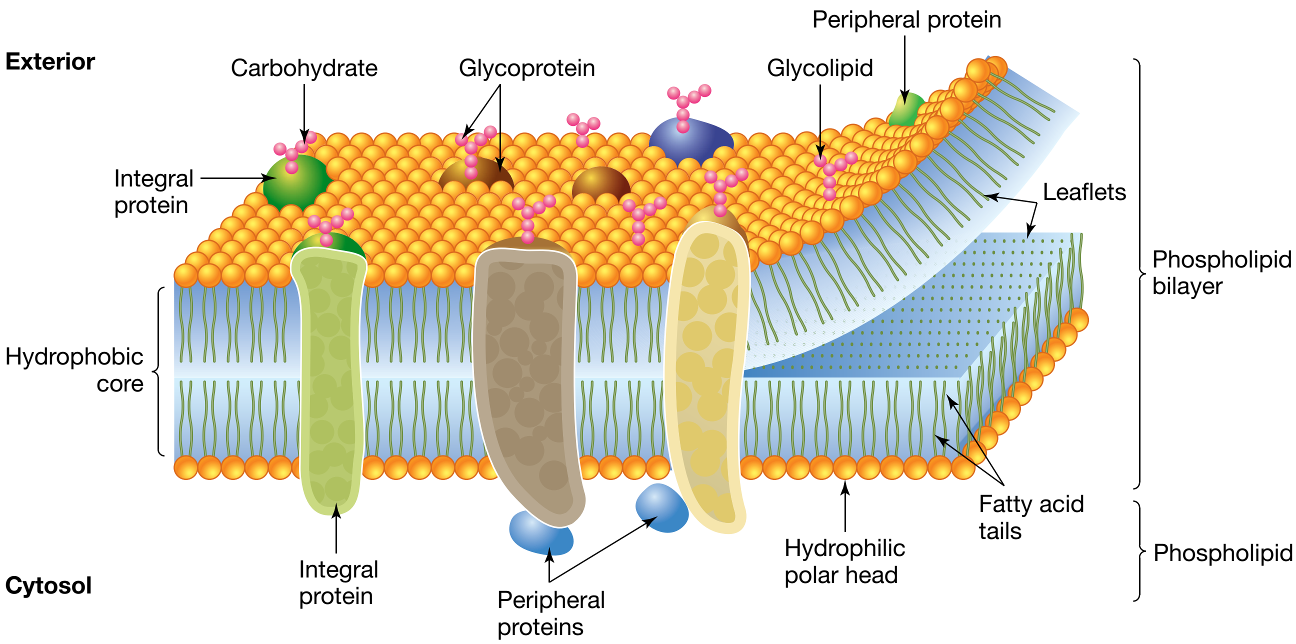



**FIGURE 1.14** Diagram showing the fluid mosaic model of membrane structure. Note the bilayer of phospholipids. What are the two components of the phospholipids? Note the integral proteins, some of which extend through the bilayer and are exposed at both the outer and the inner surface of the membrane. Do any of these proteins have carbohydrate chains attached to their exposed regions? Note the peripheral proteins that are more loosely associated with the membrane.

## KEY IDEAS

- **A major role of the plasma membrane of a cell is to act as a gatekeeper that controls the entry and exit of materials into and out of the cell.**
- **The major structural component of plasma membrane is a bilayer of phospholipid molecules, each with a hydrophilic head and hydrophobic tail.**
- **The fatty acid chains within the plasma membrane confer flexibility on the plasma membrane.**
- **Proteins comprise the other essential component of the plasma membrane; these are both integral proteins and peripheral proteins.**
- **The fluid mosaic model of membrane structure is currently accepted as the best description of the structure of the plasma membrane (and other cellular membranes).**

## QUICK CHECK

6 What are the two major components of a plasma membrane?
7 Identify whether each of the following statements is true or false:
  a The plasma membrane is present as a boundary in all living cells.
  b The plasma membrane consists of layers of proteins in which phospholipids are embedded.
  c A key role of the plasma membrane is the control of transport of materials into or out of cells.
  d Trans-membrane proteins span the width of the plasma membrane.
8 What is a glycoprotein?
9 What part of a plasma membrane is responsible for its flexibility?
10 Briefly outline the fluid mosaic model of the plasma membrane.

# Functions of the plasma membrane

The plasma membrane carries out several important functions for a cell. The plasma membrane:

1. is an active and selective boundary
2. denotes cell identity
3. receives external signals
4. transports materials.

## The active boundary

The plasma membrane forms the active boundary of a cell, separating the cell from its external environment and from other cells; it allows the passage of some substances only. The plasma membrane forms the boundary of a compartment in which the internal environment of a living cell can be held within a narrow range of conditions that are different from those of the external environment.

Within the cell, similar membranes form the active boundaries of cell organelles, including the nucleus, the endoplasmic reticulum, the Golgi apparatus and **lysosomes**. In other cell organelles, such as mitochondria and chloroplasts, membranes form both the external boundary and part of the internal structure. Because of the presence of their membrane boundaries, membrane-bound cell organelles can maintain internal environments that differ from those in the surrounding cytosol and can perform different functions. This is a critical role of cell membranes.

## Cell identity

**ODD FACT**

The human blood group A and B antigens are present on red blood cells as glycolipids and as glycoproteins — they differ by one sugar group.

Glycoproteins on the outer plasma membrane function as **cell surface markers**, also known as antigens or cell identity tags. Each cell type has a different combination of surface markers. In mammals, these markers enable the immune system to identify these cells as 'self' and distinguish them from foreign cells. Glycolipids on the plasma membrane play a role in tissue recognition.

## Receiving external signals

Cells receive signals from their external environment. In the case of a multicellular organism, the signal may originate from another cell within that organism. In the case of unicellular organisms, the signal may come from other organisms in its neighbourhood or from its external environment. These external signals are often chemical compounds, for example, hormones.

Cell signalling is treated in detail in chapter 5.

Trans-membrane proteins on the outer surface of the plasma membrane are the **receptors** for these signals, and each cell has many different kinds of receptor protein. The signal binds to the receptor protein and this binding alters the shape of the receptor protein and starts a specific response in the cell. For example, *Saccharomyces cereviseae* is a single-celled yeast used in winemaking, brewing and bread making. During one stage of the yeast life cycle, haploid yeast cells exist in one of two different mating types, designated 'a' and '$\alpha$' (alpha). When one type is ready to mate, it releases a small chemical, called a mating factor, into its environment. The mating factor is a signal to nearby yeast cells of the other mating type that it is ready to mate (see figure 1.15). The signal from an a-type yeast cell can be received by receptors on the surface of the plasma membrane of yeast cells of the other mating type. The signal binds to a specific receptor and causes a change in the behaviour in the receiver yeast cell — it changes its shape and moves towards the source of the signal. The originator of the signal also changes shape in response. The end result is the fusion of the two haploid yeast cells to form one diploid yeast cell.

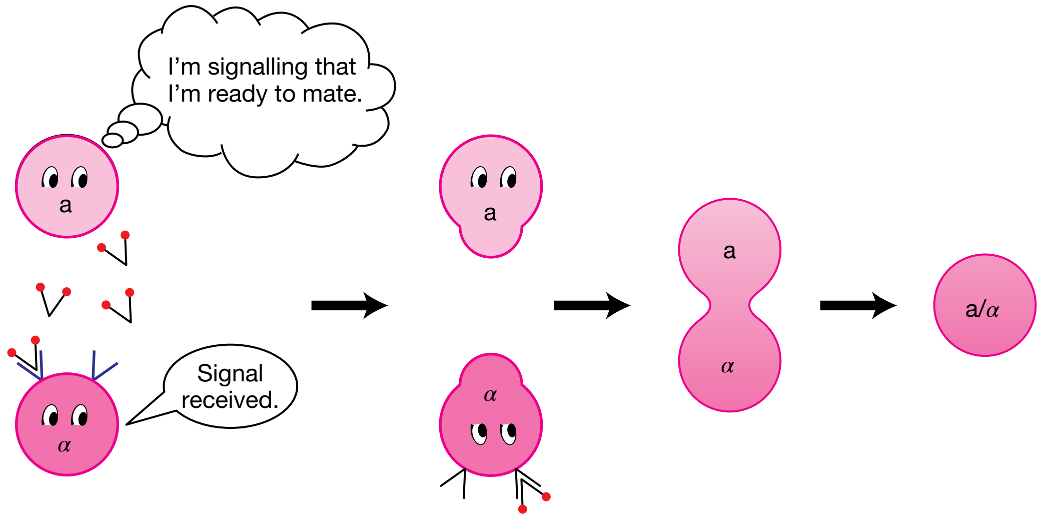



**FIGURE 1.15** Two mating types (a and α) of haploid cells occur during the life cycle of yeast. Chemical signals released by a yeast cell of one mating type can travel to surface protein receptors on neighbouring cells of yeasts of the other mating type. This signal is an invitation to mate. Reception of the signal causes a change in the behaviour of the receiver cell.

**eBook*plus***

**eLesson**
Mechanisms of membrane transport
**eles-2463**

## Transport

All cells must take in or expel a range of substances and the plasma membrane forms a selectively permeable barrier between a cell and its external environment. An impermeable barrier allows no substances to cross it; a fully permeable barrier allows all substances to cross it, while a selectively permeable barrier allows some substances to cross it but precludes the passage of others.

Some substances can cross the hydrophobic phospholipid bilayer of the plasma membrane. Other substances can cross the plasma membrane, but only with the assistance of special trans-membrane proteins, collectively called transporters, that are embedded in the plasma membrane. In the next section, the various modes by which molecules can be transported into and out of cells will be explored, including those that involve transporter proteins.

### KEY IDEAS

- The plasma membrane performs a range of cellular functions.
- Functions involving the plasma membrane include creating a compartment that separates the cell from its external environment, receiving external signals as part of communication between cells, providing cell surface markers that identify the cell and transporting materials across the plasma membrane.
- Transport proteins in the plasma membrane enable movement of substances that cannot cross the lipid bilayer of the membrane.

## QUICK CHECK

**11** Is the plasma membrane impermeable, selectively permeable or fully permeable?
**12** Identify two functions of the plasma membrane.
**13** What kind of proteins act as cell identity tags?
**14** What advantage might result from creating several membrane-enclosed compartments within a cell?
**15** What is the role of receptor proteins in the plasma membrane?
**16** Give an example of a problem that arises from the malfunction of a protein transporter in the plasma membrane.

# Crossing the plasma membrane

Generally, substances entering or exiting a cell are in aqueous solution. Several factors determine whether or not dissolved substances can diffuse down their concentration gradients across the phospholipid bilayer of the plasma membrane (see table 1.3).

**TABLE 1.3** **Factors affecting the ease with which substances can cross the plasma membrane by simple diffusion**

| | |
|---|---|
| molecular size | Smaller molecules cross more easily than larger molecules; however, very large molecules (macromolecules), such as proteins and nucleic acids, cannot cross the plasma membrane. |
| presence of net charge (+ or −) | Gases, such as $CO_2$ and $O_2$, and small uncharged molecules, such as urea and ethanol, can cross the plasma membrane. In contrast, mineral ions such as $Na^+$, $K^+$, $Cl^-$ cannot cross because they are repelled by the hydrophobic lipid component of the plasma membrane. |
| solubility in lipid solvents | Lipophilic molecules can cross easily, but hydrophilic molecules, such as glucose, cannot cross because they are repelled by the hydrophobic lipid component of the plasma membrane. |
| direction of concentration gradient | Movement down a gradient (from a region of higher to a region of lower concentration of a substance) does not require an input of energy and can occur by diffusion. Movement against a gradient cannot occur by diffusion. |

From this table, it may be seen that the smaller the molecule and the more lipophilic (lipid-loving) it is, the more easily it can diffuse across the plasma membrane down its concentration gradient. This is summarised in figure 1.16.

It is apparent that many substances are prevented from crossing the plasma membrane because they are repelled by the hydrophobic lipid component of the plasma membrane and/or because their concentration gradients are in the wrong direction. However, at any time, many substances are entering or leaving cells. These include substances that *cannot* cross the hydrophobic lipid bilayer of the plasma membrane, such as glucose, sodium ions and calcium ions. Yet movement of these substances across the plasma membrane is essential for life. **So, it may be concluded that simple diffusion down a concentration gradient across the phospholipid bilayer *cannot* be the only means by which**

**substances cross the plasma membrane. Instead, additional means of entry to and exit from cells must exist**. These additional means of transport for dissolved substances, such as facilitated diffusion and active transport, involve the proteins of the plasma membrane (see the following section).

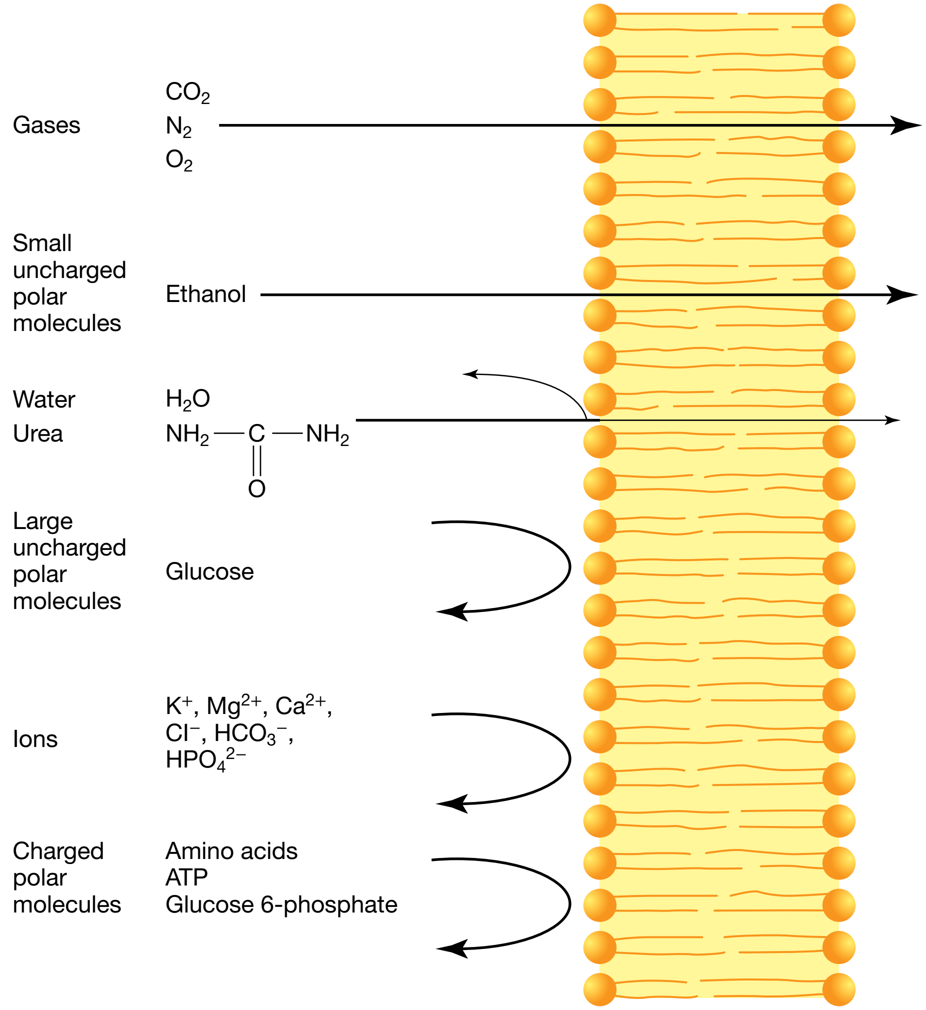



**FIGURE 1.16** Diagram showing the semipermeable nature of a phospholipid bilayer membrane. The membrane is fully permeable to some substances, partially permeable to others and impermeable to yet other substances. Note that the term 'polar' refers to molecules with an unequal distribution of electrons such that one side of a molecule has more electrons (and so is more negative) than the other side that has fewer electrons (and so is more positive). This is different from ions that have lost or gained an electron and so have a net electric charge.

## Various ways of crossing the boundary

Movement of substances across the plasma membrane into or out of cells can occur by several mechanisms:

1. **Simple diffusion** is the means of transport of small lipophilic substances. Water can also move across the plasma membrane by diffusion; this is a special case of diffusion known as **osmosis**.
2. **Facilitated diffusion** involves protein transporters and is the means of transport of dissolved hydrophilic substances down their concentration gradients.
3. **Active transport** involves protein transporters known as pumps and is the means of transport of dissolved hydrophilic substances against their concentration gradients.
4. **Endocytosis/exocytosis** are the means of bulk transport of macromolecules and liquids.

studyon

Unit 3 | AOS 1 | Topic 2 | Concept 4

**Movement across the membrane: simple diffusion**
Summary screen and practice questions

## Simple diffusion

Simple diffusion is the movement of substances across the phospholipid bilayer from a region of higher concentration to one of lower concentration of that substance; that is, *down* its concentration gradient (see figure 1.17a).

Movement down a concentration gradient by simple diffusion does *not* require any input of energy. It is the gradient that drives the diffusion (like letting a ball roll down a slope). The end point of simple diffusion is reached when equal concentrations of the substance are reached on both sides of the plasma membrane.

Substances that move easily across the plasma membrane by simple diffusion are small lipophilic molecules that can dissolve in the lipid bilayer. Among these substances are steroid hormones, alcohol and lipophilic drugs.

Figure 1.17b shows the stages in simple diffusion of a dissolved substance (X) across a plasma membrane. Its molecules are in constant random motion, some colliding with the membrane. If the concentration of substance X outside the cell is greater than that inside the cell, more movement of X into the cell will occur compared with movement in the opposite direction. This will produce a net movement of substance X into the cell. Net movement stops when collisions on both sides of the membrane equalise. This occurs when the concentrations of X on both sides of the membrane are equal.

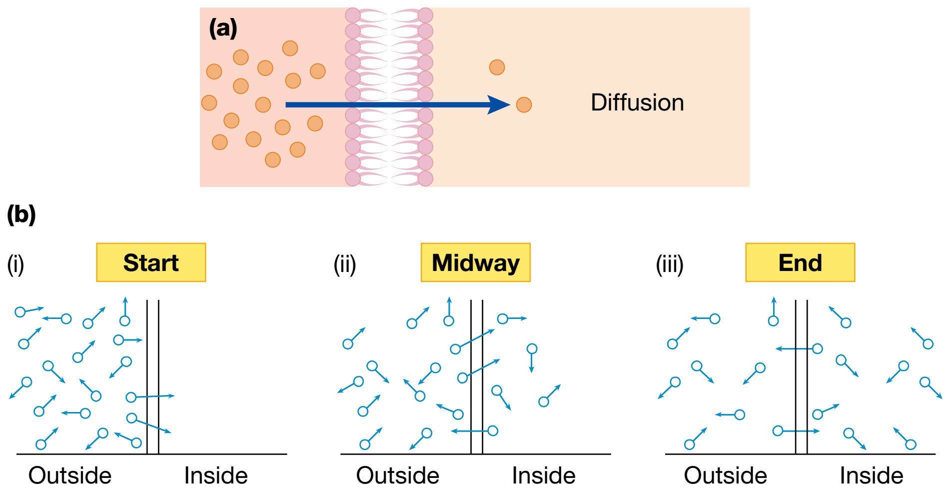



**FIGURE 1.17** **(a)** Simple diffusion involves the movement of substances through the phospholipid bilayer of the plasma membrane. The direction of movement is down the concentration gradient of the diffusing substance (from high to low concentration). Does this process require an input of energy? **(b)** Stages of simple diffusion: (i) At the start, substance X starts to move into the cell because of random movement that results in some collisions with the membrane. (ii) Midway, molecules of substance X are moving both into and out of the cell, but the net movement is from outside to inside. (iii) When the concentration of X is equal on each side of the membrane, the number of collisions on either side of the membrane is equal and the net movement of molecules of substance X stops. Does this mean that collisions of molecules of substance X with the membrane stop?

## Osmosis: a special case of diffusion

Osmosis is a special case of diffusion that relates to the movement of solvents and, in biological systems, that solvent is water.

Osmosis can be defined as the net movement of water across a semipermeable membrane from a solution of lesser solute concentration to one of greater solute concentration. A condition for osmosis is that the membrane must be permeable to water but not to the solute molecules. Solutions that have a high concentration of dissolved solute have a lower concentration of water, and vice versa. **The net movement of water molecules in osmosis from a solution of high water concentration to one of lower concentration is known as osmotic flow.**

solute = substance that is dissolved

solvent = liquid in which a solute dissolves

solution = liquid mixture of the solute in the solvent

When an external solution is compared with the dissolved contents of a cell, the external solution may be found to be either:

- **hypotonic** — having a lower solute concentration than the cell contents
- **isotonic** — having an equal solute concentration to that of the cells
- **hypertonic** — having a higher solute concentration than the cell contents.

Osmosis can be seen in action when cells are immersed in watery solutions containing different concentrations of a solute that cannot cross the plasma membrane. Remember that as the concentration of solute molecules increases, the concentration of the water molecules decreases.

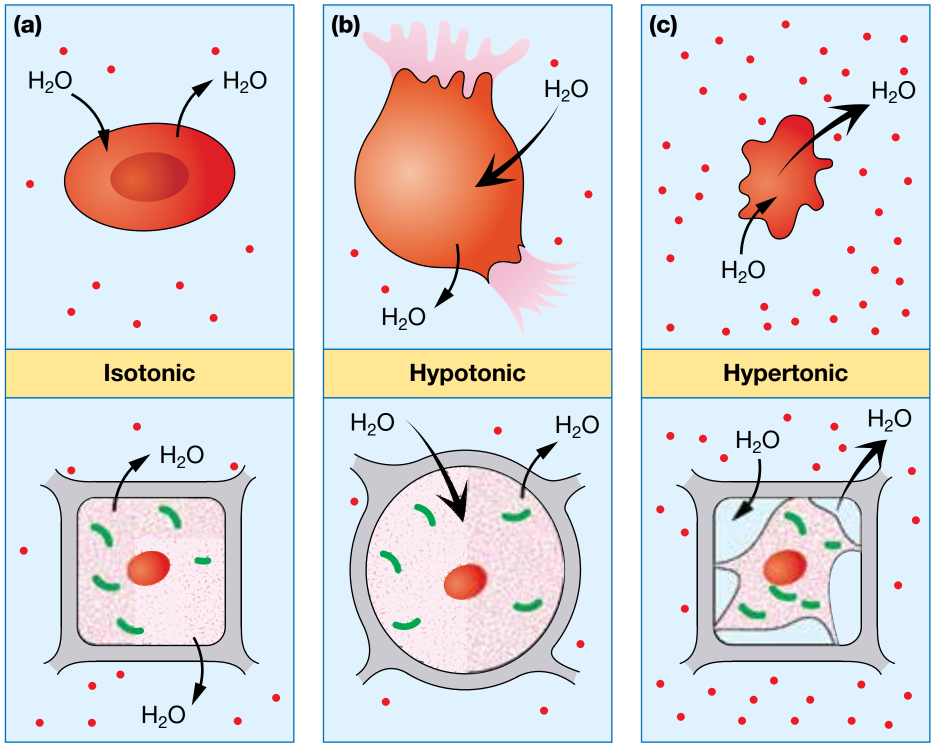



**FIGURE 1.18** Osmosis in action: behaviour of animal cells (above) and plant cells (below) in solutions of different concentrations of a dissolved substance (solute molecules). Solute molecules are denoted by red dots. Note the presence of a cell wall outside the plasma membrane in the plant cells. The solute molecules are too large to cross the plasma membrane, but water molecules will move down their concentration gradient in a process called osmotic flow. Does the movement of water by osmosis require an input of energy?

Look at figure 1.18a. The water molecules of the external isotonic solution are at the same concentrations as those in the cell contents. Since there is no concentration gradient, no net uptake of water molecules occurs in either cell; in a given period, the same number of water molecules will diffuse into the cell as will diffuse out.

Now look at figure 1.18b. The water molecules of the external hypotonic solution are more concentrated than those of the cell contents. Water molecules will diffuse down their concentration gradient from the hypotonic solution into the cell, resulting in a net uptake of water by the cell. As the red blood cell takes up the water molecules, it continues to swell until its plasma membrane bursts, dispersing the cell contents. The plant cell also takes up water, swells until it becomes rigidly swollen (turgid), but the cell does not burst because of the thick cell wall that lies outside the plasma membrane. The cell wall acts as a pressure vessel preventing the plasma membrane from swelling to a point of bursting. Net entry of water molecules into the plant cell finally stops as a result of the increasing outward pressure of the cell contents that opposes the net inward flow of water.

Finally look at figure 1.18c. The water molecules of the external hypertonic solution are at a lower concentration than those in the cell contents. Water molecules will diffuse down their concentration gradient from the cells into the external solution, resulting in a net loss of water from the cells. The red blood cell shrinks, becoming crenated. The plant cell within its plasma membrane shrinks away from its cell wall.

**ODD FACT**

Each day about 2 litres of fluid are taken into the gut in food and drinks. This is increased by a further volume of about 7 litres from secretions, including those of the salivary glands, stomach, liver and pancreas.

### The dehydrated person

A person suffering from a prolonged bout of diarrhoea is severely dehydrated and may need to be admitted to hospital. Normally, the cells lining the small and large intestine absorb the large volume of fluid and dissolved salts that enter the gut daily (see Odd fact). Several bacterial infections, including *Staphylococcus* sp., can inhibit this absorption. As a result, most of the fluid and the dissolved salts pass into the large intestine and are expelled from the body in large volumes of watery diarrhoea, resulting in dehydration and salt loss. Worsening dehydration is serious because it produces a decrease in blood volume that, if untreated, may lead to cardiovascular failure.

Initial treatment of severe dehydration is replacement of the lost fluid and salts. The rehydration therapy involves an isotonic solution of saline (salt) solution plus glucose. This solution can be administered either orally (by mouth) or by direct infusion into the bloodstream (by intravenous infusion, or IV).

Why does this treatment act as a rehydration therapy? Glucose stimulates the absorption of sodium by the gut cells. The uptake of sodium and glucose from the gut into the extracellular fluid creates a hypertonic internal environment that causes water to move by osmosis from the gut across the cells lining the gut, and so returns water to the extracellular fluid and from there to the body cells.

## Facilitated diffusion

studyon

Unit 3 | AOS 1 | Topic 2 | Concept 5

**Movement across the membrane: facilitated diffusion**
Summary screen and practice questions

Facilitated diffusion is an example of protein-mediated transport. Facilitated diffusion is so named because the *diffusion* across the membrane is enabled or *facilitated* by special protein transporters in the plasma membrane.

Like simple diffusion, facilitated diffusion does *not* require an input of energy. Like simple diffusion, facilitated diffusion moves substances down their concentration gradients. **However, facilitated diffusion of dissolved substances requires the action of protein transporters that are embedded in the cell membrane.**

**Facilitated diffusion enables molecules that cannot diffuse across the phospholipid bilayer to move across the plasma membrane through the agency of transporter proteins.** These transporters are either **channel proteins** or **carrier proteins**. Are transporters required in simple diffusion? (Refer to figure 1.19.)

In the case of ions, which carry either + or – charge(s), it is more correct to say that, rather than moving down their concentration gradients, they move down their electrochemical gradients.

### Channel proteins

One group of transport proteins involved in facilitated diffusion are the channel proteins. Each channel protein is trans-membrane and has a central water-filled pore through which dissolved substances can pass down their concentration gradient. Different channel proteins are specific for the diffusion of charged particles or polar molecules.

Each channel protein consists of a narrow water-filled pore in the plasma membrane (see figure 1.19b). By providing water-filled pores, channel proteins create a hydrophilic passage across the plasma membrane that bypasses the phospholipid bilayer and facilitates the diffusion of charged particles, such as sodium and potassium ions, and small polar molecules.

**ODD FACT**

Channel proteins known as **aquaporins** are trans-membrane proteins that are specific for the facilitated diffusion of water molecules.

### Carrier proteins

Another group of membrane proteins involved in facilitated diffusion are carrier proteins. Carrier proteins are specific, with each kind of carrier enabling the diffusion of one kind of molecule across the plasma membrane. **After binding to its specific cargo molecule, the carrier protein undergoes a change in shape as it delivers its cargo to the other side of the plasma membrane** (see figure 1.19c).

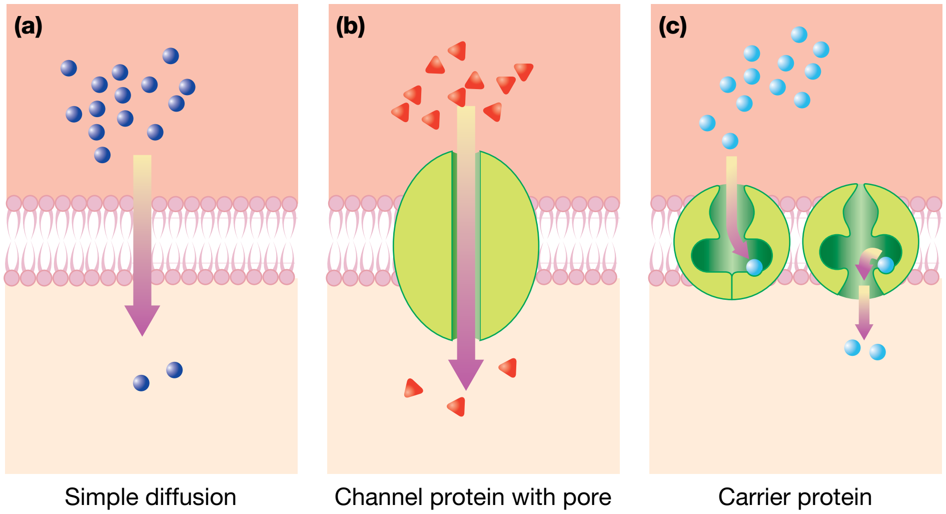



**FIGURE 1.19** **(a)** Simple diffusion. In contrast, facilitated diffusion requires either **(b)** channel proteins or **(c)** carrier proteins to enable certain dissolved substances to diffuse down their concentration gradients. Note the change in shape of the carrier protein as it operates.

**eBook*plus***

**Weblink**
Diffusion

Carrier proteins are important in the facilitated diffusion of hydrophilic uncharged substances, such as glucose and amino acids. In the absence of carrier proteins, these hydrophilic molecules cannot cross the plasma membrane directly.

### Rates of diffusion

In the previous sections, two types of diffusion of dissolved substances across plasma membranes have been explored: simple diffusion and facilitated diffusion. Do these two types of diffusion differ in the rate at which substances can move down their concentration gradients in either direction across a plasma membrane?

In simple diffusion, the rate at which substances move across the plasma membrane by simple diffusion is determined by their concentration gradient, that is, the difference in the concentration of the substance inside and outside the cell. **The higher its concentration gradient, the faster a substance will move by simple diffusion down this gradient and across a plasma membrane.** (Think of this as like rolling a ball downhill — the steeper the incline, the faster the ball rolls.)

In facilitated diffusion, the rate of movement of a substance is also influenced by the steepness of its concentration gradient on either side of the plasma membrane. The steeper the concentration gradient, the faster the rate of facilitated diffusion, *but only up to a point* (see figure 1.20).

Why the difference between simple diffusion and facilitated diffusion? Facilitated diffusion requires the involvement of transporters, either channel proteins or carrier proteins, for the movement of substances. These

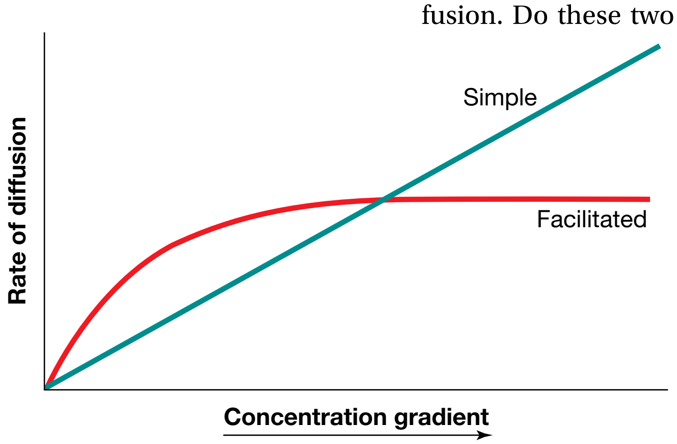



**FIGURE 1.20** Graph showing rates of simple and facilitated diffusion with increasing concentration gradient of the diffusing substance. Note that the rate of simple diffusion continues to increase as the concentration gradient increases. In contrast, the rate of facilitated diffusion is initially linear, but begins to taper off and finally reaches a plateau. The maximum rate is reached when all the transporters are fully occupied.

**studyon**

Unit 3 | AOS 1 | Topic 2 | Concept 6

**Movement across the membrane: active transport**
Summary screen and practice questions

channel and carrier proteins are present in limited numbers on the plasma membrane. Because their numbers are limited, eventually a concentration of the diffusing substance will be reached when all the transporters are saturated (fully occupied). So, as the concentration gradient increases, the rate of facilitated diffusion of a substance will at first increase, then become slower, and finally will reach a plateau. The plateau is the maximum rate of facilitated diffusion. When this is reached, all the transporter molecules are fully occupied.

## Active transport

Active transport is the process of moving substances across the plasma membrane against the direction that they would travel by diffusion; that is, **active transport moves dissolved substances from a region of low concentration to a region of high concentration of those substances**. **Active transport can occur only with an input of energy**, and the energy source is typically adenosine triphosphate (ATP).

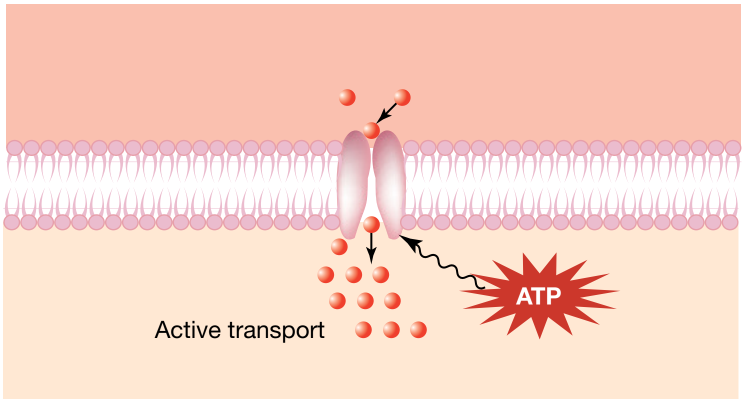



**FIGURE 1.21** Diagram showing a simplified representation of active transport. A pump is a trans-membrane protein that is both a carrier and an ATPase enzyme. The enzyme component of the pump catalyses the energy-releasing reaction that powers active transport.

Special transport proteins embedded across the plasma membrane carry out the process of active transport. The proteins involved are called **pumps** and each different pump transports one (or sometimes two) specific substance(s). Important pumps are proteins with both a transport function and an enzyme function. The enzyme part of the pump catalyses an energy-releasing reaction:

$$\text{ATP} \rightarrow \text{ADP} + \text{P}_i + \textit{energy}$$

The transport part of the pump uses this energy to move small polar molecules and ions across the plasma membrane *against* their concentration gradients. During this process, the protein of the pump undergoes a shape change (see figure 1.21).

Cells use pumps to move materials that they need by active transport. Active transport is essential for the key function of cells including pH balance, regulation of cell volume and uptake of needed nutrients. Examples of active transport include:

- uptake of dissolved mineral ions from water in the soil by plant root hair cells against their concentration gradient
- production of acidic secretions (pH of nearly 1) by stomach cells that have a low internal concentration of hydrogen ions (H+) but produce secretions (gastric juice) with an extremely high concentration of hydrogen ions
- uptake of glucose from the small intestine into the cells lining the intestine against its concentration gradient, using the glucose–sodium pump
- maintenance of the difference in the concentrations of sodium and potassium ions that exist inside and outside cells (see table 1.4) by action of the sodium–potassium pump, which actively transports these ions against their concentration gradients (see below).

Some pumps actively transport a single dissolved substance against its concentration gradient. Other pumps transport two substances simultaneously, for example, the **sodium–potassium pump**. The importance of the sodium–potassium pump is highlighted by the fact that, in the human body overall, about 25% of the body's ATP is expended in keeping the sodium–potassium pump operating. For brain cells, the figure is even higher, about 70%.

Why is this pump needed? Table 1.4 provides a clue to the answer.

**ODD FACT**

Compared with the concentration of $H^+$ ions in the contents of stomach cells, the gastric juice in the stomach has about a concentration three million times higher. This is achieved by active transport of these ions out of the stomach cells against their concentration gradient.

**TABLE 1.4** Approximate concentrations of sodium and potassium ions in the cytosol of cells and in the surrounding extracellular fluid

| Ion | Inside cell | Outside cell |
|---|---|---|
| sodium ($Na^+$) | 10 mM | 142 mM |
| potassium ($K^+$) | 150 mM | 5 mM |

**ODD FACT**

The Danish scientist Jens Skou was awarded the Nobel Prize in chemistry for his discovery of the ion-transporting enzyme $Na^+$, $K^+$ ATPase, otherwise known as the sodium–potassium pump.

Table 1.4 shows the concentrations of sodium and potassium ions inside and outside cells. Note that these concentrations differ greatly. In what direction would sodium ions tend to flow passively? What about potassium ions? How are these concentration differences maintained? The different concentrations are maintained by the sodium–potassium pump, which is present in all animal cells and constantly expends energy in active transport, pushing sodium ions out of the cell and pulling potassium ions in. For each ATP molecule expended, the pump pushes three sodium ions out and drags two potassium ions in. This process of active transport compensates for the constant passive diffusion of sodium ions into cells and of potassium ions out of cells, both down their concentration gradients.

The sodium–potassium pump plays a key role in excitable cells, such as nerve cells and muscle cells. During the transmission of a nerve impulse, sodium ion channels open and sodium ions rapidly flood into the nerve cell by facilitated diffusion. After the nerve impulse has passed, the sodium channels close and the sodium–potassium pump then restores the concentrations of sodium and potassium ions to their resting levels by actively pushing sodium ions across the membrane out of the cell and dragging potassium ions into the cell (refer back to table 1.4). Restoring these concentrations involves active transport against the concentration gradients of these ions.

**studyon**

Unit 3
AOS 1
Topic 2
Concept 6

**Do more**
Movement across membranes

## When transport goes wrong . . .

The importance of transport proteins in moving substances becomes apparent if they do not operate as expected, as may be seen in cystic fibrosis and in cholera infections.

### Cystic fibrosis

**ODD FACT**

Before accurate genetic testing for cystic fibrosis became available, an early test was the so-called 'sweat test' that measured salt levels in a baby's sweat.

Symptoms seen in people affected by the inherited disorder cystic fibrosis result from a defect in one transporter protein on the plasma membrane of cells. This protein is a channel that normally allows chloride ions ($Cl^-$) to move out of cells. Cystic fibrosis affects various organs, including the lungs, the pancreas and the skin. A faulty chloride ion channel protein, such as occurs in cystic fibrosis, blocks the movement of chloride ions. This affects the various organs as follows:

- *Lungs*. Normally, the inner surfaces of a person's lungs are covered with a thin layer of mucus. This mucus is important in a healthy lung because it traps microbes and other particles, and it is constantly removed from the lungs by the beating action of hair-like projections (cilia) that line the airways. Cough or clear your throat and that mucus and those trapped particles are gone. In the lungs, the chloride ion channel normally moves chloride ions out of lung cells.

  In cystic fibrosis, however, the defect in the transporter stops
  - i the movement of $Cl^-$ ions out of the cells into the lung cavity
  - ii the consequential flow of sodium ($Na^+$) ions that move in response to the electrochemical gradient created by the movement of the negative chloride ions into the lung cavity
  - iii osmotic flow of water into the lungs that normally thins the mucus.

**FIGURE 1.22** Capsules and tablets containing the missing pancreatic enzymes (lipase, protease, amylase) are taken by people affected by cystic fibrosis. Note the enzyme granules in the capsules. (Image courtesy of Marjory Martin.)

These effects of the faulty chloride ion channel mean that the mucus in the lungs is abnormally thick, sticky and difficult to move and, rather than being easily cleared, the mucus remains in the lungs affecting breathing and acting as a potential source of infection.

- *Sweat glands in skin*. In the skin, the chloride ion channel is normally involved in reabsorbing salt (NaCl) from fluid within cells of the sweat glands before it is released as sweat. When the chloride ion ($Cl^-$) transporter is blocked, reabsorption does not occur and very salty sweat is produced.
- *Pancreas*. Pancreatic enzymes are normally involved in the digestion of some foods. In cystic fibrosis, these enzymes are unable to enter the gut because abnormally thick mucus blocks the narrow duct that connects the pancreas to the small intestine. Without these enzymes, food cannot be fully digested. However, replacement enzymes in the form of tablets, powders or capsules can replace the enzymes normally released by the pancreas (see figure 1.22).

## Cholera

The pores of some channel proteins are permanently open, while the pores in other channel proteins open only in response to a specific signal. The chloride ion channel on the plasma membranes of cells lining the intestine is *not* always open. Normally chloride ions are retained within the cells lining the intestine. Only when this channel is opened can chloride ions move from the cells into the cavity of the intestine.

Cholera results from a bacterial infection of *Vibrio cholera*. A toxin produced by these bacteria causes the chloride ion channels in the cells lining the intestine to be locked in the 'open' position. This results in a flood of chloride ions into the intestinal space that is followed by a flow of sodium ions (down the resulting electrochemical gradient that is created). In turn, the increased concentration of salt in the gut creates a hyperosmotic environment that draws water into the gut by osmosis. The continuous secretion of water into the intestine causes the production of large volumes of watery diarrhoea. If left untreated the water loss caused by this diarrhoea can be fatal within hours. Cholera epidemics have caused many deaths (see figure 1.23). Cholera can be spread by water that is contaminated by contact with untreated sewage or by the faeces of an infected person. Cholera can be spread by food that is washed in or mixed with water contaminated by cholera bacteria, or by food that is inappropriately handled by a person infected with cholera.


**FIGURE 1.23** Cholera epidemics and pandemics have taken many lives. During the period 1899 to 1923, a cholera pandemic that began in India spread across the globe and reached as far as Russia and Eastern Europe. This French publication from 1912 illustrates the public fear and the widespread deaths that this cholera pandemic caused.

## Bulk transport of solids and liquids

To this point, we have been concerned with movement of dissolved substances across the plasma membrane. In addition, small solid particles and liquids in bulk can be moved across the plasma membrane into or out of cells. Figure 1.24 gives a summary of how bulk material can enter cells.

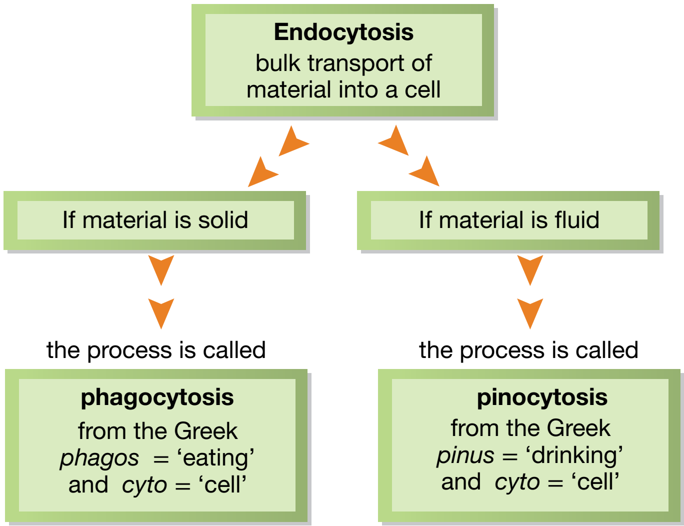
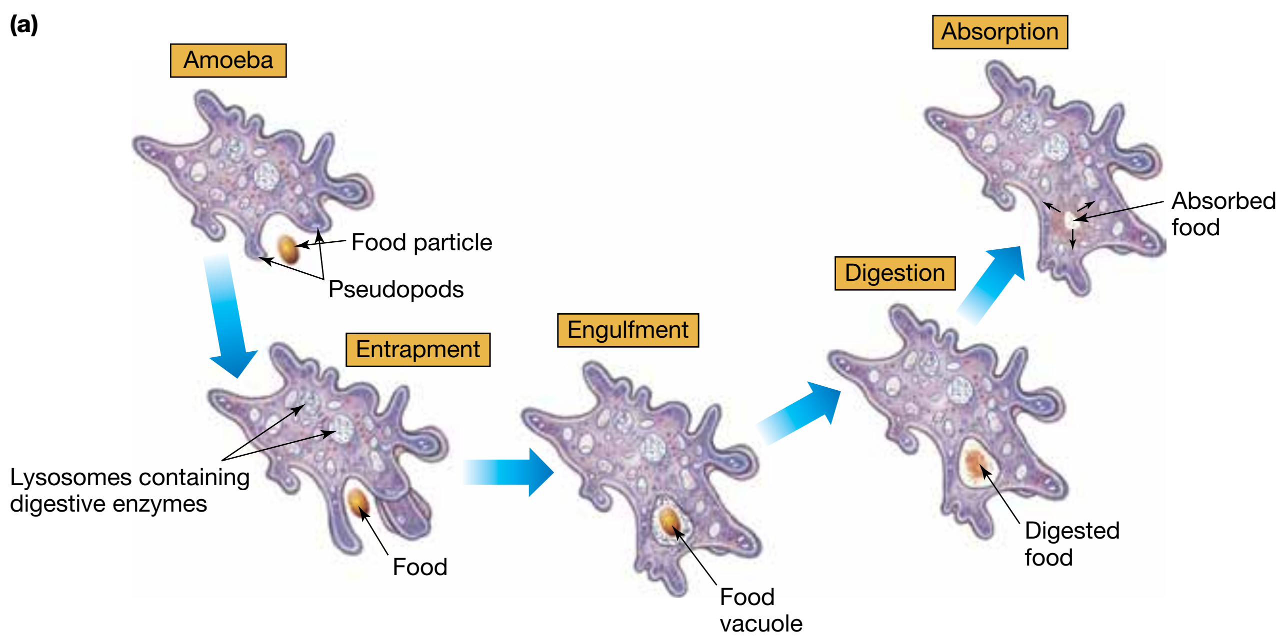



**FIGURE 1.24** Endocytosis — a summary

### Endocytosis: getting in

Solid particles can be taken into a cell. For example, several cells of the immune system can engulf disease-causing bacteria cells, enclosing them within lysosomal sacs where they are destroyed. Unicellular protists, such as *Amoeba* and *Paramecium*, obtain their energy for living in the form of relatively large 'food' particles, which they engulf and enclose within a sac where the food is digested (see figure 1.25a).



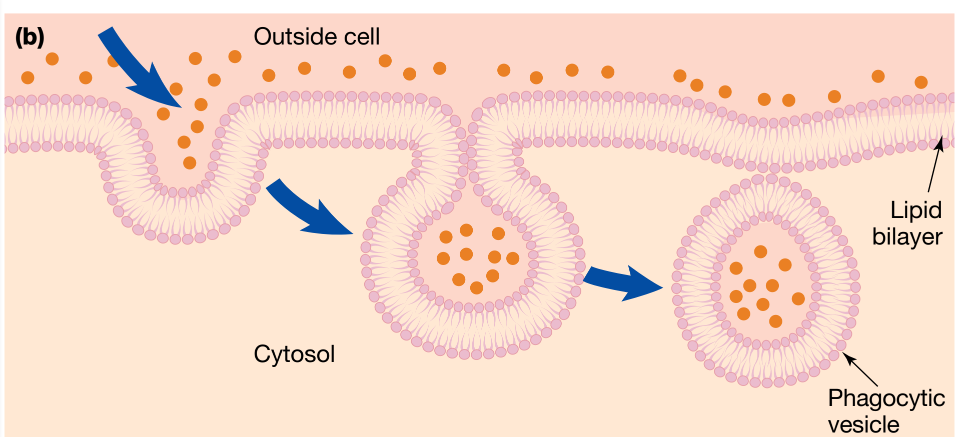




**FIGURE 1.25** **(a)** Transport of a solid food particle across the membrane of an amoeba **(b)** Endocytosis occurs when part of the plasma membrane forms around food particles to form a phagocytic vesicle (or phagosome). This vesicle then moves into the cytosol where it fuses with a lysosome, a bag of digestive enzymes. The same process is also part of the body's immune defence against infectious microbes.

Note how part of the plasma membrane encloses the material to be transported and then pinches off to form a membranous **vesicle** that moves into the cytosol (figure 1.25b). This **process of bulk transport of material into a cell** is called endocytosis. When the material being transported is a solid food particle, the type of endocytosis is called **phagocytosis**.

Although some cells are capable of phagocytosis, most cells are not. Most eukaryotic cells rely on **pinocytosis**, **a form of endocytosis that involves material that is in solution** being transported into cells. The process of endocytosis is an energy-requiring process and requires an input of ATP.

**study on**

| | |
|---|---|
| Unit 3 | **Bulk transport across membranes: endocytosis** |
| AOS 1 | Summary screen and practice questions |
| Topic 2 | |
| Concept 10 | |

## Exocytosis: getting out

Bulk transport out of cells (such as the export of material from the Golgi complex, discussed in chapter 2) is called exocytosis. **In exocytosis, vesicles formed within a cell fuse with the plasma membrane** before the contents of the vesicles are released from the cell (see figure 1.26). If the released material is a product of the cell (e.g. the contents of a Golgi vesicle), then 'secreted from the cell' is the phrase generally used. If the released material is a waste product after digestion of some matter taken into the cell, 'voided from the cell' is generally more appropriate. The process of exocytosis requires an input of energy in the form of ATP.

**study on**

| | |
|---|---|
| Unit 3 | **Bulk transport across membranes: exocytosis** |
| AOS 1 | Summary screen and practice questions |
| Topic 2 | |
| Concept 9 | |

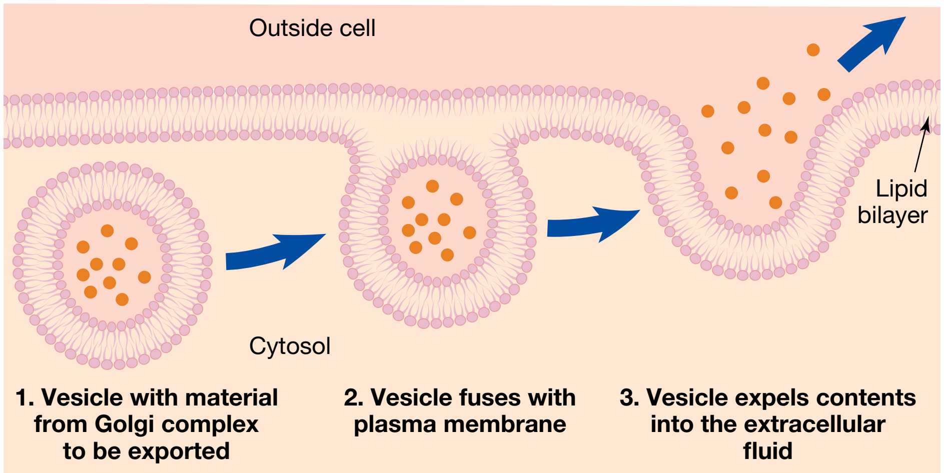



**FIGURE 1.26** Exocytosis (bulk transport out of cells) occurs when vesicles within the cytosol fuse with the plasma membrane and vesicle contents are released from the cell.

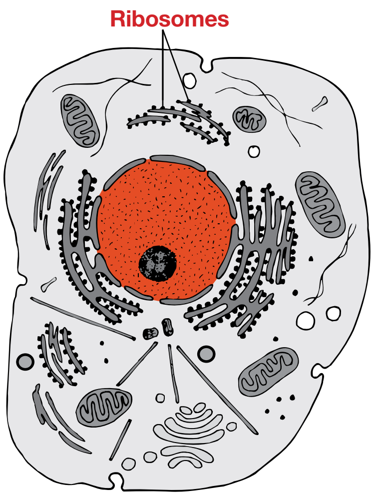
## KEY IDEAS

- Simple diffusion moves dissolved substances across the plasma membrane down their concentration gradient and requires no input of energy.
- Osmosis is a special case of diffusion, being the movement of water across the plasma membrane down its concentration gradient.
- Facilitated diffusion moves dissolved substances across the plasma membrane down their concentration gradients, but this movement occurs through involvement of transport proteins, either channel or carrier proteins, and requires no input of energy.
- Active transport moves dissolved substances across the plasma membrane *against* the concentration gradient, a process that can occur only via the action of protein pumps.
- Active transport requires an input of energy that commonly comes from ATP, catalysed by the ATPase enzyme that is part of some protein pumps.
- Endocytosis is the bulk transport of material into cells; if solids are being moved, the process is termed phagocytosis and, if liquids, the process is termed pinocytosis.
- Exocytosis is the bulk movement of materials via secretory vesicles out of cells.

## QUICK CHECK

**17** What is the process by which bulk materials are exported out of cells?
**18** Consider passive diffusion and facilitated diffusion:
  **a** Identify one difference between these processes.
  **b** Identify one similarity that they share.
**19** Identify one difference between diffusion and active transport.
**20** Which transport process relies on the involvement of either a carrier or a channel protein?
**21** By which process do cells of the stomach lining manage to move hydrogen ions out of the cells to produce a highly acidic gastric secretion?
**22** What process is involved in the movement of water down its concentration gradient and across a layer of cells from outside the body to inside?

# The role of different organelles in the export of protein from the cell




## Ribosomes: protein factories

Cells make a range of proteins for many purposes: for example, manufacture of haemoglobin, an oxygen-transporting protein by developing human red blood cells in the bone marrow; manufacture of the contractile proteins actin and myosin by the muscle cells; and manufacture of the hormone insulin and digestive enzymes including lipases by different cells of the pancreas. Insulin is an example of a protein that is exported from the cell in which it was manufactured.

**Ribosomes** are the site in the cell where proteins are made. It is on the ribosomes that amino acids are assembled and joined into polypeptide chains or proteins. The diameter of a ribosome is only about 0.03 μm. Because of their very small size, ribosomes can be seen only through an electron microscope (see figure 1.27a). However, ribosomes are present in very large numbers in a cell. Ribosomes in animal and plant cells are composed of about two-thirds RNA and one-third protein.

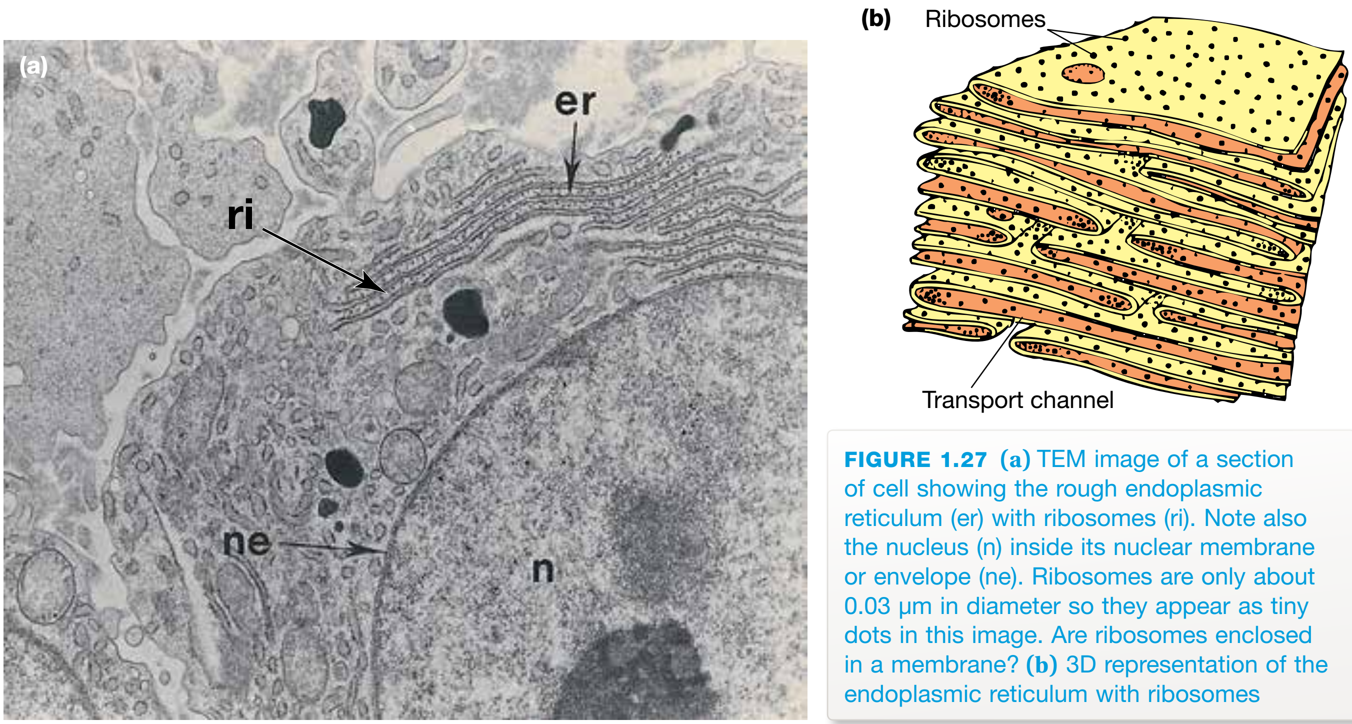
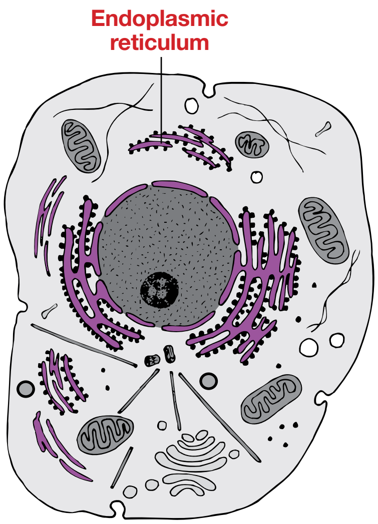



**FIGURE 1.27** **(a)** TEM image of a section of cell showing the rough endoplasmic reticulum (er) with ribosomes (ri). Note also the nucleus (n) inside its nuclear membrane or envelope (ne). Ribosomes are only about 0.03 μm in diameter so they appear as tiny dots in this image. Are ribosomes enclosed in a membrane? **(b)** 3D representation of the endoplasmic reticulum with ribosomes

**ODD FACT**

Ribosomes can join amino acids into a protein chain at the rate of about 200 per minute.




**ODD FACT**

An estimated 13 million ribosomes are attached to the rough ER in a typical human liver cell.

Within a cell, many ribosomes are attached to membranous channels known as the endoplasmic reticulum (see next section). Other ribosomes are found free in the cytosol. Proteins made by 'free' ribosomes are for local use within the cell.

## Rough endoplasmic reticulum

The **endoplasmic reticulum (ER)** is an interconnected system of membrane-enclosed flattened channels. Figure 1.28 shows part of the channels of the endoplasmic reticulum in a eukaryotic cell. Refer also to figure 1.27a. Where the ER has ribosomes attached to the outer surface of its channels, it is known as **rough endoplasmic reticulum**. Proteins produced by ribosomes on the endoplasmic reticulum are generally exported from the cell.

**FIGURE 1.28** False coloured scanning electron micrograph of part of a eukaryotic cell showing the channels of rough ER (pink). Note the many tiny bead-like structures attached to the outside of the ER channels. What are these structures?

studyon

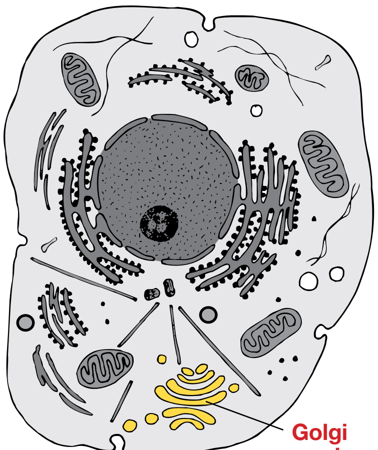
Unit 3
AOS 1
Topic 2
Concept 7

**Role of endoplasmic reticulum in transport**
Summary screen and practice questions

### Rough ER

Through its network of channels, the rough ER is involved in transporting some of the proteins to various sites within a cell. Proteins delivered from the ribosomes into the channels of the rough ER are also processed before they are transported. The processing of proteins within the rough ER includes:

- attaching sugar groups to some proteins to form glycoproteins
- folding proteins into their correct functional shape or conformation
- assembling complex proteins by linking together several polypeptide chains, such as the four polypeptide chains that comprise the haemoglobin protein.




## Golgi complex: the exporter

Some cells produce proteins that are intended for use outside the cells where they are formed. Examples include the following proteins that are produced by one kind of cell and then exported (secreted) by those cells for use elsewhere in the body:

- the digestive enzyme pepsin, produced by cells lining the stomach and secreted into the stomach cavity
- the protein hormone insulin, produced by cells in the pancreas and secreted into the bloodstream
- protein antibodies, produced in special lymphocytes and secreted at an area of infection.

How do these substances get exported from cells? The cell organelle responsible for the export of substances out of cells is the **Golgi complex**, also known as the Golgi apparatus. The Golgi complex has a multi-layered structure composed of stacks of membrane-lined channels (see figure 1.29).

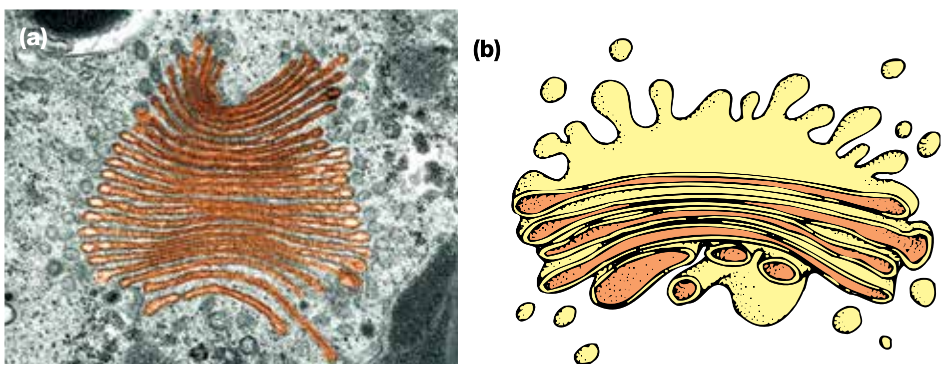



**FIGURE 1.29** **(a)** False coloured TEM image of the Golgi complex (orange). Note the stacks of flattened membrane-lined channels with their wider ends that can break free as separate vesicles. **(b)** 3D representation of the Golgi complex. Note the vesicles breaking off from the ends of the Golgi complex membranes. What is their role?

**ODD FACT**

The Golgi complex is named after Camillo Golgi, who, in 1898, first identified this cell organelle.

Proteins from the rough ER that are intended for export must be transferred to the Golgi complex. Figure 1.30 outlines the pathway followed. Because there is no direct connection between the membranes of the ER and the Golgi complex, the proteins are shuttled to the Golgi complex in membrane-bound transition vesicles. The vesicles are taken into the Golgi complex where the proteins are concentrated and packaged into secretory vesicles that break free from the Golgi complex. The secretory vesicles move to the plasma membrane of the cell where they merge with it, discharging their protein contents.

studyon

Unit 3
AOS 1
Topic 2
Concept 8

**Role of Golgi apparatus in packaging**
Summary screen and practice questions

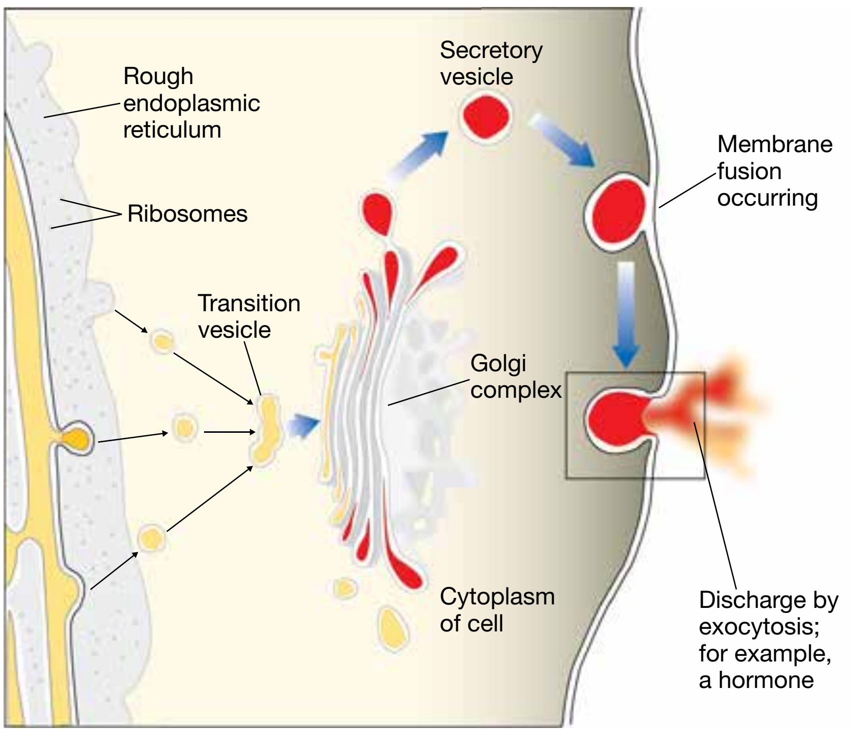



**FIGURE 1.30** The secretory export pathway for proteins. Packages of protein are transferred from the rough ER in transition vesicles to the Golgi complex where they are taken in. From the Golgi complex, the secretory vesicles with their protein cargo move to the plasma membrane of the cell, merge with it and discharge their contents.

## KEY IDEAS

- **Ribosomes are cell organelles where proteins are manufactured.**
- **The endoplasmic reticulum (ER) is made of a series of membrane-bound channels.**
- **Rough ER is so named because of the presence of ribosomes on the external surface of its membranes.**
- **Rough ER is involved in processing of proteins and in their transport.**
- **The Golgi complex packages substances into vesicles for export from a cell.**

## QUICK CHECK

**23** Identify whether each of the following is true or false:
- **a** The RNA of the ribosomes is made in the nucleolus.
- **b** The folding of a protein into its functional 3D shape takes place on the ribosomes.
- **c** Ribosomes are membrane-bound organelles that form part of the cell cytoplasm.
- **d** The channels of the Golgi complex are connected to those of the ER.

**24** A scientist wishes to examine ribosomes in a liver cell.
- **a** Where should the scientist look: in the nucleus or the cytoplasm?
- **b** What kind of microscope is likely to be used by the scientist: a light microscope or a transmission electron microscope? Explain.

## BIOLOGY IN THE WORKPLACE

### Drew Berry — animator specialising in biomedical science

I am a 3D animator based at the Walter and Eliza Hall Institute of Medical Research in Melbourne. The goal of my animation is to visually explain scientific discoveries to the public on national news, science and current affairs programs. I also create animations for documentaries, museum exhibits and art gallery installations. A recent project involved creating visualisations of the nasty bugs smallpox, ebola and anthrax for a National Geographic documentary on bioterrorism.

I love what I do, yet when I finished school I intended to follow a very different career path. I wanted to be a marine biologist and study sharks. I was inspired by the documentaries of Jacques Cousteau and David Attenborough, and I loved the films by Australian couple Ron and Valerie Taylor. The most memorable for me was seeing Valerie Taylor put on a chain-mail diving suit and shove her arm into the gaping mouth of a hungry shark — that looked like fun!

At the University of Melbourne, I studied all the subjects for marine science. One subject was Cell Biology, a topic that didn't interest me. However, the professor who taught Cell Biology, Jeremy Pickett-Heaps, gave the most amazing and entertaining lectures. His specialty was filming living cells using time-lapse techniques and his passion for cell biology and extraordinary footage hooked everyone in his classes. I decided to do a BSc Honours year in Jeremy's lab and went on to begin a PhD on filming cells and conducting research into how cells create their shapes (morphogenesis).

A couple of years into my PhD, I realised that I loved science and found it fascinating but didn't want a career doing experiments at a lab bench. I wrote up my thesis and submitted it as a Masters degree instead. At the same time, an opportunity came up to work in an advertising company writing copy (text) for magazine ads. I wasn't very good at it and the company moved me onto Photoshop work. This was a time of high-pressure, relentless work but I was also gaining many new skills in design and visual communication.

I joined the Walter and Eliza Hall Institute as their 'Photoshop guy', preparing scientific images for publication. Because of my skills from advertising, I was fast and efficient with Photoshop and usually finished my work by morning tea. I began playing around with 3D software and started to create animations for education videos. Around that time, major discoveries were made at the institute about malaria and I created some animations that explained how the parasite infects red blood cells and causes disease. The malaria animation proved pivotal in my career as it has been popular with many TV programs and museum exhibitions. On the basis of its success, I was able to transform my job into working full time creating scientific animations.

For me it is the perfect job. I read journals and other scientific literature, discuss ideas with scientists and think about the concepts and discoveries at the cutting edge of science. Once I have a clear understanding and mental picture of the science, I access the raw data wherever possible and import it into my animation system. The next phase involves an enormous amount of problem solving, creative design and visual storytelling, which offers unlimited scope for exploring new ideas and techniques.

The animation software, Maya, is the same type used for blockbuster feature films. The fact that it allows a special-effects artist to create the amazing creatures in *Lord of the Rings* and a scientist to build realistic and accurate visualisations with the same set of tools is a credit to the flexibility and power of Maya.

Creating 3D animation is not for everyone. You must be confident with computers, able to troubleshoot the frequent technical problems and have the patience and perseverance to work through the design challenges. If it does appeal to you, there is an endless amount of science waiting to be explained to the public!

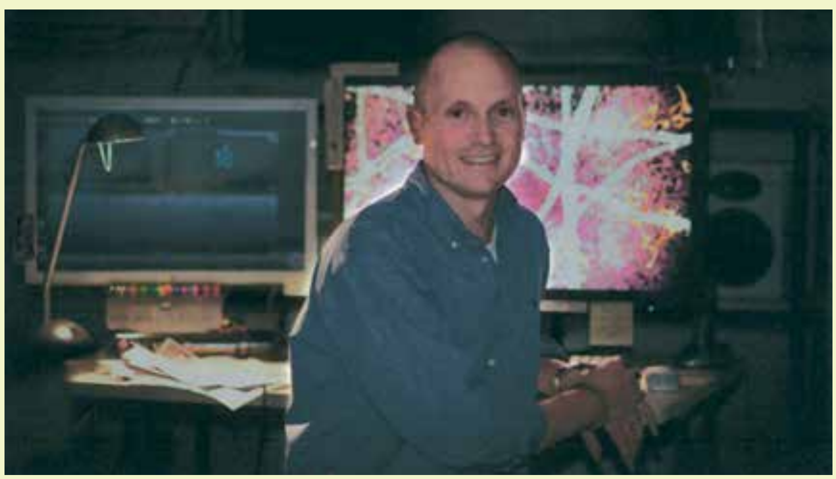


**FIGURE 1.31** Drew Berry at work on some new animations (Image courtesy of the Walter and Eliza Hall Institute.)

# BIOCHALLENGE

## Exploring the plasma membrane

1 The plasma membrane has been described as being like a 'train track'. This was because the first images of the plasma membrane showed it as two dark lines separated by a lighter region. Figure 1.32 shows part of the plasma membranes of two adjoining cells. The plasma membranes have been sectioned so that their surfaces are oriented horizontally at right angles into the plane of this page.

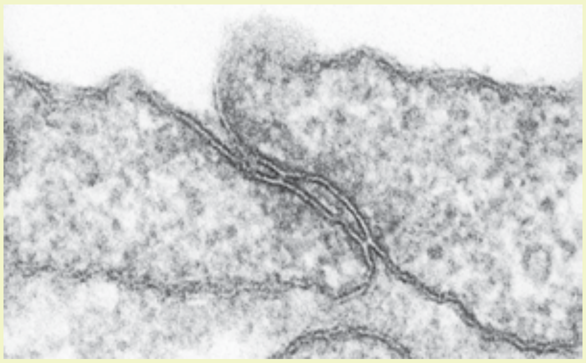


**FIGURE 1.32** Plasma membrane

- a How thick is the plasma membrane in nanometres? In micrometres?
- b What kind of microscope was needed to produce the image in figure 1.32?
- c What are the 'rails' of the train track composed of?
- d What is present in the space between the rails?

2 Key information about the nature of the plasma membrane came from an experiment carried out in 1925 by two Dutch scientists. They took a known number of red blood cells and, based on the average size of these cells, they estimated their combined surface area. Then they extracted only the lipid from the plasma membrane of these cells and allowed it to spread out on a water surface where it formed a monolayer or single layer of molecules. (Remember, lipids will not mix with water!) To their surprise, the scientists found that the area of the lipid monolayer on the water surface was twice the combined surface area of the red blood cells that were the source of the lipid.

Consider this finding and suggest what key information this result provided about the structure of the plasma membrane.

3 In 1970, Frye and Edidin carried out an experiment in which they took a human cell and a mouse cell and fused them to form a human–mouse hybrid cell. They showed the distribution of the surface proteins on the plasma membrane of each cell by using anti-human and anti-mouse antibodies labelled with a different fluorescent dye. A red dye showed the positions of the surface proteins on the membrane of the human cell. A green dye showed the positions of the surface proteins on the membrane of the mouse cell.

Figure 1.33a shows the initial observation immediately after the fusion of the two cells. After 40 minutes, the researchers carried out a second observation and their findings are shown in figure 1.33b.

From the results of this experiment, which of the following is it reasonable to conclude?

- a Surface proteins are fixed in position on the plasma membrane.
- b Surface proteins from each cell type have fused.
- c Surface proteins can move laterally across the plasma membrane.

4 True or false?

The results of this experiment provide support for the fluid mosaic model of membrane structure. Briefly explain.

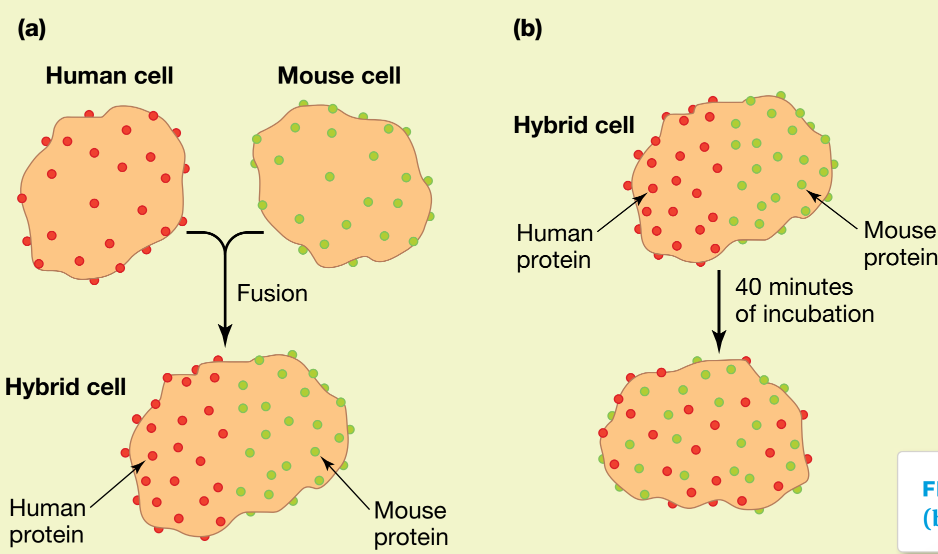



**FIGURE 1.33 (a)** Start of experiment **(b)** 40 minutes later

# Chapter review

studyon
Unit 3 | AOS 1 | Topic 2
The plasma membrane
Practice questions

## Key words

active transport
aquaporins
carrier proteins
cell membrane
cell surface markers
channel proteins
endocytosis
endoplasmic reticulum (ER)
eukaryotes
eukaryotic
exocytosis
facilitated diffusion
fluid mosaic model
glycoprotein
Golgi complex
hydrophilic
hydrophobic
hypertonic
hypotonic
integral proteins
isotonic
lysosome
nuclear envelope
osmosis
peripheral proteins
phagocytosis
phospholipids
pinocytosis
plasma membrane
prokaryotes
prokaryotic
proteins
pumps
receptors
ribosomes
rough endoplasmic reticulum
selectively permeable
semipermeable
simple diffusion
sodium–potassium pump
trans-membrane
vesicle

## Questions

**1** **Making connections** ➔ The key words listed above can also be called concepts. Concepts can be related to one another by using linking words or phrases to form propositions. For example, the concept 'compound light microscope' can be linked to the concept 'lenses' by the linking phrase 'contains at least two' to form a proposition. An arrow shows the sense of the relationship: when several concepts are related in a meaningful way, a concept map is formed. Because concepts can be related in many different ways, there is no single, correct concept map. Figure 1.34 shows one concept map containing some of the key words and other terms from this chapter.

Use at least six of the key words above to make a concept map relating to the movement of substances across a cell membrane. You may use other words in drawing your map.

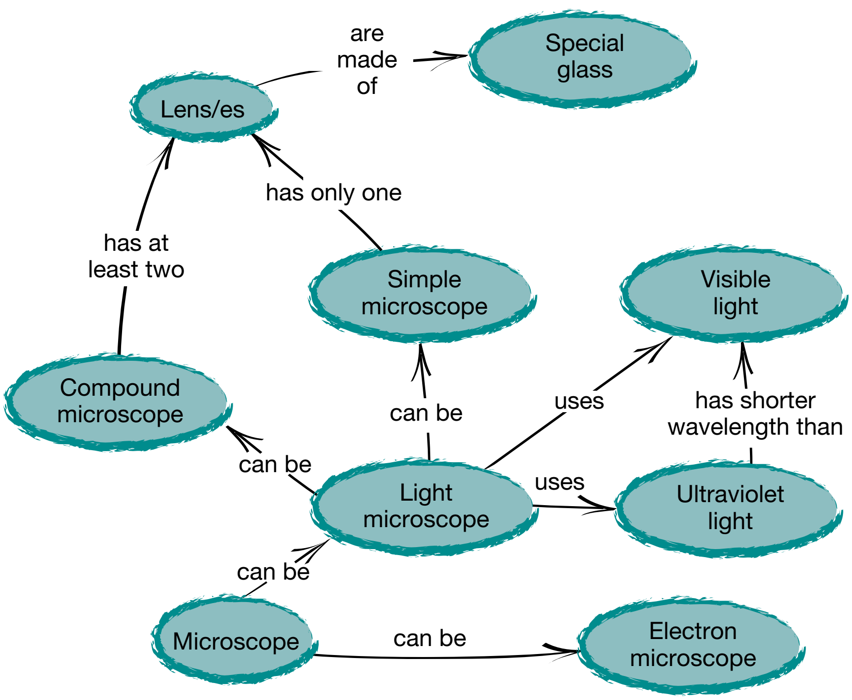



**FIGURE 1.34** Example of a concept map

2 **Applying your understanding** → Consider the information in table 1.5.

**TABLE 1.5** Data for three different shapes, each having the same volume. (Where necessary, figures have been rounded.)

| Cell | Shape | Dimensions | Surface area | Volume | SA:V ratio |
|---|---|---|---|---|---|
| A | flat sheet | 10 × 10 × 0.1 | 204 | 10 | 20.4 |
| B | cube | 2.15 × 2.15 × 2.15 | 28 | 10 | 2.8 |
| C | sphere | diameter: 1.67 | 22 | 10 | 2.2 |

a If these shapes represented cells, which cell (A, B or C) would be most efficient in moving required materials into and removing wastes from the cell? Explain.
b Which cell would be least efficient? Explain.
c Can you suggest a biological consequence of your conclusion?
d Identify one way in which a cell might retain its overall shape, but greatly increase its surface area with a minimal increase in volume. (Clue: This strategy is used by cells involved in absorption of material, such as those lining the small intestine.)

3 **Analysing information and drawing conclusions** → Consider the data in table 1.6.

**TABLE 1.6** Data for two sets of cells of identical shape but of decreasing sizes.

| Cell | Shape | Dimensions | Surface area | Volume | SA:V ratio |
|---|---|---|---|---|---|
| P | flat sheet | 10 × 10 × 0.1 | 204 | 10 | 20.4 |
| Q | flat sheet | 5 × 5 × 0.05 | 51 | 1.25 | 40.8 |
| R | flat sheet | 1 × 1 × 0.01 | 2.04 | 0.01 | 204 |
| H | sphere | diameter: 10 | 314.2 | 523.6 | 0.6 |
| K | sphere | diameter: 5 | 78.5 | 65.5 | 1.2 |
| L | sphere | diameter: 1.0 | 3.14 | 0.52 | 6 |

*Note:* relative to the first shape in each set, the dimensions of other members of the set are scaled down by a factor of 2 and by a factor of 10.

a A student stated the same shape scaled down should retain the same surface-area-to-volume ratio, the student's reason being 'the shapes stay the same'. Do you agree with this student? Explain your decision.
b With regard to the information in table 1.6, identify how scaling a shape (up or down) affects the SA:V ratio of a given shape by completing the following sentences:
 i If the size of a given shape is doubled, its SA:V ratio is . . .
 ii If the size of a given shape is halved, its SA:V ratio is . . .
c A particular shape has an SA:V ratio of 10.
 i What would happen to this ratio if this shape were scaled up by a factor of 5?
 ii What would happen to this ratio if this shape were scaled down by a factor of 2?
d Sphere (M) has a diameter of 0.5 units. Refer to table 1.6 and predict its SA:V ratio.
e Consider a different shape, such as a cube or a pyramid, that is changed in scale. Would its SA:V ratio be expected to follow a similar or a different pattern to that shown by the flat sheets and the spheres?

4 **Communicating understanding** → Two cells (P and Q) have the same volume, but the surface area of cell P is 10 times greater than that of cell Q.
a Placed in the same environment, which cell would be expected to take up dissolved material at a greater rate? Why?
b What might reasonably be inferred about the shapes of these two cells?
c Which measure — surface area or volume — determines the rate at which essential materials can be supplied to a cell?
d Which measure — surface area or volume — determines the needs of a cell for essential materials?
e Briefly explain why the surface-area-to-volume ratio provides a clue as to why cells are microscopically small?

5 **Analysing information and drawing conclusions** → Identify the following statements as true or false. For (d) only, briefly justify your choice.
a Osmotic flow of water occurs from a region of high to low solute concentration.
b Simple diffusion does not require the involvement of transporter proteins.
c Facilitated diffusion requires the involvement of a protein pump.
d The movement by diffusion of charged ions, such as $Na^+$ and $Cl^-$, across the plasma membrane is blocked by the lipid bilayer in the middle of the plasma membrane.
e Water is moved into and out of cells by active transport.
f Solid particles cannot cross the plasma membrane.
g Plant cells immersed in a hypertonic solution would be expected to burst.

6 **Applying your understanding** → Sucrose cannot cross the plasma membranes of red blood cells,

but glucose can. Red blood cells are immersed in the following solutions:

- a hypertonic sucrose solution
- a hypertonic glucose solution
- a hypotonic sucrose solution
- a hypotonic glucose solution.

**a** Which solution would be expected to cause the greatest water loss and shrinkage of the red blood cells? Explain.

**b** Which solution, if any, might cause the red blood cells to burst? Explain.

**7 Making valid comparisons →**

**a** Name the two types of diffusion that are involved in the movement of dissolved substances across the plasma membrane.

**b** Identify two similarities in these two types of diffusion.

**c** Identify how these two types of diffusion differ.

**d** A student stated 'Surely two types of diffusion are unnecessary.' Indicate whether you agree or disagree with this statement and give a reason for your decision.

**e** Identify one key difference between diffusion and active transport of a substance.

**8 Analysing information and drawing conclusions →** Suggest a possible explanation for the following observations:

**a** Proteins can move laterally across the plasma membrane.

**b** A person with cystic fibrosis is at high risk of lung infections.

**c** Lipophilic substances cross the plasma membrane by simple diffusion, but not charged particles.

**d** A baby with cystic fibrosis produces abnormally salty sweat.

**e** People infected with cholera suffer severe diarrhoea.

**9 Making valid predictions based on your knowledge →** An artificial membrane, composed of a phospholipid bilayer only, was manufactured. Its behaviour was compared with that of a natural plasma membrane.

Predict if these two membranes might behave in a similar or a different manner when tested for their ability to allow the following dissolved substances to cross them:

- small lipophilic substances
- charged particles, such as sodium ions
- glucose
- proteins.

Briefly justify each of your decisions.

**10 Performing calculations →** The width of an average head hair from a Caucasian is about 0.6 mm. Refer back to figure 1.2a and estimate about how many red blood cells could fit across the width of such a hair.

**11 Making an estimation →** Use your knowledge of the structure of the plasma membrane to suggest which of the following dimensions is most likely to designate the width and length of one phospholipid molecule:

- 0.9 × 3.4 nm
- 0.9 × 3.4 μm OR 0.9 × 3.4 mm.

Briefly explain your choice.

**12 Applying your knowledge and understanding →** Nerve impulses involve several movements of sodium ions in different directions across the plasma membrane of a nerve cell as follows:

**a** Before a nerve impulse occurs, sodium ions are more concentrated in the extracellular fluid outside the cell than inside the cell. By what means does the cell maintain this difference?

**b** During transmission of the nerve impulse, sodium ions flood into the nerve cell from the extracellular fluid. By what means do these ions enter the cell?

**c** After the impulse has passed, the original concentration of sodium ions is restored to its high concentration outside the cell by a process that moves sodium ions out of the cell. By what means does this restoration occur?

**13 Discussion question →** In Haiti, a disastrous earthquake in January 2010 killed and injured hundreds of thousands of people, left even more homeless, destroyed buildings and damaged infrastructure including roads, telecommunications, water supplies and water treatment plants. Homeless people sought shelter in camps that soon became overcrowded. Cholera infections broke out and developed into an epidemic. By March 2016, more than 770 000 cases of cholera have been reported, resulting in more than 9000 deaths, and the epidemic continues.

**a** What is the causative agent of cholera?

**b** What causes the particular effects of a cholera infection?

**c** By what means is cholera spread?

**d** What possible health consequences would be expected from the destruction of the water supplies and the re-housing of large numbers of people in temporary camps?

**e** In an attempt to halt the spread of cholera, measures have been introduced, including:

- repair of water treatment plants
- distribution of water purification tablets and hygiene kits to families

- distribution of oral rehydration salts packs to treatment centres
- introduction of community health education programs
- introduction of oral vaccines for cholera
- introduction of rapid diagnosis tests to distinguish cholera from diarrhoea.

Consider each of these measures in turn and discuss how each might contribute to slowing and stopping the spread of cholera.

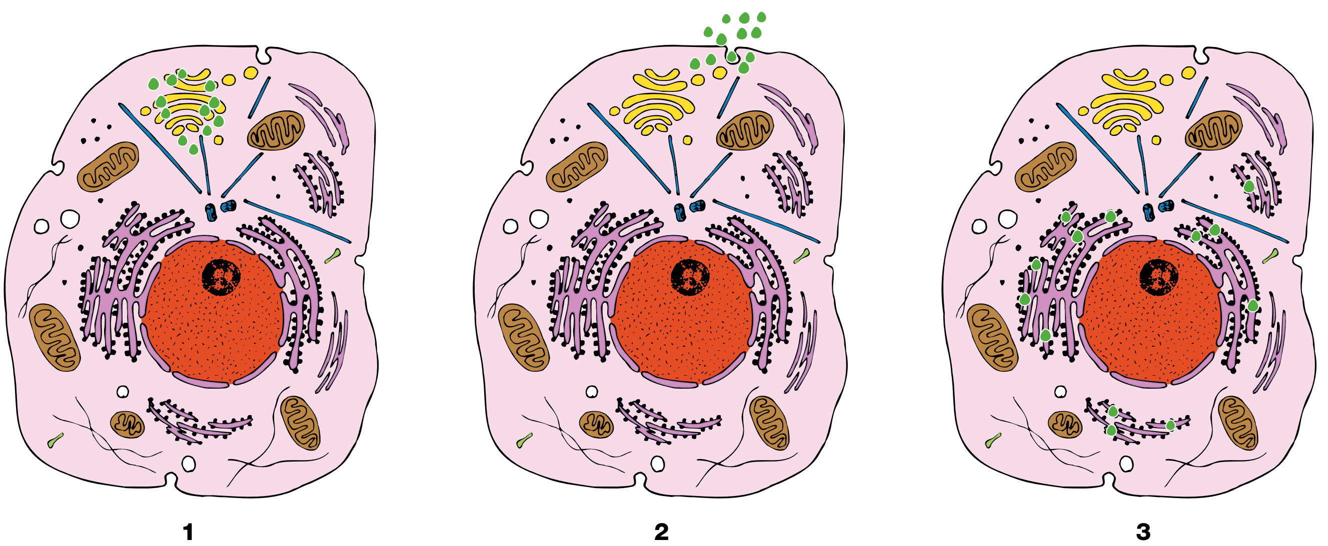
**14 Analysing information and drawing conclusions ➔**

A scientist carried out an experiment to determine the time it took for a cell to manufacture proteins from amino acids. The scientist provided the cell with radioactively labelled amino acids and then tracked them through the cell to establish the time at which protein synthesis commenced. He monitored the cell 5 minutes, 20 minutes and 40 minutes after production started in order to track the proteins from the site of synthesis to a point in the cell from which they were discharged from the cell.

The scientist made an image of the cell at each of these times but forgot to mark each image with its correct time. The images are shown below. Location of the radioactivity is shown by the green spots.

**a** Which cell corresponds to each of the particular times of viewing? List the correct order according to time of viewing.

**b** On what grounds did you make your decision?



CHAPTER

# 2 Nucleic acids and proteins: a review

## KEY KNOWLEDGE

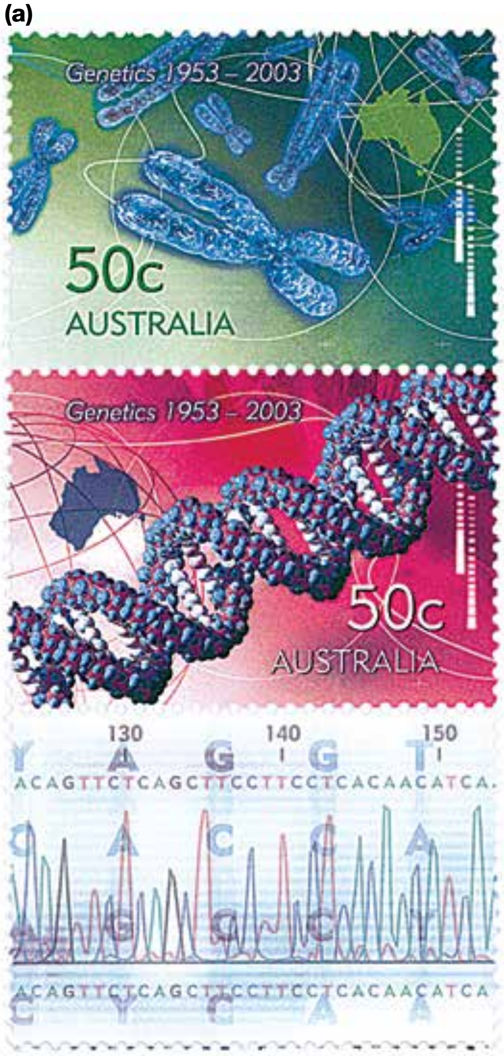
This chapter is designed to enable students to:

- enhance their knowledge and understanding of nucleic acids
- enhance their understanding of the structure of proteins and polymerisation
- develop an understanding of the structure of DNA, RNA and triplet codes
- explain the differences between structural genes and regulatory genes
- understand the lac operon.

**FIGURE 2.1** In celebration of the 50-year anniversary of the discovery of the structure of DNA. **(a)** Australian stamps depicting, from top to bottom, a chromosome, a DNA molecule and the base sequence of a DNA molecule. (Image courtesy of the Australian Postal Corporation. Designer: Rod Oliver, Fragile Design.) **(b)** An English two-pound sterling coin showing the double helix structure of DNA. (Image courtesy of The Royal Mint, UK.)

The year 2003 marked a significant celebration. It was the 50-year anniversary of the discovery of the structure of DNA by James Watson and Francis Crick. The discovery was a giant step forward in the field of science and opened the way to extensive research that has given us the molecular biology knowledge, understanding and applications that we have today. Many countries around the world celebrated the anniversary through the issue of special stamps. Australia was one of those countries. The country in which Watson and Crick did their work, England, also issued a coin.

(a)




(b)

# Organic molecules

Many organic molecules are made of large numbers of smaller sub-units that are linked together by specific covalent bonds. Nucleic acids, proteins and polysaccharides are examples of this type of organic molecule. **Compounds formed in this way are called polymers. The sub-units are called monomers.** The joining of monomers involves the release of a water molecule. Reactions of this kind are termed condensation reactions. We classify polymers on the basis of the kind of sub-unit they contain (refer to figure 2.2). These polymers are important to the functions of living organisms and, in the case of polysaccharides and proteins, also form part of their structures.

In addition, other organic compounds are involved in the structure and function of living organisms, with one important group being lipids. Lipids, which include fats, oils and phospholipids, are not regarded as polymers because they are not linked together in a chain of many repeating identical sub-units.

**ODD FACT**

Condensation (water-producing) reactions assemble polymers, and hydrolysis (water-consuming) reactions break down polymers.

**ODD FACT**

The outer coat of pollen grains is composed of an organic polymer called sporopollenin, which is highly resistant to chemical breakdown and can last for millions of years. It has been termed '*one of the most extraordinarily resistant materials in the organic world*'.

**eBook*plus***

**Interactivity**
Organic compounds
**int-3026**

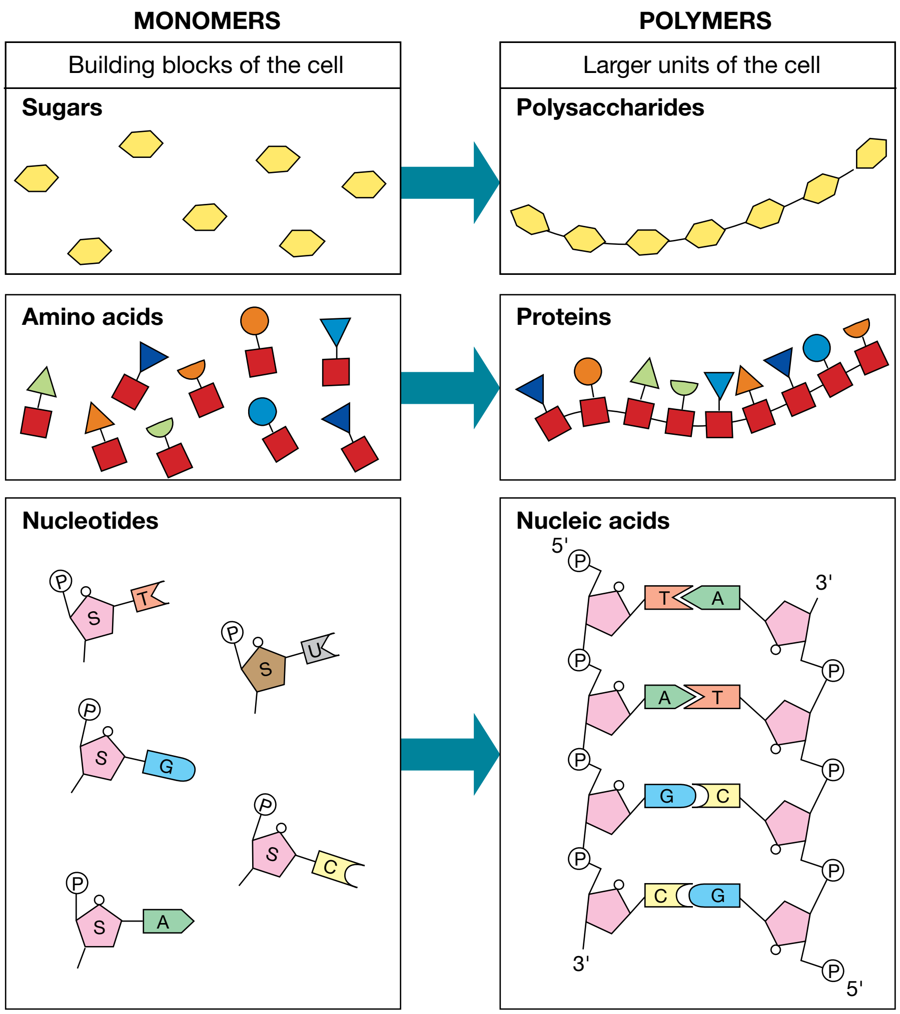



**FIGURE 2.2** Three main groups of polymeric organic molecules present in cells. Note the type of monomer that makes up each type of polymer. Examples of polysaccharides include starch, glycogen and chitin; examples of protein groups include enzymes, immunoglobulins and contractile proteins; examples of nucleic acids include DNA and RNA.

The kinds of organic molecules that we will consider are **proteins** and **nucleic acids**. For each of these, we will examine:

- the basic unit of structure
- how the units combine to form complex molecules
- where each kind of molecule is found in a cell
- the functions of the molecules.

# Nucleic acids

There are two kinds of nucleic acid:

- **deoxyribonucleic acid (DNA)**, which is located in chromosomes in the nucleus of eukaryotic cells (see figure 2.3). It is the genetic material that contains hereditary information and is transmitted from generation to generation.
- **ribonucleic acid (RNA)**, which is formed against a template strand of DNA.

## DNA — an influential molecule

The genetic material **deoxyribonucleic acid is a polymer of nucleotides**. Each nucleotide unit, or monomer, has:

- a sugar (deoxyribose) part
- a phosphate part
- an N-containing base.

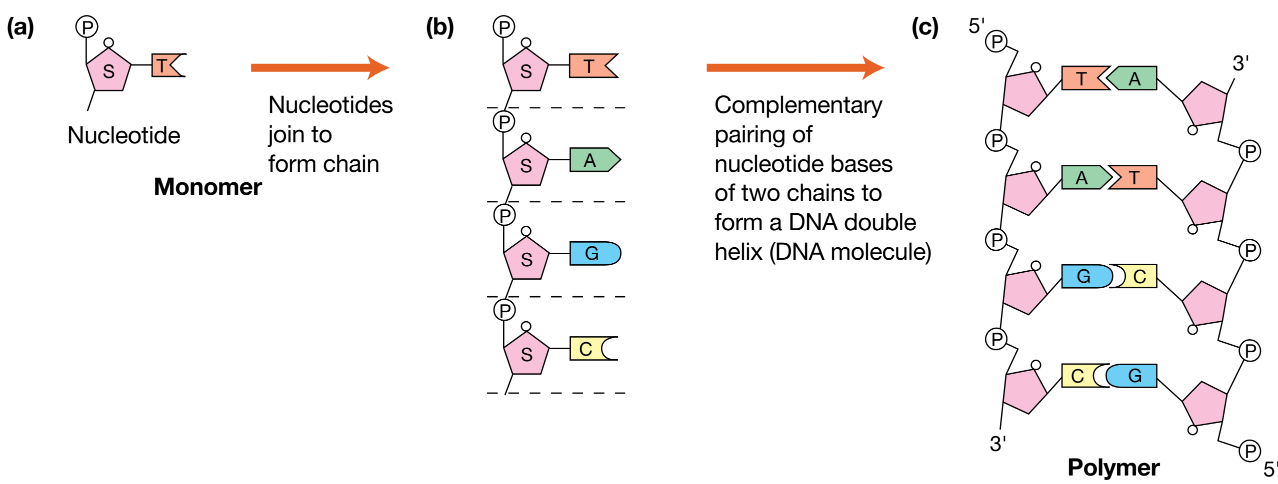



**FIGURE 2.3** Deoxyribonucleic acid is made from nucleotide sub-units. Each DNA molecule is made of two complementary chains of nucleotides.

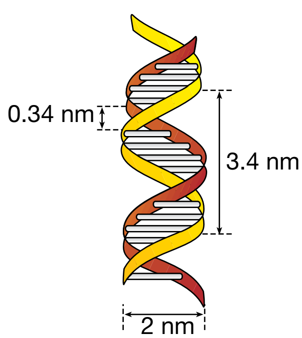



**FIGURE 2.4** The double helix structure and dimensions of DNA. The two chains are held together by hydrogen bonds between complementary bases.

The sugar and phosphate parts are the same in each nucleotide.

There are four different kinds of nucleotides because **four different kinds of N-containing bases are involved:** adenine (A), thymine (T), cytosine (C) and guanine (G). Examine figure 2.3. The nucleotide sub-units (a) are assembled together to form a chain (b) in which the sugar of one nucleotide is linked to the phosphate of the next nucleotide in the chain. Each DNA molecule contains two chains (c) that form a double helix, with the bases in one strand pairing with the bases in the other strand. **The base pairs between the two strands, namely A with T, and C with G, are said to be complementary pairs.** The two chains form a double-helical molecule of DNA (see figure 2.4).

The DNA double helix combines with certain proteins, **in particular histones, as it condenses to form a chromosome** (see figure 2.5a). **As the DNA winds around clusters of histone proteins, it forms structures called nucleosomes** (see figure 2.5b).

studyon

Unit 3 | AOS 1 | Topic 3 | Concept 4

**DNA structure**
Summary screen and practice questions

studyon

Unit 3 | AOS 1 | Topic 3 | Concept 4

**See more**
DNA structure

studyon

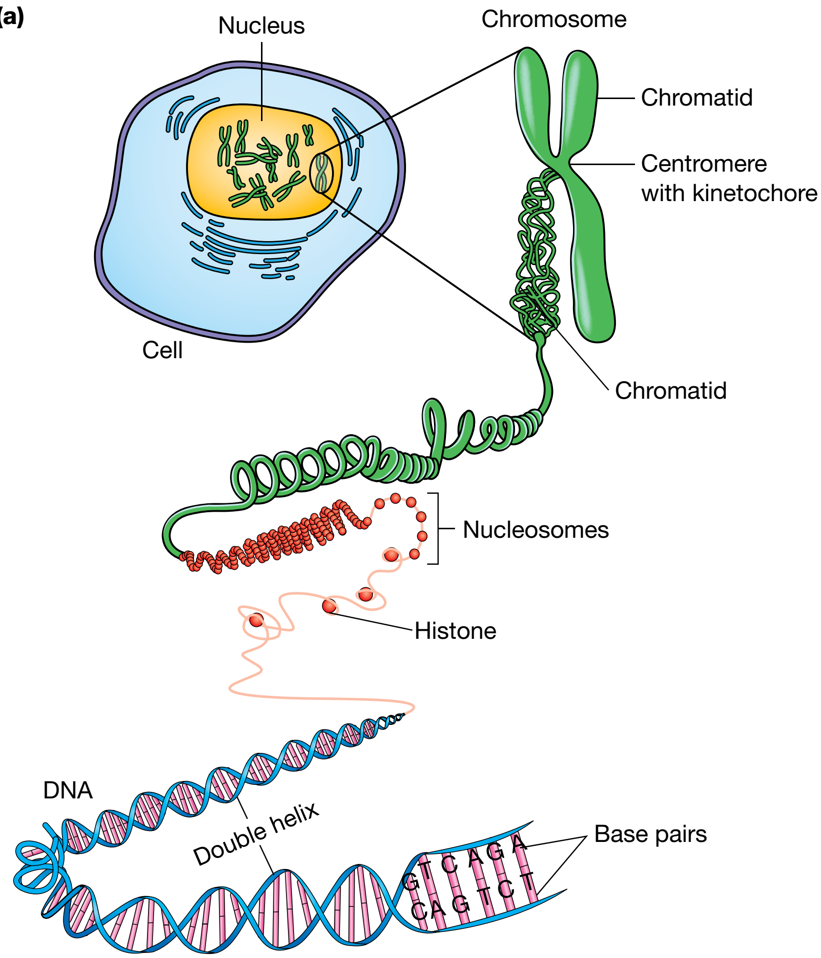
Unit 3 | AOS 1 | Topic 3 | Concept 5

**DNA function**
Summary screen and practice questions

**ODD FACT**

The total length of the DNA double helix molecule in an 'average' human chromosome is about five centimetres. By coiling and supercoiling, this long DNA molecule becomes compressed into a microscopic chromosome.

(a)




(b)

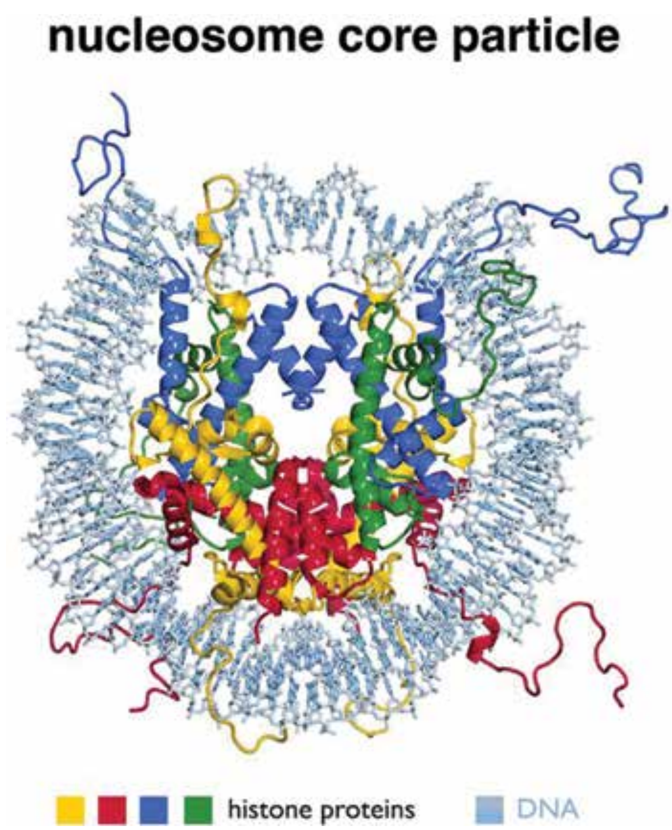
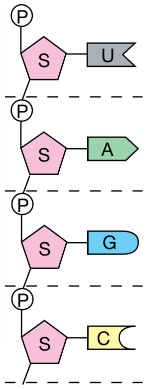



**FIGURE 2.5** **(a)** Diagram showing the coiling and supercoiling of one molecule of a DNA double helix to form a eukaryotic chromosome. Note that a key process involves the coiling of DNA around histone proteins to form structures called nucleosomes that then become supercoiled. **(b)** A model of a nucleosome showing the DNA double helix (grey) coiling around a cluster of histone proteins (shown in colours). (Image (b) courtesy of Dr Song Tan, Pennsylvania State University)

**FIGURE 2.6** The four nucleotide sub-units, uracil, adenine, guanine and cytosine, from which the three forms of RNA are constructed.




How does DNA control functions within cells? All metabolic reactions in cells are controlled by enzymes that, almost without exception, are proteins built of one or more polypeptide chains — chains of amino acids. Hence, **the DNA in the nucleus of a eukaryotic cell controls all the metabolic processes in a cell, through the polypeptide chains for which the DNA dictates production.**

## Ribonucleic acid

Ribonucleic acid (RNA) is also a polymer of nucleotides (see figure 2.6). It differs from DNA in that it is **an unpaired chain of nucleotides** that contain the sugar ribose. RNA nucleotides include four different bases, three of which — adenine, guanine and cytosine — are identical to those in DNA. **The fourth nucleotide is uracil,** which is capable of pairing with A.

**FIGURE 2.7** A transmission electron micrograph (TEM) of a fragment of an mRNA translation unit from the salivary gland cell of a midge (*Chironomus* sp.). Ribosomes (blue) attach to the mRNA strand (pink) and read its code. A tRNA molecule carrying a corresponding amino acid binds to the ribosome. As the ribosome moves onto the next bases along the mRNA, a protein (green) grows from the ribosome.

The three different forms of RNA are all produced in the nucleus alongside DNA as a template:

- **messenger RNA** (mRNA), which carries the genetic message to the ribosomes where the message is translated into a particular protein (see figure 2.7)
- **ribosomal RNA** (rRNA), which, together with particular proteins, makes the ribosomes found in cytosol
- **transfer RNA** (tRNA), molecules that carry amino acids to ribosomes where they are used to construct proteins.

The strand of nucleotides in each of the forms of RNA is folded in a different way.

**study on**

| | |
|---|---|
| Unit 3 | **RNA structure** |
| AOS 1 | Summary screen and practice questions |
| Topic 3 | |
| Concept 6 | |

### KEY IDEAS

- Proteins, polysaccharides and nucleic acids are polymeric organic molecules built out of a very large number of repeating sub-units.
- The nucleic acids, double-helical DNA and single-stranded RNA are built out of a very large number of repeating sub-units called nucleotides.
- Each nucleotide consists of a sugar, a phosphate and an N-containing base, with the sugar in DNA being deoxyribose and that in RNA being ribose.
- Each DNA molecule consists of two chains of nucleotides that are complementary to each other and held together by hydrogen bonds.
- Each RNA molecule consists of a single strand of nucleotides.

**study on**

| | |
|---|---|
| Unit 3 | **RNA function** |
| AOS 1 | Summary screen and practice questions |
| Topic 3 | |
| Concept 7 | |

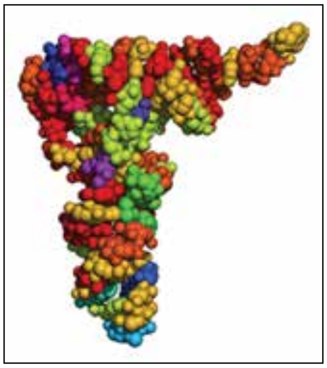
### QUICK CHECK

1 What do the initials DNA and RNA stand for?
2 What are the four kinds of nucleotides (a) in DNA and (b) in RNA?
3 Where in a cell would you find DNA and what is its function?
4 There are three different kinds of RNA molecules. What are they and where is each found in the cell?
5 In what ways do the kinds of RNA differ from one another?

**studyon**

Unit 3
AOS 1
Topic 3
Concept 7

**Do more**
RNA function



# Gene structure

## The template strand and its partner

A gene consists of a particular part of a double-helical molecule of DNA. **Only one of the two chains contains the information present in a particular gene, and this is called the template strand**. The complementary chain is sometimes called the nontemplate strand, also called the mRNA-like strand or the coding strand.

## Representations of DNA

DNA cannot be seen with a light microscope. However, a technique known as scanning tunnelling microscopy (STM) allows DNA molecules to be visualised (see figure 2.8).

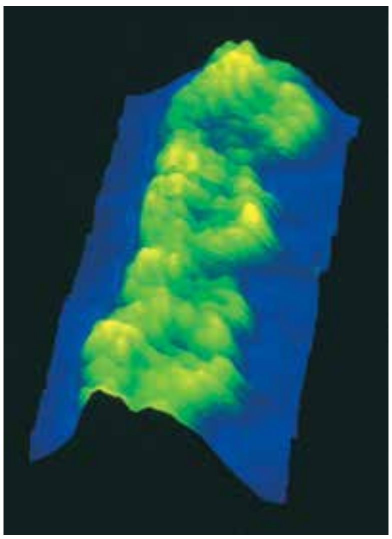


**FIGURE 2.8** Part of a DNA double helix revealed through scanning tunnelling microscopy.

Part of a single chain of DNA could be shown as follows:

. . . -nucleotide-nucleotide-nucleotide-nucleotide-nucleotide- . . .

OR it could be shown as:

```
. . . -P-sugar-P-sugar-P-sugar-P-sugar-P-sugar- . . .
         |       |       |       |       |
        base    base    base    base    base
```

OR the specific bases in the nucleotides in one chain could be shown as:

5′ . . . A T T A G C T T G A G G C G . . . 3′

Which representation is correct? All are correct. Which representation is the most informative? The third is the most informative because it gives the information about the order of nucleotides, the only variable part of the genetic material. What information does the second representation provide?

DNA is not always represented in diagrams as a double helix. Figure 2.9 shows some of the many ways of representing DNA. The representation used will depend on the purpose of the diagram. Each provides different information about DNA. For example, (a) gives information about the coding regions (exons) and the noncoding regions (introns) within a gene, while (c) gives the base sequence.

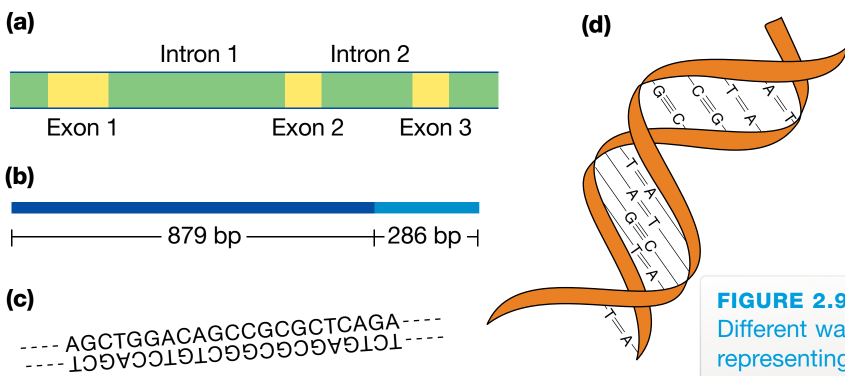
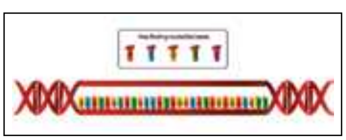



**FIGURE 2.9** Different ways of representing DNA.

**studyon**

Unit 3
AOS 1
Topic 3
Concept 5

**Do more**
Complementary DNA



# Gene sequencing

What is this?

A T G G T G C A C C T G A C T C C T G A G G A G A A

This is part of the **nucleotide sequence** of the template strand of the human **HBB** gene, which encodes the information for one of the protein chains found in haemoglobin. What is the sequence of the complementary strand? **When the order of the nucleotides in a gene is identified, the gene is said to be**

FIGURE 2.10 A scientist examining the sequence of bases in DNA.

**'sequenced'**. The order in many genes from animals, plants and bacteria has been identified, and data continue to be added.

**Gene sequencing** involves the process of identifying the order of nucleotides along a gene. Figure 2.10 shows a scientist examining some sets of bands arranged in columns. Each band represents one nucleotide and the order of the bands down the column corresponds to the **gene sequence**. New techniques of sequencing are described in the box below.

Are gene sequencers used only for human genes? No! The genetic material of all organisms is DNA and the structure of that DNA is identical, regardless of whether it comes from wheat, jellyfish, ducks, *Bacillus* bacteria, insects or people. **In all organisms, genes are built of the same 'alphabet' of four letters**, namely, the nucleotides A, T, C and G of DNA. In contrast, non-cellular agents, such as hepatitis C and Ebola viruses, include many types that have RNA as their genetic material rather than DNA. Refer to chapter 6, page 249 for further detail.

So the genetic instruction kit to 'make a human being' or 'make an oak tree' or 'make a white shark' consists of thousands of instructions, each consisting of DNA with different base sequences.

## DNA SEQUENCERS

(a)

(b)

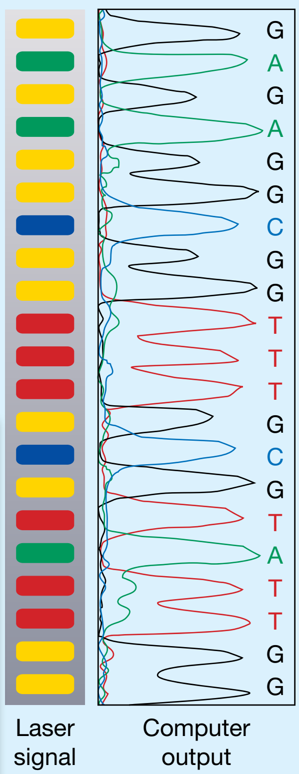



**FIGURE 2.11** **(a)** Applied Biosystems 3500 genetic analyser by Thermo Fisher Scientific. (Image courtesy of Applied Biosystems.) **(b)** Output from a DNA sequencer showing the laser signals and the output from the computer, which identified the base sequence in part of a DNA fragment. What is signalled by a red band?

The process of gene sequencing has now been automated and is done using instruments known as **DNA sequencers** (see figure 2.11a). One automated system, known as the Sanger method, involves the use of four different coloured fluorescent dyes, each of which binds to a specific base (A, T, C or G) in DNA. The DNA chain is sequenced using a stepwise procedure that makes a complementary copy using the unknown DNA as the template, with each copy being one nucleotide longer than the previous one as shown below:

| | |
|---|---|
| **Unknown DNA:** | **CTCTCCGCCAAACGCATAACC** |
| 1st copy | G* |
| 2nd copy | GA* |
| 3rd copy | GAG* |
| 4th copy | GAGA* |
| etc. | |
| 21st copy | GAGAGGCGGTTTGCGTATTGG* |

In each case, the nucleotide at the end of each copy becomes attached to the specific fluorescent dye (shown as *). The copies move in turn, shortest first, past a scanning laser that activates the dye so that it emits a fluorescent signal, which is captured by a detector. This detector transfers the signal to a microcomputer, which determines the entire base sequence. The output from a DNA sequencer shows the base sequences as a series of coloured signals (see figure 2.11b) with a yellow peak denoting G, a red peak denoting T, a green peak denoting A and a blue peak denoting C.

The latest developments in sequencing are the next-generation DNA sequencers (NGS). These sequencers enable **millions of DNA fragments from tiny samples to be strung together *at the same time*, in contrast to the one-at-a-time method of the Sanger technique**.

**ODD FACT**

In October 2001, the cost of DNA sequencing was almost US$5000 per 1000 bases (1 megabase). By 2015, this cost had dropped to about 0.33 cents per raw megabase. (Data from National Human Genome Research Institute.)

## Comparing DNA from different organisms

Databases hold information about the base sequences in the DNA of genes of many organisms. Table 2.1 shows part of various genes from different organisms — a duck, a *Bacillus* bacterium, a corn plant and a human being. Could you pick the human gene? It's not possible. The genes look similar because the genetic language of all living things is written in a common language based on an 'alphabet' of four letters (A, T, C and G) that correspond to the nucleotides (and bases) present in DNA.

**TABLE 2.1** Part of the sequences of different genes from various organisms. Numbers are placed above the sequences for ease of locating a particular nucleotide.

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | | 10 | | | 20 | | | 30 | | | | 40 | |
| **Organism P** | ATG | GCT | ACC | AAG | ATA | TTA | GCC | CTC | CTT | GCG | CTC | CTT | TCC | CTT | TTA |
| | 1 | | | 10 | | | 20 | | | 30 | | | | 40 | |
| **Organism Q** | ATG | AAG | TGT | AAT | GAA | TGT | AAC | AGG | GTT | CAA | TTA | AAA | GAG | GGA | AGC |
| | 1 | | | 10 | | | 20 | | | 30 | | | | 40 | |
| **Organism R** | ATG | ACG | CTG | ACT | CAA | GCT | GAG | AAG | GCT | GCC | GTG | ATC | ACC | ATC | TGG |
| | 1 | | | 10 | | | 20 | | | 30 | | | | 40 | |
| **Organism S** | ATG | AGG | CTC | TTG | TGG | TTG | CTT | TTC | ACC | ATT | GGG | TTC | TGC | TGG | GCT |

P: corn plant
Q: *Bacillus* bacterium
R: duck
S: human being

### How do genes differ?

The total human DNA contains a 'make a human being' instruction kit; the total geranium DNA contains a 'make a flowering plant, geranium type' instruction kit. The human genetic instruction kit consists of tens of thousands of separate instructions, such as 'make the hair protein, keratin' and 'make growth hormone'. **Different genetic instructions within and between species are due to different nucleotide sequences in the genes.**

**ODD FACT**

Global databases with publicly available DNA, RNA and protein sequences include the National Center for Biotechnology Information (NCBI) database and the European Molecular Biology Laboratory (EMBL) database.

**KEY IDEAS**

- The length of a double-helical DNA molecule can be expressed as the number of base pairs (bp) it contains, and one chain of this DNA as the numbers of nucleotides.
- Each human chromosome contains one long molecule of double-stranded DNA with millions of base pairs.
- A typical gene consists of tens of thousands of nucleotides.
- The estimated total number of human genes is 21 000.
- Of the two DNA chains in a gene, the one containing the genetic information is known as the template strand of DNA, while its complementary chain is called the nontemplate strand.
- Genetic instructions are coded in an 'alphabet' of four letters only: (the nucleotides) A, T, C and G.
- Identification of the order of nucleotides along a length of DNA is called DNA sequencing.
- Different genes vary in the nucleotide sequences along their DNA.

**QUICK CHECK**

**6** A piece of DNA comprises 20 000 bp. Is this more likely to be a whole chromosome or a whole gene? Explain.

**7** If genes were isolated from a cat, a cyanobacterium and a cauliflower, what similarity would be seen?

**8** If the two human genes for making blood-clotting factor and salivary amylase were compared:

**a** in what way would they be similar

**b** in what way would they be different?

# Nature of the genetic code

## Codes and more codes

The genetic instructions for all organisms are written in a code that uses an 'alphabet' of four letters only, namely A, T, C and G. What is a code?

Figure 2.12 shows some codes. Codes contain or encode information. To translate or decode a coded message, the information held in the code elements must be known. When **decoded**, this information may be verbal (words) or numerical (numbers) or auditory (musical sounds) or may specify an object.

If the coded information exists in a permanent form, the information can be decoded at any time by returning to read the code. For example, you can play a CD many times; each time the CD player decodes (translates) the information that is permanently **encoded** as microscopic pits on the disk surface.


**FIGURE 2.12** Various ways of representing information codes. Can you think of others?

## The genetic code

The DNA of genes is an information-carrying molecule. Genetic information is encoded using just four elements: A, T, C and G. Before DNA was identified as the genetic material, many biologists thought that DNA was too simple a molecule to contain complex genetic instructions. How can a large amount of information be encoded by a code that has a small number of elements?

A code consisting of a few elements can encode a large amount of information. The morse code has two elements only, a dot (.) and a dash (–). By using various combinations of these elements, morse code can convey very complex information, such as all the words in this chapter.

The discussion on this page refers to those genes that encode proteins. Some genes have functions other than encoding proteins.

### Information encoded in the genetic code

The **genetic code** in the DNA of protein-encoding genes typically contains information for joining amino acids to form polypeptides. We can show this as:

| **Coded information** | **Decoded information** |
|---|---|
| nucleotide sequences in DNA template strand | order of amino acids in polypeptides |

A polypeptide on its own may be a functional protein or, in other cases, a functional protein may be built of several different polypeptide chains.

Proteins have many functions, and the various types include:

- structural proteins, which occur in connective tissues and in cell membranes
- contractile protein of muscle and myofilaments
- enzymes that regulate chemical processes
- proteins of the immune system, such as the antibodies
- oxygen-carrying proteins, such as haemoglobin
- hormonal proteins, such as insulin and growth hormone.

**By encoding the sets of instructions on how to make the various types of proteins, genes control the structure and the biochemical and physiological functioning of an organism.** The estimated 21 000 genes of a human organism contain all the instructions on 'how to make a human organism' that, if printed as the base sequences, would fill 1000 volumes of an encyclopedia.

## Organisation of the genetic code

Consider two observations:

1. Genes typically contain coded information for assembling amino acids to form polypeptides.
2. Polypeptides are made of combinations of 20 different amino acid sub-units.

From these observations, it may be inferred that the genetic code must have at least 20 different instructions or pieces of information.

Examine table 2.2.

**TABLE 2.2**

| Number of nucleotides in one instruction | Total number of different instructions possible |
|---|---|
| 1 (e.g. T) | 4 |
| 2 (e.g. AA, AT, GA) | 16 |
| 3 (e.g. TTA, GCC, AAA) | 64 |
| 4 (e.g. GGGA, TGCA, AATG) | 256 |

In discussing the genetic code, the terms *base sequence* and *nucleotide sequence* are used interchangeably.

In fact, one genetic instruction consists of a group of three bases, such as AAT, GCT and so on. Because of this, **the genetic code is referred to as a triplet code**. This form of code is sufficient to account for the pieces of information that must be encoded.

Consider a piece of template-strand DNA with the base sequence T A C A A A C A A G C T C C T A C T . . . This DNA has six coded instructions (shown underlined) that are decoded or translated as follows:

1. TAC = start building a protein, commencing with the amino acid *met*
2. AAA = now add the amino acid *phe*
3. CAA = now add the amino acid *val*
4. GCT = now add the amino acid *arg*
5. CCT = now add the amino acid *gly*
6. ACT = now stop.

## Cracking the genetic code

The genetic code was originally unknown and had to be broken. In 1961, the first piece of the genetic code was broken when the three-base triplet 'AAA' in DNA was decoded as 'Add the amino acid *phe* into a protein being constructed' (see figure 2.13). We will see later in this chapter that the translation of each DNA triplet occurs through an intermediate molecule, messenger RNA (mRNA). The AAA triplet in DNA is transcribed into a complementary UUU codon in mRNA that is then translated into the amino acid phenylalanine (*phe*). By 1966, all 64 pieces of the genetic code had been deciphered.

Refer to table 2.5 on page 63 to see the full translation of the genetic code in terms of the 64 mRNA codons. Note that some codons contain the instructions 'START adding amino acids' and 'STOP adding amino acids'. Refer to the Appendix to see the structures of the amino acids and their three-letter and single-letter abbreviations.

The main features of the genetic code are:

- Pieces of information in the genetic code consist of triplets or three-base sequences (e.g. TCA in DNA, or UGT in mRNA).
- The code is non-overlapping. So, a fragment of DNA consisting of 12 bases contains four pieces of information or instructions.
- The code is essentially the same in bacteria, in plants and in animals — it is said to be universal (but see the Odd Fact at left).
- The code is said to be redundant because, in many cases, more than one triplet of bases codes for one particular amino acid.
- The information encoded in DNA is the set of instructions to assemble amino acid sub-units into polypeptides.
- The information in the DNA template strand also includes a START instruction (TAC) and three STOP instructions (ATT, ATC or ACT).

**ODD FACT**

The genetic code is not completely universal. Some triplets that code for one instruction in most organisms code for a different instruction in a few other organisms. The code is also different for mitochondrial DNA — TCT is a STOP signal, not a code for *arg*. The code in mitochondrial DNA also shows slight variations.

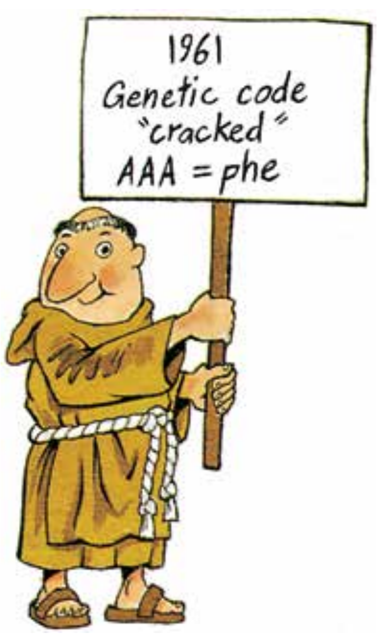



**FIGURE 2.13** In 1961, Nirenberg and Matthaei 'cracked' the first element of the genetic code.

**KEY IDEAS**

- DNA contains information encoded in the base sequence of its template strand.
- Genes contain coded instructions for joining specific amino acids into polypeptides.
- The genetic code in DNA is a non-overlapping triplet code consisting of groups of three bases.
- One piece of genetic code typically contains the information to add one amino acid to a protein.

**QUICK CHECK**

**9** Give an example of a code.
**10** In what form is information held in a DNA molecule?
**11** Which of the following statements is most accurate? Explain your choice.
  **a** DNA is converted to the amino acid sub-units of proteins.
  **b** DNA contains the coded information for joining amino acids to form polypeptides.
  **c** DNA turns into polypeptides.
**12** How many instructions (for adding amino acids) are present in the base sequence TTAGGG?
**13** Which code means 'START joining amino acids to form a polypeptide'?
**14** What are the meanings of the following codes in DNA: CAA and ACT?

# Proteins

Although water is the main compound in living cells, **more than half of the remainder, about 18 per cent is protein**. There are thousands of different proteins in each cell and many of these control all metabolic processes within cells. **All proteins contain nitrogen (N), hydrogen (H), carbon (C) and oxygen (O). Some also contain sulfur (S)**. They are large complex compounds as indicated by the following examples:

- zian from Indian corn $C_{736}H_{1161}N_{184}O_{208}S_3$
- gliadin from wheat $C_{685}H_{1068}N_{196}O_{211}S_5$
- casein from milk $C_{708}H_{1130}N_{180}O_{224}S_4P_4$.

Some plants contain granules of solid protein that they store as reserve food. Animals, as we will see in the next section, have most of their energy stores in fat.

## Amino acids: the building blocks of proteins

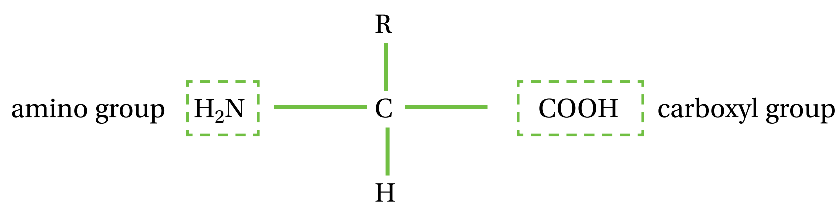
These nine amino acids cannot be made by human cells but they are essential and must be obtained in food eaten:

| | |
|---|---|
| threonine | isoleucine |
| methionine | histidine |
| lysine | phenylalanine |
| valine | tryptophan. |
| leucine | |

**Proteins are large molecules built of sub-units called amino acids.**

There are 20 naturally occurring amino acids (refer to the Appendix). Humans are unable to make all 20 amino acids and must rely on their food for the nine they are unable to make (see the note at left). Not all plants can make all 20 amino acids so a vegetarian diet should be well planned to ensure a balanced intake of appropriate amino acids. Generally, animal proteins are a better source of amino acids for humans because animal protein is more like that of humans. Different proteins contain different numbers and proportions of each of the amino acids.

The general formula of an amino acid is:

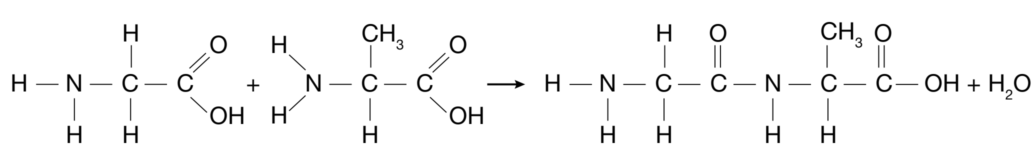


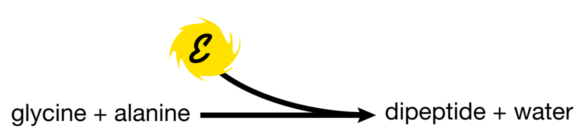

studyon

| | |
|---|---|
| Unit 3 | **Proteins** |
| AOS 1 | Summary screen and practice questions |
| Topic 3 | |
| Concept 1 | |

Each amino acid has one part of its molecule different from other amino acids. **The R group in the general formula is the part that varies.** Refer to the Appendix and note that the R group of the amino acid cys contains sulfur (S). When several cys amino acids are present, disulfide bonds (-S-S-) can form (see figure 2.19).

Two amino acids join together as a dipeptide when a peptide bond forms between the amino group of one amino acid and the carboxyl group of another amino acid and a water molecule is released (see below and figure 2.14). **When a number of amino acids join in this way, a polypeptide is formed.** Each type of protein has its own particular sequence of amino acids. Polypeptide chains become folded in different ways depending on their amino acid sequences.





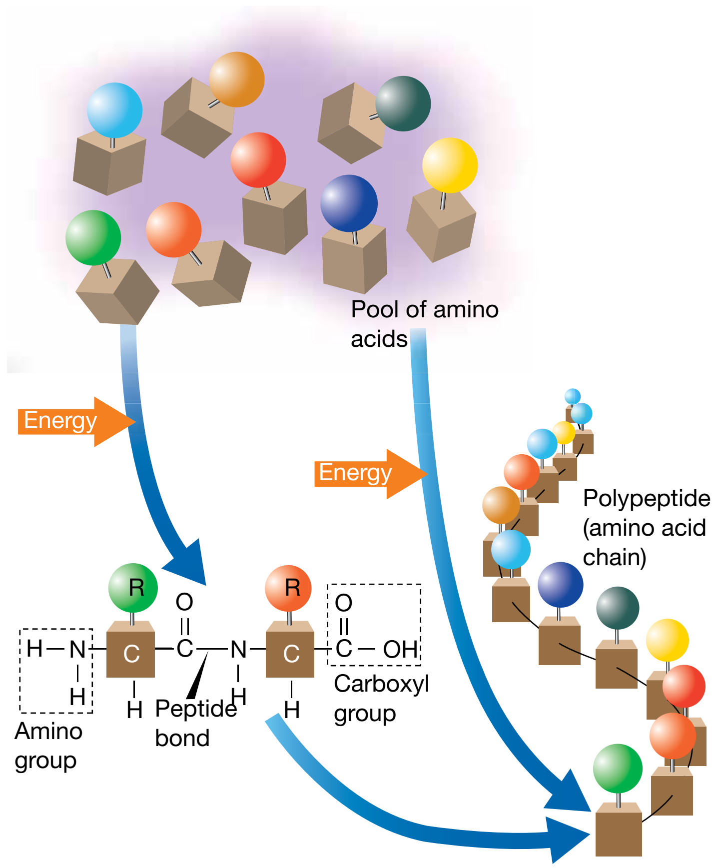



**FIGURE 2.14** Proteins are assembled from amino acids that are joined by peptide bonds. Each peptide bond forms by the linkage of an amino group from one amino acid and a carboxyl group of another amino acid. A number of amino acids joined by peptide bonds form a polypeptide chain. Note that the different R groups are represented by different colours. The remainder of each amino acid molecule is identical.

**studyon**

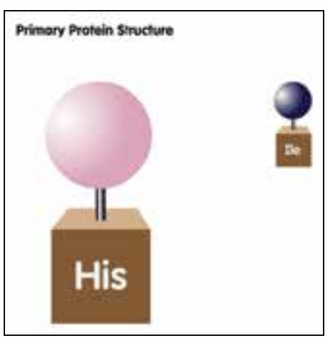
Unit 3 | AOS 1 | Topic 3 | Concept 2

**Primary structure of proteins**
Summary screen and practice questions

**studyon**

Unit 3 | AOS 1 | Topic 3 | Concept 2

**See more**
Primary structure of proteins




**studyon**

Unit 3 | AOS 1 | Topic 3 | Concept 3

**Higher levels of protein structure**
Summary screen and practice questions

## The structure and shape of proteins

Proteins are very large molecules; some contain thousands of amino acids. These large molecules fold and organise into different shapes. Protein structure is described at four different levels of organisation (see figure 2.15).

### Primary structure

The primary structure of a protein is the specific linear sequence of amino acids in the protein. Different proteins have different primary structures and hence have different functions. The sequence of amino acids in a protein is determined by the genetic material in the nucleus (this is discussed further in the next section).

### Secondary structure

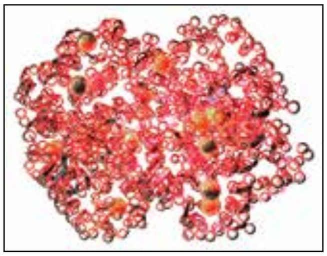
Amino acid chains fold in three different ways (figure 2.15b). Hydrogen bonds form between segments of the folded chain that have come close together and help stabilise the three-dimensional shape. The following are some examples of secondary structure:

1. The major proteins of wool are keratins that have a spiral secondary structure, known as an **alpha helix**. If the fibre is stretched and the hydrogen bonds are broken, the fibre becomes extended. If the fibre is then 'let go,' the hydrogen bonds reform and the fibre returns to its original length.
2. The major protein of silk is fibroin, which is fully extended and lacks the coiling found in the structure of wool. The silk molecules form a **beta-pleated sheet** (see figure 2.15b). The polypeptide chains of silk are already extended and cannot be extended further.
3. The secondary structure of portions of a protein is called **random coiling** if the portions do not conform to the shape of an alpha helix or a beta-pleated sheet.

**study on**

Unit 3
AOS 1
Topic 3
Concept 3

**Do more**
Protein structure



## ODD FACT

The secondary structure of myoglobin, the oxygen-binding protein of muscle, consists mainly (75%) of a coiled alpha helix structure.

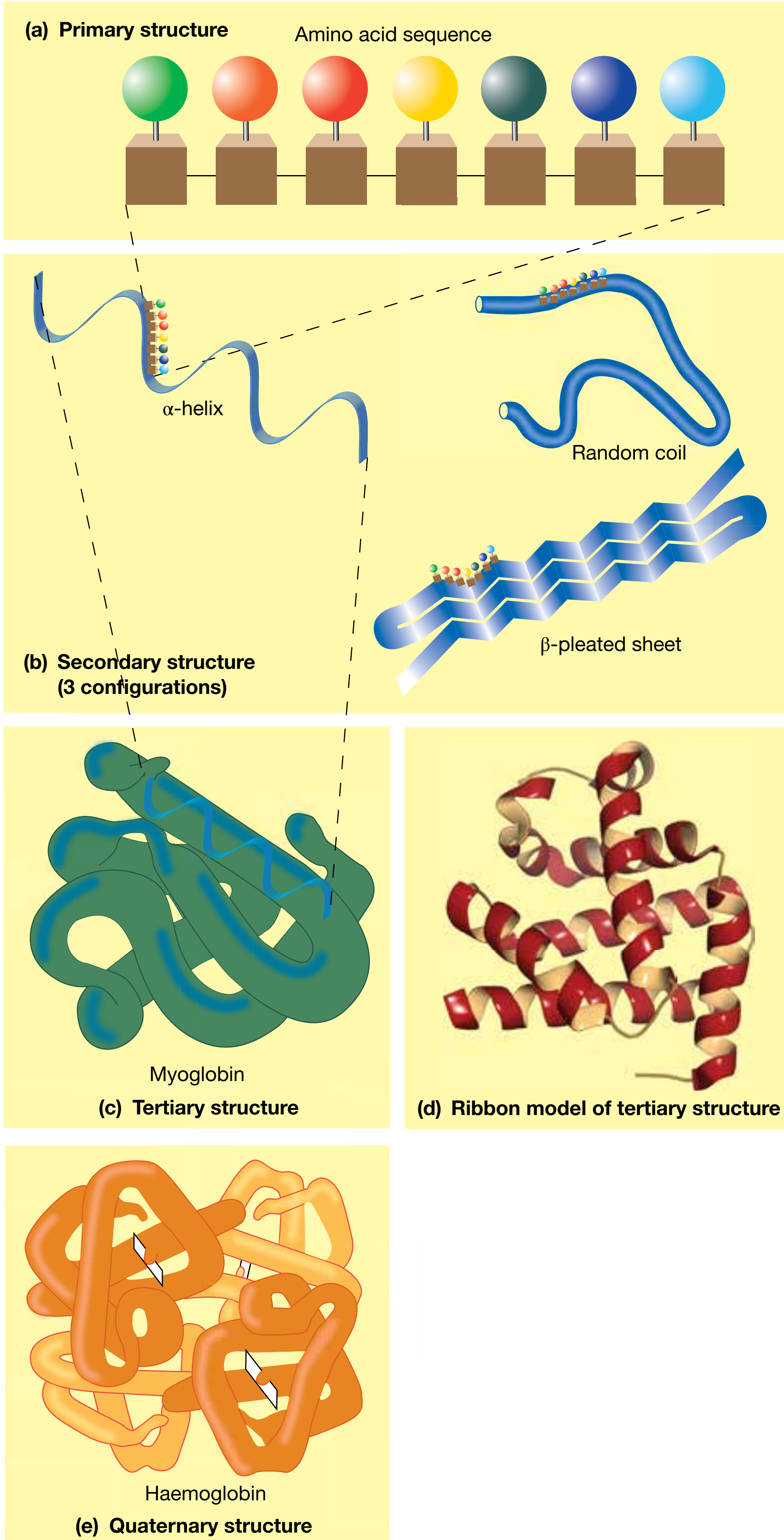



**FIGURE 2.15** The four levels of protein structure: **(a)** primary — order of amino acids in the molecule; **(b)** secondary — folding of some portion of the amino acid chain (note the three different modes); **(c)** tertiary — describes the shape of the entire polypeptide chain; **(d)** tertiary ribbon model — note that some of the secondary structure of this protein can also be seen, namely the regular coils, each showing an alpha coil or an alpha helix; **(e)** quaternary — some proteins comprise a number of polypeptide chains.

The forces that maintain the tertiary structure of proteins are:
1. hydrogen bonds
2. ionic attractions between charged R groups
3. interactions between hydrophobic R groups in the protein interior
4. covalent disulfide cross links.

## Tertiary structure

The tertiary structure refers to the total irregular folding held together by ionic or hydrogen bonds forming a complex shape such as that of myoglobin. The bonds form between side chains of amino acids to form a complex internal structure.

The 3D shape that constitutes the tertiary structure of a protein is critical for its function. For example, if the shape of an enzyme is changed, particularly at its active site, the protein can no longer function as an enzyme.

The regular folding of alpha coils or beta-pleated sheets that is present in the secondary structure of some proteins may be seen when you examine the overall tertiary structure of the protein. Figure 2.15d shows a ribbon model of the tertiary structure of the muscle protein myoglobin. Note the regular coils that form part of this protein — each of these is an alpha helix or alpha coil that is part of its secondary structure. Other regions of the myoglobin molecule do not have a regular secondary structure — these regions are shown as thin lines with random shapes.

## Quaternary structure

Quaternary structure describes a structure in which two or more polypeptide chains interact to form a protein. The resulting structure can be, for example, **globular** as in haemoglobin (figure 2.15e and figure 2.16) or **fibrous** as in collagen, the most common protein in skin, bone and cartilage.

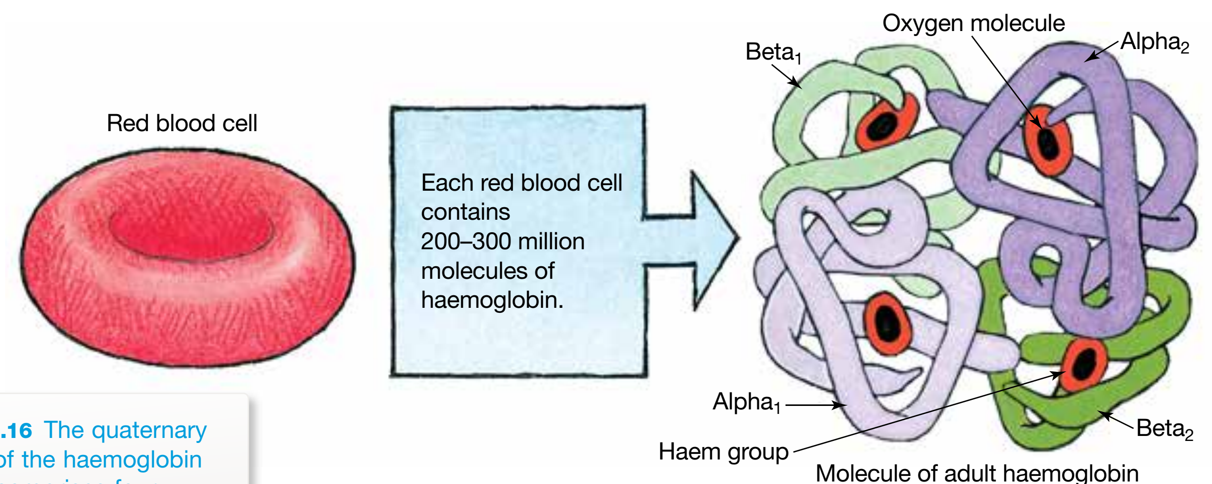



**FIGURE 2.16** The quaternary structure of the haemoglobin molecule comprises four chains: two alpha chains and two beta chains. Note how many of these molecules are present in each red blood cell.

Table 2.3 lists a number of different types of protein and their functions. Some examples of proteins are also shown in figure 2.17.

**TABLE 2.3** Examples of proteins and their functions.

| Type of protein | Function | Example |
|---|---|---|
| structural | fibrous support tissue in skin, bone, tendons, cartilage, blood vessels, heart valves, and cornea of the eye | collagen, keratin |
| enzyme | catalyse reactions | ATP synthase |
| contractile | muscle movement | myosin, actin |
| immunoglobulin | defence against disease | antibodies |
| hormone | regulate body activity | insulin |
| receptor | respond to stimuli | insulin receptors |
| transport | carry other molecules | haemoglobin |

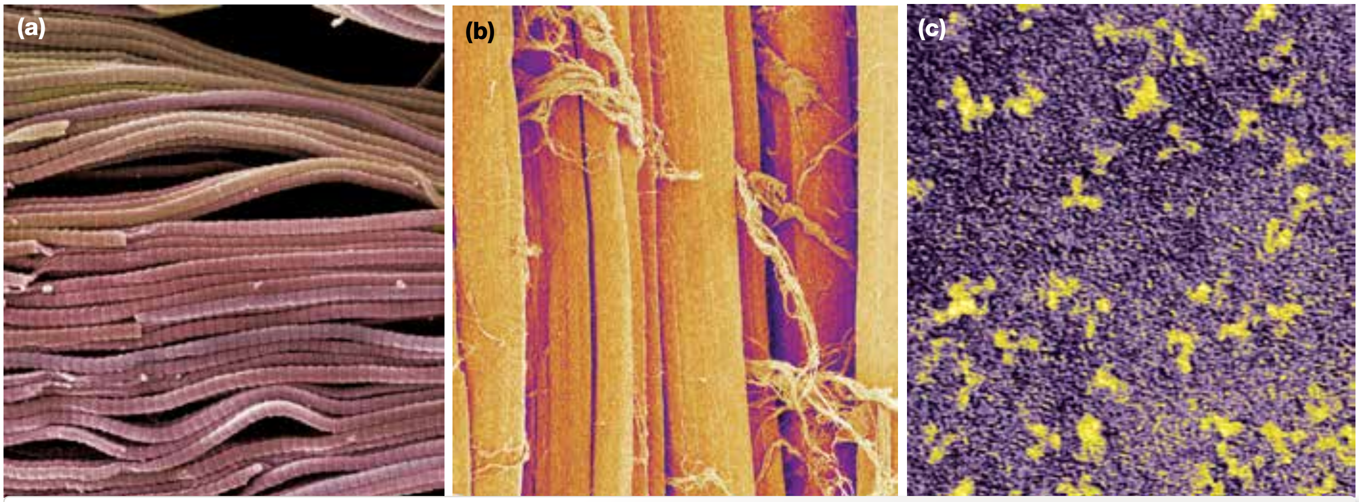



**FIGURE 2.17** **(a)** A scanning electron micrograph (SEM) of collagen bundles from connective tissue that wraps around and supports nerve fibres (notice the characteristic banding of collagen fibres). **(b)** A scanning electron micrograph (SEM) of skeletal muscle fibre showing the thick filaments that are made up of myosin. **(c)** A transmission electron micrograph (TEM) showing Y-shaped structures (yellow), which are molecules of the immunoglobulin G antibody, an important part of the immune system.

## FORENSICS — DETECTING HAEMOGLOBIN

The popularity of television shows about forensics has brought some knowledge about the detection of blood into many homes. Turn off the light, spray a chemical at the crime scene and examine the scene for signs of luminescence. The luminescence produced is because the chemical used reacts with any haemoglobin present (figure 2.18).

Although the chemical Luminol has been the star in many television shows, a new star has been developed. BLUESTAR® Forensic is an improved product that gives a brighter luminescent result visible in normal light. It is also nontoxic and causes no change in the DNA so that minute quantities that are present can also be used for DNA-typing.

Although the presence of blood is indicated, further testing is required to confirm that the substance luminescing is in fact blood, because other compounds also react with the chemical. The important factor is that the position of minute blood remains are indicated even though no blood was visible to the naked eye. Without this, testing for blood would not be possible.

**FIGURE 2.18** Images from a crime scene investigation. Although no blood was visible to the naked eye, use of BLUESTAR® Forensic detected the presence of minute levels of haemoglobin. The luminescence indicates possible blood. Additional tests can prove it to be so. (Images courtesy of BLUESTAR® Forensic)

eBook*plus*

**eLesson**
eles-1558
**Forensic scientist**

## Conjugated proteins

Many of the proteins we have described above are *simple* proteins — the final molecule contains amino acids only. **With some proteins, the chains of amino acids conjugate with other groups.** This is particularly the case for those proteins in the nucleus. They are mostly **nucleoproteins** — they comprise a molecule containing both protein and nucleic acid, as, for example, nucleosomes that are complexes of DNA coiled around a group of eight histone proteins (refer back to figure 2.5b).

Another conjugated protein is haemoglobin. Each tertiary structure component produced associates with a haem group. The quaternary molecular structure comprises four chains: two alpha chains and two beta chains. The amino acid sequence in a protein is important. If the order of amino acids in either chain is altered, a defective chain results. An individual inherits, from each parent, the DNA that encodes the beta chain. If a defect in this DNA is inherited from both parents, an individual is unable to produce any normal haemoglobin and has the genetic disorder beta thalassaemia.

## Inactive to active molecule

Insulin, produced by beta cells in the pancreas, is a protein hormone. It controls the level of glucose in the blood by facilitating the uptake of glucose from the blood into tissue cells. When this occurs the level of blood glucose declines. An insulin molecule comprises two chains of amino acids held together by disulfide bonds (figure 2.19a). This is the active state of the hormone.

When initially produced, insulin is inactive as a hormone. It is produced as a single chain of amino acids with folds that are held together by three disulfide bonds (figure 2.19b). A section of this molecule is removed by an activating enzyme leaving the active hormone as two chains of amino acids held together by the three disulfide bonds.

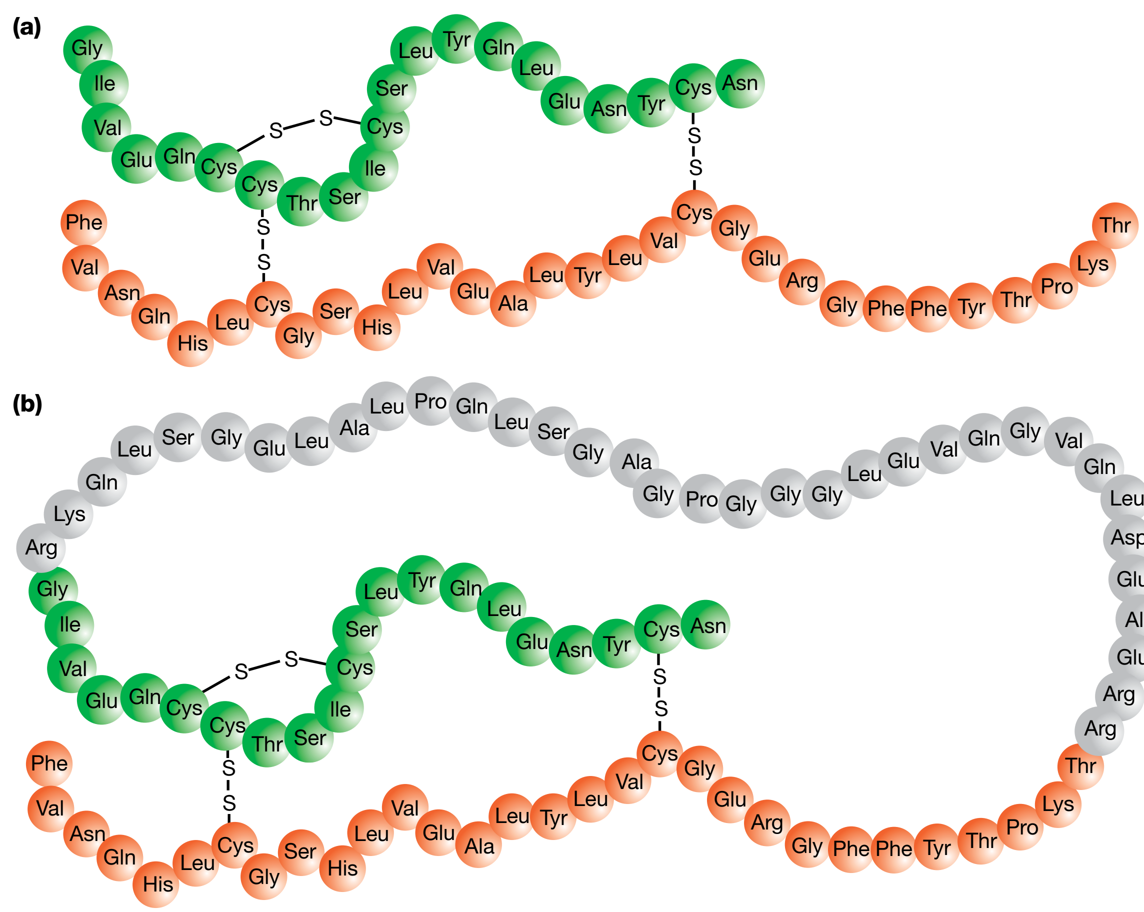



**FIGURE 2.19** **(a)** A molecule of insulin comprises two chains of amino acids held together by three disulfide bonds. This is the active state of the molecule. **(b)** Inactive insulin is produced as a single chain of amino acids folded on itself and held together by the three disulfide bonds.

ODD FACT

There has also been a change in emphasis regarding the genetic material. The focus on single genes has been replaced with a global analysis of the *genome*, the total genetic material. This study is *genomics*.

Hence, although a molecule may be made from a number of molecules linked together by sulfide or other bonds, they may be derived from the same initial inactive protein. Many enzyme proteins are produced in an inactive form and become active only if the appropriate enzyme is present to convert them for active service. Can you think of a situation that you considered in your previous studies that involved an enzyme being converted from the inactive to the active state?

## What is the proteome?

In living organisms, proteins are involved in one way or another in virtually every chemical reaction. They may be the enzymes involved, they may be the reactants or the products, or they may be all three. **The complete array of proteins produced by a single cell or organism in a particular environment is called the proteome** of the cell or organism. The study of the proteome is called **proteomics**.

Scientists are moving away from investigating single proteins because no protein acts in isolation from other proteins. They are now exploring the total pattern of proteins produced by a cell and analysing these patterns to compare them with patterns from different kinds of cells. What are the differences? What are the similarities? What is the proteome profile of diseased tissue or even the fluids surrounding the tissue? In what ways do they differ from the healthy state?

There is also an emphasis on structural proteomics (see figure 2.20). **Knowing that a protein exists is different from knowing how it operates.** Knowing how some operate and knowing about their structures may enable testable predictions about the roles of other proteins on the basis of their structures.

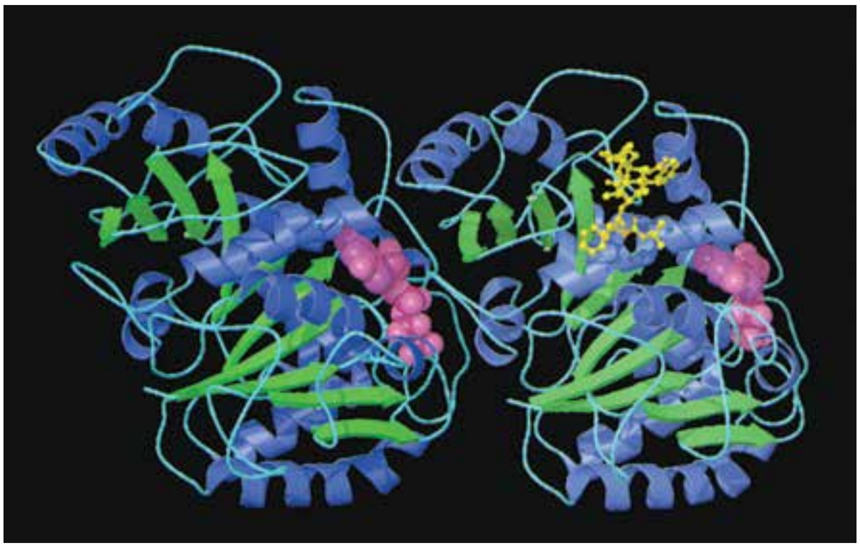


**FIGURE 2.20** Knowledge of three-dimensional structures of well-known proteins may give insights into their functions. This three-dimensional model shows the overall tertiary structure of two tubulin molecules. Tubulin, a globular protein, is a component of microtubules that are found, for example, in the cilia of the human lung. The secondary structure of tubulin is denoted by the blue regions showing the alpha helices, the green regions showing the beta-pleated sheets, and the thin lines denoting regions of random coiling. (The pink denotes the nucleotide GTP, which binds to the protein.) (Image courtesy of Lawrence Berkeley Laboratory.)

KEY IDEAS

- All proteins contain N, C, H and O, and some contain S.
- The structure of all proteins can be identified at the primary, secondary and tertiary levels, while proteins that consist of more than one polypeptide chain have an additional quaternary level of structure.
- Proteins can also be classified on the basis of their different functions.
- The proteome of an organism is the complete array of proteins produced by that organism.

QUICK CHECK

**15** What is the basic formula of an amino acid molecule?
**16** How is a peptide bond formed?
**17** List the four basic structures of protein molecules and draw an example of each.
**18** Give an example of (a) a structural protein, (b) a contractile protein and (c) a conjugated protein.
**19** Why is proteomics considered important?

study on

Unit 3
AOS 1
Topic 4
Concept 3

**Exons and introns**
Summary screen and practice questions

# A closer look at a gene

**The part of a gene that contains the coded information for making a protein is called the coding region** of a gene. The regions on either side of the coding region of a gene are called **flanking regions**. The flanking region before the start of the coding region is called the upstream region. The flanking region after the end of the coding region is called the downstream region.

The box below provides further information about these regions.

## Introns: just an interruption or two

An unexpected discovery about the genes in eukaryotes was made in 1977. Until then, the coding region of a gene was thought to be continuous (see figure 2.21a).

Instead, the coding region is 'interrupted' by other segments of DNA. **Each segment of the coding region of a gene is called an exon**. The exons are separated by lengths of DNA that do not contain instructions relating to the protein chain. These **noncoding segments are called introns** (see figure 2.21b).

To avoid confusing the two terms *exon* and *intron*, think about INterruption and INtron and that will help you remember the functional difference. Introns are transcribed but not translated.

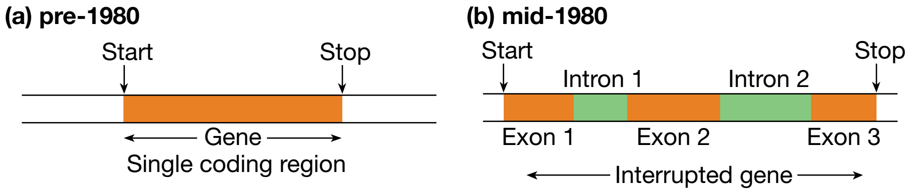



**FIGURE 2.21** **(a)** Pre-1980 view of a gene, with a continuous coding region. **(b)** Actual situation with coding region (orange) broken into segments (exons) by noncoding segments (green).

### CODING AND FLANKING REGIONS

If it were possible to travel in a miniaturised vehicle along a gene, what would we see?

#### The coding region

The coding region of a gene is the segment of DNA double helix that includes the DNA template strand, which encodes the information that will later be translated into the amino acid sequence of a polypeptide. This region of a DNA template strand begins with a start signal (TAC) and, some distance away, there is a stop signal (ATT or ATC or ACT).

#### Upstream from the coding region

The region of DNA on the template strand upstream from the coding region contains some particular base sequences. One part of the upstream region is rich in As and Ts and is often called the **TATA box**, because the sequence TATA AA (or similar) occurs there. Another sequence commonly found further upstream is called the CAT or CAAT box.

***Role of upstream sequences***

Consider the following observations:

- Upstream sequences are invariably found in all organisms. It is reasonable to suggest that these upstream sequences serve an important function since they have been maintained during evolution.
- If upstream sequences are altered by mutation, the activity of the coding region of the gene may be reduced or even become inactive. The absence of the correct upstream signal is a cause of some inherited human diseases. One form of thalassaemia is due to a missing TATA group in the upstream region of the DNA of both copies of the specific gene in the people concerned.
- The upstream region includes segments of DNA to which hormones can attach. The fact that some hormones can bind to DNA provides one clue as to how hormones can influence the action of genes.

These observations support the conclusion that specific DNA sequences upstream of the coding region of a gene initiate transcription, the process by which the encoded information in the DNA coding region is transcribed into mRNA. **Promoters** also act as sites where proteins called transcription factors can bind and regulate the expression of genes.

*(continued)*

### Downstream from the coding region

The DNA following the end of the coding region is referred to as the 'downstream' region (see figure 2.22). About 20 bases downstream, the sequence AATAAA is usually found. If this sequence is altered, the gene action is altered. The downstream region includes an 'end transcription signal', which terminates the process of transcription of mRNA from the DNA template.

Critical sites where mutations can affect normal functioning of a gene include:
- the upstream promoter region
- the START and STOP signals.

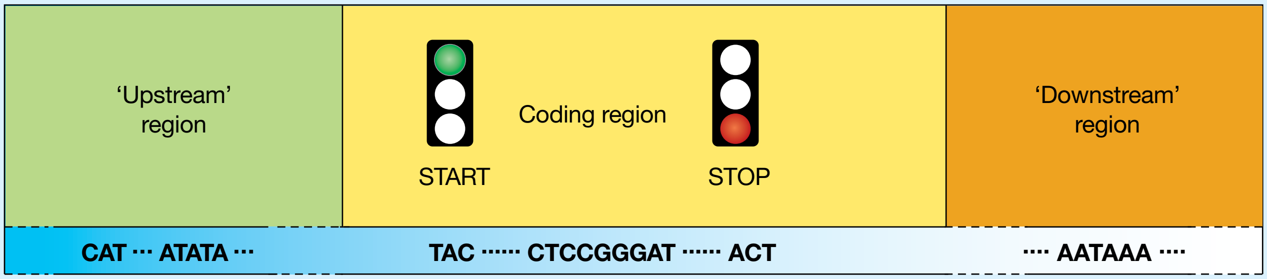



**FIGURE 2.22** Regions of the template strand of a typical gene.

So, eukaryote genes are not like nursery rhymes in a book, where the reader starts at the beginning and reads through to the end. The information in genes is broken up into segments and the sections in between are filled with other printed material that is unrelated to the rhyme.

Here is an interrupted rhyme:

'<u>Hey, diddle diddle the cat and the fid</u> HERE IS AN INTERRUPTION <u>dle, the cow jumped over the</u> AND HERE IS ANOTHER INTERRUPTION <u>moon. The little dog laughed to see such fun and the dish</u> HERE'S ANOTHER <u>ran away with the spoon</u>.'

If this interrupted rhyme were thought of as a gene, how many exons and how many introns would it contain? The underlined portions are like exons, and there are four of them. The interruptions are like introns, and there are three of them; they are removed from the mRNA before translation.

The number of exons and introns in genes varies. The DNA making up the **HBB** gene, which controls the production of one chain of the haemoglobin molecules, consists of three exons and two introns. The **F8C** gene, which controls the production of factor VIII, which assists in blood clotting, consists of 26 exons and 25 introns.

## KEY IDEAS

- Each gene in eukaryote organisms contains a coding region, and also includes flanking regions upstream and downstream of the coding region.
- The coding region of a gene typically consists of several exons separated or interrupted by introns.

## QUICK CHECK

**20** Using words or diagrams, distinguish between the members of each of the following pairs:
   **a** intron and exon
   **b** coding region and flanking region.

**21** True or false? All genes contain the same number of exons.

studyon

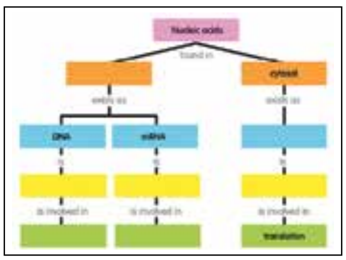
Unit 3 | AOS 1 | Topic 3 | Concept 8

**Protein synthesis**
Summary screen and practice questions

studyon

Unit 3 | AOS 1 | Topic 3 | Concept 8

**Do more**
Protein synthesis



**ODD FACT**

Various kinds of RNA occur in cells, including messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA) and small nuclear RNA (snRNA). By far the most common is rRNA since it forms part of every ribosome, a cell organelle that is present in large numbers in the cell cytoplasm. rRNA is also found in the nucleolus.

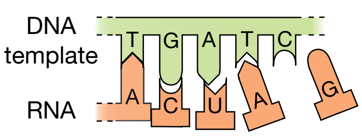



**FIGURE 2.23** Because of pairing of complementary bases in DNA and RNA, one DNA chain can act as a template to build an RNA chain with a predictable nucleotide sequence. What does an A in DNA put in place in RNA?

# Protein synthesis

The genetic instructions for producing proteins are found within the DNA of the chromosomes. A gene is a segment of DNA that codes for a protein. In eukaryote organisms how do genetic instructions get from the nucleus to the ribosomes?

When a gene becomes active, it first makes a mobile copy of the coded instruction that it contains. This occurs by a process known as **transcription**. This mobile copy of a genetic instruction can leave the nucleus and move to the cytoplasm where the instruction is decoded. This occurs by a process known as **translation**. So, in the case of protein-encoding genes, **gene action** involves two processes: transcription, which occurs in the nucleus, and translation, which occurs on ribosomes in the cytoplasm.

# Transcription: copying the original

The nucleus of a eukaryote cell is like a safe that contains the genetic masterplan in the form of DNA. The genetic masterplan containing the entire set of instructions for an organism is like the complete plan for the construction of a complex structure, such as a jumbo jet. One gene or instruction for a particular protein is like the plan for making one component of the jet, such as a wing flap.

The workers at the site where the wing flaps are made do not work directly from the complete masterplan; instead, they have copies of the relevant section of the plan. Likewise, **before a genetic instruction in DNA is decoded, that instruction is copied (transcribed) from the genetic masterplan, which remains in the nucleus.** This copy is encoded in a different nucleic acid called ribonucleic acid (RNA). Because the role of this particular RNA is to carry a copy of a genetic instruction from the nucleus to the cytoplasm, it is known as **messenger RNA (mRNA)**.

Pairing or hybridisation can occur between the bases in one DNA strand and complementary bases in an RNA strand as follows:

| | | |
|---|---|---|
| A in DNA | pairs with | U in RNA |
| T | pairs with | A |
| C | pairs with | G |
| G | pairs with | C. |

This pairing means that a DNA chain can act as a template to guide the construction of RNA with a complementary base sequence (see figure 2.23). This means that the genetic information in DNA can be accurately copied into RNA during the process of transcription.

Consider a DNA template with the base sequence:

DNA template . . . **ATGCCTGAAT** . . .

This DNA can act as a template to guide the formation of an RNA molecule with the complementary base sequence as follows:

mRNA transcript (copy) . . . **UACGGACUUA** . . .

The base sequence of the mRNA primary transcript is not identical to that of the template DNA strand; instead, the mRNA has a complementary sequence. However, the mRNA base sequence matches that of the complementary nontemplate DNA strand, except that U replaces A. For this reason, the nontemplate DNA strand is also called the coding DNA strand.

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Nontemplate DNA: | A | T | C | C | C | G | G | T | A | A | C | A |
| Template DNA: | T | A | G | G | G | C | C | A | T | T | G | T |
| mRNA transcript: | A | U | C | C | C | G | G | U | A | A | C | A |

studyon

| | |
|---|---|
| Unit 3 | **Transcription** |
| AOS 1 | Summary screen and practice questions |
| Topic 3 | |
| Concept 9 | |

## From DNA to mRNA: step by step

A particular gene in the nucleus is switched on at a specific stage of development. In the case of the **HBB** gene, this gene becomes active during late fetal development in certain bone marrow cells. **Transcription takes place in a series of steps** (see figure 2.24).

1. An enzyme, known as **RNA polymerase**, attaches to a specific promotor sequence of DNA in the upstream region of the template strand. The double-stranded DNA of the gene unwinds and exposes the bases of the template strand.
2. The base sequence of the DNA template guides the building of a complementary copy of the mRNA sequence. The RNA polymerase enzyme moves along the DNA template in a 3′ to 5′ direction and, as it moves, complementary nucleotides are brought into place and, one by one, are joined to form an RNA chain.
3. After the RNA polymerase moves past the coding region and into the downstream region of the gene, transcription stops and the mRNA molecule is released from the template.

The result of this process is a single-stranded molecule, called pre-messenger RNA. The base sequence in the pre-mRNA molecule is complementary to the base sequence of the DNA of the template strand.

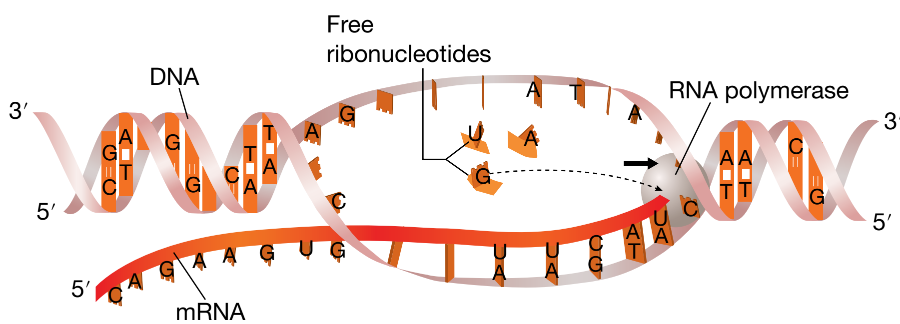



**FIGURE 2.24** Transcription occurs in the cell nucleus. The enzyme RNA polymerase moves along the DNA template building an mRNA molecule at the rate of about 30 bases per second.

**ODD FACT**

Splicing of pre-mRNA is carried out by a complex known as the spliceosome. This complex consists of protein and snRNA.

## Pre-mRNA is modified after transcription

The primary product of translation is pre-mRNA, also known as the primary transcript. The sequence of bases in the pre-mRNA is complementary to all the DNA bases of a gene, both introns and exons. The primary mRNA transcript then undergoes a process termed **post-transcription modification** (see figure 2.25) before it leaves the nucleus.

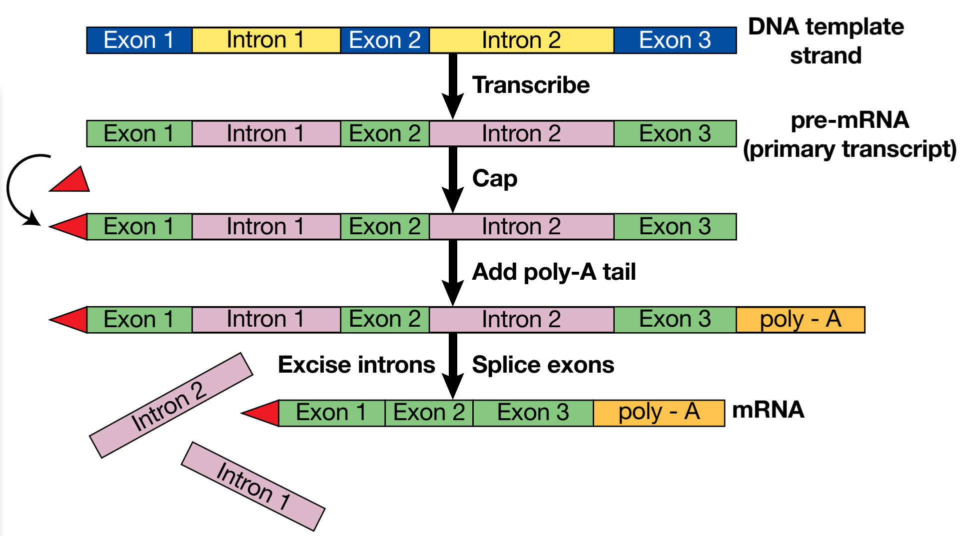
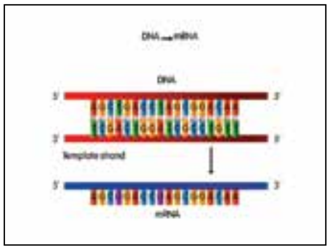



**FIGURE 2.25** Various processes in the post-transcription modification of pre-mRNA (primary transcript) produce mRNA. Post-transcription modification occurs in the nucleus. The **HBB** gene comprises three exons and two introns and, in the post-transcription modification of the primary transcript, the two introns are excised and the three exons are spliced to form the mRNA.

studyon

Unit 3 | AOS 1 | Topic 3 | Concept 9

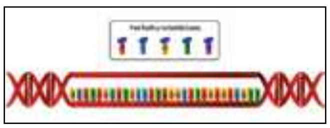
**See more** Transcription



studyon

Unit 3 | AOS 1 | Topic 3 | Concept 9

**Do more** Construct mRNA molecule



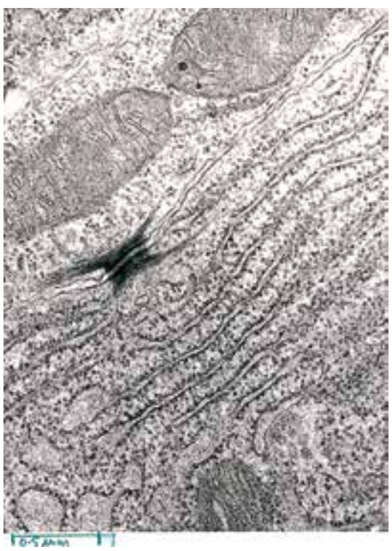


**FIGURE 2.26** Ribosomes are composed mainly of a type of ribonucleic acid known as ribosomal RNA (rRNA). In this electron photomicrograph, ribosomes appear as dark dots located on the endoplasmic reticulum. (Image courtesy of Dr Maret Vesk.)

Post-transcription modification includes the following processes:

- *Capping:* The 5′ end of the pre-mRNA is capped with an altered guanine (G) base (methyl guanosine). The cap protects the pre-mRNA from enzyme attack and contributes to its stability.
- *Adding a tail:* The primary transcript is clipped at a specific point downstream of the coding region and a poly-adenine (A) tail, with up to 250 As, is then added at the 3′ end. The tail contributes to the stability of the mRNA.
- *Splicing:* The regions in the pre-mRNA that correspond to the introns are cut out and the remaining exons are spliced together. This cutting and splicing is done by **spliceosomes**, which recognise specific base sequences at the ends of the introns: GU at the 5′ end, and AG at the 3′ end.

Post-transcription occurs within the nucleus. The final mRNA product now moves across the nuclear membrane into the cytosol, carrying with it a copy of the information originally encoded in the DNA of the gene.

In the next section, we will examine how the genetic information that is copied into mRNA is decoded (translated) into a particular protein chain.

**KEY IDEAS**

- Coded genetic instructions are located in the DNA of the nucleus of eukaryote organisms.
- DNA and RNA are both nucleic acids, but differ in several ways.
- During transcription, the information in the template strand of the DNA of a gene is copied into a pre-RNA molecule.
- The final mRNA molecule results when post-transcription modification is complete, including intron excision and exon splicing.

**QUICK CHECK**

**22** Where does transcription occur? What is the end product of this process?

**23** A template strand of DNA includes the base sequence:
. . . T A T C G G C A T . . .
Write the base sequence of the complementary nontemplate strand.

**24** A strand of mRNA includes the base sequence:
. . . A U G U A U C C G . . .
Write the base sequence of the DNA coding strand.

**25** Where does transcription occur in a cell?

**26** List two differences between pre-mRNA and mRNA.

# Translation: decoding genetic instructions

The decoding of the genetic instructions occurs through the process of translation, which takes place in the cytoplasm. By the end of this process, the genetic instructions carried in mRNA have been decoded and translated into a protein chain built of amino acids. The coded instruction in the mRNA is not changed in this process, just as the plan of a jumbo jet part is unaltered after the part is made.

Translation involves the combined action of several agents (see table 2.4).

The mRNA moves from the nucleus to the cytoplasm where it attaches to submicroscopic organelles known as **ribosomes** (see figure 2.26).

**TABLE 2.4** Players and places in translation. (Many enzymes are also involved.)

| Agents | Analogy |
|---|---|
| DNA in the nucleus | masterplan with complete set of instructions |
| mRNA | working copy of one instruction |
| ribosomes | construction site |
| tRNA | carriers of raw material |
| amino acids | raw material |
| protein chain | end product |

The construction of a protein according to the coded instructions in mRNA involves the assembly of amino acid sub-units. The various **amino acids** are present in solution in the cytosol. How are the correct amino acids selected from this solution?

Each amino acid is brought to the mRNA on the ribosomes by a carrier molecule called **transfer RNA (tRNA)**. Each tRNA molecule consists of a single strand of 76 nucleotides coiled and paired with themselves. At one end of each tRNA molecule are three bases that make up an **anti-codon**. At the other end of a tRNA molecule is a region that attaches to one specific amino acid (see figure 2.27). An enzyme, amino acyl tRNA synthetase, catalyses the linking of each amino acid to its specific tRNA carrier.

In this text, the term *protein chain* is used to refer to a single chain built of amino acid sub-units. Polypeptide is a term for a short chain consisting of amino acids.

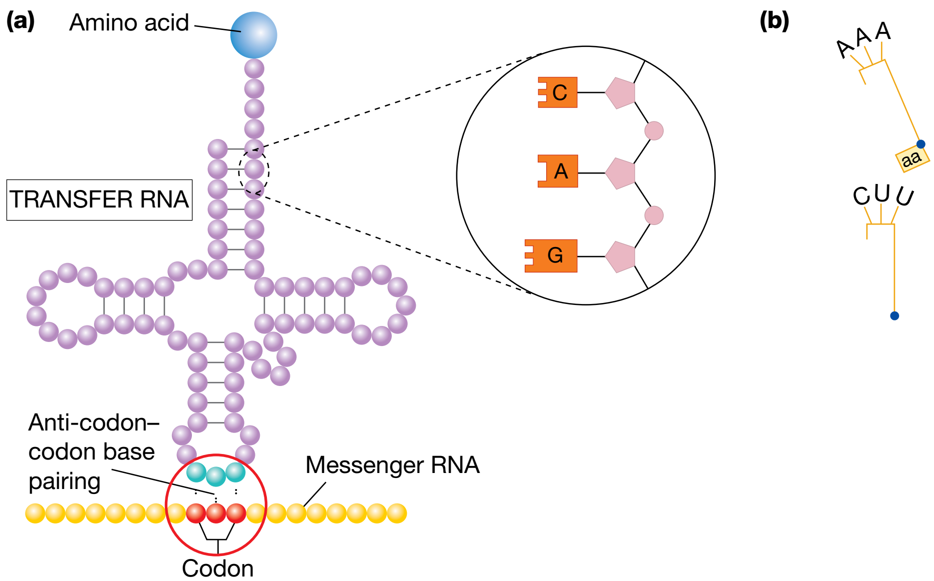



**FIGURE 2.27** **(a)** Transfer RNA (tRNA) molecule. At one end is an amino acid attachment site and, at the opposite end, an anti-codon region. How many bases make up an anti-codon? **(b)** A simpler representation as used in this textbook.

## From mRNA to protein: step by step

The information in mRNA is present in coded form as sets of three bases or triplets. These triplets, such as AGG and UCU, are called **codons**. **Most codons contain the information to add one specific amino acid to a protein chain.** In addition, one codon (AUG) is a START TRANSLATION instruction, and three different codons (UAA, UAG and UGA) are STOP TRANSLATION instructions (see table 2.5).

The instructions in an mRNA molecule are decoded three bases (or one codon) at a time. Translation begins at the 'start adding amino acids' signal (AUG codon) (see figure 2.28a). This codon both starts the process of building a protein chain and puts the amino acid *met* into place as the first amino acid in the chain. Then, the next three bases are translated by adding the next amino acid to the growing protein, and so on (see figure 2.28b).

The colour coding in table 2.5 groups amino acids according to their side chains; for example, those with a positive charge are shown in blue.

**TABLE 2.5** Genetic code shown as the 64 mRNA codons and information they specify. (See Appendix for the full names of the amino acids.) The codon AUG is a start signal and it also codes for the amino acid *met*. Note the three stop signals.

| mRNA codon | Amino acid | mRNA codon | Amino acid | mRNA codon | Amino acid | mRNA codon | Amino acid |
|---|---|---|---|---|---|---|---|
| UUU | *phe* | UCU | *ser* | UAU | *tyr* | UGU | *cys* |
| UUC | | UCC | | UAC | | UGC | |
| UUA | *leu* | UCA | | UAA | STOP | UGA | STOP |
| UUG | | UCG | | UAG | | UGG | *trp* |
| CUU | | CCU | *pro* | CAU | *his* | CGU | *arg* |
| CUC | | CCC | | CAC | | CGC | |
| CUA | | CCA | | CAA | *gln* | CGA | |
| CUG | | CCG | | CAG | | CGG | |
| AUU | *ile* | ACU | *thr* | AAU | *asn* | AGU | *ser* |
| AUC | | ACC | | AAC | | AGC | |
| AUA | | ACA | | AAA | *lys* | AGA | *arg* |
| AUG | START /*met* | ACG | | AAG | | AGG | |
| GUU | *val* | GCU | *ala* | GAU | *asp* | GGU | *gly* |
| GUC | | GCC | | GAC | | GGC | |
| GUA | | GCA | | GAA | *glu* | GGA | |
| GUG | | GCG | | GAG | | GGG | |

**study on**

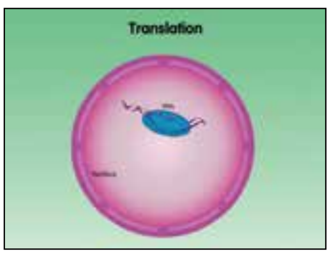
Unit 3
AOS 1
Topic 3
Concept 10

**Translation**
Summary screen and practice questions

**study on**

Unit 3
AOS 1
Topic 3
Concept 10

**See more**
Translation



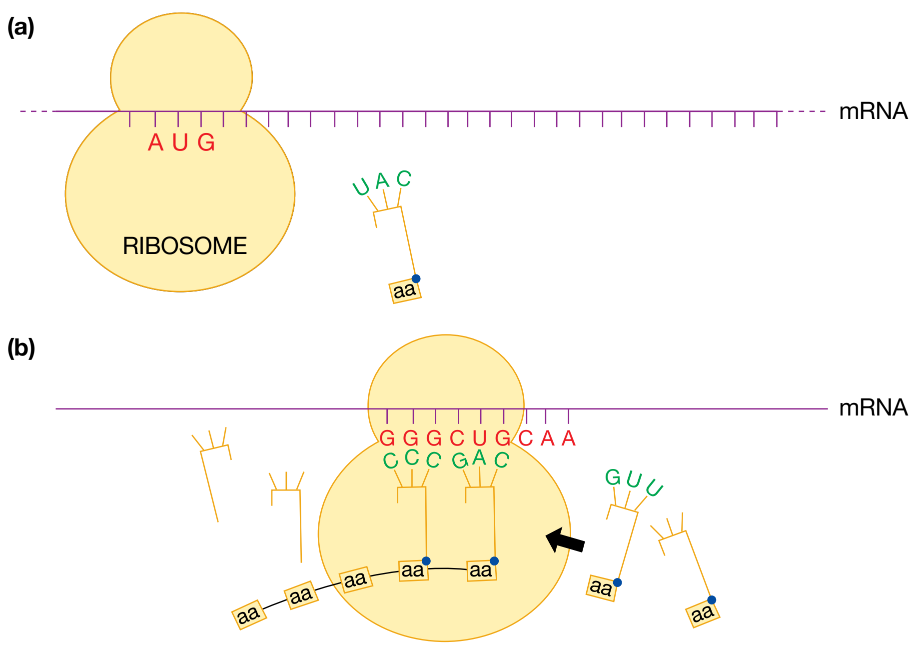




**FIGURE 2.28** **(a)** The mRNA molecule attaches to a ribosome. In turn, as the ribosome moves along the mRNA molecule, each codon pairs with the tRNA with the complementary anti-codon. **(b)** The amino acids carried by each tRNA molecule are joined to form a chain. The final product is a protein consisting of a chain of amino acid sub-units, which, if short, is termed a polypeptide.

**ODD FACT**

Messenger RNA (mRNA) formed during gene transcription has a short life. This is in contrast to ribosomal RNA (rRNA), which forms part of the ribosomes and is very stable.

**As each codon is translated, the tRNA molecule with the complementary anti-codon pairs momentarily with the mRNA.** The pairing between bases in codons and the complementary bases in an anti-codon is as follows:

A pairs with U
U pairs with A
C pairs with G
G pairs with C.

So, when the mRNA codon UUU is reached, the tRNA carrier molecule that has the anti-codon AAA comes into place with its specific cargo of the amino acid *phe*. The amino acid carried by that tRNA is brought into the correct position to be joined into the growing protein chain. Amino acids continue to be added until a STOP signal is reached, which stops the addition of amino acids to the protein chain.

What is the mRNA codon for STOP? Will this have a corresponding anti-codon on a tRNA carrying an amino acid?

## Putting it all together: transcription and translation

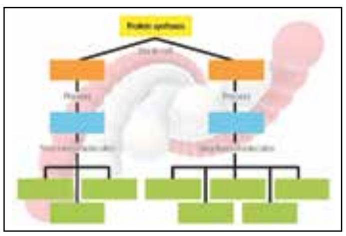
The production of a particular protein, such as beta chains of haemoglobin, starts in the cell nucleus. It is here that pre-mRNA molecules are produced by transcription from a template DNA chain. The mRNA leaves the nucleus and moves to the cytosol where it becomes attached to ribosomes. It is here that protein chains are formed by translation of mRNA.

Figure 2.29 shows **a summary of the processes of transcription and of translation.** For simplicity, this diagram omits one important step. Can you identify the missing step? By referring to the genetic code shown in table 2.5, identify the amino acid that is about to be added to the growing protein chain in this diagram.

**study on**

Unit 3
AOS 1
Topic 3
Concept 10

**Do more**
Comparing transcription and translation



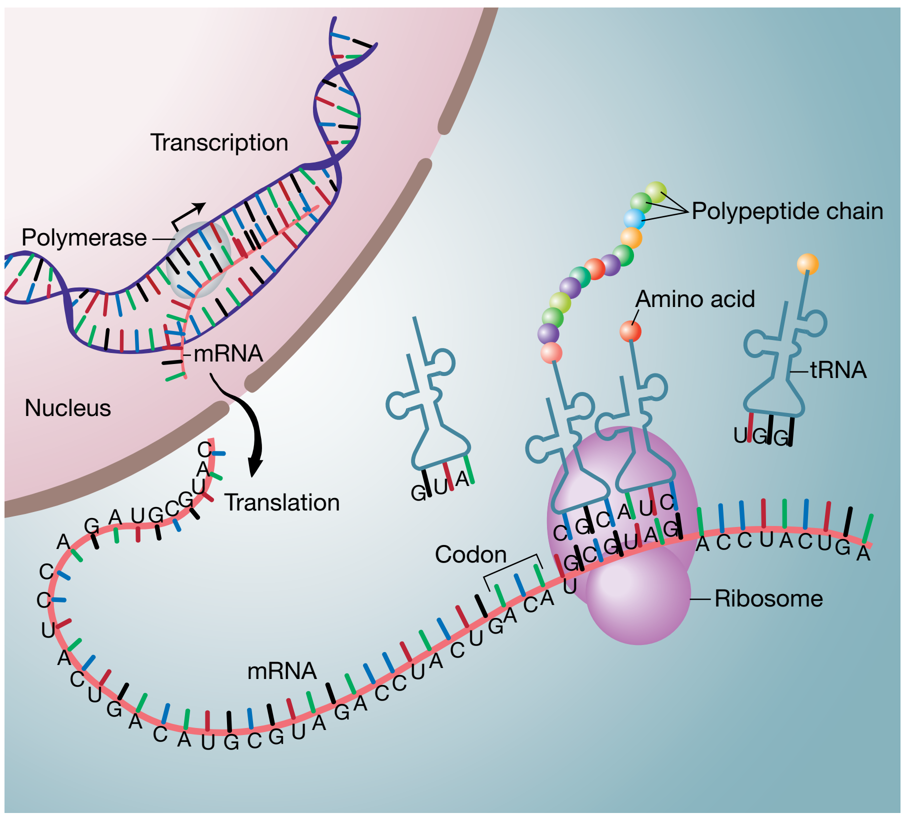



**FIGURE 2.29** Representation of the processes of transcription and translation in a eukaryote cell. (Image courtesy of the National Human Genome Research Institute.) Which of these processes involves the action of three different kinds of RNA?

**ODD FACT**

One scientist has suggested that each gene is like a Swiss army knife; it can do several jobs, depending on how it is handled (or, in the case of a gene, how it is regulated).

## Alternative splicing of pre-mRNA

The human genome contains only about 21 000 genes and this range is typical of other mammals. In the past, it was accepted that one gene had a single function as, for example, producing one particular protein — this is the 'one gene, one polypeptide' concept. The question remains: How can a relatively small number of genes produce the complexity of structure and function of a living mammal?

Research is now revealing that **one gene can be regulated in different ways so that it can produce more than one protein**. This means that:

- one gene could produce one protein at one stage of development but a different protein at another stage of development
- one gene could produce a particular protein in one tissue but a different protein in another tissue.

It is estimated that more than 60 per cent of human genes are regulated to operate in this way. The complexity of mammals is not in the number of genes that they have, but in the complex processes by which their genes are regulated.

How might one gene produce different protein products at different developmental stages or in different tissues? One way is through **alternative splicing** of the pre-mRNA molecules from a single gene.

Alternative splicing can occur during post-transcription modification of pre-mRNA and is the major mechanism enabling the same gene to produce different proteins in different tissues. Usually, alternative splicing involves **exon juggling**, where different exons are combined to form several kinds of mRNA, each with a different base sequence (see figure 2.30).

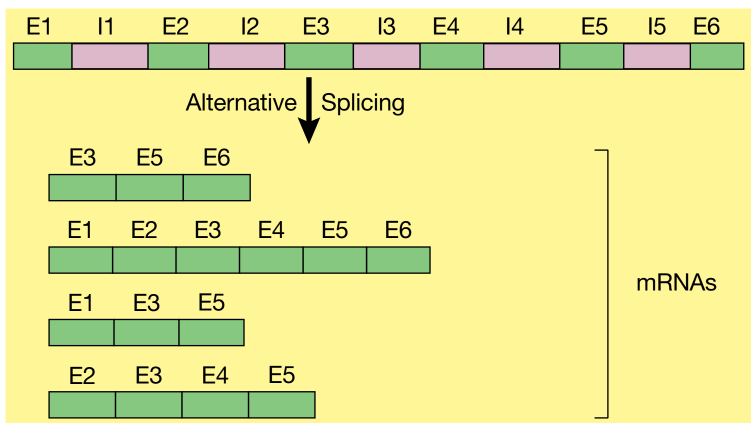



**FIGURE 2.30** Exon juggling of pre-mRNA (E = exon and I = intron). Juggling of exons generates mRNAs that differ in their base sequences and so code for different proteins. Are any mRNAs possible other than those shown?

These different mRNAs are then translated into proteins that differ in their amino acid sequences. For example, the human **TPM** gene has 11 exons. Juggling of these exons produces different mRNAs in different tissues with the result that each tissue has a different form of the protein (see table 2.6).

**TABLE 2.6** Result of exon juggling of the pre-mRNA from one human **TPM** gene in different tissues. The resulting proteins differ in their amino acid sequences. Why?

| Tissue | Exons in mRNA | Excised exons |
|---|---|---|
| skeletal muscle | exons 1, 3–11 | exon 2 |
| smooth muscle | exons 1, 2, 4–9, 11 | exons 3, 10 |
| liver | exons 1, 4–6, 8, 9, 11 | exons 2, 3, 7, 10 |
| brain | exons 1, 4–9 | exons 2, 3, 10, 11 |

**ODD FACT**

Another means of alternative splicing is intron retention. Instead of all introns being excised during post-transcription modification of mRNA, an intron can be retained, leading to another form of the protein.

Alternative splicing means that the number of proteins from the genetic instructions (genes) in a genome is far greater than the number of genes. Identifying how one gene can be regulated in different tissues, or at different times, to produce different products is an exciting area of ongoing research.

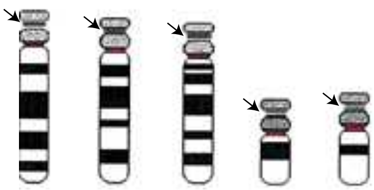
# Genes have various functions

## Structural and regulator genes

Genes vary in the functions that they carry out in the cells of an organism. **Some genes produce proteins that become part of the structure and the functioning** of the organism. These genes are termed **structural genes**.

**Some genes produce proteins that control the action of other genes.** These genes are termed **regulator genes** and their actions determine whether other genes are active ('on') or not ('off') and, if active, the rate at which their products are made.

(a)




(b)

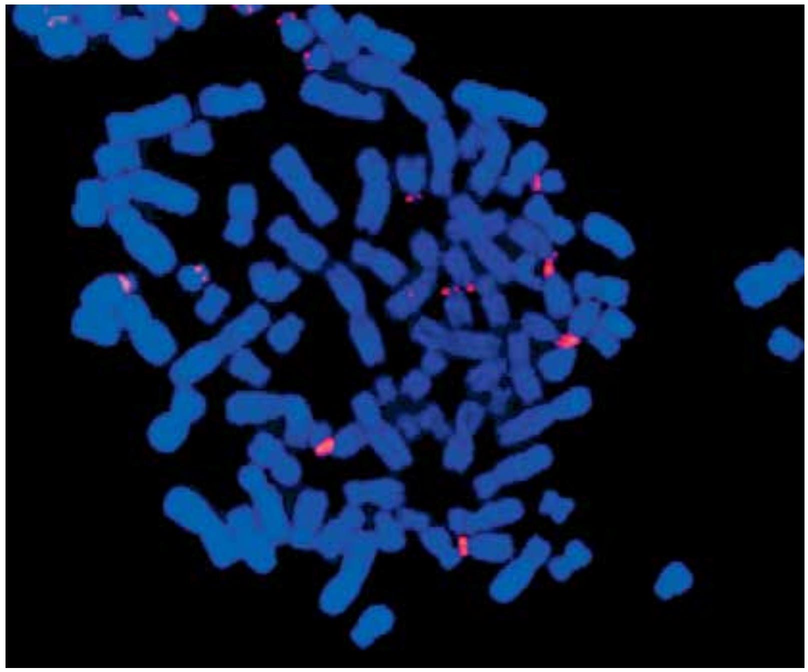


**FIGURE 2.31** **(a)** Human chromosomes 13, 14, 15, 21 and 22. Note the secondary constriction (arrowed) on the short arm. This forms the nucleolar organiser region (NOR) and is where the ribosomal genes are located. **(b)** Photomicrograph of human chromosomes hybridised with a probe that is specific for the gene that controls production of ribosomal RNA (rRNA). The rest of the chromosomes are counterstained with a blue dye. Note the multiple locations of this gene. (Image (b) courtesy of Dr Brian McStay)

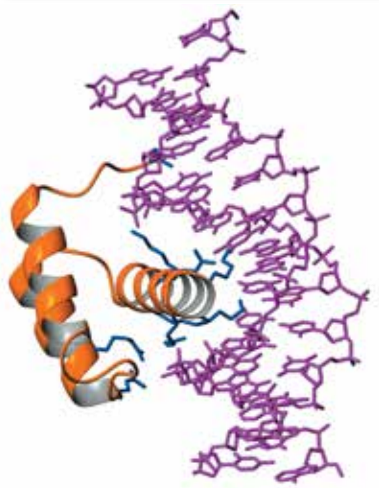


**FIGURE 2.32** Homeotic genes produce a polypeptide with a region that carries + + charges that can bind to DNA. Here, the DNA-binding region of the polypeptide is shown bound to DNA (purple).

Regulator genes switch other genes 'on' or 'off' by producing proteins that act in one of two different ways:

1. Some proteins, known as **DNA-binding proteins**, bind to regions of nuclear DNA near genes and directly switch these genes on or off. These proteins carry a net positive charge that enables them to bind to DNA, which has a net negative charge. By binding to the DNA near their target genes, these proteins can switch these genes on and off (see figure 2.32).
2. Some proteins bind to receptors on the membrane of cells in their target tissue and trigger a series of intercellular reactions that switch genes on or off; these are **signalling proteins**.

**Genes that control embryonic development** in insects and in vertebrates are master genes, known as **homeotic genes**, and are examples of regulator genes. These genes act by producing DNA-binding proteins. Homeotic genes control the action of hundreds of other genes that are needed to build the various parts of an animal body in their correct locations. In insects, for example, when a particular homeotic gene malfunctions, the result is an insect that has, instead of antennae, legs on its head (see figure 2.33). If another homeotic gene does not operate normally, the affected insect may have an antenna where it would normally have a wing.

Homeotic genes that organise the body plan of mammals are known as **HOX** genes. Mutations in these genes reveal how they control the pattern of the body plan during embryonic development. For example, in mice, mutation of the HOXC8 gene results in an extra pair of ribs. In humans, mutation of the HOXD13 gene results in limb abnormalities, with an extra digit between the third and fourth digits and with all three digits fused. If the genes at one end of the HOXD complex are removed by a chromosomal deletion, severe limb and genital deformities result.

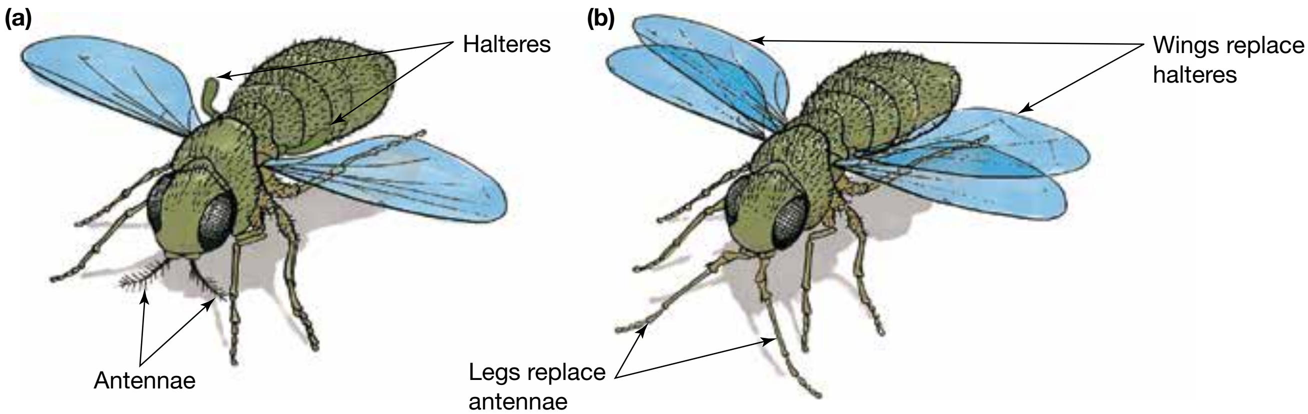



**FIGURE 2.33** Mutations in 'master' or homeotic genes result in the appearance of body parts in unexpected locations. **(a)** Normal fly. **(b)** Fly has mutations in homeotic genes that result in the appearance of a second pair of wings instead of halteres and legs instead of antennae.

**ODD FACT**

High doses of vitamin A (retinoic acid) appear to affect the action of some of the HOX genes, resulting in major malformations in developing embryos involving heart, limbs, central nervous system and facial bones.

Homeotic genes produce DNA-binding proteins. By binding to the DNA near other target genes, these proteins can switch these genes on and off.

**KEY IDEAS**

- All active genes produce RNAs of some kind.
- Most genes transcribe mRNAs that are then translated into proteins.
- Some genes produce other kinds of RNA, such as tRNA and rRNA, and these are the end products.
- Genes can be classified as structural genes or as regulator genes.
- Regulator genes control the activity of other genes, switching them on or off, either directly through the action of DNA-binding proteins, or indirectly through the action of signalling proteins.

**studyon**

Unit 3 | AOS 1 | Topic 4 | Concept 1

**Genes**
Summary screen and practice questions

**QUICK CHECK**

**27** List one difference between a structural gene and a regulator gene.
**28** What is a homeotic gene?

# Time and place for everything

## When are genes active?

**Genes vary in the time of their action.** Some genes are active in making mRNA and proteins only during a short period of the life span of a person, while other genes are active throughout a person's life.

The **HBZ** gene on the number-11 chromosome controls production of one kind of chain found in a type of haemoglobin that occurs only in mammalian embryos. After the first few weeks of embryonic development, this gene is switched off and remains silent during fetal development and after birth.

A baby boy with the genotype ***d*** (Y) for the **DMD** gene on the X chromosome is healthy at birth and for several years after birth. However, the signs of Duchenne muscular dystropy gradually appear during childhood, so that, by about 10 years old, he is confined to a wheelchair.

Some genes are not expressed in the phenotype until a person is well into adulthood. A person with the genotype ***Hh*** has the allele for Huntington's disease on the number-4 chromosome. The ***H*** allele remains inactive often until middle age when the person shows the first signs of this devastating disease.

The **AD1** gene, on the number-21 chromosome, is involved in a familial form of Alzheimer disease. The expression of the specific allele (***A***) typically occurs after the onset of middle age when the person shows the first signs of familial Alzheimer disease.

Many genes remain active throughout the life of a person. These genes include the genes responsible for controlling the production of enzymes that are essential for cellular respiration, such as the **SDH** gene located on the number-1 chromosome. This gene controls production of a sub-unit of one of the enzymes (succinic dehydrogenase) essential for cellular respiration.

## Where are genes active?

**Some genes are active only in the cells of specific tissues:**

- The **HBB** gene on the number-11 chromosome, which controls the production of the beta chains of haemoglobin, is active only in those cells of the bone marrow destined to become red blood cells.
- The **DMD** gene on the X chromosome, which controls production of the protein dystrophin, is active only in skeletal muscle tissue.
- The **GH** gene on the number-17 chromosome, which controls production of growth hormone, is active only in the pituitary gland at the base of the brain.

In contrast, genes involved in controlling production of enzymes involved in cellular respiration are active in all living cells.

The term *gene action* refers to the processes of gene transcription and gene translation. The end products of gene action typically are proteins, such as beta protein chains of the **HBB** gene or amylase enzyme of the **AMY** gene. However, the final phenotype of an organism is more complex than just proteins produced by the genes.

## Identifying active genes

Of the approximately 21 000 genes in a human cell, only some genes are expressed or 'switched on' at a given time. A 'switched on' gene is one that is transcribing the mRNA, while a 'switched off' gene is one that is not producing mRNA. But which genes are active? The pattern of gene activity or gene expression differs between:

- cells of different tissues in the same organism
- cells of the same tissue but at different stages of development
- normal and abnormal cells, such as cancerous cells
- cells under different environmental conditions, such as cells exposed to chemical pollutants or drugs.

In the past, it was possible to examine whether only one or a few genes were active. Through a new technology involving **microarrays** (also known as **DNA arrays** and **gene chips**), **it is now possible to study large numbers of**

**ODD FACT**

Commercially available microarrays include one with 32 996 spots covering the mouse genome and one with 32 878 spots covering the human genome. In both cases, each spot comprises a segment of DNA, 60 bases long, from a different gene.

**genes simultaneously or even the entire genome**. This means that it is now possible to:

- identify which genes are active and which genes are switched off
- compare gene expression in different cell types
- compare active genes in the same cells under different conditions.

A microarray consists of a glass slide (or other solid surface) on which thousands of tiny spots of different DNA molecules are present. Each spot is in a precise location and consists of a short segment of single-stranded DNA from one particular gene (see figures 2.34 and 2.35). The spots are put in place using a robotic instrument to ensure their precise locations. Typically, the DNA spots in a microarray are organised into segments, with each often having up to 20 columns and 20 rows.

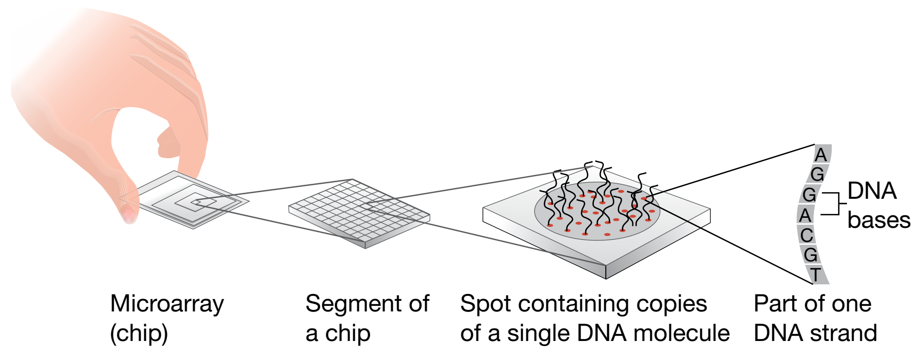



**FIGURE 2.34** Diagram showing a microarray (at left), one segment from this microarray (middle) and one spot from a segment (at right). Note that the DNA in each spot contains single-stranded DNA from one specific gene. The DNA spots are typically 150 to 200 micrometres (μm) in diameter and the centres of the spots might be 300 μm (0.3 mm) apart. So, one segment with 20 rows and 20 columns would fit in a square with a side of 6 millimetres.

**FIGURE 2.35** DNA microarrays enable the simultaneous identification of the active genes in a cell. This figure shows the printing tool that applies the spots to a microarray.

Each of the thousands of spots in the microarray acts as a probe to detect the activity of one different gene. How? Active genes produce mRNAs that can bind at the spots where the complementary single-stranded DNA is located. (If a gene is not active, no mRNA is being transcribed so no binding occurs.)

### Applications of microarray technology

The ability to recognise which genes are active can assist our understanding of the role of genes in various diseases. For example, Swedish researchers used microarrays to identify the active genes in breast cancer cells from 159 patients who were followed up after surgery to remove the cancer. They found that patients whose cancer returned or who died during the study period shared 64 active genes in common, but this set of genes was not active in other patients where the cancer did not reoccur. This study showed how microarrays might be used to predict how people with breast cancer will respond to treatments and these findings may lead to new treatments for cancer. The following box outlines how microarrays work.

## HOW DO MICROARRAYS WORK?

Imagine that some research scientists want to identify how gene activity changes when mouse cells are exposed to a particular drug. Remember that mRNA production signals an active gene. To do this, the scientists need:

1. a 'mouse' microarray with the single-stranded DNA of the spots coming from genes of the mouse genome
2. two cell samples, one from cells that have been exposed to the drug treatment and the second from normal cells that have not been exposed to the drug treatment. (Why two cell samples?)

The next step is to separately extract all the mRNA from each cell population. This mRNA is used as a template to make single-stranded DNA known as complementary DNA (cDNA). The unknown cDNA molecules from the two cell samples are labelled with a coloured fluorescent marker so that they can later be detected, with a different colour being used for the control cells (green label) and the drug-treated cells (red label).

Next, the mixtures of unknown cDNA molecules from both samples are exposed to the microarray. When a cDNA sequence is complementary to one of the spots on the microarray, the cDNA with its fluorescent label binds to this probe. Possible results for one gene are shown in table 2.7.

**TABLE 2.7** Results expected when cDNA from untreated (control) cells with a green fluorescent label and drug-treated cells with a red label are probed using a microarray

| Control cells | Drug-treated cell | Result on microarray |
|---|---|---|
| Gene ON | Gene OFF | GREEN spot |
| Gene ON | Gene ON | YELLOW spot |
| Gene OFF | Gene ON | RED spot |
| Gene OFF | Gene OFF | no colour |

Finally, the microarray is scanned using a confocal laser microscope and the locations of the fluorescent spots are identified and displayed on a computer screen (see figure 2.36).

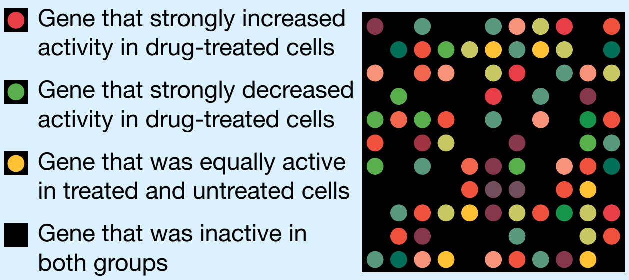



**FIGURE 2.36** Because of the large amount of data involved, computer software is used to analyse the microarray results. Here we see one segment of a microarray and the information that it provides. Does this result provide evidence that the drug treatment changed the activity of genes in the treated cells compared with the control cells?

studyon

Unit 3 | AOS 1 | Topic 4 | Concept 2

**Structural and regulator genes**
Summary screen and practice questions

# Gene regulation in bacteria

In 1961, the French scientists Jacques Monod (1910–1976) and François Jacob (1920–2013) made a significant discovery — they were the first to document a gene regulation system. Rather than producing the same enzymes all the time, Jacob and Monod found that the bacteria produced enzymes only when needed — for example, they discovered that the bacteria *Escherichia coli (E. coli)* produced the enzymes needed to metabolise the sugar lactose only when lactose was present and when glucose was not available. In other words, the bacterial genes that encode the enzymes to metabolise lactose are expressed only in the presence of lactose. Glucose is the favoured metabolite of these bacteria, and, when glucose is available, the genes governing lactose metabolism are silent.

Clearly, there was a mechanism in place to control the activity of the genes encoding enzymes involved with lactose. What Jacob and Monod discovered was how the activity of structural genes that encode particular enzymes are regulated by other genes known as regulator genes. They introduced **the term operon to describe a group of linked structural genes with a common promoter and operator, that is transcribed as a single unit**. The expression of operons is controlled by regulator genes that produce repressor proteins.

studyon

Unit 3 | AOS 1 | Topic 4 | Concept 4

**Gene regulation**
Summary screen and practice questions

This was a significant discovery because it revealed, for the first time, the existence of a new class of genes, **regulator genes**, that control the activities of structural genes, switching them on and off as conditions require. For their discovery, Jacob and Monod shared the Nobel Prize in Physiology or Medicine in 1956.

Since the discovery of the **lac operon** in bacteria, other operons have been identified in bacteria, such as the *trp* operon.

Let's look at some details of the lac operon in *E. coli*.

## The lac operon

Figure 2.37 shows the structure of the lac operon. Also shown is the regulator gene that controls the operon.

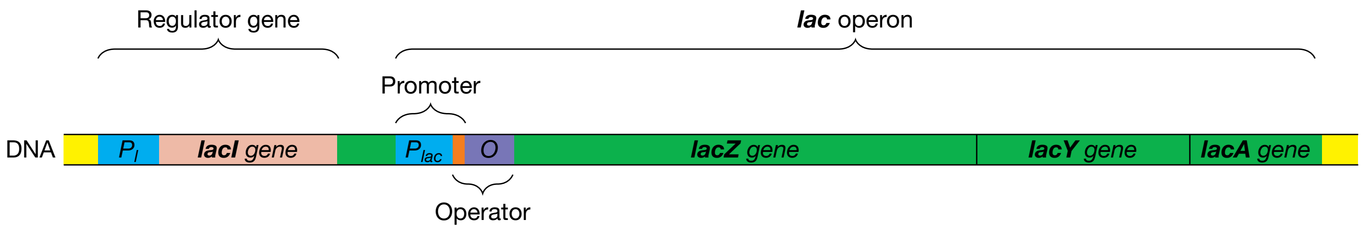



**FIGURE 2.37**

studyon

Unit 3 | AOS 1 | Topic 4 | Concept 5

**The lac operon**
Summary screen and practice questions

Check out this figure and note that the lac operon consists of:

- Three structural genes concerned with the metabolism of lactose as follows:
  - **lac Z** encodes the enzyme galactosidase, which breaks down lactose into glucose and galactose
  - **lac Y** encodes the enzyme permease, which enables lactose to enter cells
  - **lac A** encodes the enzyme acetyltransferase, which also has a role.
- A **promoter** (**P**$_{lac}$): The promoter is a short DNA segment where RNA polymerase can attach and start transcription of the three downstream lac genes. The three lac genes are transcribed as a single entity with one long mRNA transcript being produced.
- An **operator** (**O**): An operator is a short DNA segment that provides a binding site for a repressor.

ODD FACT

An mRNA molecule that contains the transcripts of more than one gene is called a polycistronic mRNA.

Regulator genes may be located adjacent to an operon or be at a distance from the operon or may even be within it. There is no fixed location, unlike the requirement that the DNA sequences of the promoter and operator must be just upstream from the structural genes of the operon.

This operon is controlled by a regulator gene as follows:

- The **lac I** gene (with its upstream promoter): The lac I gene is a regulator gene that encodes a repressor. The repressor is a protein that, when active, down-regulates gene expression by preventing gene transcription.

***What happens when lactose is absent?*** When lactose is absent, the repressor protein is active and binds to the operator, physically blocking RNA polymerase from attaching to the promoter. Transcription cannot start and the lac operon is repressed (see figure 2.38a).

***What happens when lactose is present?*** **When lactose is present, it binds to the repressor protein changing its shape and inactivating it so that it cannot bind to the operator.** This means that RNA polymerase can attach to the promoter and start transcription of the structural genes (see figure 2.38b). Then the long mRNA transcript is translated to produce the enzymes needed to metabolise lactose. So, in the presence of lactose, the enzymes needed to transport it into cells and break it down into galactose and glucose are synthesised.

The operon system enables the bacteria to adapt rapidly to their environment and to use their resources economically by synthesising enzymes only when they are needed: a just-in-time strategy, rather than just-in-case.

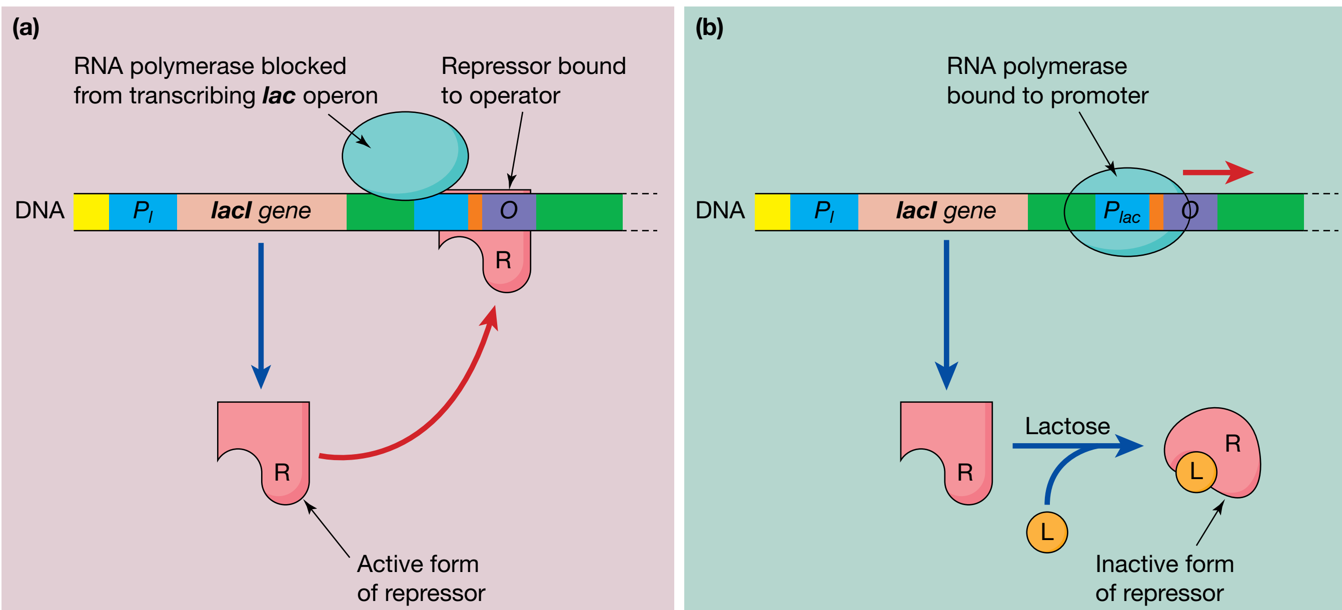
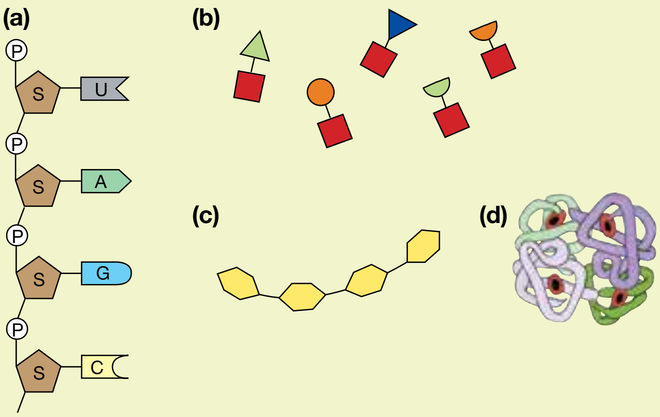



**FIGURE 2.38** Details showing a regulator gene and part only of the lac operon it controls. **(a)** When an active repressor (R) is bound to the operator (O) of the operon, the RNA polymerase is blocked from attaching to the promoter ($P_{lac}$) and transcription of the structural genes of the operon cannot occur. **(b)** The repressor protein is inactivated by binding with lactose (L) and cannot attach to the operator. RNA polymerase can now attach to the common promoter of the operon ($P_{lac}$) and transcription can occur so that the genes of the lac operon can be expressed.

# BIOCHALLENGE

1 Examine each of the molecules in the diagram. For each one, answer the following:
   a Decide whether it is a monomer or a polymer and, if a polymer, name the monomers of which it is made.
   b Classify the molecule as comprehensively as possible with regard to its organic molecule category.
   c Name where, in either a plant or an animal cell, you could expect to find such a molecule.
   d When the sub-units shown in (b) join to form a larger unit,
      i What kind of bond is formed?
      ii Is this a condensation reaction or a hydrolysis reaction?




Thousands of years ago, woolly mammoths (*Mammathus primigenius*), now extinct, lived in frigid regions including present-day Siberia (see figure 2.39). The woolly mammoth evolved from a line that originated in Africa and diverged to give rise also to the African elephant (*Loxodonta africana*) and the Asian elephant (*Elephas maximus*). The woolly mammoth shows features that equip it for life in very cold climates.

**FIGURE 2.39** Woolly mammoth, an extinct mammal, closely related to Asian and African elephants.

2 The DNA sequences of the coding strand of the **HBA** and the **HBB** genes of the woolly mammoth were compared with those of the African and the Asian elephants. Unique mutations in the exons of the **HBB** gene of the woolly mammoth were found:

| Exon | Base number* | Mutation |
|---|---|---|
| 1 | 37 | A → G |
| 2 | 259 | G → T |
| 2 | 304 | G → C |

(*Base number is given relative to the ATG start codon, which begins at base number 1.)

Refer to the genetic code, shown as mRNA codons on page 63, to answer the following:
   a In the woolly mammoth, the A → G mutation at base number 37 in the DNA coding strand produces an amino acid substitution in the beta-globin chain of haemoglobin, with *thr* being replaced by *ala*. Identify what bases number 38 and number 39 could be in the DNA coding strand.
   b In the woolly mammoth, the G → T mutation at base number 259 produces a new sequence (TCC instead of GCC) in the DNA coding strand of the **HBB** gene. What effect will this have on the amino acid sequence in the beta-globin chain of haemoglobin?
   c A mutation present in the exon 3 of the Asian elephant is a G → A substitution at the third position of the 119th codon. This mutation, however, is silent as it does not change the amino acid in the beta-globin chain. Give another example of a silent mutation in the coding strand of the DNA of a gene.
   d A codon consists of a three-base sequence, either in the coding strand of DNA or in the mRNA strand. A single mutation in a codon may affect the first, second or third base in the codon. In terms of changing the amino acid sequence, which position of a codon has the most impact?

# Chapter review

studyon

Unit 3 Nucleic acids and proteins
AOS 1 Gene structure and regulation
Topics 3&4 Practice questions

## Key words

alternative splicing
amino acids
anti-codon
coding region
codons
conjugate
decoded
deoxyribonucleic acid (DNA)
DNA arrays
DNA-binding proteins
DNA sequencers
encoded
exon
exon juggling
flanking regions
gene action
gene chips
gene sequence
gene sequencing
genetic code
homeotic genes
introns
lac operon
messenger RNA (mRNA)
microarrays
monomers
nucleic acids
nucleoproteins
nucleotides
nucleotide sequence
operon
polymers
polypeptide
post-transcription modification
promoters
proteins
proteome
proteomics
regulator genes
ribonucleic acid (RNA)
ribosomes
RNA polymerase
signalling proteins
spliceosomes
structural genes
TATA box
template strand
transcription
transfer RNA (tRNA)
translation
triplet code

## Questions

1 **Applying your knowledge and understanding** →
- a When two monomers such as amino acids join together, a molecule of water is produced. Use two amino acid molecules to explain where this water comes from.
- b How many water molecules would be required to completely hydrolyse a carbohydrate polymer that contained 100 monomers?

2 **Applying understanding and drawing conclusions** → Before the introduction of genetically engineered insulin for use by people with diabetes, the protein hormone was extracted from beef or pig pancreas. Explain how you would expect the sequence of amino acids in the beef and pig insulin to compare with that of humans.

3 **Analysing data and drawing conclusions** → A particular small polypeptide contains nine amino acids. The polypeptide has been fragmented in various experiments by breaking particular peptide bonds. The fragments obtained were:

ser – cys – his – pro – arg – cys
pro – arg – cys
X – gly – met – cys
his – pro – arg – cys
X – gly – met – cys – ser – cys

X is known to be the first amino acid in the polypeptide. What is the primary structure of the polypeptide?

4 **Analysing data and drawing conclusions** → Assume that the nitrogen base sequence in one of the strands in a DNA molecule is:

A-T-C-G-A-C-A-T-G-G-A-A-T-A-C-C-T-C

- a What is the base sequence in the complementary strand?
- b How many amino acids does this piece of DNA code for?

5 **Making connections** → Use at least eight of the key words from this chapter to form a map for the concept 'gene action'. In drawing your map, add any other concepts that you wish.

6 **Demonstrating understanding** → The following is part of the nucleotide sequence in the template strand of part of a gene:

. . . T A T G G G C A T G T A A T G G G C . . .

- a Identify the base sequence in each of the following:
  - i the complementary DNA strand
  - ii the mRNA that would be transcribed from this template.
- b How many codons are present in this mRNA?
- c List the anti-codons that correspond with each codon.

7 **Applying knowledge** → Refer to the genetic code (see table 2.5 on page 63) and answer the following questions:
- a Which codons in mRNA control the addition of the amino acid *gly*?
- b How many codons contain the information to add the amino acid *lys* to a protein? For each codon, write the complementary anti-codon.
- c When the mRNA codon UUU is translated, which amino acid is added to a protein chain?

8 **Demonstrating understanding** → A protein includes the following amino acids in part of its structure:

. . . - *val* - *thr* - *lys* - *pro* - . . .

- a How many codons are needed for the instruction to put these amino acids into place?

**b** Write this instruction in genetic code, as it would appear in mRNA. Would you predict that your code would be identical to that written by all your fellow students? Explain.

**9** **Analysing information and drawing conclusions →** A mutation in one gene (G1) affects just one kind of protein produced by a cell. A mutation in another gene (G2) affects a large number of different kinds of proteins produced by that cell. One of these genes carries the coded instructions to make one kind of tRNA. The other carries coded information to make one kind of salivary enzyme.

Which is more likely to be the gene for the tRNA: **G1** or **G2**? Explain.

**10** **Demonstrating understanding →** A segment of mRNA has the base sequence:

. . . C A U A A G A A U C U U G C. . .

**a** Write the base sequence of the DNA template strand.

**b** Write the amino acids that would be translated from this mRNA segment.

**c** Assume that a base substitution occurs in the original DNA so that the third base (U) of the mRNA is replaced by a G:

. . . C A G* A A G A A U C U U G C. . .

Write the amino acid sequence that would result from this change.

**d** Assume that a base addition occurs in the original DNA so that a G is added between the third and fourth bases:

. . . C A U G*A A G A A U C U U G C. . .

Write the amino acid sequence that would result from this change.

**e** On the basis of this information, what kind of change in DNA has a more extensive effect on the protein resulting from gene translation — a base substitution or a base addition? Explain.

**11** **Applying knowledge →**

**a** *Where is it?*

Where would you find the following:

i a codon

ii gene transcription in action

iii an anti-codon

iv gene translation in action

v nucleolar organiser region?

**b** *What is it?*

Suggest possible identities for the following:

i a STOP codon

ii a START codon

iii an amino acid that has six codons

iv a self-replicating molecule that carries information in coded form

v the cell organelle to which an mRNA molecule attaches for translation.

**12** Undertake some online research to find the gene that causes cystic fibrosis (the search term OMIM can be useful).

Answer the following questions about the gene that causes cystic fibrosis:

**a** How many exons are there in the DNA of this gene?

**b** The pre-mRNA is longer than the mRNA. Explain.

**c** The coding sequence within the mRNA is shorter than the mRNA. Explain.

**d** What relationship exists between the number of bases in the coding sequence of the mRNA and the number of amino acids in the protein product?

**13** *E. coli* bacteria have a requirement for amino acids, including tryptophan (*trp*). These bacteria can take up *trp* from their environment, but if it is not available, *E. coli* can synthesise this amino acid. The five genes involved in the synthesis of tryptophan are part of a system called the ***trp* operon**. (This is similar to the lac operon.)

**a** As well as the five structural genes needed to synthesise *trp*, what other DNA segments form part of the *trp* operon?

**b** Draw a rough line diagram showing the essential components of the *trp* operon.

Consider a situation in which tryptophan is present in the environment in which the *E. coli* bacteria are growing.

**c** Under these conditions, do the *E. coli* need to synthesise *trp*?

**d** Under these conditions, would you expect that the *trp* operon would be repressed or be activated?

If tryptophan is *not* available from the environment in which the bacteria are growing, they will manufacture it themselves.

**e** Under these conditions, would you expect that the *trp* operon would be repressed or be active?

When the *trp* operon is repressed, the structural genes that encode the various enzymes needed are silent and the *trp* operon is said to be repressed.

**f** Identify a possible means by which the *trp* operon might be repressed.

When the *trp* operon is activated, the structural genes that encode the various enzymes needed for its synthesis and transport from cells are transcribed and translated and the *trp* operon is said to be activated.

**g** Identify a possible means by which the *trp* operon might be activated.

CHAPTER

# 3 Enzymes and biochemical pathways

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- recognise that enzymes play an essential role as biological catalysts
- develop awareness of the distinguishing characteristics of enzymes
- distinguish between monomeric and oligomeric enzymes
- understand the mode of operation of enzymes
- recognise the factors that affect enzyme activity
- gain knowledge and understanding of the concept of enzyme inhibition, both reversible and irreversible
- become familiar with the role of cofactors, coenzymes and prosthetic groups in enzyme activity.

**FIGURE 3.1** The search for enzymes begins within days of birth. A blood sample from this baby is tested for the presence of several key enzymes. If the test reveals a missing enzyme, treatment is begun. Treatment can prevent the destructive mental retardation of phenylketonuria (PKU) that can result when the enzyme phenylalanine hydrolase is missing. Treatment can prevent the serious consequences of galactosemia when the enzyme galactose-1-phosphate uridylyltransferase is missing. In this chapter, we will explore enzymes, their structure and their function in life-sustaining reactions.

# The plague of the sea

As the long ocean voyage continued with no landfalls, the majority of the ship's crew became increasingly ill and exhibited a range of unpleasant symptoms — their gums became swollen and purple in colour, teeth became loose in their sockets, skin discoloured and spontaneous haemorrhages occurred. The suffering of the sailors was compounded because even minor wounds failed to heal. Then followed swelling, ulceration and, eventually, death. This illness described as '*the plague of the sea*' came into prominence in the fifteenth century when explorers first ventured on extended ocean voyages or on expeditions to remote terrestrial regions in high latitudes. We now know this disease as **scurvy**.

Read a description of scurvy written by Jacques Cartier, an expedition leader who sailed from France in 1535 to present-day Canada and travelled up the Saint Lawrence River: '*Some could not stand on their feet. Others also had their skins spotted with spots of blood of a purple colour then it did ascend up their ankles, knees, thighs, shoulders, arms and necks. Their mouths became stinking, their gums so rotten, that all their flesh did fall off, even to the roots of their teeth which did also almost all fall out.*'

**ODD FACT**

In 1520, Ferdinand Magellan lost more than 80 per cent of his crew while crossing the Pacific. In the 1740s, Commodore George Anson lost 1300 men from the crews of his fleet, the majority of them dying from scurvy.

## What is scurvy?

Scurvy is a deficiency disease, resulting from a lack of vitamin C (ascorbic acid).

However, most animals can produce their own supplies of vitamin C by a multistep biochemical pathway that converts glucose to ascorbic acid. In contrast, human beings and some other primates, guinea pigs and a few other animals cannot synthesise their own supplies of vitamin C. Instead, these animals must acquire vitamin C through their dietary intake. So, while the sailors on diets of dry beef and biscuits on long voyages during the fifteenth to the eighteenth centuries were at extreme risk of death by scurvy, the rats on these ships were scurvy-free because they could synthesise their own supplies of vitamin C.

Figure 3.2 shows the multistep pathway from glucose to ascorbic acid that occurs in the liver of almost all mammals (and in the kidneys of birds).

**ODD FACT**

All primates except the lemurs, lorises and aye-ayes have lost the functional enzyme that catalyses the last step in the synthesis of vitamin C. So, people, apes, monkeys and tarsiers cannot manufacture their own vitamin C.

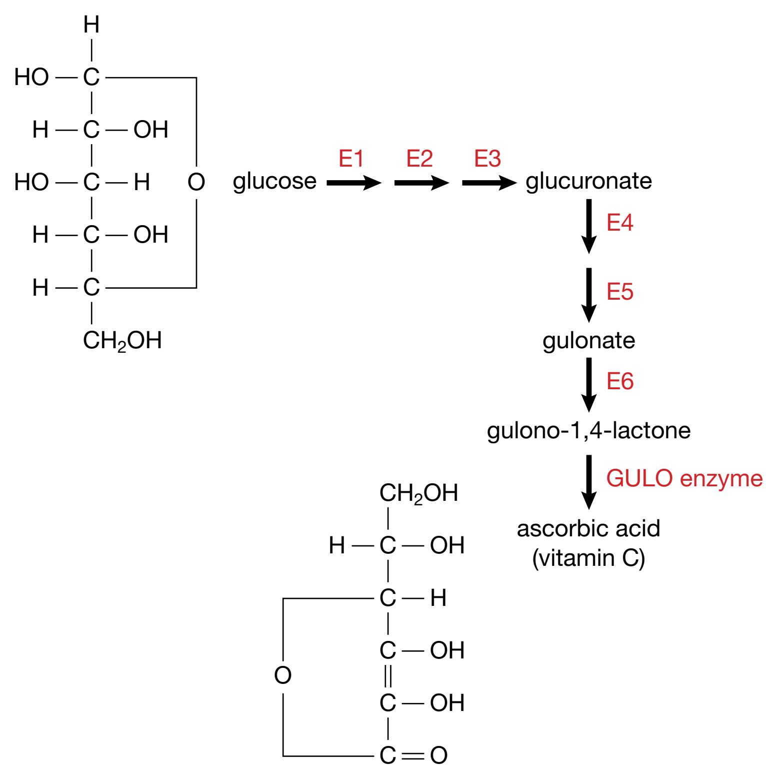



**FIGURE 3.2** Multistep biochemical pathway from glucose to vitamin C (ascorbic acid). Each step in this pathway is catalysed by a specific enzyme. Almost all mammals can synthesise vitamin C by this pathway, except for human beings and some other primates, guinea pigs, and bats of the genus *Pteropus.*

**FIGURE 3.3** Examples of foods that are rich sources of vitamin C (ascorbic acid).

Those few mammals that cannot synthesise their own vitamin C are missing the terminal enzyme in the biosynthetic pathway of glucose to ascorbic acid. The gene encoding this enzyme, gulono-lactone oxidase, has undergone substantial mutation so that it no longer produces a protein that can function as an enzyme. This missing enzyme is the reason that sailors on long sea voyages during the fifteenth to the eighteenth century suffered from and were at risk of death from scurvy.

James Lind (1716–1794), a Scottish physician, was the first to recognise that scurvy could be prevented by the juice of citrus fruits. In his *Treatise of the Scurvy*, published in 1753, Lind recommended that citrus juice be distributed to sailors to prevent scurvy. In 1795, when the Royal Navy finally adopted the practice of issuing lime juice to sailors, scurvy disappeared from the crews of all their ships.

Many fresh fruits and vegetables are rich sources of vitamin C, as, for example, cantaloupe, cherries, kiwi fruit, and cabbage, cauliflower, bean sprouts, red and green peppers, and potatoes (see figure 3.3).

## The missing enzyme in scurvy

Those few mammals, including humans, that cannot synthesise their own vitamin C are unable to produce an enzyme called **gulono-lactone oxidase (GULO)**. This enzyme is needed to catalyse the last step in the pathway that manufactures ascorbic acid from glucose. In humans, the **GULO** gene, located on the number-8 chromosome, is defective.

When the defective **GULO** gene is compared to that of the rat, several major differences are seen: the functional **GULO** gene of the rat consists of a long sequence of DNA that includes 12 exons. In contrast, the homologous **GULO** gene in humans is missing most of these exons; it has only five. Not surprisingly, the major deletions of DNA coding sequences mean that the protein product translated from this gene is non-functional as a GULO enzyme. Figure 3.4 shows a diagrammatic representation of the rat **GULO** gene and the corresponding gene in humans. In addition, compared to the rat **GULO** gene, the DNA of the remaining five exons in the human gene shows other mutations including two single nucleotide deletions, one triple nucleotide deletion, and one single nucleotide insertion.

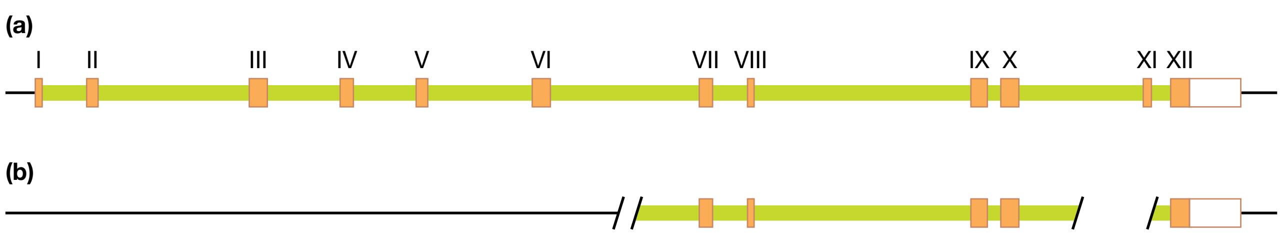



**FIGURE 3.4** Diagram showing a comparison of **(a)** the **GULO** gene that encodes the active GULO enzyme in rats, and **(b)** the defective **GULO** gene of humans, which encodes a non-functional protein. The loss of exons in humans is also seen in the genome of most primates.

Scurvy is just one example of the critical role of enzymes in maintaining the life of an organism. Every living organism has a myriad of enzymes that operate continuously to ensure that the cellular processes essential for life occur at rates sufficiently fast to maintain the living state.

At every moment, enzymes are jump-starting essential chemical reactions in cells and organisms, as, for example, the chemical reactions that:

- enable some microbes to convert nitrogen in the air to essential amino acids
- enable plants to convert the energy of sunlight to sugars
- supply the energy needs of all cells, through cellular respiration
- break down food of animals and fungi to forms that can be absorbed and utilised through digestion
- synthesise macromolecules that are essential to the structure and function of mammals, including:
  - proteins, such as contractile proteins that enable movement, antibody proteins that provide immunity against infectious diseases, signalling proteins that enable communication between cells, proteins that transport oxygen, and proteins that form the structure of bones, cartilage and connective tissue
  - nucleic acids that determine aspects of heredity.

Table 3.1 gives some examples of enzymes essential to these life-maintaining processes.

**TABLE 3.1** Examples of some enzymes involved in various biological processes. Processes involve multistep biochemical pathways, with each step catalysed by a specific enzyme.

| Process | Example of enzyme |
|---|---|
| ***cellular respiration*** | |
| glycolysis | hexokinase, phosphofructokinase |
| Krebs cycle | succinic dehyrogenase |
| electron transport chain | cytochrome c oxidase |
| ***nucleic acid functions*** | |
| DNA replication | DNA polymerase |
| transcription of mRNA from DNA | RNA polymerase |
| mRNA translation | amino acyl tRNA synthetase |
| joining DNA segments | ligase |
| cutting DNA | endonucleases |
| ***digestion*** | |
| breakdown of starch | amylase |
| breakdown of proteins | trypsin, pepsin |
| breakdown of lipids | lipase |
| breakdown of galactose | galactose-1-phosphate uridylyltransferase |
| breakdown of lactose | lactase |
| ***metabolic pathways*** | |
| formation of lactose | lactose synthase |
| formation of uric acid (as in gout) | xanthine oxidase |
| formation of melanin (pigment) | tyrosinase |
| formation of collagen | lysyl oxidase |
| formation of ascorbic acid (vit C) | gulono-lactone oxidase |

## ODD FACT

The term 'enzyme' was introduced by Wilhelm Kuhne (1837–1900) in 1877. In 1926, James Sumner (1887–1955) was the first person to isolate an enzyme when he crystallised pure samples of the enzyme urease.

## ODD FACT

The exception to the 'All enzymes are proteins' statement was discovered by Thomas Cech (b. 1947) and Sidney Altman (b. 1939), who discovered ribozymes. Ribozymes are catalytic RNA molecules that can cut themselves out of long RNA sequences.

**FIGURE 3.5** Ribbon model of lysozyme, a monomeric enzyme, showing the overall globular shape of its tertiary structure. The different colours denote the various amino acids that form its polypeptide chain. Note the coiled regions of the enzyme, which are the alpha helixes of its secondary structure. (*Source:* Tsujino, S., Tomizaki, T., RCSB Protein Data Bank)

# What is an enzyme?

The key features of enzymes may be summarised as follows:

- **Enzymes are biological catalysts:** Catalysts are substances that speed up the rate of chemical reactions without themselves being used up in the reactions. The basic function of enzymes is to act as catalysts by increasing the rate of almost all the chemical reactions in living organisms, and to do this within the prevailing conditions of temperature and pH within cells. Later in this chapter (see page 88), we will explore how enzymes manage to speed up the rates of chemical reactions in cells.

  Why is there a need to speed up chemical reactions in living organisms? You cut your hand on a piece of jagged metal and start bleeding. Enzymes are immediately activated, causing the blood in the damaged region to clot, plugging the wound and stopping the bleeding. Without the action of these enzymes, bleeding could continue, creating a life-threatening situation. The steak that you recently ate was broken down in the acidic conditions of your stomach at body temperature over a period of several hours. This breakdown was the result of the catalytic action of digestive enzymes such as pepsin and trypsin acting on the protein of the steak. To do the same without enzymes also requires acidic conditions, but would require a few days (not hours) and a temperature of 100 °C (not body temperature). These two examples give a clue to the importance of enzymes in maintaining life. Without enzymes, the speed of biochemical reactions would be far too slow to sustain the living state.
- **Enzymes are proteins:** **Almost without exception, enzymes are soluble globular proteins (in contrast to insoluble fibrous protein, such as collagen and silk)**. Globular proteins are spherical or globe-like in overall shape, and are typically water-soluble. Water solubility results from the fact that, in a globular protein, the amino acids with polar side chains are exposed at the surface of the protein where they can interact with nearby polar water molecules and dissolve. In contrast, amino acids with nonpolar side chains tend to point inwards. (Refer to the Appendix to check out the different side chains of amino acids.)
- **Enzymes consist of one or more polypeptide chains:** A minority of enzymes are termed **monomeric enzymes**, with each enzyme consisting of a single polypeptide chain formed of amino acids joined by peptide bonds. The typical number of amino acids in the polypeptide chain is in the range of 100 to 300. The folding and twisting of the polypeptide chain of such an enzyme produces its tertiary structure, which is stabilised by the formation of other bonds, including disulfide bonds, ionic bonds and hydrogen bonds.

  Examples of monomeric enzymes include lysozyme (in saliva), trypsin (in the gastric juices of the stomach), papain (in pineapple) and ribonuclease. Figure 3.5 shows a ribbon model of the lysozyme enzyme with 12 different colours denoting the various amino acids, either singly (grey = *gly*) or a pair (red = *asp* and *glu*, and blue = *lys* and *arg*) or, in one case, a trio (green = *leu*, *val* and *ile*).

  The majority of enzymes are **oligomers**, with each enzyme consisting of two or more polypeptide chains that, in this context, are called sub-units. In some oligomeric enzymes, the sub-units or polypeptide chains are identical, while, in others, the sub-units differ. Examples of oligomeric enzymes include lactose synthase (in mammary glands) with two different sub-units, and phosphofructokinase with four identical sub-units (see figure 3.6).

  Because they consist of two or more sub-units, oligomeric enzymes have a quaternary structure that is stabilised by covalent bonds. Oligomeric enzymes with their multiple sub-units have properties that are not present

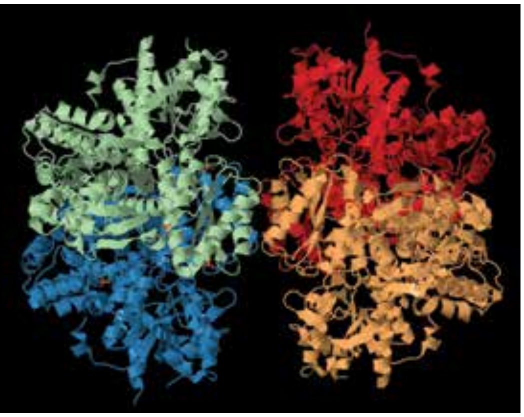


**FIGURE 3.6** Ribbon model of the oligomeric enzyme phosphofructokinase, showing its four identical sub-units, distinguished by different colours. This enzyme catalyses one step in the glycolysis pathway of cellular respiration. Note that the majority of this enzyme displays an alpha-coiled secondary structure. Does this enzyme have a quaternary structure? (*Source:* Kloos, M., Straeter, N., RCSB Protein Data Bank)

in monomeric enzymes, as, for example, the presence of multiple binding sites.

The functional enzyme lactate dehydrogenase (LDH) consists of four sub-units, each of which may be either an H or an M sub-unit. LDH present in heart muscle consists of four H sub-units, while the LDH in skeletal muscle consists of four M sub-units. Other tissues have other forms of LDH, including $H_3M$, $H_2M_2$, and $HM_3$. These five forms of LDH are termed isozymes, and each has slightly different properties. Note that just two sub-units can produce five different isozymes.

- **Every enzyme has an active site:** Typically, a small pocket or cleft in an enzyme is an **active site**, also known as the **substrate binding site**, which is essential for the enzyme to function. The **substrate** of an enzyme is the molecule that the enzyme acts on or the input to an enzyme-catalysed reaction. The output of an enzyme-catalysed reaction is the **product** or products. The active site of the enzyme that comes into contact with the substrate is small in comparison to the rest of the enzyme.

The active site of each enzyme is formed by the folding of the polypeptide chain(s) to form a small region with a unique stereospecific 3D shape. The active site can recognise and bind to its specific substrate because of the complementary shapes of the active site and the substrate.

Figure 3.7a shows a diagrammatic representation of a substrate interacting with the active site of an enzyme. Typically, an active site consists of three to twelve amino acids organised to form a particular shape. If the shape of the active site of an enzyme is altered by any chemical or physical treatment, the ability of the enzyme to function is reduced or permanently or temporarily lost. Figure 3.7b shows the active site of the enzyme aspartate transaminase, with six amino acids involved in the active site. The cytoplasmic aspartate transaminase enzyme is an oligomer with the functional enzyme consisting of two identical sub-units, each containing 412 amino acids.

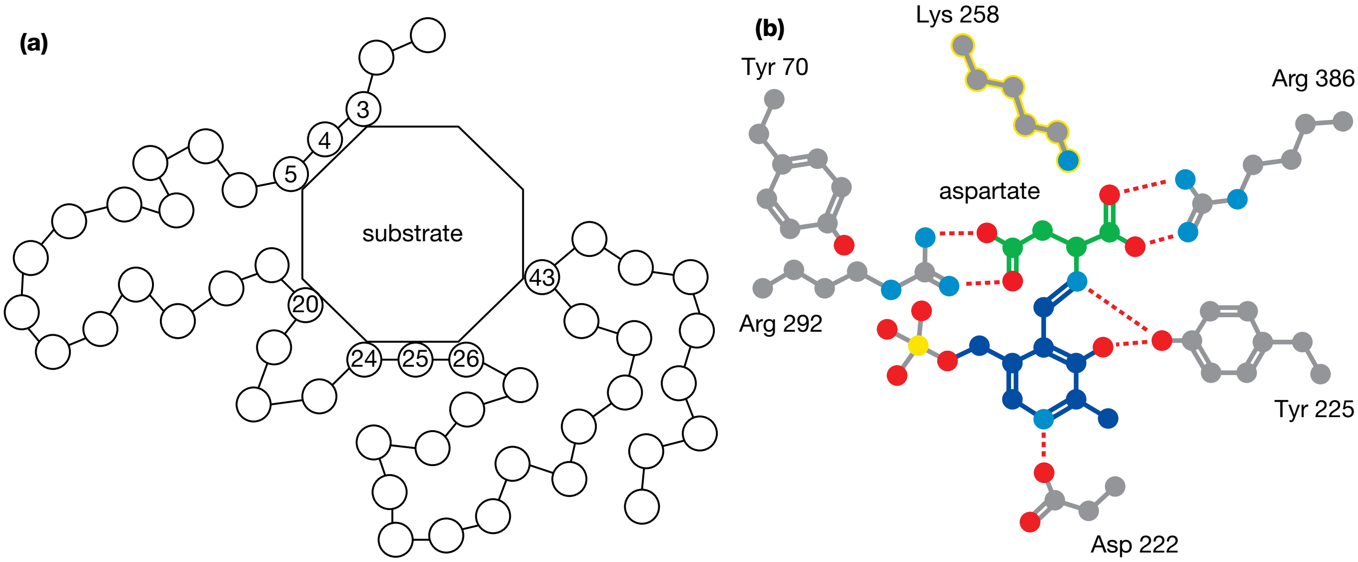



**FIGURE 3.7** **(a)** Diagram showing a 2D representation of a substrate molecule in contact with the active site of an enzyme. **(b)** Diagram showing the active site of the enzyme aspartate transaminase, which includes six amino acids. The substrate molecule, glutamate (shown in green), and a cofactor (shown in dark blue) are held in place in the active site of the enzyme by weak bonds, such as hydrogen bonds and ionic bonds.

In fact, the active site of an enzyme is a 3D space. Figure 3.8a shows a space-filling model of the enzyme hexokinase, which is involved in the glycolysis pathway. The active site of hexokinase is at the bottom of the cleft that can be

studyon

Unit 3
AOS 1
Topic 5
Concept 3

**Chemical nature of enzymes**
Summary screen and practice questions

studyon

Unit 3
AOS 1
Topic 5
Concept 3

**See more**
Structure and specificity of enzymes

seen on the left-hand side of the model. Figure 3.8b shows a stylised diagram of an active site of an enzyme with its substrate identifying the weak bonds that hold the substrate and the active site together. These weak bonds are hydrogen bonds (H—H), and ionic bonds between groups on the substrate and groups on the active site that carry opposite charges. These weak bonds are sufficient to stabilise the **enzyme-substrate (E-S) complex**, but they can easily be broken, allowing the products of the reaction to be released, and freeing the active site to be occupied by another substrate molecule. (Later in this chapter we will see that, if a molecule forms a strong covalent bond with a group that is part of the active site of an enzyme, the enzyme will be inactivated.)

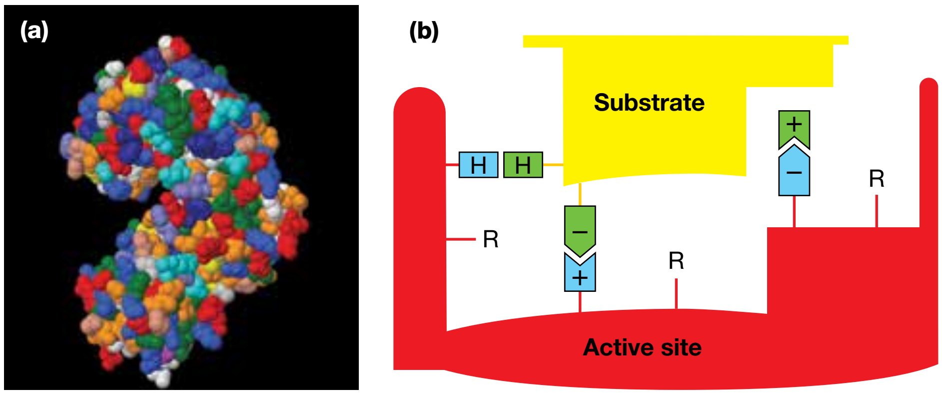



**FIGURE 3.8 (a)** Space-filling model of a hexokinase enzyme with its active site at the base of the cleft shown at the left-hand side of the model. Different colours denote the various amino acids that form the single polypeptide chain of this enzyme. (*Source:* Feng, J., Zhao, S. and Liu, L., RCSB Protein Data Bank) **(b)** Schematic diagram showing the weak bonds, such as hydrogen bonds and ionic bonds, that temporarily stabilise the enzyme-substrate (E-S) complex at the active site of an enzyme.

- **Each enzyme has a specific activity:** The specificity of enzymes is demonstrated in various ways, including:
  - absolute or substrate specificity, meaning that an enzyme can act on one substrate only; for example, the enzyme lactase can only hydrolyse the milk sugar lactose to form glucose and galactose.
  - bond specificity, meaning that an enzyme can act on one kind of chemical bond; for example, peptidase enzymes act specifically on the peptide bonds between any two amino acids, and the enzyme amylase acts specifically on the 1,4 glycosidic bonds that link the glucose monomers of starch and glycogen.
  - group specificity, meaning that an enzyme can only act on molecules with particular functional group(s) surrounding a bond; for example, the digestive enzyme trypsin can only act on peptide bonds that are adjacent to amino acids with basic side chains, namely lysine and arginine.

## A time and place for enzyme action

Where and when does enzyme action occur? Most enzymes act on substrates in solution within cells of particular human tissues. So, most enzymes are intracellular.

However, some enzymes take action in extracellular locations, such as lysozyme in tears, which are part of the body's first line of defence against microbial infection (see chapter 7, page 287), the digestive enzymes such as amylase, which are produced in salivary glands, and the clotting enzyme thrombin. The digestive enzymes of the stomach, the pancreas and the small intestine are also examples of extracellular enzymes.

Many enzymes exist in cells as inactive **proenzymes** (also termed zymogens), which are only activated in response to relevant signals. This strategy prevents enzyme activity occurring at inappropriate times or places and causing damage. For example:

- The proteolytic enzyme pepsin is secreted by stomach cells as an inactive proenzyme, called pepsinogen. Once in the extracellular space of the stomach cavity, the pepsinogen molecules are activated by the acidic conditions (pH 2 to 3) in the stomach to form pepsin. This activation involves the removal of a peptide (44 amino acids long) from pepsinogen that masks the active site. Once this is removed, active pepsin is produced (see figure 3.9).

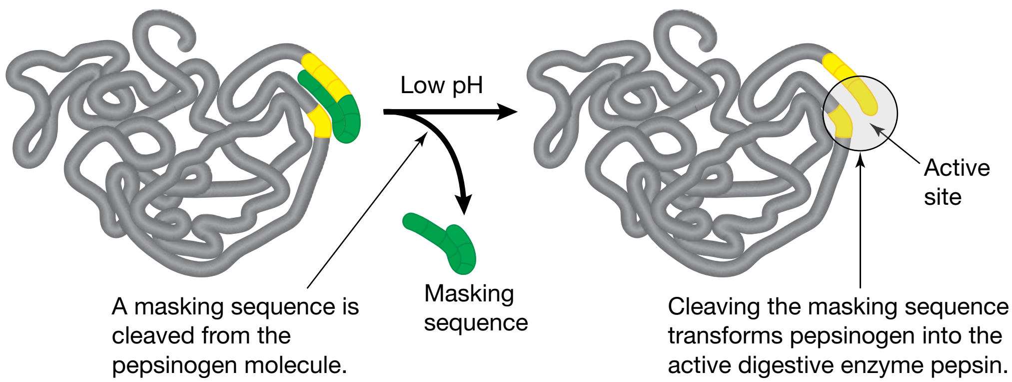



**FIGURE 3.9** Diagram showing the activation of the inactive proenzyme pepsinogen to the active proteolytic enzyme pepsin, which is initiated by the low pH conditions of the stomach. These low pH conditions cause the removal of a masking sequence of a 44 amino-acid peptide from the active site, transforming pepsinogen into pepsin.

- The clotting enzyme thrombin is normally present as an inactive proenzyme called prothrombin. Prothrombin is converted to active thrombin only when blood vessels are damaged, bleeding occurs, and there is a need for clot formation. In the bloodstream is a soluble monomeric protein called fibrinogen. Thrombin catalyses the conversion of the soluble blood protein fibrinogen to the insoluble polymeric protein fibrin, which aggregates and forms a clot, stopping the loss of blood (see figure 3.10).

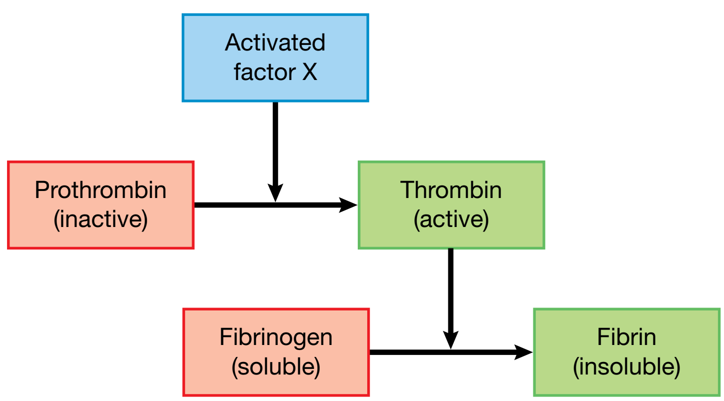



**FIGURE 3.10** An extract of part of the clotting cascade that shows the conversion of inactive prothrombin to active thrombin, which, in turn, catalyses the conversion of soluble fibrinogen to insoluble fibrin aggregates, forming a clot and trapping platelets and blood cells.

## NAMING ENZYMES

Usually an enzyme has two names, as follows:

1. a recommended or **trivial name** that is short, often ending in –ase; in many cases, the substrate of the enzyme is included in this name: for example, tyrosin**ase**, lact**ase**, sucr**ase**, creatine kin**ase**. (Recommended names are the enzyme names that are used throughout this chapter.) What reaction do you think the enzyme succinic dehydrogenase carries out?
2. a **systematic name** that is long and identifies the exact reaction that is catalysed by the enzyme; for example:
   - amylase, present in human saliva, has the systematic name *4-alpha-D-glucan glucanohydrolase*
   - glucose oxidase has the systematic name *β-D-glucose:oxygen 1-oxidoreductase*
   - aspartate aminotransferase has the systematic name *L-aspartate:2-oxoglutarate aminotransferase*
   - tyrosinase has the systematic name *L-tyrosine, L-dopa:oxygen oxidoreductase.*

In addition, each enzyme has an EC (Enzyme Commission) number. The EC numbers group enzymes into one of six groups identifying the types of enzyme-catalysed reaction that are carried out. Table 3.2 shows these six groups. Each enzyme has an EC identifier that consists of four numbers, such as:

> Tyrosinase is identified as EC 1.14.18.1; glucose isomerase is identified as EC 5.3.1.5; aspartate aminotransferase is identified as EC 2.6.1.1.

The first number identifies the major reaction type that the enzyme can catalyse. So, any enzyme with the first EC number of 2 must be a transferase enzyme, which catalyses a particular transferase reaction.

**TABLE 3.2** The Enzyme Commission (EC) groups. Every enzyme belongs to one of these six groups, which identify the types of chemical reactions that are catalysed by enzymes.

| Group | Reaction type catalysed | Example of enzyme |
|---|---|---|
| EC 1 *Oxidoreductases* | oxidise/reduce molecules by transfer of O and H atoms | succinic dehydrogenase |
| EC 2 *Transferases* | transfer a group from substrate to product | aspartate transaminase |
| EC 3 *Hydrolases* | cleave chemical bonds by hydrolysis | amylase |
| EC 4 *Lyases* | cleave bonds in a substrate (not by hydrolysis) | aldolase |
| EC 5 *Isomerases* | change arrangement of atoms in a molecule | glucose isomerase |
| EC 6 *Ligases* | join two molecules by forming a new bond | glutamate-cysteine ligase |

### ODD FACT

The enzyme known as pepsin (EC 3.4.23.2) was given its name in 1836 by Theodor Schwann (1810–1882), who, along with Matthias Schleiden (1804–1881), formulated the Cell Theory in 1839.

### KEY IDEAS

- **Enzymes are biological catalysts composed of protein.**
- **Catalysts speed up the rate of chemical reactions.**
- **Some enzymes are monomeric and consist of a single polypeptide chain, while most are oligomers.**
- **The active site of an enzyme has a 3D shape that is complementary to the shape of its specific substrate.**

**QUICK CHECK**

1. Briefly explain the meaning of the following statement: *Enzymes are biological catalysts*.
2. Identify a key difference between the members of the following pairs:
   - a monomeric enzyme and oligomeric enzyme
   - b proenzyme and enzyme
   - c substrate specificity and bond specificity
   - d hydrolase and transferase.
3. Identify the following statements as true or false:
   - a The typical shape of an enzyme is globular.
   - b The enzyme lysozyme consists of a single polypeptide chain.
   - c Most of the structure of an enzyme forms part of its active site.
   - d Most enzymes are oligomers.
   - e Enzymes speed up the rate of chemical reactions in the body.
   - f Each enzyme can catalyse just one specific reaction involving a single substrate.
   - g The E-S complex is stabilised by strong covalent bonds.

# Explaining enzyme specificity

Let's now look at why enzymes have a high level of specificity for one substrate, or for a functional group or side chain, or for a bond between particular groups. To do this, let's use models. What is a model? Models are used to represent biological entities, such as biological processes, in a way that makes them conceptually easier to understand or enables predictions to be made about their operation. A model may be an analogy, such as the **lock-and-key model** of enzyme-substrate interactions described below. A model may be a mathematical representation, such as the Hardy-Weinberg equation to represent a population gene pool.

## Lock and key or induced fit?

The essential prelude to the catalytic action of an enzyme is the binding of the substrate(s) to the active site of the enzyme. Two models have been put forward in an attempt to explain this specificity:

1. the lock-and-key model
2. the induced fit model.

The earliest explanation for the specificity of enzymes for particular substrates was proposed by Emil Fischer (1852–1919) in 1894. Fischer explained the specificity of enzymes using the analogy of a lock and key. In this lock-and-key model, the enzyme acts as the lock and the substrate as the key. Only the correctly shaped key (substrate) can fit the lock (the active site of the enzyme). This model correctly identified the importance of the complementary shapes of the active site of an enzyme and its specific substrate and recognised that enzyme specificity is determined by the active site of the enzyme. The lock-and-key model assumes that the active site of an enzyme (the lock) is rigid and fixed.

A second model was premised on the active site of an enzyme having a defined shape, but one that has a degree of flexibility. This is the **induced fit model**, which was proposed by Daniel Koshland in 1958. Koshland proposed that, just as a glove changes its shape when a hand slips into it, so does the active site change its shape to fit tightly around its substrate. In this model, the binding site of an enzyme is not viewed as being exactly complementary to the substrate. When the reaction is complete, the empty active site returns to its relaxed resting state.

Figure 3.11 shows the difference between these two models.

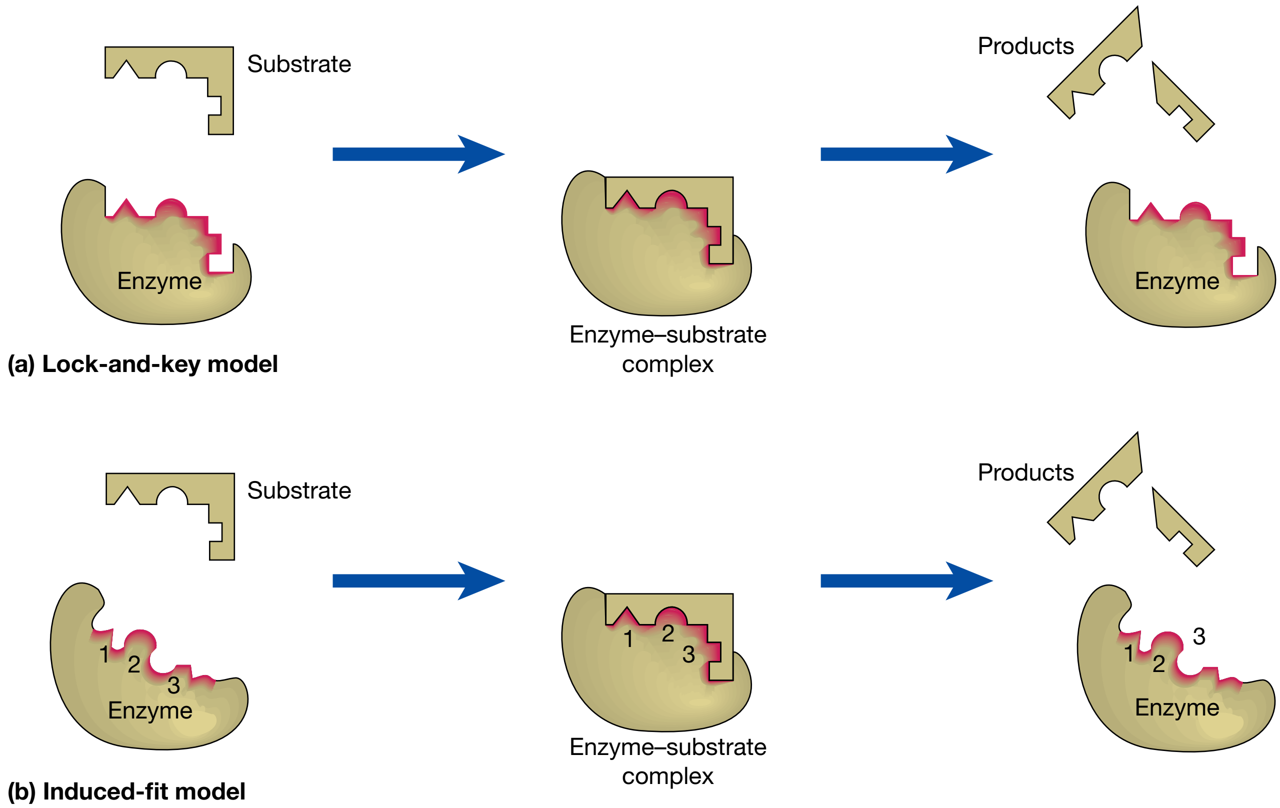



**FIGURE 3.11** Two models to explain enzyme specificity are shown. **(a)** The 'lock-and-key' model and **(b)** the 'induced fit' model. Examine these figures carefully. In which model is the shape of the active site (shown in red) exactly complementary to the shape of the substrate? In which model does the shape of the active site change slightly when it binds to the substrate, making a tight fit?

# Mode of action of enzymes

At any minute, hundreds of chemical reactions are taking place in living organisms, and this multitude of linked chemical reactions comprises the **metabolism** of an organism. Most of these chemical reactions occur inside cells, but a few are extracellular in location. Almost without exception, every step in a metabolic pathway is catalysed by a specific enzyme. In a person, enzyme-managed reactions are involved in growth, repair, cellular respiration, reproduction, digestion, breathing, clotting of blood, and immunity, with several thousand different enzymes involved.

In metabolism, some compounds are broken down with an associated release of energy, while other compounds that are needed by cells are built up or synthesised in a process that requires an input of energy. These two aspects of metabolism are recognised in the following classification:

- **catabolism** = series of cellular reactions that releases useful energy from the breakdown of complex molecules
- **anabolism** = series of cellular reactions in which the complex molecules required by cells are synthesised from simpler building blocks, a process that requires an input of energy.

Metabolic reactions that release energy are said to be **exergonic**, while reactions that require energy are termed **endergonic**. Figure 3.12 shows how the energy released in catabolic reactions drives the energy-requiring anabolic reactions. This illustrates the principle of **coupling** of exergonic and endergonic reactions

studyon

| | |
|---|---|
| Unit 3 | **Metabolic reactions** |
| AOS 1 | Summary screen and practice questions |
| Topic 5 | |
| Concept 1 | |

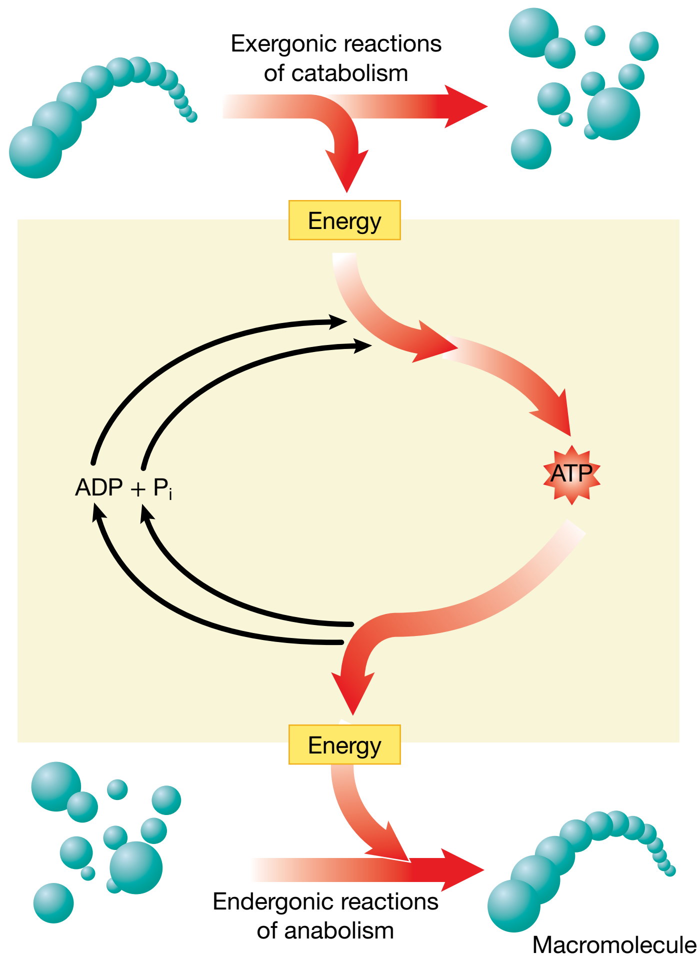



**FIGURE 3.12** Diagram showing interrelationship between energy release by catabolism and energy usage by anabolism.

Metabolic reactions are neither random nor unrelated. Instead, each reaction is part of a coordinated sequence of enzyme-driven reactions or pathways. Several mechanisms operate to ensure that the chemical reactions operate in orderly and balanced sequences. For example, end-product inhibition (see pages 94 and 95) prevents the build-up of an excess product of a reaction and ensures that product(s) of reactions are formed only as and if required. (In bacteria, another way of achieving this is through operons and their regulation — refer to chapter 2.) Interference between biochemical pathways is minimised by having particular pathways occurring within different membrane-bound compartments in a cell. We have already seen how the production of inactive proenzymes, which must be activated to functional enzymes, can also control when and where enzymes are active.

Let's look now at how enzyme-catalysed reactions take place.

studyon

Unit 3 | AOS 1 | Topic 5 | Concept 2

**Enzymes as catalysts**
Summary screen and practice questions

studyon

Unit 3 | AOS 1 | Topic 5 | Concept 2

**See more**
Enzymes as catalysts

## Mode of enzyme action

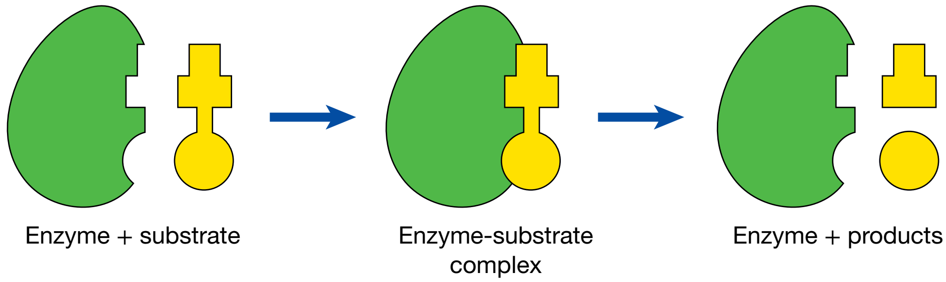
Enzymes are biological catalysts that accelerate chemical reactions in cells. The molecules that are inputs to enzymic action are termed substrates, and the outputs of enzymic action are termed products. We can show a general enzyme-catalysed reaction as follows:

$$\mathbf{E + S \rightleftharpoons ES \rightleftharpoons P + E}$$

that is: **Enzyme + Substrate ⇌ Enzyme-Substrate complex ⇌ Product + Enzyme**

Figure 3.13 is a diagram illustrating the critical step in an enzyme-catalysed reaction, namely the formation of the enzyme-substrate complex at the active site of the enzyme.

**FIGURE 3.13** Diagram showing the stages in an enzyme-catalysed reaction. This shows a catabolic reaction. How would this diagram change for an anabolic reaction?




Look at the equation and diagram, and note that:

1. The reaction involves the formation of an enzyme-substrate complex (E-S). The formation of the E-S complex means that the enzyme specifies the substrate it will bind to, the reaction it will catalyse, and the product(s) that will be formed.
2. The enzyme is not consumed in the reaction. Once the product is formed, the enzyme is free to interact and bond with another substrate molecule. A consequence of this is that only a small amount of enzyme is needed to catalyse a reaction and produce large quantities of product.

For any chemical reaction to occur, the reactants must be in an activated state. The energy level that must be attained is termed the **activation energy**. If the reactant molecules do not have sufficient energy, no reaction can occur.

**Enzymes lower the energy level that is needed to activate reactant molecules.** Enzymes do this by binding to the reactant molecules and positioning or bending them in a way that enables the breaking of existing bonds and the formation of new ones to occur more readily. As a result, more reactant molecules have sufficient energy to react, and so the reaction proceeds much more rapidly than a non-catalysed reaction. In reality, enzymes speed up chemical reactions in the body by a factor of between $10^6$ to $10^{12}$ times so that essential chemical reactions can occur at body temperature.

By providing an alternative reaction pathway with a lower activation energy requirement, enzymes produce reaction rates that are much faster than those for the non-catalysed pathway. Figure 3.14 shows a graphical representation of an enzyme-catalysed reaction and a non-catalysed reaction.

**ODD FACT**

The enzyme catalase has one of the highest turnover numbers of all enzymes. One catalase molecule can convert approximately five million molecules of its substrate (hydrogen peroxide) to water and oxygen each minute.

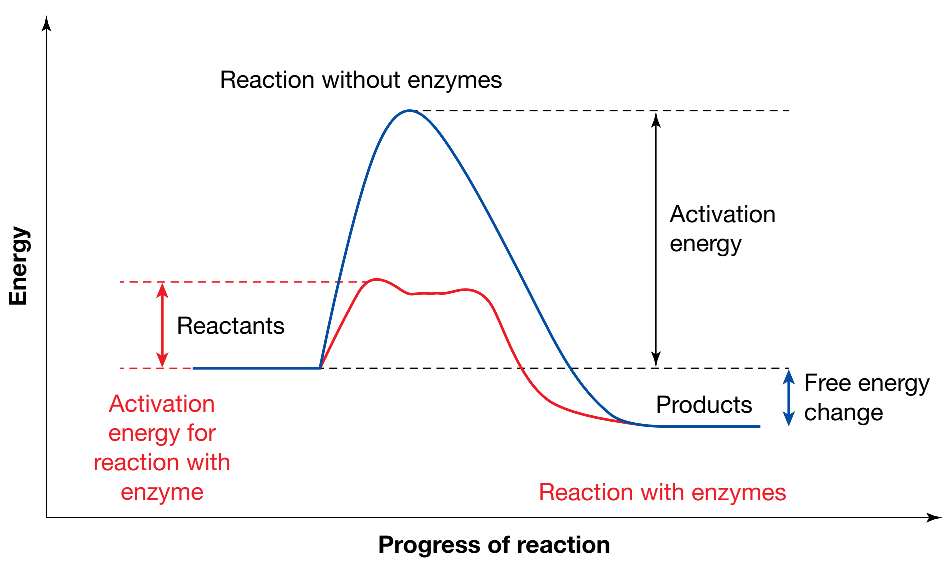



**FIGURE 3.14** Diagram showing the different ‘routes’ taken by an enzyme-catalysed reaction and a non-catalysed reaction. In both cases, the overall reaction is the same. However, the enzyme-catalysed reaction involves the formation of an intermediate enzyme-substrate complex with a lower activation energy than for the non-catalysed reaction.

studyon

| | |
|---|---|
| Unit 3 | **Factors affecting enzyme activity** |
| AOS 1 | Summary screen and practice questions |
| Topic 5 | |
| Concept 4 | |

studyon

| | |
|---|---|
| Unit 3 | **Do more** |
| AOS 1 | Identify the enzyme |
| Topic 5 | |
| Concept 4 | |

# Factors affecting enzyme activity

Various factors that have an impact on enzyme activity include:

- temperature
- enzyme concentration
- pH
- substrate concentration.

## Effect of temperature on reaction rate

For most chemical reactions, an increase in temperature causes an increase in the reaction rate. As the reactants are heated, the particles move faster and are more likely to collide with sufficient energy to overcome the activation energies of the reactants. However, beyond a certain point, that relationship no longer holds.

Each enzyme has a temperature at which it operates most effectively, and, for most human enzymes, that optimum temperature is about 37 °C or normal body temperature. On either side of that optimum is a range of temperatures at which enzymes will still operate, but at decreasing effectiveness.

As the temperature increases further, thermal denaturation of proteins begins to occur. Heat denaturation permanently changes the tertiary shape of enzyme molecules so that the shape of the active site is altered and the enzyme is inactivated. Fevers occur when the body temperature rises above normal. A high fever, defined as a core body temperature of 41.5 °C and higher, is extremely dangerous in part because of the harmful effect of this temperature on the thousands of enzymes in the human body.

If the temperature falls too low, the enzyme and substrate molecules do not have sufficient energy to interact. Enzymes that are inactivated because of a low temperature can become active again when the temperature rises. Because low temperatures do not permanently inactivate enzymes, enzymes are often stored in laboratories in the cold; for example, enzymes can be stored in buffered 50 per cent glycerol, which remains liquid at temperatures down to –35 °C.

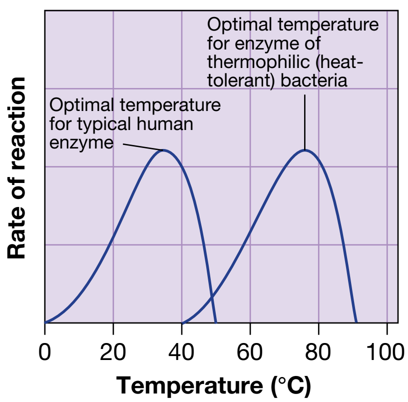



**FIGURE 3.15** Graph showing the effect of temperature on the activity of a typical human enzyme compared with the effect on the activity of an enzyme taken from heat-tolerant bacteria.

Figure 3.15 shows the effect of temperature on the activity of a typical human enzyme compared with an enzyme from a heat-tolerant bacterial species. Note the sharper fall off in reaction rate when the temperature exceeds the optimum. Contrast this with the more gentle drop off in rate when the temperature falls below the optimum.

Consider the following experiment: tubes containing the same amount of enzyme and the same volume of solutions of different concentrations of substrate (glucose) were incubated under identical conditions. The rate of breakdown of glucose was measured and plotted against the original concentration of the substrate. What was the maximum rate of the reaction?

Not all organisms have enzymes with optimal temperatures similar to that of the enzymes of the human body. Table 3.3 shows some temperature minima, maxima and optima for the growth of some microbes, both bacteria and archaeans. While these figures relate to growth and not to a particular enzyme, they are indicators of the temperature tolerance of the enzymes of these various microbes.

**TABLE 3.3** Temperature (°C) ranges for a number of bacteria and archaea.

| Microbe | Minimum | Optimum | Maximum |
|---|---|---|---|
| *Listeria moncytogenes* | 1 | 30–37 | 45 |
| *Staphylococcus aureus* | 10 | 30–37 | 45 |
| *Streptococcus pyogenes* | 20 | 37 | 40 |
| *Thermus aquaticus* | 40 | 70–72 | 79 |
| *Sulfolobus acidocaldarius* | 70 | 75–85 | 90 |
| *Pyrobacterium brackii* | 80 | 102–105 | 115 |

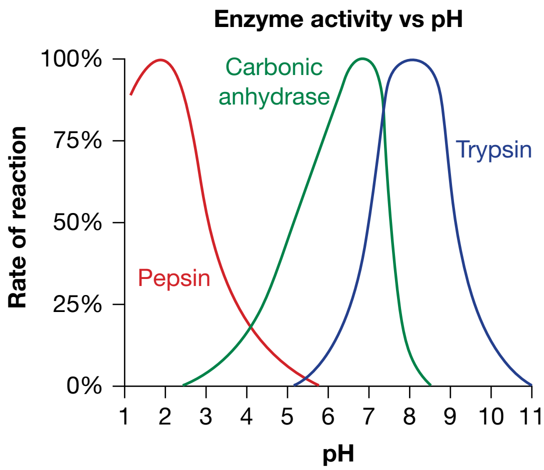



**FIGURE 3.16** Graph showing the effect of pH on the reaction rate of three human enzymes. Can you suggest why the pH optima of these enzymes are so different from one another?

Enzymes synthesised by bacteria and archaea with optimal growth temperatures greater than 80 °C are termed hyperthermophilic enzymes. In contrast to the enzymes of most organisms, these enzymes are resistant to irreversible denaturation at high temperatures, and are said to be thermostable.

## Effect of pH on reaction rate

The pH (hydrogen ion concentration) at which enzymes can operate varies, typically according to the environment in which the enzyme normally operates (see figure 3.16). Unlike the normal human core body, which is regulated within a narrow range, various compartments of the human body have very different prevailing pH values, such as the stomach with a pH of 2 to 3, and the small intestine with a pH around 5.

Enzymes can operate as catalysts within a range that is defined by a maximum pH and a minimum pH, with the highest activity (in terms of substrate molecules catalysed per unit time) occurring at the optimal pH for the enzyme concerned. What is the pH optimum for the enzyme alkaline phosphatase?

Usually, an enzyme works within quite a narrow pH range. At the optimal pH, the reaction rate is greatest. Changes in pH can alter intra- and intermolecular bonds, changing the shape of the enzyme and affecting its efficiency

The pH of a solution is a measure of the concentration of hydrogen ions per litre of solution. The pH of a neutral solution, such as pure water, is 7.0. A pH value below 7 indicates an acid — the lower the number, the more acidic the solution. A pH value above 7 indicates a basic solution. Most biological fluids have a pH between 6 and 8; for example, human blood is maintained at about pH 7.4. There are a few extremes beyond this range, such as gastric juice in the stomach, which has a pH of about 2.

Each enzyme acts best at a particular pH, which is known as the optimal pH for the enzyme (see figure 3.16). Pepsin, which is secreted in the human stomach, has an optimal pH of 2. Trypsin, secreted in the human small intestine, has an optimal pH of about 8.0. Carbonic anhydrase, an enzyme found in human blood, has an optimal pH of 7.4. Enzymes that act within human cells generally have an optimal pH of about 7.6. A change in pH from the optimum can change the shape of an enzyme and affect its ability to combine with its substrate. Because the enzyme is less able to combine with its substrate, it is unable to act and the rate of the metabolic reaction declines. An enzyme becomes less efficient if the variable value is greater or less than the optimal.

**study on**

| Unit 3 | **The rate of a reaction** |
|---|---|
| AOS 1 | Summary screen and practice questions |
| Topic 5 | |
| Concept 5 | |

## Effect of substrate concentration on reaction rate

The addition of more substrate to an enzyme solution initially increases the rate of the reaction, provided that some active sites of the enzyme are not yet occupied. However, a set amount of enzyme molecules is present, and the rate of the reaction tapers off once all the active sites of the enzyme molecules become occupied.

So, for a given enzyme concentration, the rate of reaction increases with increasing substrate concentration. But only up to a point. Beyond this, any further increase in substrate concentration produces no significant change in reaction rate (see figure 3.17). At this point, all the active sites of the enzyme molecules at any given moment are occupied by substrate molecules. The E-S complexes (enzyme/substrate complexes) have to dissociate before the active sites are free to accommodate more substrate.

On such a graph, two values are typically noted:

$V_{max}$ = maximum reaction rate or velocity. $V_{max}$ is reached when all the active sites of the enzyme are saturated with substrate molecules. $V_{max}$ is directly proportional to enzyme concentration.

$K_M$ = substrate concentration at half the maximum velocity. $K_M$ is inversely proportional to the affinity of an enzyme for its substrate, such that the lower the $K_M$, the higher the affinity of the enzyme for its substrate.

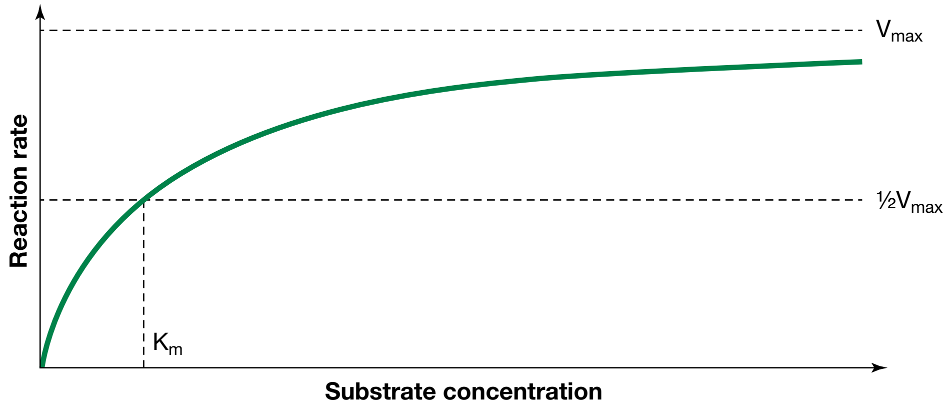



**FIGURE 3.17** Graph showing reaction rate plotted against increasing substrate concentration.

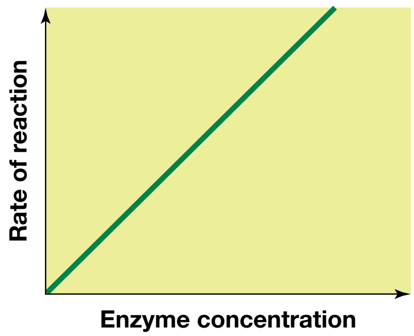



**FIGURE 3.18** Provided substrate concentrations are high and conditions are kept constant, increases in enzyme concentration are expected to produce an increase in reaction rate.

### Effect of enzyme concentration on reaction rate

Enzymes are not consumed in reactions and so are usually needed in small amounts only. For this reason, the question of the effect of increasing enzyme concentration on reaction rate is not a common concern. Provided that the substrate concentration is high and that temperature and pH are kept constant, the rate of reaction is proportional to the enzyme concentration, as shown in figure 3.18. In the most unusual situation of a very high concentration of enzyme, the substrate concentration may become rate-limiting and, if so, the reaction rate will stop increasing and will flatten out.

## Enzymes can be inhibited

An **enzyme inhibitor** is a molecule that binds to an enzyme, reducing its activity by interfering with the enzyme in some way.

### Inhibition can be irreversible or reversible

**Irreversible inhibition** of an enzyme occurs when a compound binds covalently to one or more amino acids and alters the structure of the enzyme, affecting its active site. This permanently inhibits the enzyme because the inhibitor-enzyme bond is so strong that the inhibition cannot be reversed by the addition of excess substrate. Not surprisingly, compounds that are irreversible inhibitors of enzymes are termed poisons. Examples of enzyme inhibitors that are irreversible include cyanide and nitrogen mustard gas.

For a discussion of cyanide poisoning, refer back to *Nature of Biology Book 1 Fifth Edition*, chapter 3, page 84.

- Cyanide is an irreversible inhibitor of cytochrome c oxidase, a key enzyme in the electron transport chain. Inhibition of this enzyme stops the production of ATP by aerobic respiration, resulting in death.
- Another irreversible inhibitor is the poison malonate, which binds covalently to the active site of the Krebs cycle enzyme, succinate dehydrogenase.
- DFP (diisopropylfluorophosphate), a nerve gas, is an irreversible enzyme inhibitor. DFP binds to a serine amino acid at the active site of the enzyme, acetylcholine esterase (ACE) (see figure 3.19). ACE plays an essential role at nerve synapses (see figure 5.22 on page 177) in deactivating neurotransmitters after a nerve impulse has passed. By inhibiting the action of ACE, DFP is a powerful neurotoxin, and small quantities of this poison cause continuous muscle contraction, constriction of the bronchial tubes of the lungs, seizures and death.

Irreversible inhibitors have applications in medicine. The drug penicillin acts as an irreversible inhibitor of the enzyme transpeptidase, which is needed by bacteria to form their cell walls, and so prevents them from reproducing. The anti-inflammatory and analgesic properties of aspirin are due to the fact

**studyon**

Unit 3 | AOS 1 | Topic 5 | Concept 6

**Enzyme inhibitors** Summary screen and practice questions

that this drug inhibits the activity of the platelet enzyme cyclooxygenase (COX) by reacting, chemically changing an amino acid at the active site.

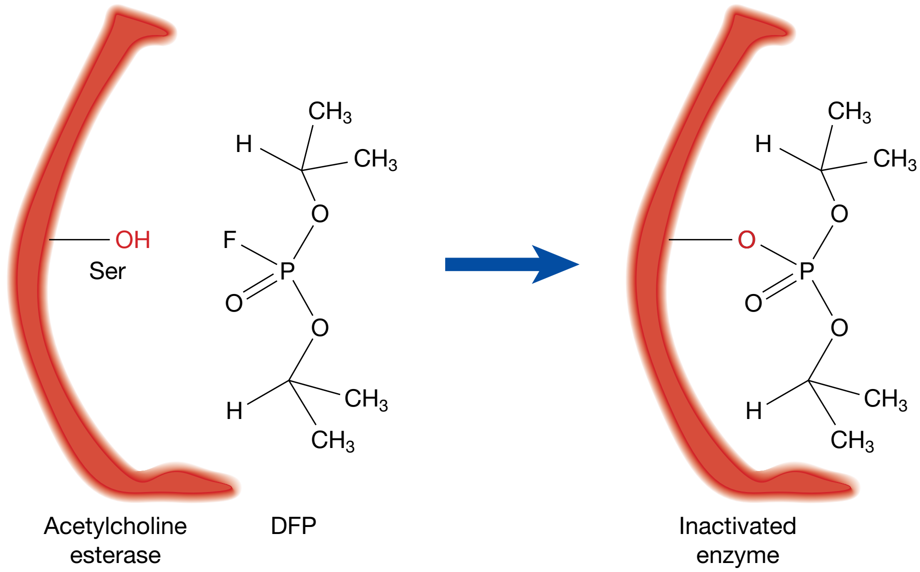



**FIGURE 3.19** Diagram showing the irreversible inhibition of acetylcholine esterase (ACE) enzyme by the nerve gas DFP. Note that DFP reacts with the side chain of a serine amino acid (Ser) at the active site of the ACE molecule, forming a strong covalent bond that blocks the active site.

In **reversible inhibition**, the enzyme is not permanently inhibited or damaged. Reversible inhibitors inactivate enzymes through noncovalent interactions that can be reversed. In contrast to the situation with an irreversible inhibitor, an inhibitor that is reversible can dissociate from the enzyme. Removal of the inhibitor reverses the inhibition.

Two types of reversible inhibition are competitive inhibition and non-competitive inhibition.

## Competitive inhibition

**Competitive inhibition:** Competitive enzyme inhibitors have a **similar shape** to the usual **substrate** molecule and so can bind to the active site of the enzyme, preventing the formation of the enzyme-substrate (E-S) complexes. As a result, over a given period of time, fewer substrate molecules bind to the active site of the enzyme so that the rate of the reaction is decreased (see figure 3.20).

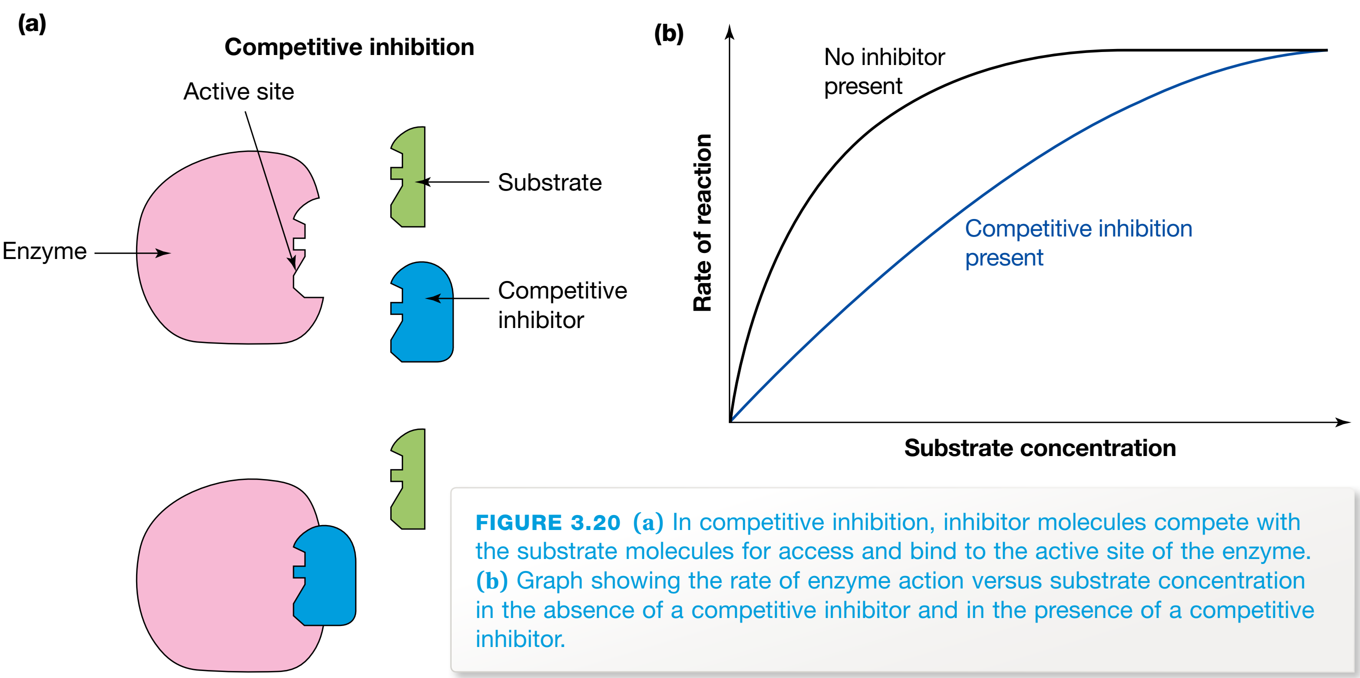



**FIGURE 3.20** **(a)** In competitive inhibition, inhibitor molecules compete with the substrate molecules for access and bind to the active site of the enzyme. **(b)** Graph showing the rate of enzyme action versus substrate concentration in the absence of a competitive inhibitor and in the presence of a competitive inhibitor.

Eventually, the inhibitor moves from the active site, so the competitive inhibition is usually temporary.

Because the substrate and the competitive inhibitor are competing for the same active sites on enzyme molecules, the degree of inhibition produced depends on the relative concentrations of substrate and inhibitor molecules. For example, adding more substrate molecules increases the chance that a random collision with the active site will be that involving a substrate molecule rather than that involving a molecule of the competitive inhibitor.

### Non-competitive inhibition

**Non-competitive inhibitors do not bind to the active site, but bind to another site on the enzyme molecule called the allosteric site**. The binding of the non-competitive inhibitor distorts the 3D shape of the enzyme so that it can no longer bind to its substrate and, therefore, can no longer catalyse the usual reaction (see figure 3.21). Adding more substrate molecules does not have any effect on this non-competitive inhibition. Can you suggest why?

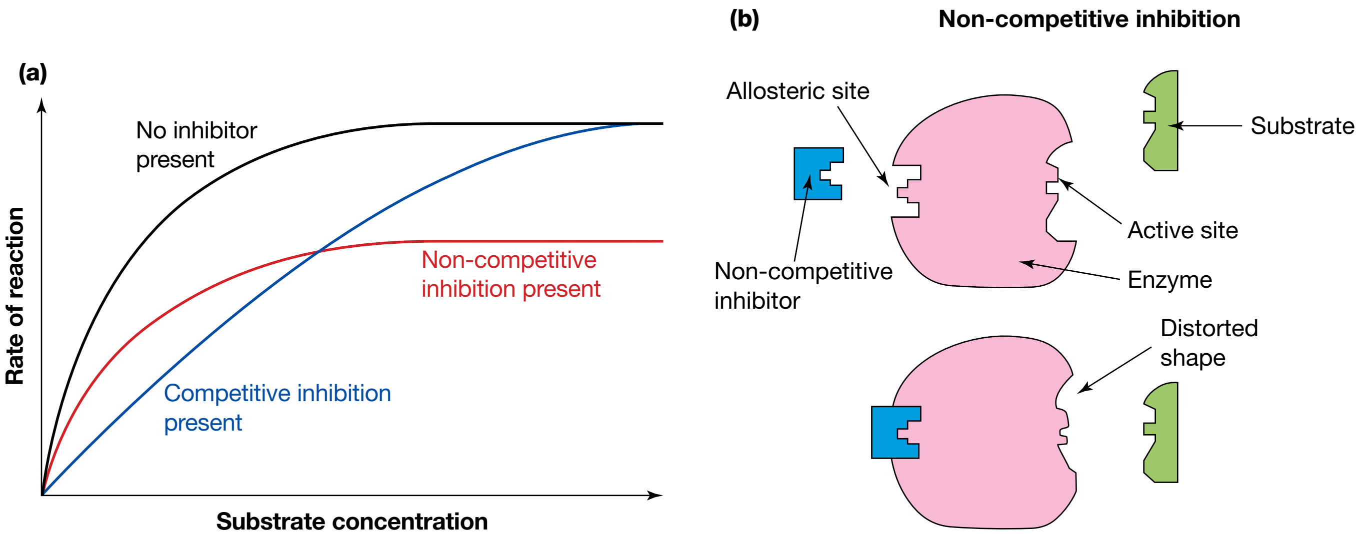



**FIGURE 3.21** **(a)** Graph showing rates of enzyme action versus substrate concentration in the absence of inhibitor (black line), in the presence of a non-competitive inhibitor (red line), and in the presence of a competitive inhibitor (blue line). These two different inhibitors bind to different sites on the enzyme molecule, with **(b)** showing the case for non-competitive inhibitors. Only in one case do the inhibitor molecules compete with substrate molecules for the same site on the enzyme. Which type of inhibition is that?

In summary, enzymes may have several binding sites, the most well known being their active site or substrate-binding site, while others may be present, including allosteric sites where binding inhibits enzyme function.

The presence of an allosteric binding site on an enzyme enables regulation of metabolic pathways to occur through a process termed **end-product inhibition**. Figure 3.22 shows an example of end-product inhibition where the end product of a metabolic pathway binds non-competitively to the first enzyme in the pathway, stopping the action of the enzyme. Regulation of this type means that products will not continue to be produced regardless of the need for them.

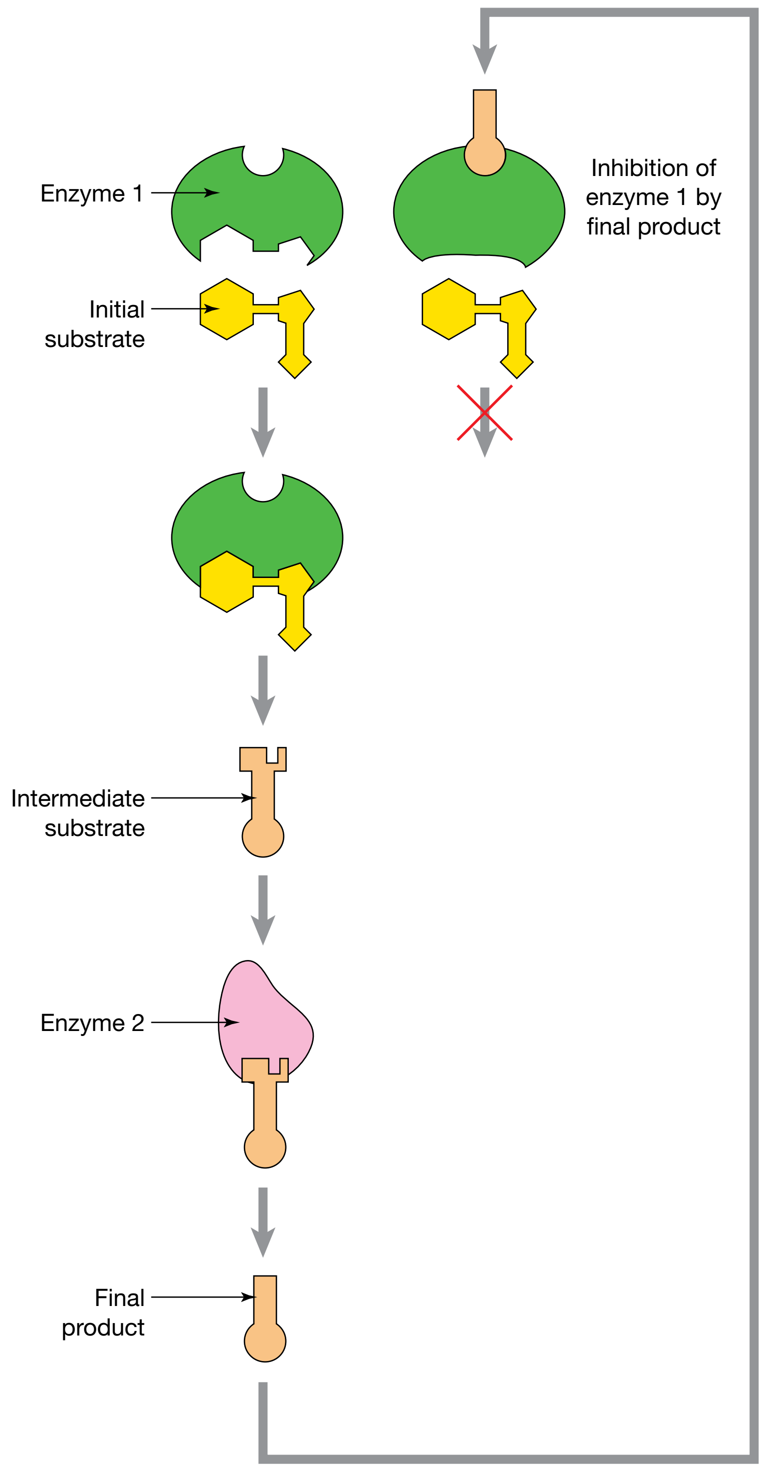



**FIGURE 3.22** Diagram showing the end-product inhibition of a simple metabolic pathway. Note that the end product does not competitively occupy the active site, but occupies a second site non-competitively. What will happen when the concentration of the end product decreases? Yes, the pathway will resume, and when the end product builds up again, the inhibition will recur.

## Enzyme's little helpers

All enzymes contain a protein backbone. For some enzymes, the protein makes up their entire structure. However, for many enzymes, additional non-protein components are present that are important contributors to both the activity and the stability of these enzymes. These additional non-protein groups needed for enzymes to function properly are termed **cofactors**.

Cofactors include both inorganic and organic molecules:

1. **Inorganic cofactors** include metal ions such as $Mg^{2+}$, $Cu^{2+}$, $Mn^{2+}$, or clusters of several ions.
   Examples of enzymes with a metal (inorganic) cofactor are shown in table 3.4.

**TABLE 3.4** Examples of enzymes with the requirement for metal ion cofactors.

| Enzyme | Ion |
|---|---|
| catalase | $Fe^{2+}$ or $Fe^{3+}$ |
| cytochrome oxidase | $Cu^{2+}$ |
| glucose-6-phosphatase | $Mg^{2+}$ |
| DNA polymerase | $Zn^{2+}$ |

2. **Organic cofactors** are sometimes divided into **coenzymes** and **prosthetic groups**, with prosthetic groups generally regarded as being those organic molecules that are bound tightly to particular enzymes.
   - ***Prosthetic groups:*** Some enzymes have a tightly bound heme group that is identified as a prosthetic group. These heme-containing enzymes include catalase and cytochrome c, and the heme molecule is a permanent part of these enzymes. Catalase is an oligomeric enzyme with four polypeptide chains and each chain has an associated heme molecule.
   - ***Coenzymes:*** Among the coenzymes are a number of key coenzymes that play a critical role in transferring functional groups in enzyme-catalysed reactions. These key coenzymes include NAD, NADP and ATP.

Let's look at NAD, which can exist in two forms:

Nicotinamide adenine dinucleotide (NAD) is found in all living cells. It is present in a **loaded form** as NADH and is a source of electrons and a helper for any enzyme that catalyses a reaction in which its substrate is reduced. When NAD unloads its electrons, it is oxidised and the enzyme substrate is reduced. NAD is also present in cells in an **unloaded form** in which NAD is a receiver of electrons and is a helper for any enzyme that catalyses a reaction in which its substrate is oxidised. When NAD loads up electrons, it is reduced and the substrate is oxidised. As NAD transitions between its loaded and unloaded forms, it switches from being a helper for the group of enzymes that catalyse reduction reactions to being a helper for those enzymes catalysing oxidation reactions.

Figure 3.23 shows the loaded and unloaded ('empty') forms of the NAD coenzyme.

So, NAD is a versatile molecule involved in a large number of redox reactions, sometime donating electrons and sometimes accepting them. It never rests . . . whatever form it is in, either loaded or unloaded, NAD can serve as a coenzyme for groups of enzymes involved in metabolic pathways. A similar story applies to the coenzyme NADP, which plays a comparable role in plants.

Let's now look at ATP as a coenzyme that transports chemical energy within cells for use in metabolic pathways. This coenzyme exists in the energy-rich form of ATP (adenosine triphosphate), which can provide energy to drive anabolic reactions when one of its phosphate groups is hydrolysed and it is converted to ADP (adenosine diphosphate).

So, ATP is consumed in cells by energy-requiring (endergonic) processes, but ATP is regenerated by energy-releasing (exergonic) processes. Through these processes, ATP transfers energy between various metabolic pathways occurring in cells. There are many ATP-dependent enzymes in cells and these include ATP-dependent DNA ligases, ATP-dependent restriction enzymes, ATP-dependent deoxyribonucleases, and ATP-dependent phosphofructokinases.

**studyon**

Unit 3 | AOS 1 | Topic 5 | Concept 7

**ATP as the energy currency for cells**
Summary screen and practice questions

**studyon**

Unit 3 | AOS 1 | Topic 5 | Concept 7

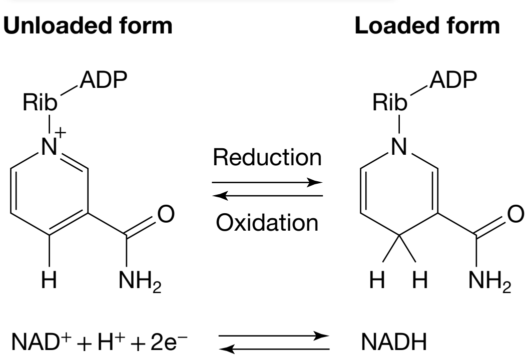
**See more**
ATP is the energy currency of cells

**studyon**

Unit 3 | AOS 1 | Topic 5 | Concept 8

**The role of NADH and NAD**
Summary screen and practice questions

**FIGURE 3.23** Diagram showing the two forms of nicotinamide adenine dinucleotide (NAD) — the loaded form, NADH, and the empty or unloaded form, NAD.



## BIOLOGIST AT WORK

### Professor Leigh Ackland — molecular biologist

Professor Leigh Ackland is a research scientist and lecturer at the Centre for Cellular and Molecular Biology, School of Life and Environmental Sciences at Deakin University. Leigh writes:

'As a research biologist, I have had a wonderful career working with stimulating and creative colleagues to unravel the mysteries of biological processes. One of my main interests is in the basic building blocks of biology: cells. To understand the functions of cells, we have developed specialised techniques to enable cells from different parts of the body to be grown outside the body: a technique termed *tissue culture* (see figure 3.24). Recent knowledge of the nutritional requirements of different types of cells has enabled us to grow cells taken from the skin, gut, breast and blood. Using this approach, we have learned much about the life of a cell, both in health and disease. In culture, we can make cells become specialised to carry out different functions, communicate with each other and their environment, and eventually die.

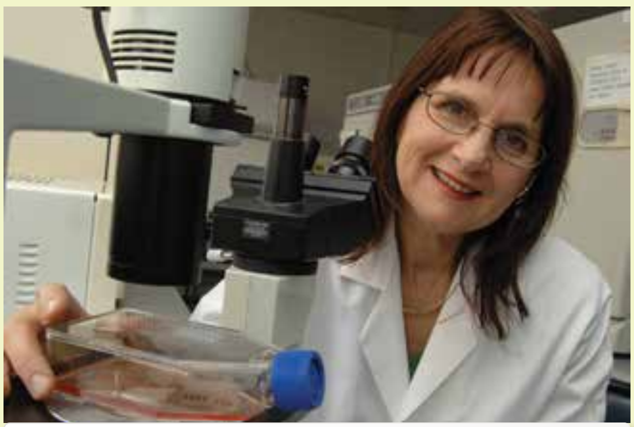


**FIGURE 3.24** Professor Leigh Ackland grows cancer cells in a transparent flask so that the live cells can be seen under the phase contrast microscope. (Image courtesy of Professor Leigh Ackland, Deacon University.)

The technique of tissue culture has provided a wealth of information about the behaviour of cancer cells. Cancer cells start their lives as normal cells but undergo changes, causing them to grow out of control and to lose contact with each other and with the substrate to which they are attached. These changes can eventually lead to cells spreading around the body.

Tissue culture has enabled us to develop a model of the human breast for studying cancer. Breast cancer arises from the glandular part of the breast, which consists of epithelia. Different epithelial cells can be identified by the types of structural proteins (cytoskeleton) they contain.

Tissue culture models and other studies indicate that breast cancer cells take on a mode of behaviour similar to that of developing cells that migrate during normal embryonic development. Cancer may thus be a recapitulation of some of the normal stages of development but is inappropriate in the context of adult function. The conversion of normal cells to those with abnormal behaviours characteristic of cancer cells is thought to occur through a process termed the *epithelial-to-mesenchymal transition*.

Our studies showed that cells that had undergone the *epithelial-to-mesenchymal transition* to become cancerous could not form proper contacts with each other or with the extracellular environment. They also had alterations in a cytoskeletal protein called vimentin. Vimentin is a hallmark of the *epithelial-to-mesenchymal transition* and is associated with the capacity of cells to invade or metastasise. Studies on cultured cells from breast cancer patients show increased levels of vimentin that correlates with the degree of invasiveness of the cancer.

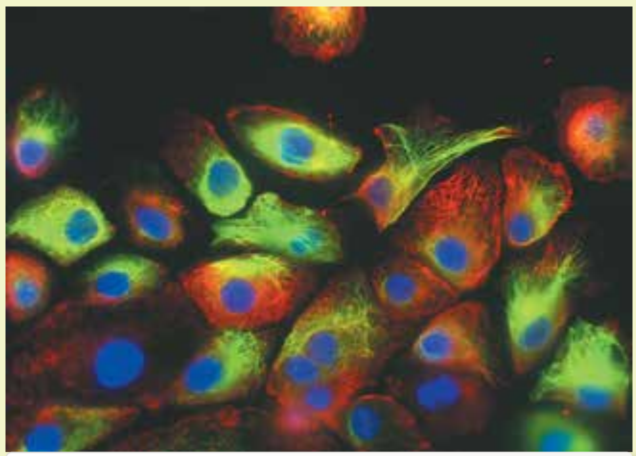


**FIGURE 3.25** Cancerous breast cells viewed with a fluorescence microscope after staining for the presence of vimentin (green) and keratin (red). (Image courtesy of Professor Leigh Ackland, Deacon University.)

The Australian Synchrotron, located in Melbourne, has provided us with novel imaging technologies for our research. This amazing facility generates accelerating electrons that produce beamlines with different properties to see inside cells. We have used the X-ray fluorescent beamline to detect differences between normal and diseased tissue.

My interest in science was kindled by my maternal grandfather, a chemistry teacher, who took me on expeditions to places such as museums and questioned the science behind what we saw. His inspiration stimulated me to study science at school and then at university.'

## KEY IDEAS

- Two models to explain enzyme specificity have been proposed.
- Enzymes lower the energy level needed to activate reactant molecules as compared with the non-enzyme catalysed situation.
- Enzymes have an active site where their specific substrate binds, and typically have an allosteric site where inhibitors bind.
- The rate of an enzyme-catalysed reaction is affected by several factors, mainly temperature, pH and substrate concentration.
- Enzymes can be inhibited either irreversibly or reversibly.
- Reversible inhibition includes competitive and non-competitive inhibition.
- End-product inhibition is a form of non-competitive inhibition that is used to regulate the operation of a metabolic pathway.
- Most enzymes have non-protein components, termed cofactors, involved in their structure and/or function.
- One key group of cofactors includes the coenzymes NADH, NADP and ATP, which move functional groups and energy between metabolic pathways in cells.

## QUICK CHECK

**4** Briefly explain the meanings of the following terms:
  **a** activation energy
  **b** heat denaturation
  **c** pH optimum.

**5** Identify one key difference between the members of the following pairs:
  **a** competitive and non-competitive inhibition
  **b** reversible and irreversible inhibition
  **c** loaded and unloaded forms of NAD
  **d** active site and allosteric site.

**6** Identify the following statements as true or false:
  **a** Enzymes change the activation energy of a reaction by providing a different reaction pathway.
  **b** Non-competitive inhibition can be reduced by increasing the concentration of substrate.
  **c** Competitive inhibition can be reduced by increasing the concentration of substrate.
  **d** Irreversible inhibition typically involves the formation of covalent bonds between the inhibitor and the enzyme with a consequent change in the shape of the enzyme.
  **e** End-product inhibition works only when the end product has a shape that matches the active site of the first enzyme in the metabolic pathway.
  **f** When in its loaded form, NAD acts as a helper to any enzymes that catalyse reactions in which their substrate will be reduced.

# BIOCHALLENGE

1 The graph immediately below (figure 3.26) shows the behaviour of enzyme X under five different conditions.
   a What variable is plotted:
      i on the vertical axis of this graph
      ii on the horizontal axis?
   b How many experimental results are shown on this graph?
   c What condition was kept constant throughout the course of one experiment?

A *dependent variable* is the *variable* being tested in a scientific experiment and its value depends on the independent variable. As scientists change the *independent variable*, they observe and record the change in the *dependent variable*.
   d In these experiments, what is:
      i the dependent variable
      ii the independent variable?
   e What conclusion(s) about enzyme X can be drawn from these results?

Enzyme Y was isolated from extremophile microbes that live in hot volcanic waters.
   f Would you predict that enzyme Y would show similar behaviour to enzyme X when subjected to the same experimental conditions? Explain.

Examine the data shown in these graphical results and answer the following questions:
   g At what temperature did enzyme X lose about half its activity after just a few minutes?
   h After one hour at 60 °C, what percentage activity of enzyme X remained?
   i At 65 °C, about how long did it take for enzyme X to lose about 80% of its maximum activity level?
   j What happened to enzyme X during the first 10 minutes of its exposure at 70 °C?
   k Explain why the behaviour of enzyme X differs at 50 °C compared with its behaviour at 70 °C.

2 Do an internet search using the term 'metabolic pathways' and ask for images. You can expect to see some complex images that show the numerous metabolic pathways that operate in living organisms. Remember that the purpose of this activity is simply to give you a sense of the complexity of metabolic pathways **and to alert you to the key role of enzymes in virtually every step in these pathways**.
   a Examine several of these images to see some detail of metabolism. Is the scope of cellular metabolism what you expected?
   b One site states that these metabolic pathways are 'daunting in their majesty beauty'. Do you agree?
   c Many of the images may show the various substrates and their products in metabolic pathways, but omit the enzymes involved in the various steps.

Can you locate a metabolic map that also gives the enzymes involved?

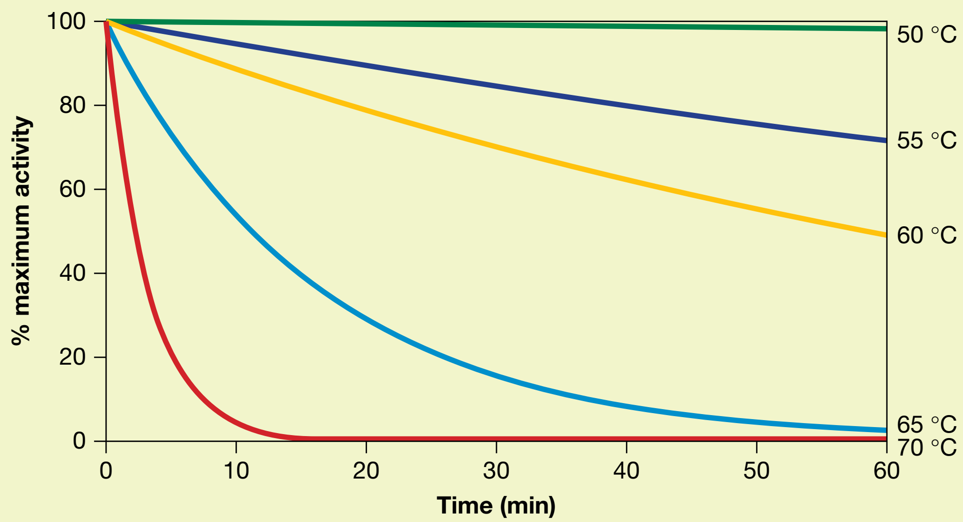



FIGURE 3.26

# Chapter review

studyon

Unit 3 AOS 1 Topic 5 — Structure and regulation of biochemical pathways — Practice questions

## Key words

activation energy
active site
allosteric site
anabolism
catabolism
coenzyme
cofactors
coupling
endergonic
end-product inhibition
enzyme inhibitor
enzyme-substrate (E-S) complex
exergonic
gulono-lactone oxidase (GULO)
induced fit model
inorganic cofactors
irreversible inhibition
loaded form
lock-and-key model
metabolism
monomeric enzymes
oligomers
organic cofactors
product
proenzymes
prosthetic group
reversible inhibition
scurvy
similar shape
substrate
substrate binding site
systematic name
trivial name
unloaded form

## Questions

1 **Making connections between concepts** → Use at least eight of the key words from this chapter to construct a concept map relating to enzymes.

2 **Applying knowledge and understanding** → Examine figure 3.27. Compound A is a type of protein.
  a What is the general name given to the type of protein represented by compound A?
  b What are the general names given to each of the parts labelled B, C, D and E?
  c Name the type of reaction shown in this diagram — is it catabolic or anabolic? Explain.
  d Is this reaction more likely to be endergonic or exergonic? Explain.

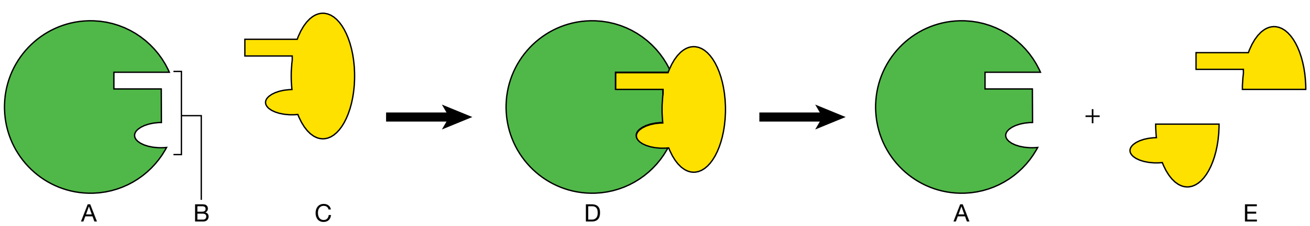
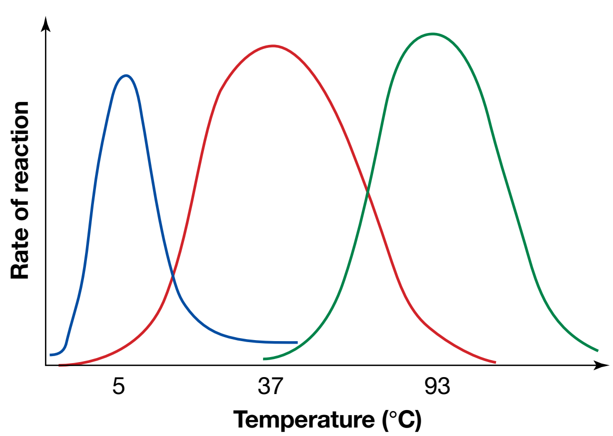



FIGURE 3.27

3 **Applying knowledge and understanding to make predictions** → The graph below shows the effect of temperature on the activity of an enzyme from three different organisms.




  a What predictions, if any, might be made about the organisms that were the sources of these enzymes? Give a brief rationale for any predictions that you make.
  b Note that the shapes of the curves are different. Suggest a reasonable explanation for these differences.

4 **Drawing conclusions** → Amino acids in a polypeptide chain are numbered from 1 to n, starting from the amino terminal; for example, the third amino acid might be Val 3 and the tenth in the sequence might be Leu 10, and so on . . .

The active site of the enzyme trypsin, a digestive enzyme, consists of three amino acid as follows: Asp 102, His 57 and Ser 195, that is, these three amino acids are the 57th, the 102nd and the 195th amino acids in the primary structure of this enzyme.
  a Identify two events that could occur at the active site of an enzyme.
  b What stabilises a substrate in the active site of an enzyme?

One student suggested that the three amino acids at the active site of trypsin had been misnumbered. The student said that there is quite a distance separating them so that they cannot be close together at the active site.

**c** Do you agree with the student? Give a clear explanation for the decision you have made.

**5 Demonstrating knowledge and communication →** Identify a key difference between the members of the following pairs:

**a** 'lock-and-key' model and 'induced fit' model
**b** a prosthetic group and a coenzyme
**c** loaded NAD and unloaded NAD
**d** competitive inhibition and end-product inhibition
**e** catabolism and anabolism
**f** pepsinogen and pepsin.

**6 Interpreting diagrammatic information →**

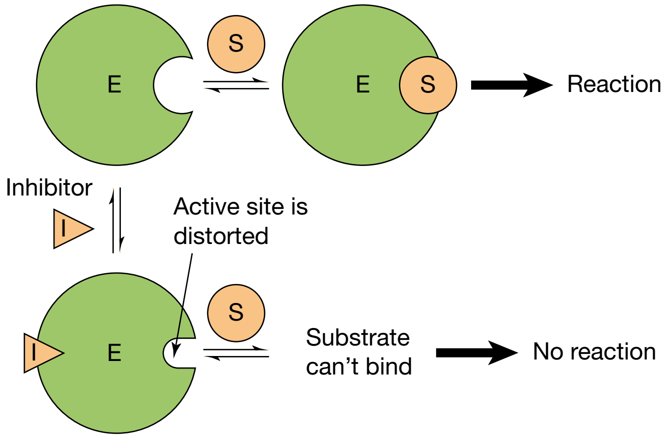
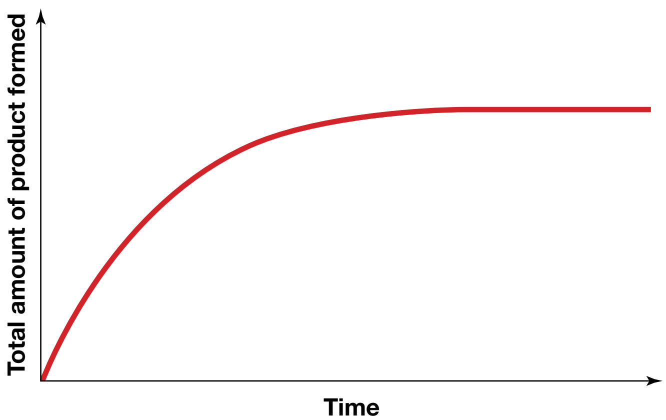



**FIGURE 3.28**

Examine figure 3.28 above.
What story does this diagram tell?

**7 Demonstrating understanding and communication →** Consider the following figure, which shows the total amount of product formed during an enzyme-catalysed reaction.



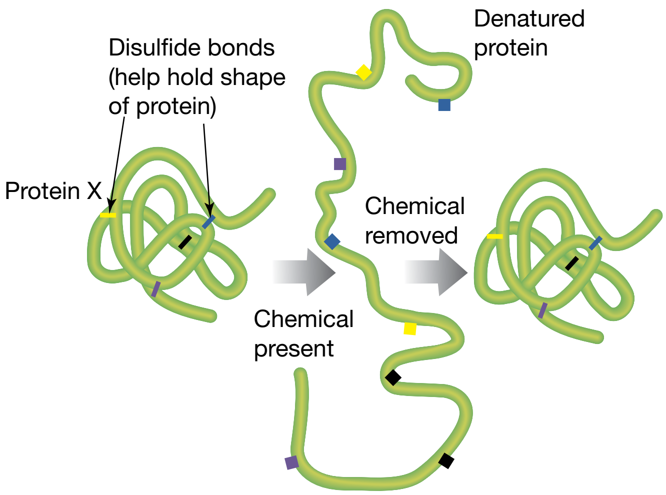

Explain the following features of this graph:

**a** the initial steep slope
**b** the slowdown in the rate of product formation
**c** the flattening of the curve to form a plateau.

**8 Developing logical explanations →** Consider each of the following observations and formulate an explanation for the observation.

**a** Salivary amylase enzyme breaks down starch. The process begins when saliva mixes with food in the mouth, but, when it reaches the stomach, the amylase enzyme no longer functions.
**b** The jelly-like white of a raw hen's egg consists principally of the protein albumin, but, after it is cooked, it forms a white solid.
**c** When an inhibitor binds to an enzyme at a site other than the active site, this inhibition cannot be overcome by increasing the concentration of substrate molecules.
**d** Enzyme-catalysed reactions operate at much faster rates than non-catalysed reactions.

**9 Demonstrating knowledge and understanding →** Protein X is an enzyme.




**a** What is the general shape of protein X (as shown in the left-hand image)?
**b** When treated with a particular chemical, the shape of protein X was radically changed, as shown in the middle image. What specific effects has the chemical had on protein X?
**c** Describe what happened when the chemical was removed.
**d** What labels would you put on the action of the chemical on protein X?
**e** How, if at all, would the ability of protein X to catalyse its specific reaction be affected when it was in the 'long snake form'?

CHAPTER

# 4 Energy transfers and transformations

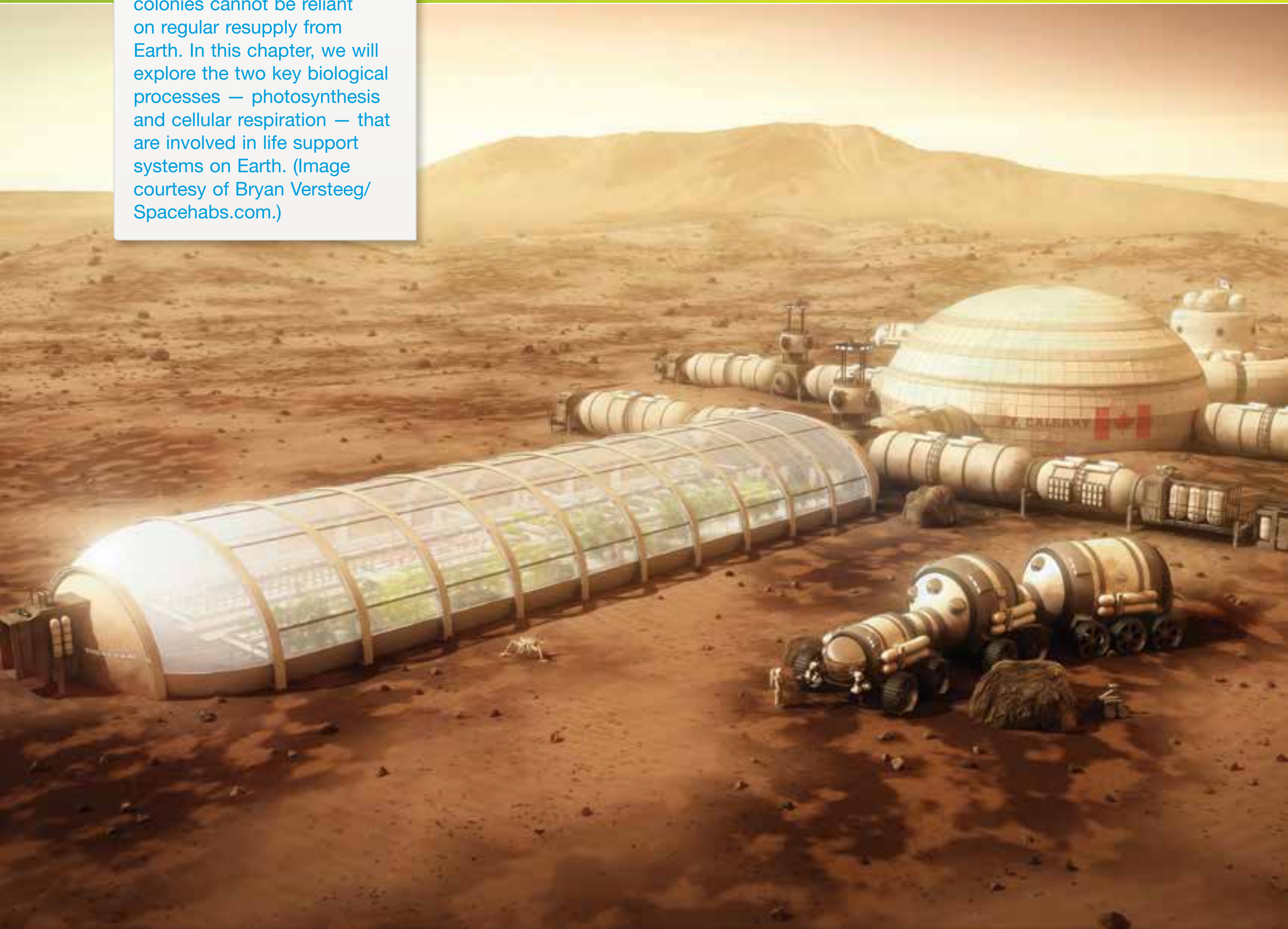


**FIGURE 4.1** Conceptual design for a human colony on Mars. Human survival on long-haul space flights and in extraterrestrial colonies will depend on robust and sustainable life support systems. In the case of long-haul space flights, taking supplies of food, water and oxygen from Earth is impractical in terms of storage and weight requirements. Similarly, extraterrestrial colonies cannot be reliant on regular resupply from Earth. In this chapter, we will explore the two key biological processes — photosynthesis and cellular respiration — that are involved in life support systems on Earth. (Image courtesy of Bryan Versteeg/ Spacehabs.com.)

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- extend their understanding of the purpose of photosynthesis
- identify the inputs and outputs of photosynthesis and its cellular site of operation
- develop awareness of factors that influence the rate of photosynthesis
- extend their understanding of the purpose of cellular respiration
- identify the inputs and outputs of aerobic and anaerobic cellular respiration and their cellular sites of operation
- extend knowledge of the origin of mitochondria and chloroplasts as endosymbionts.

# Life support in space

Humans have ventured beyond planet Earth on short-term missions to the moon. The *Apollo 11* mission, which first took humans to walk on the moon, travelled from Earth to an orbit around the moon in just over 50 hours. However, longer durations of human space travel, such as a mission to planet Mars, is confronted by major technical and biological challenges that must be addressed before the prospect can become a reality. Among these biological challenges is how to supply the basic needs of astronauts — water, food and oxygen — during long space voyages involving hundreds of days or more for the round trip.

One goal of the USA National Aeronautics and Space Administration (NASA) is for human travel to Mars in the 2030s.

Is this a case of '*No problem. Just take all the essential supplies with you in the spacecraft'*? Doing just this has been the practice for short-term space flights. All the supplies of water, food and oxygen needed to support the astronauts are carried on board the spacecraft, and wastes are stored and taken back to Earth. Under these circumstances, the spaceship is an open life support system with materials imported from and exported to the external environment.

The needs of metabolic consumables (oxygen, food and drinking water) to maintain life are estimated to be about 5 kg per day per person. So, consider a crew of six astronauts on a hypothetical space mission that, using a low-energy transfer trajectory, would take about 250 days to reach Mars. To meet the crew's needs for just the 250-day outward trip would require about 7.5 metric tons (7500 kg) of metabolic supplies with a launch cost of about US$20 000 per kg. Meeting the life support needs of crews on long-haul space flights in this manner is neither economical nor practical.

**ODD FACT**

This calculation does not include the estimated requirement for hygiene water of 20 kg per day per person that is needed for washing and the like.

Future long-haul manned space flights will need to make use of closed life support systems — that is, systems in which material neither enters nor leaves the system. In a closed system, waste is constantly recycled to generate food, water and oxygen. As well as their use in long-haul space flights, closed life support systems would be required to sustain extraterrestrial colonisation, such as a long-term human colony on the airless, waterless moon. This is even more of an issue for any future human colony on Mars. (The distance of Mars from Earth can vary, depending on their relative orbital positions, from as close as about 55 million kilometres to as far as about 400 million kilometres.)

### Project MELiSSA: towards a closed life support system

**MELiSSA**, the **M**icro-**E**cological **Li**fe **S**upport **S**ystem **A**lternative, is a project managed by the European Space Agency (ESA) that involves a consortium, including universities, research institutions and industries. The goal of this project is to create a closed life support system in which organic waste materials from astronauts are converted to life-sustaining food, oxygen and water. Dr Christophe Lasseur, the ESA project manager of MELiSSA, has commented that '*At the moment we bring from Earth all foods for our space missions, on the International Space Station, the Soyuz* [a Russian spacecraft] *and the Shuttle* [an American spacecraft]. *For longer missions to Mars, however, we cannot bring everything from Earth. That would be too heavy and require too much storage room.*'

In essence, the MELiSSA project seeks to produce a closed artificial **ecosystem** based on the photosynthetic activity of microorganisms and higher plants to supply the oxygen, food and water to sustain human life (see figure 4.2a). In turn, the microorganisms and higher plants are sustained by the organic wastes of astronauts. The ideal target of the MELiSSA project is to achieve the recycling of 100 per cent of all the chemical components and produce a fully self-sustainable ecosystem without the need for input of materials from outside the system. The only input is that of the energy of sunlight to drive the system.

An **ecosystem** is a community of interacting living organisms functioning as producers, consumers and decomposers. (Refer back to *Nature of Biology 1 5th Edition*, Chapter 8.)

Five compartments comprise the closed loop of the experimental MELiSSA artificial ecosystem. Figure 4.3 shows these compartments and the transfers of material that occur between them.



**FIGURE 4.2** **(a)** A MELiSSA researcher at the University of Nantes, France, working with a bioreactor and a light source to grow algae for use as food in space. Algae-based foods are high in protein and energy. The algae would also provide a source of oxygen as it recycles the carbon dioxide in the atmosphere — an essential part of a long-term, self-contained space mission. The ESA MELiSSA consortium consists of more than 40 partners, including universities, research centres and industries, from 13 countries. **(b)** The MELiSSA logo. (Image (b) courtesy of ESA.)

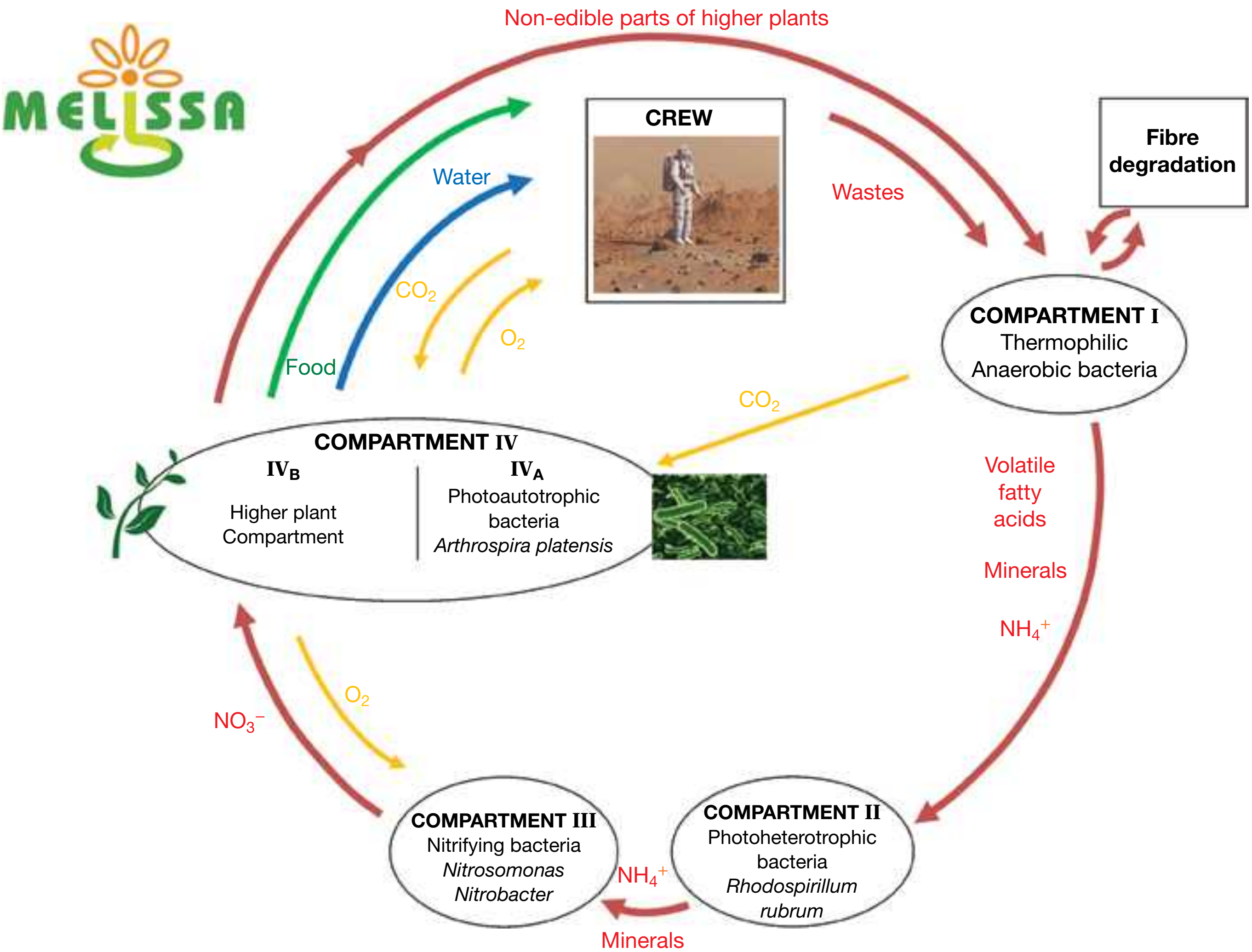



**FIGURE 4.3** The five compartments that form the closed-loop concept being researched in the MELiSSA project. (Compartment V contains the crew.) Note the various transfers of material that occur between these compartments. Between which two compartments does the greatest number of material transfers occur? The box labelled 'Fibre degradation' denotes the operation of a mechanical process that breaks down fibrous waste plant matter into smaller pieces. (Image courtesy of ESA.)

The following box provides details of each of the five closed compartments.

## THE FIVE COMPARTMENTS OF THE MELiSSA CLOSED LOOP

### Compartment I, the LIQUIFYING compartment:

***Inputs:*** All **organic waste**, including food wastes, faeces, urine, paper (from Compartment V), non-edible parts of the higher plant (straw, roots) (from Compartment IV$_B$) and the non-edible microbial biomass (from Compartment IV$_A$).

***Active occupants:*** Mixture of anaerobic bacteria, including thermophilic anaerobes and obligate anaerobic bacteria found in human faeces.

***Function:*** Microbial degradation and liquefaction of complex organic wastes.

***Outputs:*** Ammonium ions ($NH_4^+$), carbon dioxide ($CO_2$), minerals and volatile (short-chain) fatty acids, such as acetic acid and butyric acid.

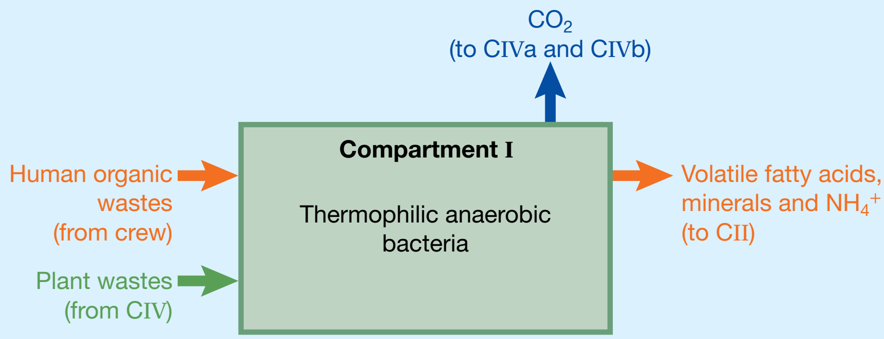



FIGURE 4.4a

### Compartment II, the PHOTOHETEROTROPHIC compartment:

***Inputs***: Fatty acids, ammonium ions, and minerals (from Compartment I).

***Active occupants***: Photoheterotrophic bacteria, *Rhodospirillum rubum*. These bacteria use sunlight as their energy source but do not use carbon dioxide as their carbon source; instead, they use fatty acids to build sugars.

***Function***: Elimination of volatile fatty acids.

***Outputs:*** Minerals and ammonium ions.

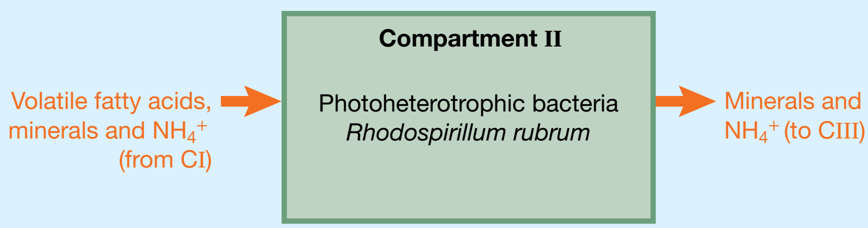



FIGURE 4.4b

### Compartment III, the NITRIFYING compartment:

***Inputs***: Minerals and ammonium ions (from Compartment II).

***Active occupants:*** Nitrifying bacterial species *Nitrosomonas* sp., which oxidise ammonium ions ($NH_4^+$) to nitrites ($NO_2^-$), and *Nitrobacter* sp. which oxidise nitrites ($NO_2^-$) to nitrates ($NO_3^-$).

***Function***: Production of nitrates ($NO_3^-$) required for plant growth.

***Outputs***: Nitrates and minerals.

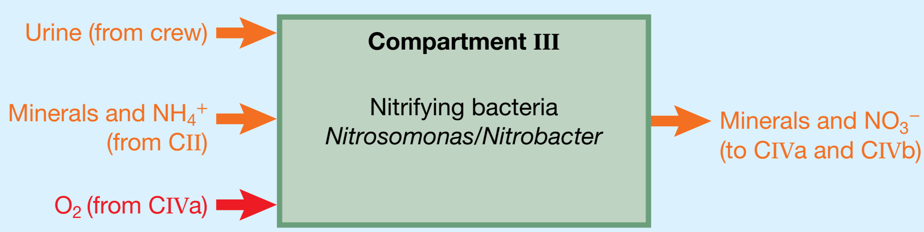



FIGURE 4.4c

### Compartment IV, the PHOTOAUTOTROPHIC compartment:

***Inputs***: Minerals and nitrates (from Compartment III) and carbon dioxide (from Compartments I and V).

***Active occupants***: This compartment (IV) is composed of two sections (Compartments $IV_A$ and $IV_B$).

- Spiral cyanobacteria, *Arthrospira platensis*, commonly known as Spirulina, occupy Compartment $IV_A$. Spirulina has been harvested for food in South America and Africa for centuries; it constitutes a protein-rich food with eight essential amino acids, minerals, essential fatty acids (omega-3 and omega-6), and essential elements (calcium, phosphorus, sodium, potassium, magnesium) and trace elements (copper, zinc, selenium, chromium).
- Higher plants occupy Compartment $IV_B$, such as food crops including wheat, rice, spinach, lettuce, soybean and tomato.

***Function***: Production of oxygen, food and water by photosynthesis.

***Outputs***: Oxygen, food and water (and non-edible plant fibrous waste).

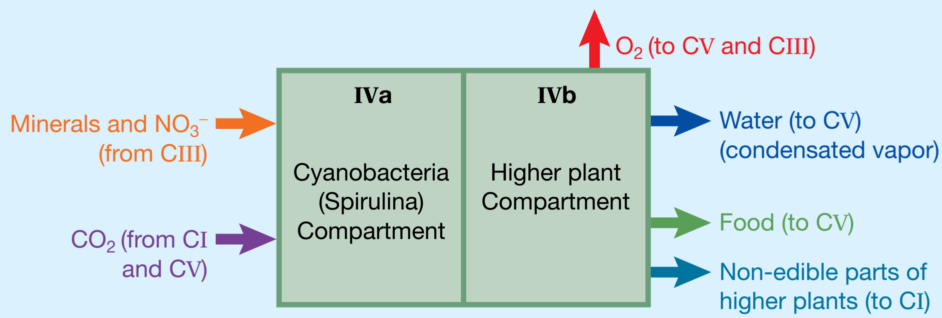



FIGURE 4.4d

### Compartment V, the CREW compartment:

***Inputs***: Water, food and oxygen.

***Occupied by***: Human crew.

***Function***: Life support input to crew.

***Outputs***: Wastes, including urine, faeces, paper (to Compartment I) and carbon dioxide (to Compartments $IV_A$ and $IV_B$).

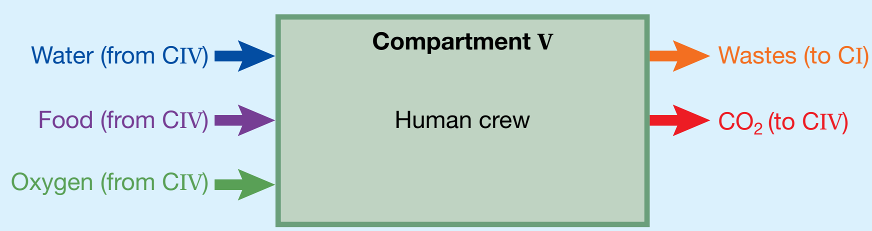



FIGURE 4.4e

## MELiSSA research and research spin-offs

The processes that operate in various compartments of the closed loop are being studied, modified and improved by research groups in universities within the MELiSSA consortium.

One research group, at the Universitat Autònoma de Barcelona in Spain, is testing the robustness and stability of a closed-loop life support system in continuous operation (see figure 4.5). Their project is monitoring the survival in a closed system of a group of 40 rats (equivalent to the cellular respiration of one person) over a two-year period. (For obvious safety reasons, 'space crew' members are rats.)

The research and development that underpins the MELiSSA project has as its primary aims the survival and well-being of astronauts on future long-term space missions. However, this research has generated new technologies that have already found application at home on Earth. For example:

- MELiSSA's grey water treatment technology is being used in Europe to treat more than 1.8 million cubic metres of water each day. While existing

technologies can convert grey water to a standard useful for flushing, the MELiSSA technology can convert grey water to a standard suitable for drinking. This technology is being used in the Concordia research station in Antarctica.
- MELiSSA's biomass sensors can directly and rapidly measure the viable biomass suspended in a medium.

**FIGURE 4.5** The MELiSSA pilot plant at the Universitat Autònoma de Barcelona in Spain, where testing of robustness and stability of a closed-loop life support system in continuous operation is being carried out. Note the green colouring in the vertical columns (at right). The green is due to the presence of the cyanobacteria *Arthrospira platensis*, commonly known as Spirulina. What role do these cyanobacteria play in the closed-loop life support system?

This type of sensor has been adapted by wine, engineering and pharmaceutical companies.

One of the key components of the MELiSSA closed-loop life support system is Compartment IV, where **photoautotrophs** produce oxygen, food and water through the process of photosynthesis. In the following section, we will explore this important process. In a later section, we will explore the key process of cellular respiration and see the link between photosynthesis and cellular respiration.

Photoautotrophs, chemoautotrophs and other styles of living are identified in the box on page 152.

## KEY IDEAS

- One of the many challenges to be overcome before extended manned space flights and extraterrestrial colonisation are possible is the provision of a robust and sustainable life support system.
- The MELiSSA project aims to produce a closed artificial ecosystem to sustain human life during extended periods away from Earth.
- The idealised concept of a closed life support system is 100 per cent recycling of waste to generate the oxygen, food and water required to support human life.

## QUICK CHECK

1. What is the key difference between a closed system and an open system?
2. Refer back to the figures showing the five compartments of the MELiSSA closed loop and identify the following statements as true or false:
   a. Most of the outputs of Compartment IV are inputs to Compartment V.
   b. All the inputs to Compartment I are human wastes.
   c. The key function of Compartment III is to convert nitrogenous compounds into nitrates, a form that can be used by plants.
3. a. An organism is described as being 'photoautotrophic'. What is a distinguishing feature of such an organism?
   b. If these bacteria replaced those in Compartment I, would this affect the operation of the closed-loop system? Explain.

# Photosynthesis: energy from the sun

Every day, a huge amount of radiant energy emitted from a star about 150 million kilometres away reaches Earth. This star is of course the sun (see figure 4.6). This radiant energy warms the oceans and the land masses of this planet, and creates winds and cyclones. It is energy for free!

studyon

| | |
|---|---|
| Unit 3 | **Describing photosynthesis** Summary screen and practice questions |
| AOS 1 | |
| Topic 6 | |
| Concept 1 | |

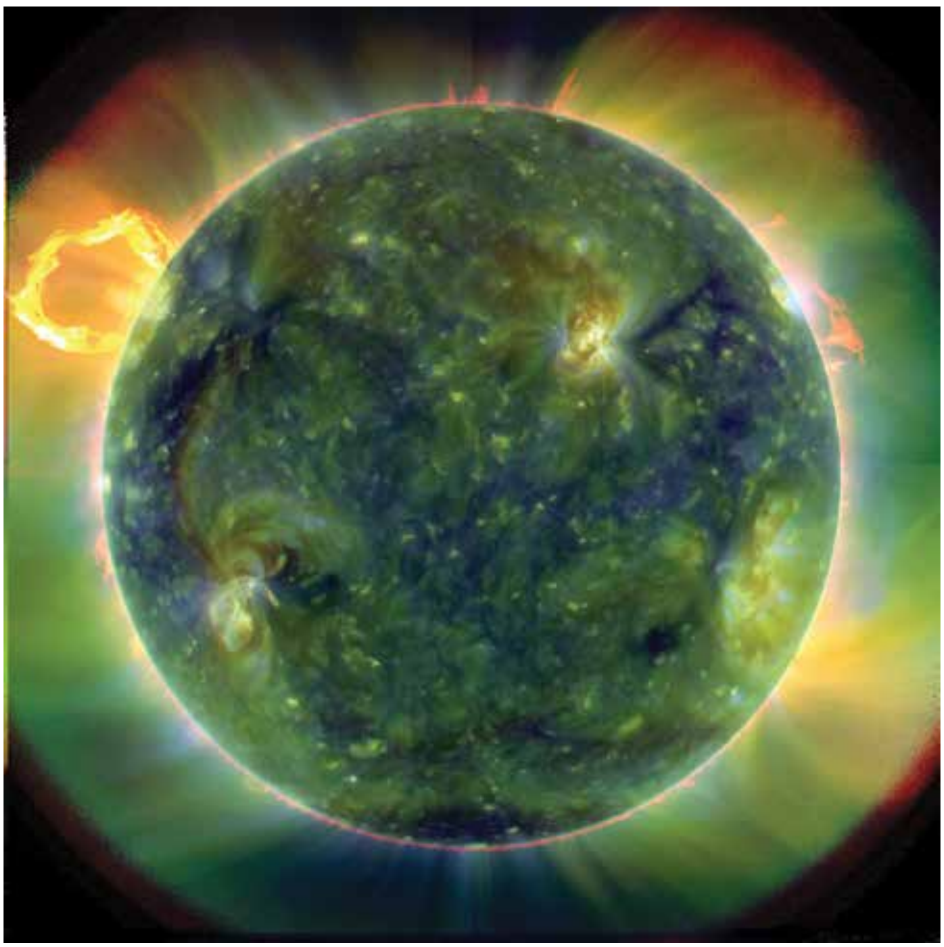


**FIGURE 4.6** An image of the sun as revealed by a small exploratory spacecraft launched in July 2013 as part of the IRIS mission. This spacecraft is equipped with a 20 cm ultraviolet (UV) telescope combined with an Interface Region Imaging Spectrograph (IRIS). Images obtained are revealing new detail of the lower layer of the sun's atmosphere. (Image courtesy of NASA.)

Being alive requires a constant input of energy. It is not surprising that many of Earth's life forms have evolved to exploit the energy provided free by the sun. There is no need for these organisms to hunt energy in the form of prey; there is no need to be mobile — they can just stand or float and expose their surfaces to the radiant energy of the sun.

## Photosynthesis: the big picture

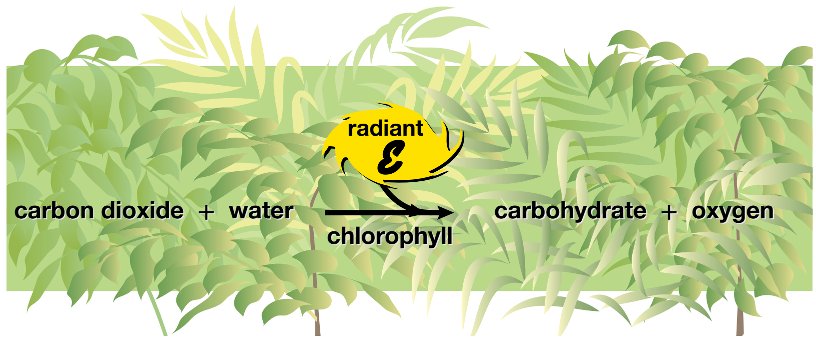
Plants, algae, some protists (such as phytoplankton) and some bacteria can use the energy of sunlight to make organic molecules, such as sugars, by photosynthesis. Organisms with this ability are described as being photoautotrophic (*photo* = light + *auto* = self + *trophe* = nourishment) because they use sunlight to manufacture **organic compounds**, such as sugars, that provide them with the energy needed for living. They are also termed **producers** because, in an ecosystem community, they are the source of food for non-photosynthetic members of the community. All other organisms, such as animals and fungi, depend, directly or indirectly, for their energy for living on the organic compounds produced by photoautotrophic organisms.

The term 'organic compound' refers to carbon-containing molecules that are (i) typically found in living organisms and (ii) contain C-H covalent bonds.

You were introduced to some aspects of photosynthesis in *Nature of Biology, Book 1*. In this section, you will learn more about the processes involved.

Photosynthesis is the process in which light energy is transformed into chemical energy stored in organic molecules, such as sugars. In a typical producer, such as a terrestrial flowering plant, the complex series of reactions in photosynthesis can be summarised as follows:

**studyon**
Unit 3 | AOS 1 | Topic 6 | Concept 1
**Do more** Photosynthesis summary equation




In fact, the complete balanced equation for photosynthesis is:

$$6CO_2 + 12H_2O \xrightarrow[\text{chlorophyll}]{\text{light}} C_6H_{12}O_6 + 6O_2 + 6H_2O$$

This is often simplified to:

$$6CO_2 + 6H_2O \xrightarrow[\text{chlorophyll}]{\text{light}} C_6H_{12}O_6 + 6O_2$$

showing net consumption of water only.

We can show these inputs to and outputs from photosynthesis in another way (see figure 4.7). Later in this chapter (starting on page 120), we will see that this simple image is underpinned by some more intricate processes.

**studyon**
Unit 3 | AOS 1 | Topic 6 | Concept 1
**See more** Photosynthesis

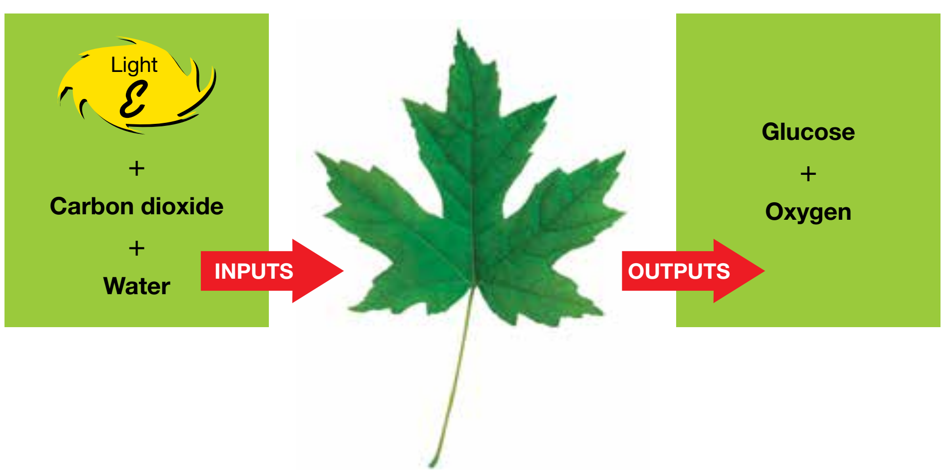



**FIGURE 4.7** Inputs and outputs of photosynthesis.

Note that the green pigment, chlorophyll, is an essential player in the process of photosynthesis, but it is not an input or an output. Chlorophyll is needed to trap the radiant energy of sunlight, bringing it within cells — the starting point for photosynthesis.

The process of converting the inorganic carbon of carbon dioxide into energy-rich organic compounds, such as sugars, is termed **carbon fixation**, and the main means by which this occurs is photosynthesis. Carbon fixation occurs on a massive scale, with an estimated 260 billion tons of carbon dioxide being fixed (or converted to organic compounds) each year. Almost half of this annual global fixation of carbon occurs in marine environments through the photosynthetic activities of phytoplankton. The remainder occurs in terrestrial environments, both natural environments — such as the tropical rainforests of northern Queensland and the hummock grasslands of Australia's sandy deserts — and farmed environments — such as the wheat belt of Victoria (see figure 4.8) and the apple orchards of Tasmania.

**FIGURE 4.8** A wheat crop ready for harvest. This crop is a physical expression of carbon fixation achieved by photosynthesis. What were the raw materials needed to produce this crop?

**ODD FACT**

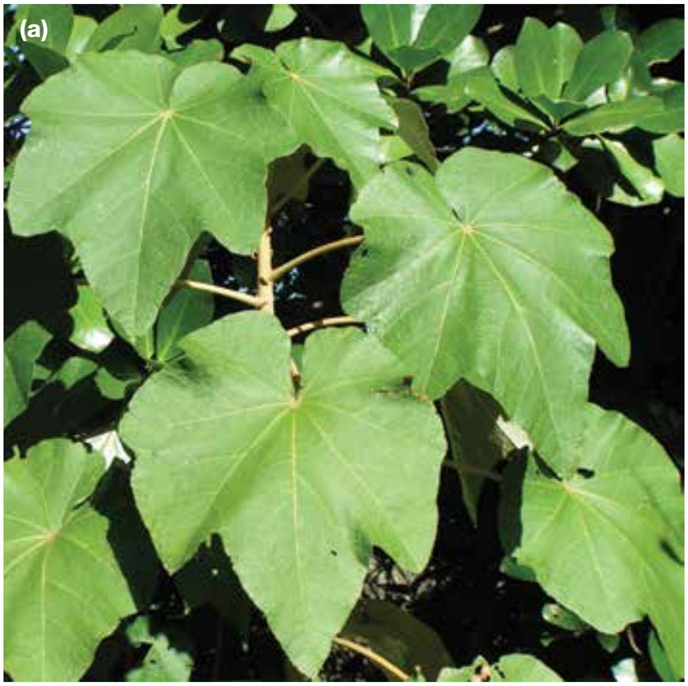
The term 'photosynthesis' was first used at a scientific meeting in 1893 by two American botanists, Charles Barnes (1858–1910) and Conway MacMillan (1867–1929).

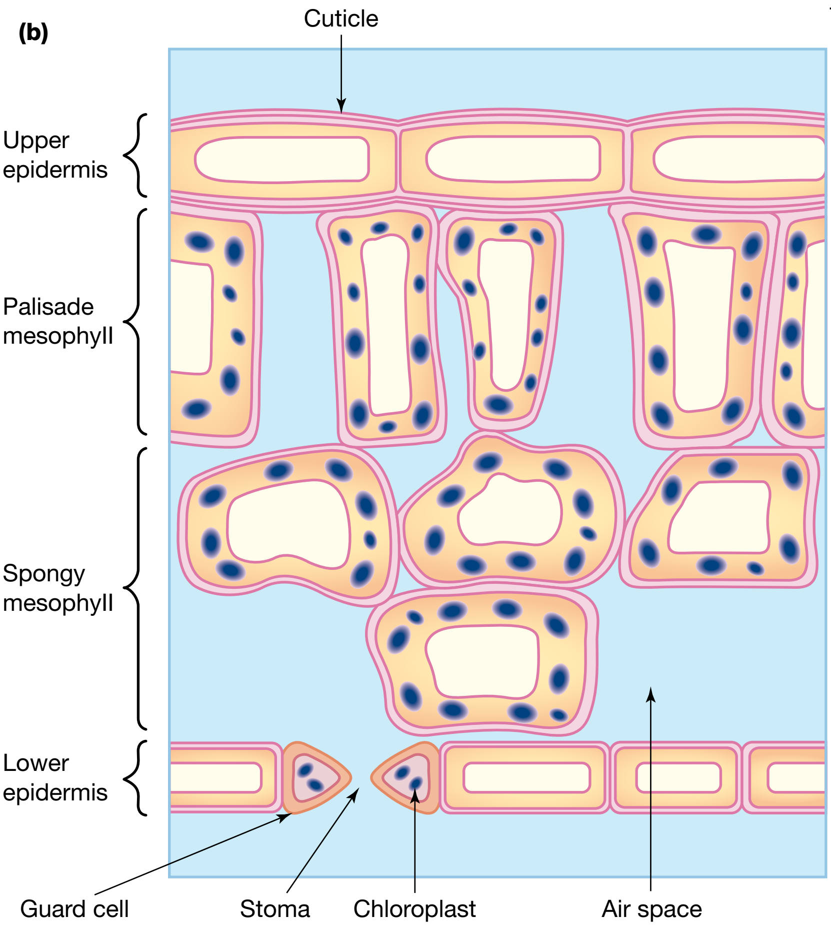
### Structural features to facilitate photosynthesis

In terrestrial flowering plants, only some cells are able to carry out photosynthesis and these are principally located in green leaves.

The shape and structure of leaves equip them to carry out photosynthesis. Check out figure 4.9 and note most of the following features:

- The flat shape of leaves provides a large surface area that can be exposed to sunlight.
- The presence of many stomata (pores) on one or both leaf surfaces provides access into the leaf for carbon dioxide.
- The presence of internal air spaces in the leaves enables the ready diffusion of carbon dioxide into the leaf tissue.
- The network of xylem vessels in the vascular tissue transports water to the photosynthetic cells.
- Each photosynthetic cell possesses many chloroplasts that contain the sunlight-trapping pigment chlorophyll.



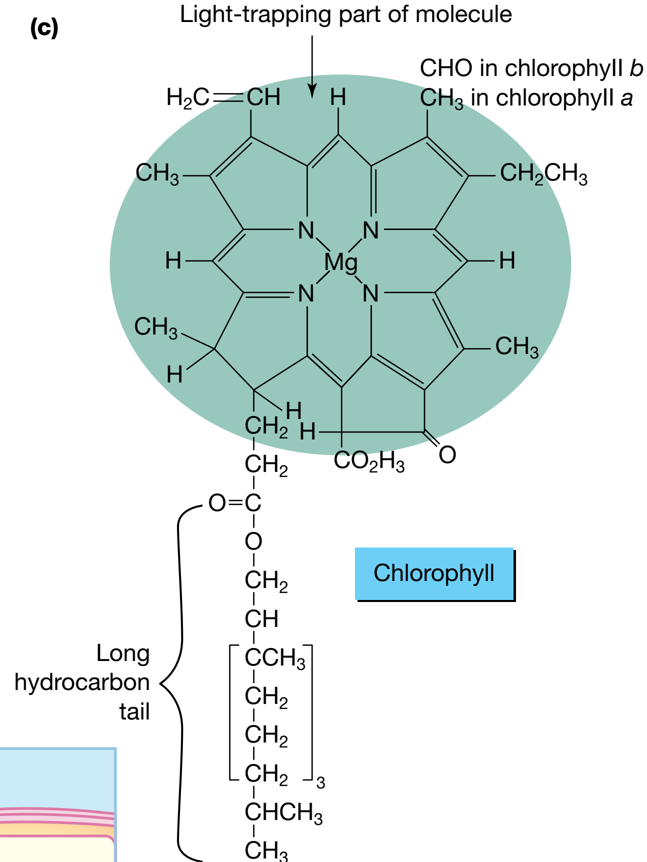




**FIGURE 4.9** **(a)** Leaves are thin, flat efficient collectors of solar energy. Note the arrangement that minimises shading of one leaf by another. (Image courtesy of Judith Kinnear.) **(b)** Drawing of cells from a transverse section through a leaf showing some of the features that equip leaves to carry out photosynthesis. (Nuclei are not shown.) Note the loose packing of the cells that enables the ready diffusion of carbon dioxide into a leaf after these molecules enter the leaf via stomata (singular: stoma). The supply of another requirement for photosynthesis is not shown in this diagram. What is this requirement and how is it supplied? **(c)** Structural formula of the light-trapping pigment chlorophyll. The hydrophilic 'head' of the molecule is where the light energy is trapped. Note the presence of the magnesium ion ($Mg^{2+}$) here. The long hydrocarbon tail is hydrophobic and enables these molecules to become embedded in the thylakoid membrane of the chloroplasts.

You will know from your previous studies that photosynthetic organisms have a range of pigments, including chlorophylls, that trap the sunlight necessary for photosynthesis and you may recall that photosynthesis is most efficient in light of red and blue wavelengths. (How was this information first obtained? See the upcoming box.)

## WHICH WAVELENGTHS OF LIGHT ARE MOST USEFUL?

The most useful wavelengths of light for photosynthesis were first identified in an experiment carried out by a German botanist, T.W. Engelmann (1843–1909), in 1881. A thin strand of the green alga *Spirogyra* sp. floating in water was exposed to visible light covering a range of wavelengths. Engelmann used a prism to produce these different wavelengths.

Large numbers of a kind of bacteria that gather in areas with a high concentration of oxygen were also added to the water. (These bacteria are termed aerophilic.) After some time, the bacteria were observed to have accumulated around the sections of the alga that were illuminated by the red and violet light (see figure 4.10).

When the alga traps sunlight and converts it to chemical energy, it produces organic matter inside its cells. As well, **it produces oxygen as a waste product, and this is released into the water.** The location of the bacteria shows where this oxygen is being released. The experimental result supports the conclusion that the biologically useful wavelengths for this alga are those of red and violet light.

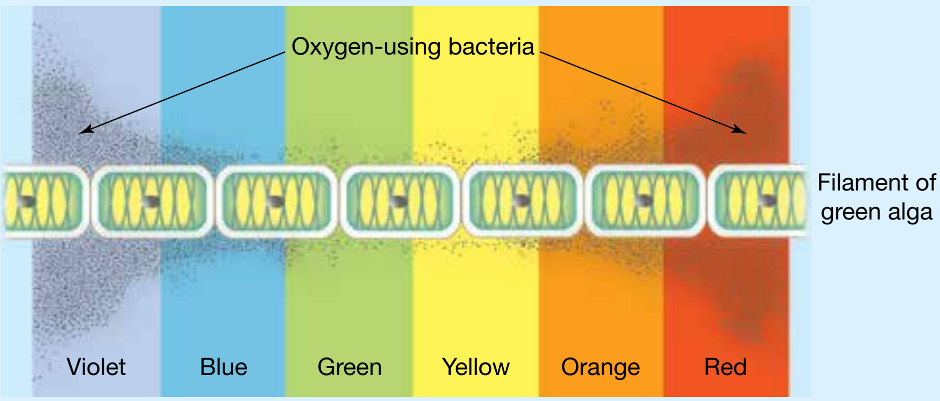



**FIGURE 4.10** Engelmann's experiment. Only certain wavelengths of visible light are trapped by producer organisms. What are the useable wavelengths?

This experiment identified the wavelengths of visible light that were trapped by the photosynthetic green alga. Today, we can directly identify the wavelengths of visible light that are trapped, not by an organism, but by individual pigments. Figure 4.11 shows the wavelengths that are trapped, that is, absorbed, by various individual pigments. The pattern for each pigment is called its absorption.

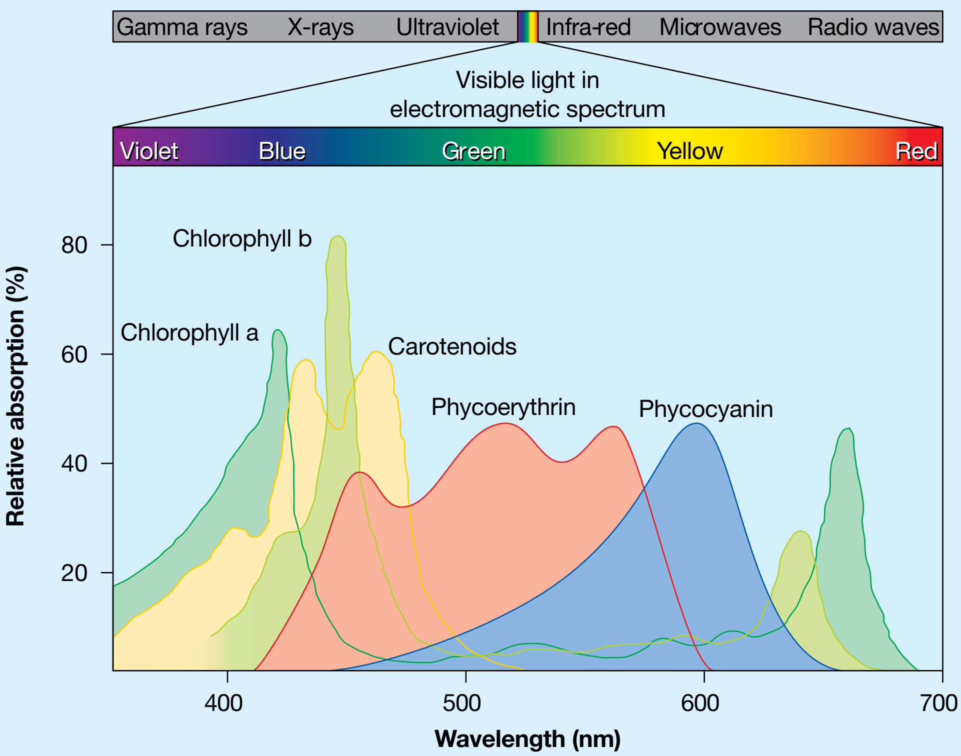



**FIGURE 4.11** Absorption of light of various wavelengths for different plant pigments. Note that both chlorophyll a and chlorophyll b have two significant regions of absorption of sunlight energy. Which pigment has its maximum absorption in the yellow region of the spectrum?

## Chloroplasts: site of photosynthesis

The chlorophyll that is needed to trap the radiant energy of sunlight is *not* present in solution in the cytosol of photosynthetic cells of plants and algae. Instead, chlorophyll is present embedded in the internal membranes of a special cell organelle called the **chloroplast**. Chloroplasts are where all the photosynthetic action takes place.

Figure 4.12a shows a photomicrograph of cells from a leaf. Note the many green chloroplasts in these cells. One estimate is that the number of chloroplasts per cell ranges from about 40 to about 200. Why might the number per cell vary?

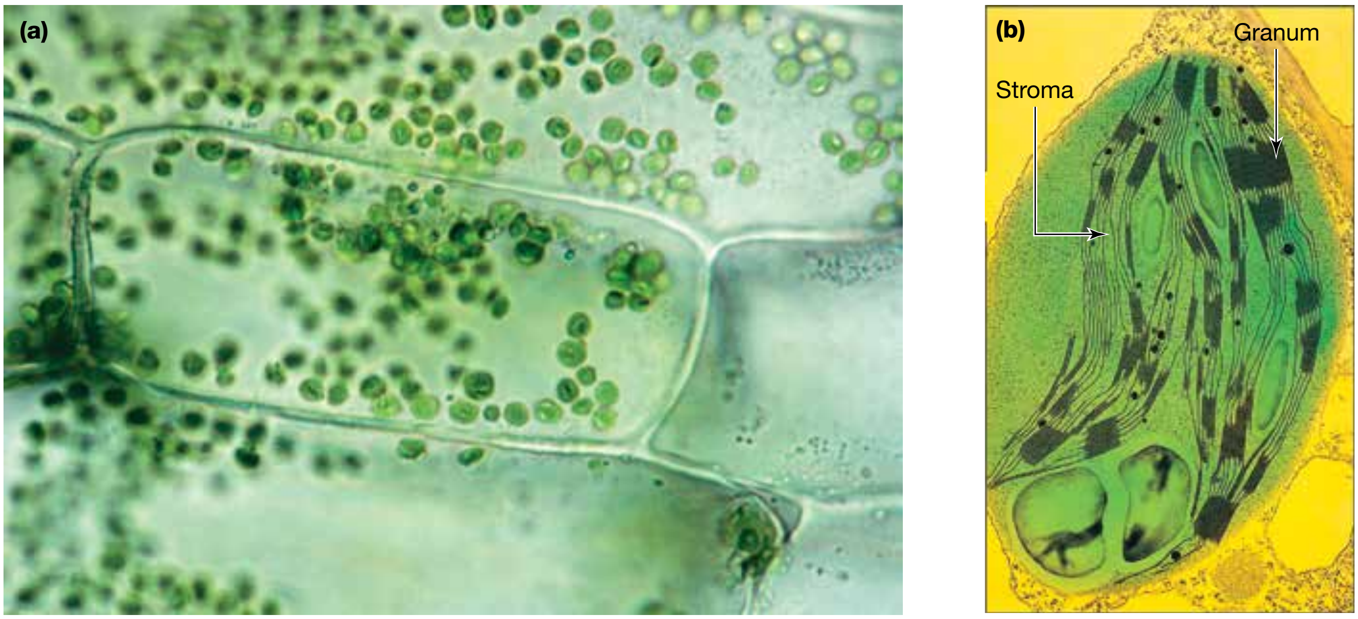



**FIGURE 4.12** **(a)** Photomicrograph of cells of leaf tissue showing the green chloroplasts present in each photosynthetic cell. These cells form the spongy and the palisade mesophyll tissues of a leaf. **(b)** Transmission electron micrograph of a chloroplast showing its internal thylakoid membranes, which in some areas are aggregated into grana.

**ODD FACT**

Some photosynthetic algae, such as the unicellular green alga *Chlamydomonas reinhartii*, have a single large chloroplast only.

Figure 4.13 shows a diagrammatic representation of the internal structure of a typical chloroplast. Note the following:

- Each chloroplast is enclosed in an envelope that is formed by a double membrane.
- Within this envelope are many membranous fluid-filled flattened discs known as **thylakoids**.
- Thylakoids are aggregated into stacks (a bit like stacks of pancakes) known as **grana** (singular: **granum**).
- The light-trapping pigment, chlorophyll, is embedded in the thylakoid membranes.
- Some of the enzymes involved in the first stage of photosynthesis are located in the thylakoid membranes.
- The **stroma** is the fluid inside the chloroplast that bathes the grana.
- The stroma contains some of the enzymes that are involved in the second stage of photosynthesis.
- The stroma contains ribosomes and DNA.

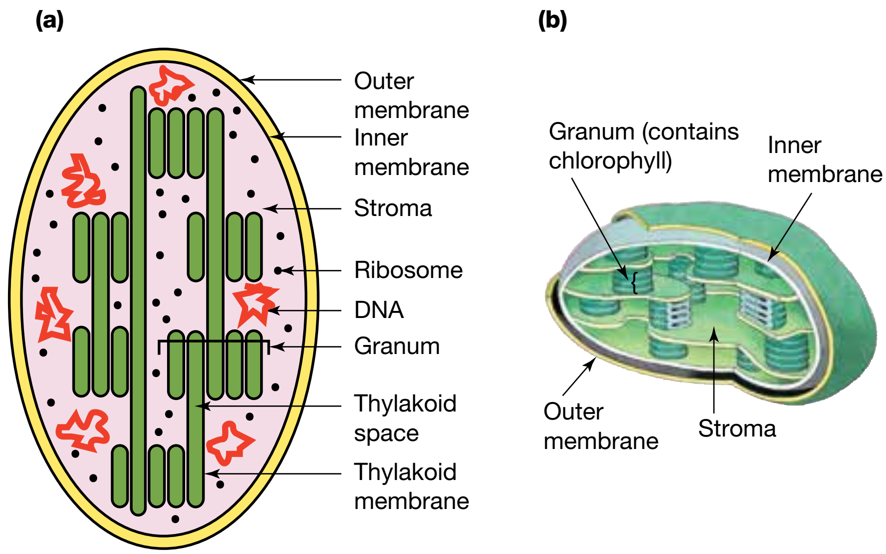



**FIGURE 4.13** **(a)** Stylised diagram of a single chloroplast showing its main features. Why are the thylakoids coloured green? Note that DNA and ribosomes are present in the stroma of the chloroplast. **(b)** Diagram showing a 3D representation of a chloroplast. Note the double membrane that forms the outer envelope of the chloroplast.

**studyon**

**Unit 3** | **AOS 1** | **Topic 6** | **Concept 2**

**Chloroplast contributions**
Summary screen and practice questions

The site of photosynthesis in plants and algae is the chloroplast, the cell organelle found in the cytosol of the cell. However, photoautotrophic bacteria, such as cyanobacteria, do not have such a cell organelle. Instead, internal membranes are the site of photosynthesis in these bacteria.

## Origin of chloroplasts

The chloroplasts of plant and algal cells are believed to have once been free-living photosynthetic microbial cells. It is postulated that, eons ago, one of these photosynthetic microbes was engulfed by another microbial cell (see figure 4.14).

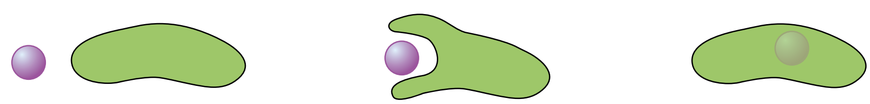


**FIGURE 4.14** A simplified diagram showing the engulfing of one microbial cell by another, where digestion does not occur. This process is termed endosymbiosis. The small cell engulfed could have been a photosynthetic microbe.

This mutual housing arrangement was of benefit to both cells — the host cell gained a source of carbohydrate (sugar) production, and the engulfed cell gained a protective environment and an additional source of energy when sunlight was not available. This special kind of symbiotic (*sym* = together; *bios* = life) relationship where one organism takes up permanent residence inside another organism is called **endosymbiosis**. This endosymbiotic event produced a cell with improved survival features and this led to the emergence of a new lineage of photosynthetic eukaryotes that is represented today by plants and algae (see figure 4.15). So, the chloroplasts that give today's plants and algae their photosynthetic ability are an ancient inheritance from light-trapping microbes of eons past. An earlier endosymbiotic event also occurred, namely the engulfing of an oxygen-using microbe by an early eukaryotic cell. In this case, the tenant cells gave rise to the mitochondria of the host cell.

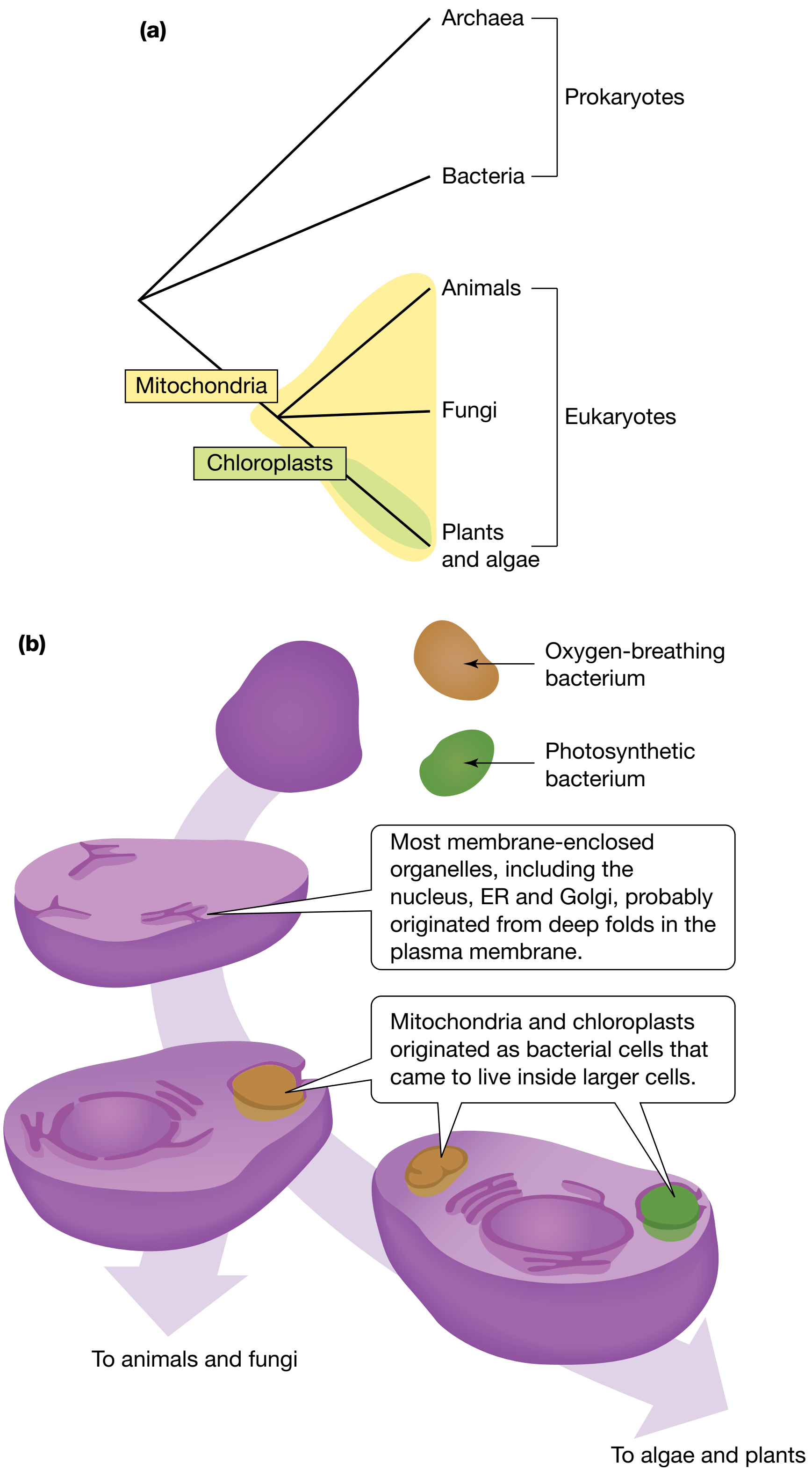



**FIGURE 4.15** **(a)** Because all eukaryotic cells have mitochondria, it is reasonable to postulate that the endosymbiosis that produced mitochondria was the first. Sometime later, a second endosymbiosis produced the eukaryotic chloroplast, which led to the emergence of photosynthetic algae and plants. **(b)** The evolution of eukaryotic cells involved infoldings of the plasma membrane of the host cell and two endosymbiotic events. Which was the first endosymbiotic event?

The free-living microbes today that show the most similarities with the chloroplasts of plants and algae are photoautotrophic cyanobacteria, such as the freshwater cyanobacterium *Nostoc punctiforme* (see figure 4.16).

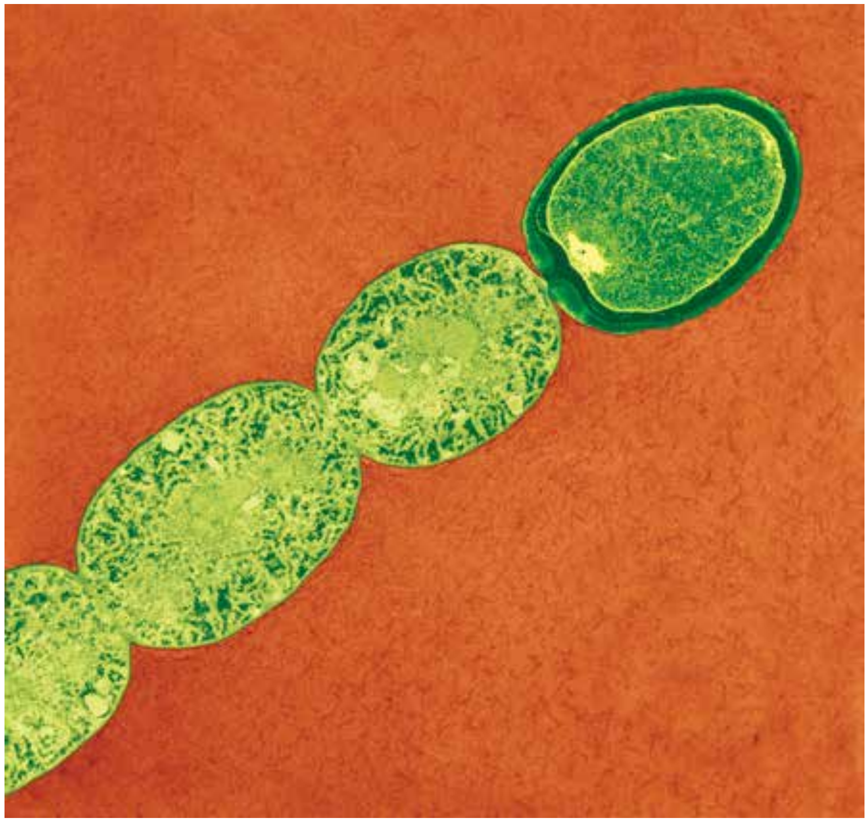


**FIGURE 4.16** Transmission electron microscope image of *Nostoc punctiforme*, a freshwater photosynthetic cyanobacterium. Chloroplasts of eukaryotic cells have the most similarities with these bacteria. The cells shown here, from the Amazon floodplain in Brazil, are magnified 26 000 times.

**ODD FACT**

Like plants and algae, cyanobacteria have chlorophyll a as their major light-trapping pigment, while other photosynthetic bacteria have a different pigment, bacteriochlorophyll, for the trapping of light energy.

It is now generally accepted that the chloroplasts of eukaryotic plant and algal cells are descendants of a cyanobacterium that took up residence within a eukaryotic cell eons ago. A key question remains: *What evidence, if any, exists to support this view?*

## Evidence for endosymbiotic origin of chloroplasts

The following observations provide evidence in support of the endosymbiotic origin of chloroplasts:

- Chloroplasts are enclosed within a double-membrane envelope; this is similar to the outer and inner membranes present in the cell wall of some bacteria (such as the Gram-negative cyanobacteria).
- They possess DNA that consists of circular double-stranded DNA and, as seen in bacterial cells, this DNA is not enclosed within a nuclear membrane.
- They divide, not by mitosis as the host cell does, but by a process of binary fission, such as occurs in bacteria.
- They have their own ribosomes that are like those found in bacteria. Chloroplast ribosomes are smaller (70S) than those present in the cytosol of their host cells (80S).
- Because they have their own DNA and ribosomes, chloroplasts can produce their own RNA and proteins.
- Their sizes fall within the size range of bacterial cells.

Gram-negative and Gram-positive bacteria are discussed in Ch 6, page 238.

- Chloroplast ribosomes are sensitive to the same antibiotics that target bacterial ribosomes. (Ribosomes in the cytosol of eukaryotic cells are not affected by these antibiotics.)

During the billion plus years since the first photosynthetic bacterium was taken into a host cell, the engulfed cell has evolved into a chloroplast and is now a cell organelle. The chloroplast has transferred most of its genes to the nucleus of the host cell and retained a smaller number that function in the light-dependent reactions of photosynthesis. The extent of this transfer of genes from chloroplasts to the nuclear genome can be seen in the following comparison:

- Genomes of plant and algal chloroplasts range from about 110 to 200 kilobases (kb) and encode information for about 100 protein-coding genes and 30 to 50 RNA-coding genes.
- Genomes of modern free-living freshwater cyanobacteria, such as *Synechococcus* sp., contain about 3000 kb and encode the information for more than 3000 genes.

Endosymbiosis is not simply a hypothetical concept. Examples exist in the living world around us; for example:

- *Rhizobium* spp. are nitrogen-fixing bacteria that live as endosymbionts in the cells of nodules on the roots of legumes, such as clover (see figure 4.17a).
- The science journal *Nature* (30 July 2010) published the report of the first confirmed case of single-celled organisms living inside the cells of a vertebrate. The vertebrate is the spotted salamander, *Ambystoma maculatum.* The endosymbiont is the unicellular green alga *Oophila amblystomatis.* Look at figure 4.17b, which shows green-coloured embryos of the spotted salamander. This colour is due to the presence of green algal cells located *inside* cells all over the salamander's body.


**FIGURE 4.17** **(a)** Image showing root nodules on a leguminous plant. Nitrogen-fixing bacteria of the genus *Rhizobium* living inside nodule cells are endosymbionts. Can you suggest a possible benefit of this arrangement to the host cells and to their bacterial boarders? **(b)** Developing embryos of the spotted salamander, *Ambystoma maculatum,* showing the green colour that is due to the presence inside their cells of a unicellular alga, *Oophila amblystomatis.* The jelly material in which each embryo is encased also shows the presence of these green algae. Can you deduce the meaning of the scientific name of the unicellular alga?

## KEY IDEAS

- Photosynthesis is the process of converting the energy of sunlight to the chemical energy in sugars.
- The raw materials of photosynthesis are carbon dioxide and water, and the products are carbohydrate and oxygen.
- Carbon fixation refers to the process of converting inorganic carbon to energy-rich organic carbon of compounds, such as sugars, which is most commonly achieved by photosynthesis.
- In photosynthetic eukaryotes, such as plants and algae, photosynthesis occurs in chloroplasts.
- The main light-trapping pigments are chlorophylls, located on the granum membranes in chloroplasts.
- Eons ago, the chloroplasts of eukaryotic cells were free-living photosynthetic bacteria that became endosymbionts living in other cells.

## QUICK CHECK

4 Do all cells of a plant carry out photosynthesis? Explain.
5 Where in a cell is light energy transformed to chemical energy?
6 Identify two structural features of a leaf that equip it to carry out photosynthesis?
7 Do photoautotrophic bacteria have chloroplasts?
8 List two lines of evidence supporting the proposal that chloroplasts were once free-living bacteria.
9 Which group of modern bacteria is believed to be most closely related to the ancient free-living bacteria that became endosymbionts?

# Photosynthesis in action

The essence of photosynthesis is to trap the energy of sunlight, convert it to energy-carrier molecules (**ATP** and **NADPH**), and use this chemical energy to make organic molecules (sugars) from carbon dioxide. Sounds complicated, but the process of photosynthesis can be broken down into two stages:

1. the light-dependent stage
2. the light-independent stage.

Let's explore these stages and identify their inputs and outputs, and find out where these stages occur in eukaryotic organisms, as, for example, in a flowering plant.

## The light-dependent stage

studyon

Unit 3 | AOS 1 | Topic 6 | Concept 3

**Light-dependent stage**
Summary screen and practice questions

This stage is called light-dependent because, in the absence of light, this stage cannot proceed and photosynthesis stops completely.

***Where does this stage occur?*** The light-dependent stage occurs in the chloroplasts found in many cells of the leaves of a flowering plant. More specifically, this stage takes place on the thylakoid membranes within the chloroplasts.

***What is its overall function?*** Overall, the light-dependent stage traps light energy and converts it to chemical energy in the form of ATP and NADPH. Both the ATP and the NADPH are used in the light-independent stage where they provide the energy and the reducing power needed to build carbon dioxide molecules into sugars.

***What are the inputs and outputs of the light-dependent stage?*** Figure 4.18 shows the inputs and outputs of the light-dependent stage of photosynthesis.

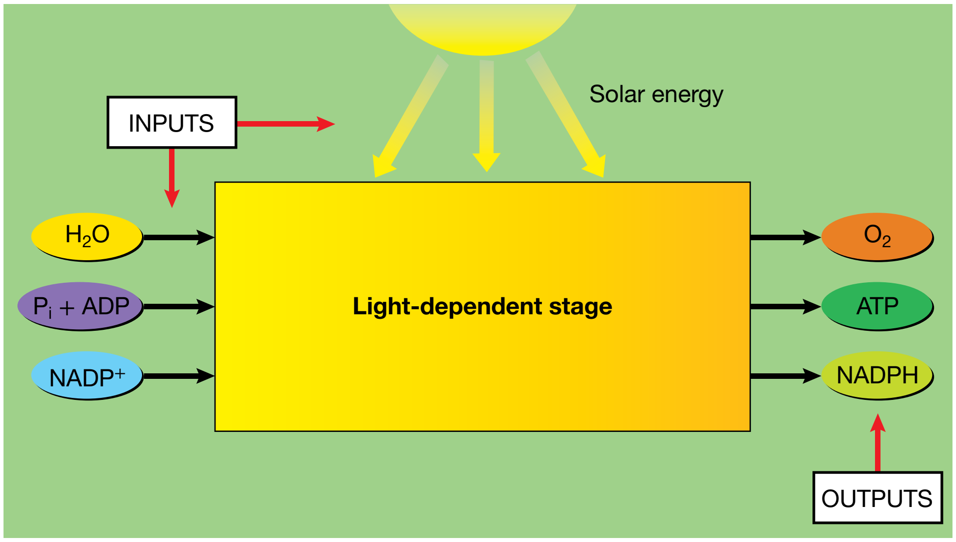



**FIGURE 4.18** Diagram showing the inputs and outputs of the light-dependent stage of photosynthesis

Note that the inputs of the light-dependent stage of photosynthesis are sunlight, water, ADP+ $P_i$, and $NADP^+$.

Let's look at the role of each of these inputs and the outputs they produce.

1. **Input: Sunlight.** Radiant energy from the sun (or from an artificial light source) can be used for photosynthesis. To be useful, however, the light energy must be trapped and converted to chemical energy. Chlorophyll molecules trap wavelengths of light in the purple-blue and the red-orange regions of the visible spectrum. Accessory pigments trap the energy of other wavelengths and transfer this energy to chlorophyll.

   The trapping of the energy of sunlight occurs in special structures, known as **photosystems**, which are embedded in the thylakoid membranes of chloroplasts. Each photosystem is composed of:
   (i) a light-harvesting system that consists mainly of chlorophyll molecules plus some accessory pigments
   (ii) a reaction centre where the light energy is converted to chemical energy, initially in the form of energised (excited) chlorophyll molecules and electrons.

Photosystems: refers to one of two kinds of structure (termed PS-I and PS-II) embedded in the thylakoid membranes of chloroplasts; each is composed of a light-harvesting complex of pigments (mainly chlorophylls) and various proteins, and a reaction centre that converts light energy to the chemical energy of energised electrons.

Figure 4.19 shows an image of the inner surface of a thylakoid membrane. Note the many rounded particles. Each particle is thought to represent one photosystem.

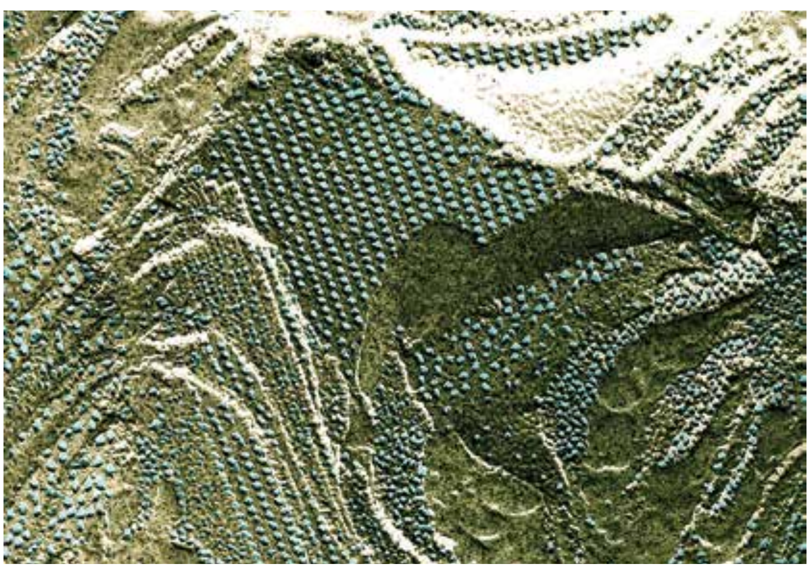


**FIGURE 4.19** Freeze-fractured transmission electron microscope (TEM) image of the inner surface of a thylakoid membrane. Each particle probably represents one photosystem (either PS-I or PS-II).

A molecule or an electron is said to be 'excited' or 'energised' when it absorbs energy and moves from a lower to a higher energy state.

When light energy is trapped by chlorophyll molecules embedded in the thylakoid membranes, the chlorophyll molecules become 'excited' (see marginal note) and they shed energised electrons. At this stage, the radiant energy of sunlight has been converted to chemical energy that appears in energised (excited) chlorophyll molecules and electrons.

2. **Input: Water.** Water molecules are split in a reaction that occurs in one of the photosystems (PS-II). The splitting of water molecules is catalysed by an enzyme at the reaction centre of this photosystem. The outputs of the splitting of water molecules are hydrogen ions, electrons and oxygen:

$$2H_2O \rightarrow O_2 + 4H^+ + 4e^-$$

What is the fate of the outputs of the splitting of water?

- The electrons ($e^-$) ultimately pass via a series of acceptors in the electron transport chain and reach the terminal electron acceptor, $NADP^+$(see below).
- The hydrogen ions ($H^+$) accumulate in the inner thylakoid space, creating a concentration gradient relative to their concentration in the stroma. This $H^+$ gradient is involved in the production of ATP (see below).
- The oxygen molecules ($O_2$) have no further use in photosynthesis and are a by-product that can diffuse out of the chloroplasts, out of the cells, and into the atmosphere. By what means do oxygen molecules exit from a plant?

(However, plant cells can use this oxygen as an input to cellular respiration (see page 136).)

Although molecular oxygen is an incidental by-product of photosynthesis, it sets the scene for the emergence of eukaryotic cells, and, later, the emergence of multicellular organisms. The atmosphere of the early Earth contained no oxygen — it was anoxic — and the prokaryotic life forms relied on anaerobic metabolic processes. However, by about 2400 million years ago, Earth's atmosphere was oxygenated. Where did this oxygen come from? Over eons, oxygen was released by photosynthetic prokaryotes that had developed the ability to exploit sunlight as an energy source. Eventually, over eons, the release of oxygen resulted in an oxygenated atmosphere. This enabled the emergence of multicellular organisms that could use aerobic cellular respiration. The oxygenation of Earth's atmosphere also allowed the formation of the ozone layer, which protects life forms from the damaging effects of solar UV radiation.

3. **Input: ADP & $P_i$.** ADP (adenosine diphosphate) combines with a molecule of inorganic phosphate ($P_i$) to produce ATP (adenosine triphosphate):

$$ADP + P_i \rightarrow ATP$$

This is an endergonic reaction, requiring an input of energy. As mentioned above, the energy to drive this reaction comes from the movement of hydrogen ions ***(protons)*** down a concentration gradient into the stroma. This movement occurs through an ion channel in a protein complex (ATP synthase) that is embedded in the thylakoid membrane (see figure 4.20). The ATP synthase protein operates both as an ion channel and as an enzyme. (A similar process occurs in mitochondria.) ATP is a high-energy molecule that acts as a coenzyme in metabolic reactions that are endergonic or energy-requiring. The light-independent stage of photosynthesis involves energy-requiring reactions that build carbon dioxide into glucose molecules, and this energy is provided by ATP. In a chloroplast, for example, one ATP synthase protein complex can produce more than one hundred ATP molecules each second.

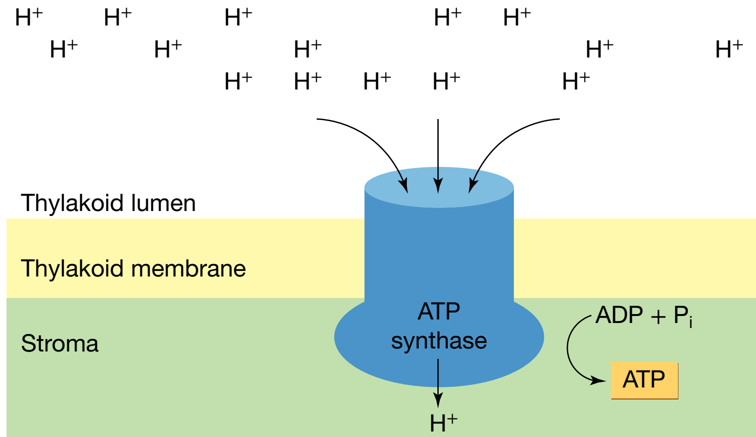



**FIGURE 4.20** The ATP synthase protein complex operates as both an ion channel and as an enzyme. It uses the energy from the movement of ions through its channel to drive the endergonic reaction of building ATP from ADP and $P_i$. This movement of hydrogen ions is from a region of high concentration (inner thylakoid space) to low concentration (stroma).

The ATP produced by this reaction is released into the stroma of the chloroplast where it is used in the light-independent stage of photosynthesis as an energy source to power the building of carbon dioxide molecules into sugars.

**4. Input: NADP⁺.** The terminal electron acceptor of the light-dependent stage of photosynthesis is **NADP⁺**. This molecule is used to accept electrons and hydrogen ions that come from the splitting of water:

$$NADP^+ + 2e^- + 2H^+ \rightarrow NADPH + H^+$$

The output of this reaction is NADPH, which is released into the stroma. This output is used in the light-independent stage of photosynthesis to reduce carbon dioxide, by supplying energised electrons and hydrogen ions.

In summary, in the light-dependent stage of photosynthesis, light energy is converted to chemical energy in the form of two compounds, namely ATP, the versatile source of energy, and NADPH, a source of energised electrons with reducing power. Note that the light-dependent stage of photosynthesis does not produce any sugars. This happens in the light-independent stage (discussed in the following section).

## KEY IDEAS

- Photosynthesis can be considered as a two-stage process — a light-dependent stage followed by a light-independent stage.
- The light-dependent stage occurs on the thylakoid membranes of chloroplasts.
- The trapping of light energy involves complexes called photocentres.
- Inputs to the light-dependent stage are sunlight energy, water, ADP and $P_i$, and $NADP^+$.
- Outputs from the light-dependent stage are ATP, NADPH and molecular oxygen.

**QUICK CHECK**

10 Identify the location within a chloroplast where light-trapping activity occurs.
11 Briefly explain the difference between ATP and ATP synthase.
12 Identify the following statements as true or false:
   a The energy of sunlight is used directly to make sugars in photosynthesis.
   b The terminal electron acceptor in the light-dependent stage of photosynthesis is $NADP^+$.
   c ATP formed in the light-dependent stage is a by-product of the splitting of water.
   d The outputs of the light-dependent stage of photosynthesis are the inputs to the light-independent stage.
   e Chlorophyll molecules absorb radiant light and transfer the energy to excited electrons.
13 In what form(s) does light energy appear at the end of the light-dependent stage of photosynthesis?

**studyon**

Unit 3 | AOS 1 | Topic 6 | Concept 4

**Light-independent stage**
Summary screen and practice questions

## The light-independent stage

The light-independent stage of photosynthesis does not depend directly on light being available. This stage of photosynthesis can occur in the absence of light. It is often termed 'the dark reaction,' but one issue with this name is it might imply that it can only occur in the dark. Not so. The light-independent stage of photosynthesis can occur in the light or the dark, provided it has a supply of its essential inputs.

***Where does this stage occur?*** The light-independent stage occurs in the chloroplasts found in many cells of the leaves of a flowering plant. More specifically, the enzymes and other factors required are present in the stroma of the chloroplasts.

***What is its overall function?*** The major purpose of the light-independent stage is to build inorganic carbon dioxide into energy-rich reduced organic molecules, such as the sugar glucose ($C_6 H_{12} O_6$). The process by which this occurs in many plants is the Calvin cycle (see below). If the glucose is not required for immediate use by a plant, it is converted to starch, a plant storage polysaccharide.

***What are the inputs and outputs of the light-independent stage?*** Figure 4.21 shows the inputs and outputs of the light-independent stage of photosynthesis.

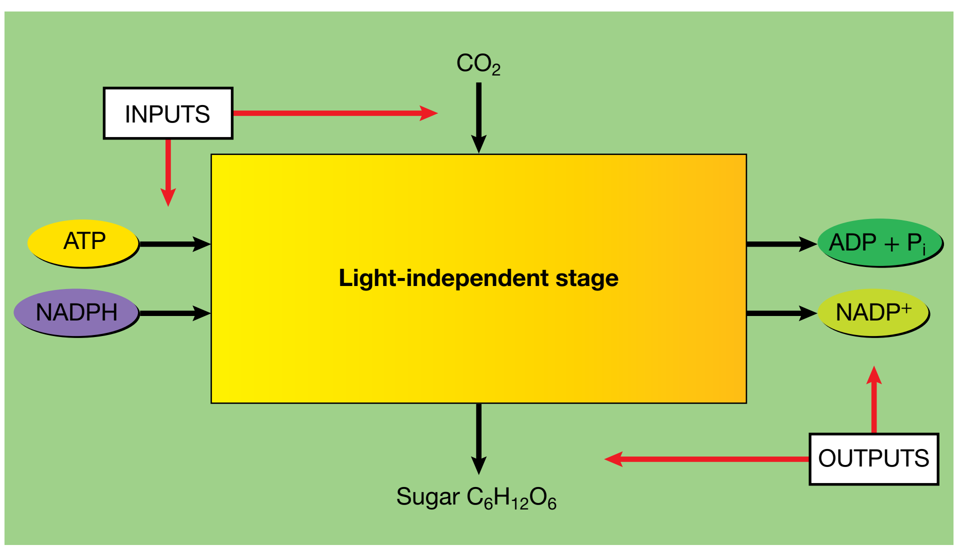



**FIGURE 4.21** Diagram showing the inputs and outputs of the light-independent stage of photosynthesis. For many plants, this stage involves the processes of the Calvin cycle.

The inputs to the light-independent stage of photosynthesis are carbon dioxide ($CO_2$), ATP and NADPH.

The inputs ATP and NADPH provide the energy and the reducing power for this stage — they are both essential inputs for the fixation of carbon. The neat arrangement is that as these energy providers are used and converted to ADP and $NADP^+$, they can be sent back to the light-dependent stage for recharging. The light-independent stage can continue as long as a supply of ATP and NADPH is available.

The role of each of these inputs and the outputs that they produce follows.

1. **Input: Carbon dioxide.** Gaseous carbon dioxide is taken up from the atmosphere by terrestrial plants, or, for aquatic plants and algae, carbon dioxide is taken up in solution in water. For flowering plants, the point of entry of carbon dioxide to the leaves is through stomata (refer back to figure 4.9b). Carbon dioxide molecules provide the carbon atoms that will be reduced and built into energy-rich glucose molecules. For many plants, this occurs through a series of reactions of the **Calvin cycle**. (See the next section for a discussion of the Calvin cycle.)
2. **Input: ATP.** ATP is one major source of the energy required to drive the endergonic reactions of the Calvin cycle. When its energy is expended, ATP is converted to ADP and $P_i$ (inorganic phosphate), which are outputs of this stage.
3. **Input: NADPH.** NADPH is the source of the energised electrons and reducing power that reduce carbon dioxide and build it into glucose. Carbon dioxide ($CO_2$) is a fully oxidised molecule. In comparison, the sugar glucose ($C_6H_{12}O_6$) is relatively reduced (see marginal note). After its reducing power is used, NADPH is converted to $NADP^+$, one of the outputs of the light-independent stage of photosynthesis.

In common usage, a molecule is said to be reduced if, in a reaction, its content of H atoms increased, and a molecule is said to be oxidised if its H atom content decreases. More strictly, in chemical terms, reduction is a net gain in electrons and oxidation is a net loss of electrons.

## The Calvin cycle: Carbon dioxide to glucose

The light-independent stage of photosynthesis in many plants involves the series of reactions known as the **Calvin cycle. The Calvin cycle was named after an American biochemist, Melvin Calvin** (1911–1997), who discovered the details of the process by which an input of carbon dioxide ($CO_2$) results in an output of $C_6H_{12}O_6$, glucose.

In simple terms, the Calvin cycle can be shown in terms of inputs and outputs only (see figure 4.22a).

Check out figure 4.22b and see if you can identify the key features of the Calvin cycle as summarised below:

***Stage A: Carbon fixation***

- The input to the cycle is carbon dioxide ($CO_2$).
- The acceptor molecule for carbon dioxide is ribulose biphosphate (RuBP), a 5-C molecule.
- The carbon dioxide that enters the Calvin cycle is fixed, forming 3-phosphoglyceric acid (3PGA), a 3-C molecule.

***Stage B: Reduction***

- 3PGA is converted to glyceraldehyde-3-phosphate (G3P), another 3-C molecule, using the energy from ATP and the reducing power of NADPH.
- The output from the cycle is G3P, which can be combined to form glucose.

***Stage C: Regeneration of RuBP***

- Using energy from ATP, five molecules of G3P are changed to form three molecules of RuBP and, in doing so, regenerate the carbon acceptor molecule needed for the cycle to continue.

As long as ATP and NADPH continue to be supplied by the light-dependent stage of photosynthesis, the Calvin cycle can continue. ATP and NADPH are both energy-rich molecules, but they do not live long and they cannot be used as energy stores by plants. The neat trick of plants is that they use the energy contained in these molecules to build long-lived and stable carbohydrate molecules, such as glucose and starch.

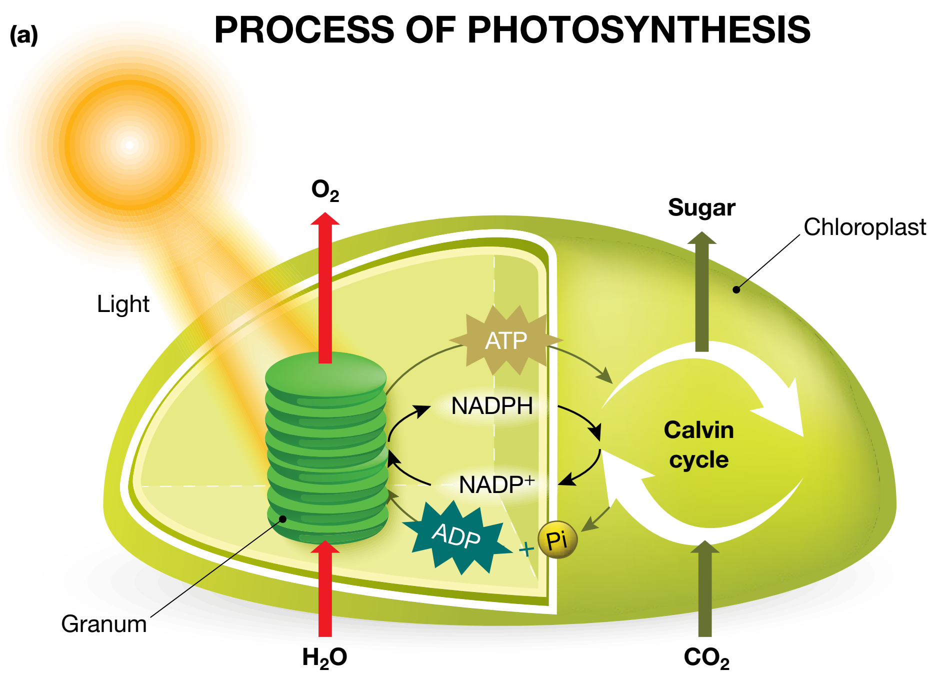
**ODD FACT**

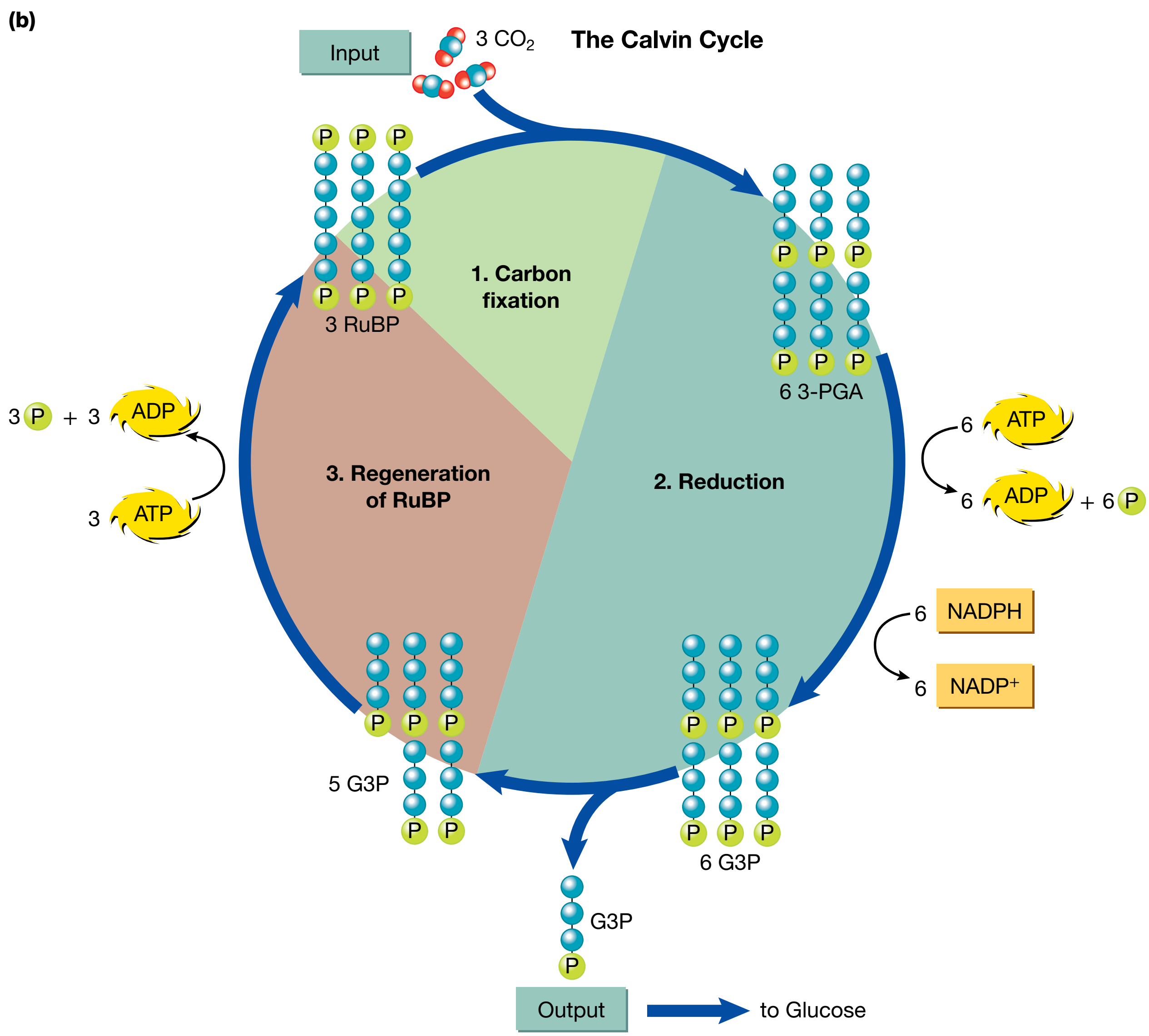
Melvin Calvin was awarded the Nobel Prize in 1961 'for his research on the carbon dioxide assimilation in plants'. Note that 'assimilation' = 'fixation'.

**ODD FACT**

The enzyme that catalyses the fixation of $CO_2$ is called **RuBisCo**, and this enzyme makes up about half the proteins found in chloroplasts. Its formal name is ribulose-1,5-bisphosphate carboxylase/oxygenase. Can you see how its nickname has been derived?

**FIGURE 4.22** **(a)** A simple representation of the Calvin cycle showing inputs and outputs only. **(b)** A more detailed diagrammatic representation of the Calvin cycle showing the input of carbon dioxide (a 1-C molecule) that is fixed as a 3-C molecule, is reduced, and is released from the Calvin cycle as 3-C sugar. Two of these 3-C molecules can then combine to form one molecule of 6-C glucose.



## Stages of photosynthesis are mutually dependent

Now, let's put the two stages of photosynthesis together and see how they interact and depend on each other. Look at figure 4.23, which shows the two stages of photosynthesis. Note that the ATP and NADPH outputs from the light-dependent stage are essential inputs to the light-independent stage. Note also that the ADP and $NADP^+$ outputs from the light-independent stage are restored to their higher energy states by being fed back into the light-dependent reactions. If the supply of either ATP or NADPH is interrupted, photosynthesis stops.

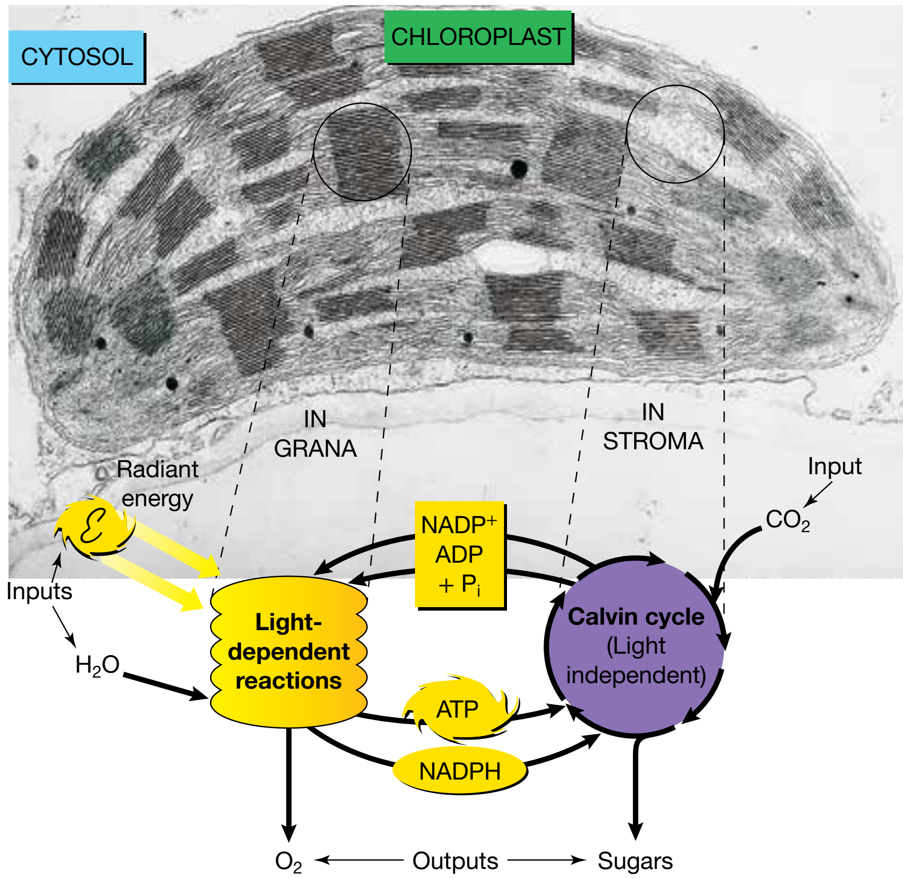



**FIGURE 4.23** Simplified diagram showing how the outputs from the light-dependent reactions in photosynthesis (ATP and the 'loaded' acceptor molecule, NADPH) are used to drive the light-independent fixation of carbon. Because water 'waste', $NADP^+$, ADP and $P_i$ are recycled, they are not considered outputs of the *total* process.

## KEY IDEAS

- The light-dependent stage of photosynthesis occurs on the thylakoid membranes of chloroplasts.
- Inputs to the light-dependent stage are sunlight, ADP & $P_i$, and $NADP^+$.
- Outputs from the light-dependent stage are ATP and NADPH.
- The light-independent stage occurs in the stroma of chloroplasts.
- Inputs to the light-independent stage are ATP, NADPH and carbon dioxide.
- Outputs from the light-independent stage are ADP & $P_i$, $NADP^+$ and glucose.
- Overall then, combining both stages, the outputs of photosynthesis are oxygen and sugars.

## QUICK CHECK

**14** Where does the trapping of light energy occur in the thylakoid membrane?
**15** Briefly explain the difference between the light-dependent and light-independent stages of photosynthesis.
**16** Identify the following statements as true or false and briefly explain why you have identified a statement as false:
- **a** The energy source for the light-dependent stage of photosynthesis is direct light energy.
- **b** The energy source for the light-independent stage of photosynthesis is the ATP and NADPH output of the light-dependent stage.
- **c** The conversion of carbon dioxide to glucose is a one-step process.
- **d** Carbon dioxide is directly acted on by sunlight energy to produce glucose.
- **e** The energy level of ADP is less than that of ATP.
- **f** NADPH provides the reducing power involved in building molecules of $CO_2$ into $(CH_2O)_6$.

# Rate of photosynthesis

It's a warm sunny day and a canola crop that is growing in a field bathed in sunlight can carry out photosynthesis at a faster rate than on the previous warm, but overcast, day. Light intensity is the factor affecting the rate of photosynthesis by the crop plants on the two days. On a bright sunny day, the illuminance or incident light on a surface is about 110 000 lux (lumens per square metre). On an overcast day, that value falls markedly to about 1000 to 2000 lux.

**ODD FACT**

The illuminance on a surface from a full moon on a clear night is about 0.25 lux.

The rate of photosynthesis in plants and algae is affected by several environmental factors including:
(i) intensity of radiant light
(ii) ambient temperature
(iii) concentration of carbon dioxide.

Can you think of other factors that could affect the rate of photosynthesis in plants and algae?

Under certain circumstances, any one of these factors may become a **limiting factor**. This occurs when the supply of the factor is in short supply such that it restricts the rate of photosynthesis. Light intensity, temperature and the availability of carbon dioxide interact, and any of them could be the limiting factor in photosynthesis.

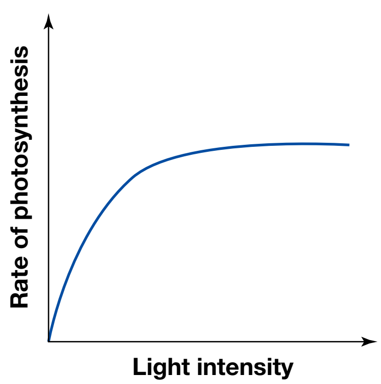



**FIGURE 4.24 (a)** Simple graph showing the change in rate of photosynthesis with increasing light intensity. (The ambient carbon dioxide concentration in which this experiment was carried out was 0.04%.) **(b)** Artificial lighting is needed at sports stadiums to illuminate the surface at night during the winter in order to increase the rate of photosynthesis of the grass to levels to support its healthy growth. This particular image shows banks of lights at Reebok Stadium, home of the Bolton Wanderers Football Club, in the UK (see over page).

## Intensity of light and rate of photosynthesis

Light is essential for photosynthesis. The rate at which photosynthesis occurs depends on the intensity of the incident light. Figure 4.24a shows a graphical representation of the change in the rate of photosynthesis with changes in the intensity of incident light. At low light intensities, the photosynthesis rate is slow because the light-dependent stage can produce only small amounts of the ATP and NADPH needed for carbon fixation (see the example of the grass at Reebok Stadium, UK, in figure 4.24b on the next page).

As light intensity increases, the rate of photosynthesis also increases, as shown by the steep sloping straight line at the left of the graph. Under these circumstances, a linear relationship exists between light intensity and photosynthetic rate.

However, a point is eventually reached when further increases in light intensity do *not* cause any increase in the rate of photosynthesis. This is shown by a flattening or plateauing of the graph. At this stage, some other factor is limiting photosynthesis — it may be the temperature or it may be the availability of carbon dioxide or it may be a combination of both factors.



| | |
|---|---|
| Unit 3 | **Factors affecting the rate of photosynthesis** |
| AOS 1 | Summary screen and practice questions |
| Topic 6 | |
| Concept 5 | |

Examine figure 4.25, which shows that carbon dioxide availability is a limiting factor. Compare the red line (carbon dioxide concentration of 0.04%) with the blue line (carbon dioxide concentration of 1%). Note that, at the one per cent concentration of carbon dioxide, the initial rate of photosynthesis is greater and the plateauing occurs later as compared to the 0.04 per cent concentration of carbon dioxide.

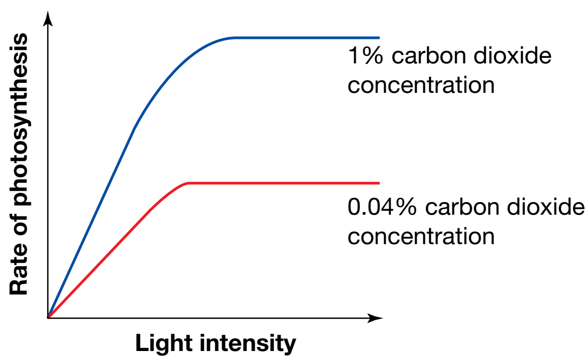



**FIGURE 4.25** Plot of rate of photosynthesis against light intensity at two different concentrations of carbon dioxide.

## Carbon dioxide concentration and rate of photosynthesis

Carbon dioxide is an essential input to the light-independent stage of photosynthesis. The rate at which photosynthesis occurs depends on the concentration of carbon dioxide. As long as the carbon dioxide molecules can be incorporated into glucose molecules, the rate will continue to increase as the concentration of carbon dioxide increases.

Figure 4.26 shows a graphical representation of the change in the rate of photosynthesis with changing carbon dioxide concentrations. Note the initial linear relationship between the increase in carbon dioxide concentration and the rate of photosynthesis. However, a point is reached when the rate of photosynthesis begins to level out, as shown by the flattening or plateauing of the graph.

One reason for this may be that the enzymes involved in carbon fixation are working at their maximum rate so that no further increase in rate is possible under the prevailing conditions. Another possible reason is that the availability of essential coenzymes, such as NADPH, may have become a limiting factor.

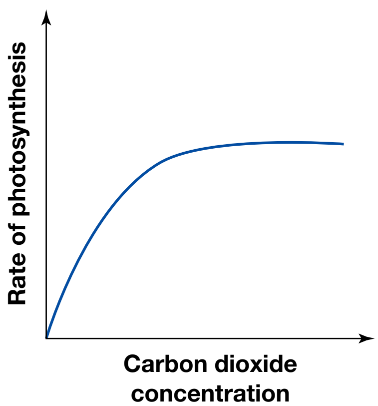



**FIGURE 4.26** Graph showing the change in the rate of photosynthesis as the concentration of carbon dioxide increases. Can you identify the point at which a limiting factor stopped the increase? What might this limiting factor be?

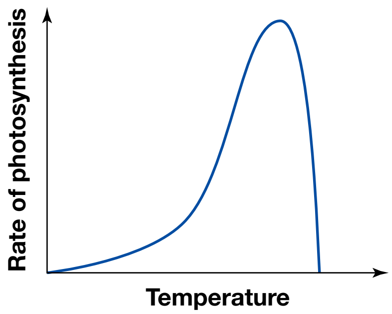



**FIGURE 4.27** Graph showing the change in the rate of photosynthesis with increasing temperature.

### Temperature and photosynthetic rate

Light-independent photosynthesis involves chemical reactions that are catalysed by different enzymes, such as RuBisCo. As the temperature rises, the rate of photosynthesis initially increases. This happens because the progress of chemical reactions depends on collisions between reactant molecules. An increase in temperature causes molecules to move faster so that the collisions between high-energy molecules occur more frequently, and reaction rates increase. As a rough guide, the rate of a reaction that normally occurs at room temperature doubles for each 10 °C increase in temperature.

Figure 4.27 shows the rate of photosynthesis plotted against increases in temperature.

Note that the increase in the rate of photosynthesis with increasing temperature does not continue indefinitely. If the optimal temperatures of the enzymes involved are exceeded, the rate of photosynthesis starts to decrease because of heat-induced changes to the tertiary structure of the enzymes. Soon after, the rate drops to zero and photosynthesis stops because the enzymes are denatured, and their structure is changed so that they can no longer function as enzymes.

#### KEY IDEAS

- Several factors can affect the rate of photosynthesis.
- These factors include light intensity, temperature and carbon dioxide concentration.
- A factor is said to be limiting if, in short supply, this factor restricts the rate of photosynthesis.

#### QUICK CHECK

**17** Briefly explain why the rate of photosynthesis does not continue to increase as the ambient temperature continues to increase.
**18** Does this same explanation apply to the fact that the rate of photosynthesis does not continue to increase as light intensity increases?
**19** What is meant by the 'plateauing' of a trend line?

## Dangers in confined spaces

A confined space is a space in an enclosed or partly enclosed structure to which a person can gain access, typically with some difficulty. Examples of confined spaces include wine vats, manure pits, sewerage tanks and bulk fuel tanks. These spaces may contain harmful fumes and may have unsafe oxygen levels.

People who enter such confined spaces may be going into potentially life-threatening situations. If workers involved in the cleaning, maintenance, or removal of large storage tanks or storage pits need to enter such spaces, they do so only when they are equipped with the appropriate personal protective equipment (PPE). Figure 4.28 shows employees of a company that was contracted to clean the insides of large above-ground storage tanks (the tanks are not shown). Note their PPE, which includes breathing apparatus that they wear to protect themselves from any residual harmful fumes inside the tanks.

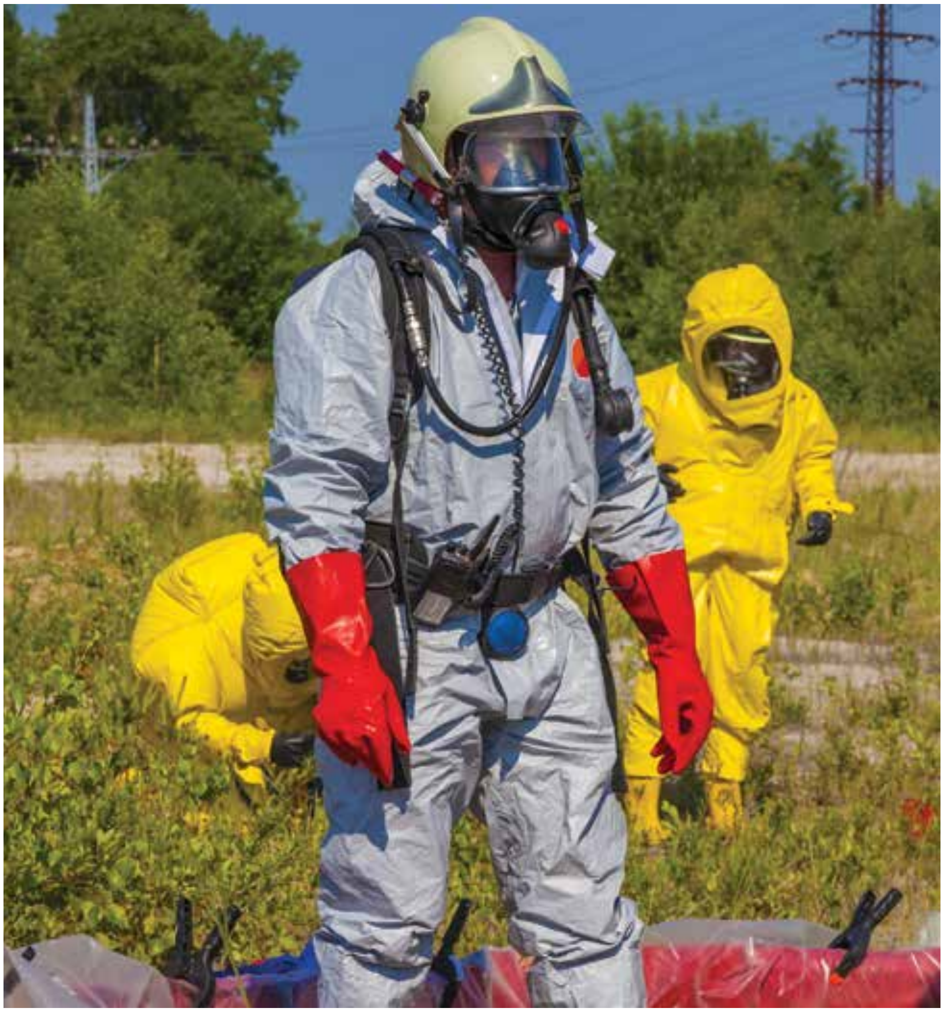


**FIGURE 4.28** Three workers equipped with personal protective equipment, including breathing apparatus, are equipped to enter the confined space of an above-ground storage tank. Maintaining an ongoing supply of oxygen is an essential safety measure.

## Storing manure poses hazards

Manure pits are another type of confined space. Manure is a waste product on a dairy farm, but, at the same time, it is a valuable fertiliser resource. Many dairy farmers, and farmers of other livestock, make use of manure pits to manage this waste, protect water quality and produce a useful fertiliser for use on their pastures and, at the same time, save money by reducing the amount of fertiliser they must buy. Storage allows manure to be safely stockpiled until conditions are environmentally safe for spreading. Modern manure storage structures on large dairy farms include walled enclosures, earthen ponds and above-ground tanks (see figures 4.29 and 4.30).

**FIGURE 4.29** A manure storage tank on a farm in Scotland.

The types of manure storage structures that pose the greatest risk are pits under buildings and closed or covered pump-out pits. The risk comes from the so-called 'manure gases', which are the products of the metabolism of microbes that gain their energy for living from the anaerobic fermentation of organic matter in the manure. The major gases produced by this type of metabolism are methane ($CH_4$), hydrogen sulfide ($H_2S$), ammonia ($NH_3$) and carbon dioxide ($CO_2$).

Prominent among these microbes are various species of archaea, known as methanogens (methane producers). These microbes carry out reactions termed methane fermentation, as for example:

acetic acid → methane + carbon dioxide
formic acid → methane + carbon dioxide + water

In cold weather, methane-producing microbes in a manure pit can produce up to 16 litres of methane per cubic metre of manure per day. In warm weather, this can increase to up to 30 litres per cubic metre of manure per day. Methane production and storage can benefit the running of a farm (see figure 4.30).

**FIGURE 4.30** Manure storage structures, such as the one in this photo, are used on 'biogas farms'. Biogas (principally methane) can be used as a fuel to heat or generate power on a farm. In this photo, the two storage tanks on the right are used to store the biogas.

The anaerobic fermentation by microbes breaks down complex organic matter in the manure to smaller organic compounds, forming a fertiliser that can be spread on pastures. To do this, the liquid slurry or fertiliser must be agitated and pumped from the storage pit. While this is happening, manure gases are released from the slurry into the air and some can remain in the pit even after it has been emptied.

Some of the manure gases, such as hydrogen sulfide, are highly toxic. Others, such as methane, are not poisonous. However, in the confined space of a manure pit, manure gases can build up and displace oxygen. If methane builds to high concentrations in a confined space and displaces oxygen, the methane can cause a person's death, not by poisoning, but by suffocation — not enough oxygen in the air in the pit to keep a person alive. Refer to table 4.1 to see the health effects associated with various levels of oxygen in the atmosphere. Entering any large confined space where oxygen may be displaced by other gases can be life-threatening for anyone without personal protective equipment (PPE). (Refer back to figure 4.28 to see PPE in use.)

**TABLE 4.1** Effects on the human body of various levels of oxygen in the atmosphere.

| % Oxygen in atmosphere | Health effects |
|---|---|
| 21 (normal oxygen content of air) | none |
| 19.5* | none |
| 16 | impaired judgment and breathing |
| 14 | faulty judgment, rapid fatigue |
| 6 | difficult breathing, death in minutes |

*Minimum concentration for safe entry to an enclosed space, such as a pit.

### Deaths on a dairy farm

Methane — you cannot see it, you cannot smell it, but it can cause instant death. Tragically, this was the case on a sunny day in July 2007 on a dairy farm in the United States.

That morning, a dairy farmer switched on a pump to transfer manure from a below-ground pit almost 3 metres deep to an above-ground holding pond. Suddenly, the outflow stopped. The pipe transferring the manure was blocked. The farmer decided to climb down a ladder into the pit and unclog the blocked pipe. To do this, the farmer effectively entered a large underground tank with a layer of about 45 centimetres of manure slurry covering the bottom of the tank. As fate would tragically have it, the farmer was entering a methane-laden, oxygen-deficient confined space where he faced death, either from suffocation or, if he lost consciousness before dying, by drowning in the slurry at the base of the pit.

On entering the pit, the farmer collapsed. A farmhand who saw this happen climbed into the pit to help but also passed out. The farmer's wife was told what had happened. She asked a neighbour to ring for help and went to the pit accompanied by two of her children. When help arrived, they found that five people were dead — the farmer, his wife, their two daughters, aged 11 and 9, and the farmhand. Sadly, the wife and two children had also entered the pit in a vain attempt to assist.

Oxygen is essential for human life. We, like many organisms, make use of **aerobic respiration** to gain the energy for living. Deprive us of oxygen for even a short period and we lose our ability to generate the energy for staying alive. In contrast, the microbes in the manure, like many other microbes, can gain the energy for living in the absence of oxygen, through **anaerobic respiration**. In the remainder of this chapter, we will examine cellular respiration, both aerobic (oxygen-requiring) and anaerobic (no requirement for oxygen).

studyon

Unit 3 | AOS 1 | Topic 7 | Concept 1

**Aspects of cellular respiration**
Summary screen and practice questions

## Purpose of cellular respiration

Cellular respiration is the process in cells that keeps all living organisms alive. There is no holiday period for cellular respiration — the use of chemical energy to stay alive is continuous and, as a result, cellular respiration is occurring all the time in every living cell. **In the simplest terms, the purpose of cellular respiration is to transfer the chemical energy stored in glucose into the chemical energy of ATP for use by cells for staying alive.**

The phrase 'staying alive' refers to all the energy-requiring (endergonic) processes that take place in living cells. In the human body, these processes include the homeostatic mechanisms that keep internal conditions within cells within narrow limits, the excretory processes that remove metabolic wastes, the production of antibodies and other activities of the immune system that defend against infection, the transmission of nerve impulses along neurons, and the contraction of muscle cells, including the cardiac muscle of the heart. Figure 4.31 shows a display of a person's heartbeat produced by electrocardiography. This display, known as an electrocardiogram, shows the electrical pattern of each heartbeat and, at right, displays the values of related parameters as follows:

- PULSE in beats per minute (BPM)
- BLOOD PRESSURE as systolic pressure (100) over diastolic pressure (65), with mean value (90)
- %$SpO_2$ is the oxygen saturation of the arterial blood
- RESPIRATION is given in respirations (breaths) per minute (RPM)
- TEMPERATURE is core body temperature (37.2 °C).

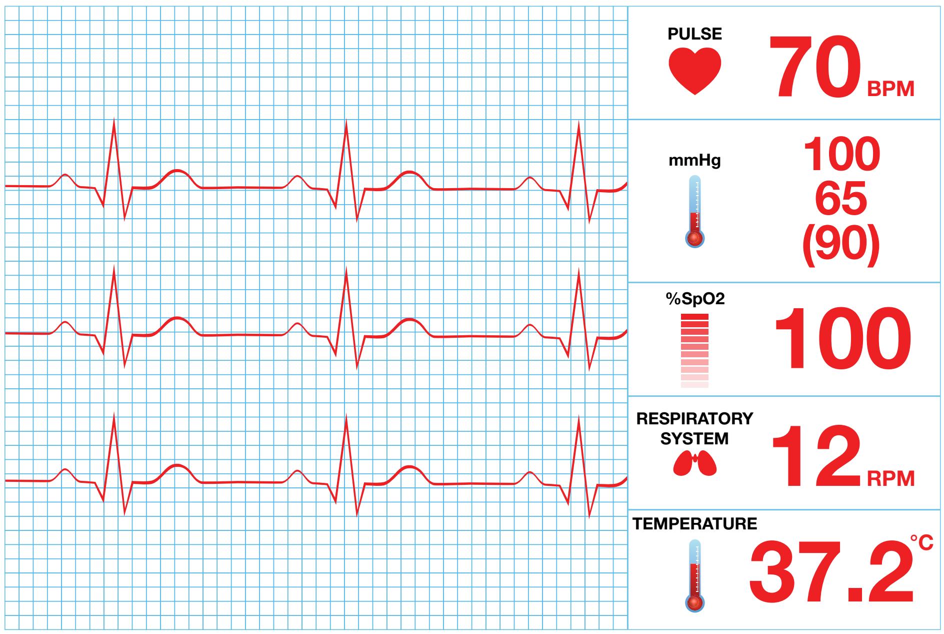



**FIGURE 4.31** An electrocardiogram (ECG) shows the electrical activity of the heart muscle as it beats. How many times per minute is this person's heart beating? What other information can be seen on the screen of an electrocardiograph machine?

- *Why use ATP for the purpose of providing energy for living?*

ATP is a useable form of energy for cells because its energy can be easily released in a single step, making energy instantly available for use by cells:

$$\mathbf{ATP \rightarrow ADP + P_i + Energy}$$

In addition, ATP has a very fast turnover rate. This means that both the breakdown of ATP to ADP and the regeneration of ATP from ADP occur very rapidly. Only a few seconds is needed to convert half of the ATP molecules in a cell to ADP and then to convert that ADP back to ATP.

- *Why not just use the chemical energy of glucose?*

The controlled release of energy from glucose is via a complex multistep pathway that is slower that the single-step release of energy from ATP. In addition, the chemical energy present in glucose is high (just under 3000 kJ/mole) — a bit like a \$100 note — while the energy content of ATP is a smaller, more useful amount (about 30 kJ/mole) — a bit like a \$1 coin.

## Cellular respiration: with or without oxygen?

Every living organism — plant, animal, fungus, bacterium, archaeon — carries out cellular respiration all the time. However, two major types of cellular respiration are recognised as differing in their dependence on an input of oxygen.

1. In most eukaryotic organisms and in most of their tissues, and in some microbes, the kind of cellular respiration that occurs is termed **aerobic respiration**, and it can only proceed in the presence of oxygen. Organisms that rely totally on aerobic respiration for their energy for living are termed **obligate aerobes**.
2. In many microbes and in some other organisms, a different kind of cellular respiration, termed **anaerobic respiration**, occurs, and it takes place in the absence of oxygen. Those organisms, mainly microbes, that rely exclusively on this form of cellular respiration are termed **obligate anaerobes**. (Later in this chapter, you will meet some versatile yeasts that fit neither pattern but, instead, can use both types of cellular respiration.)

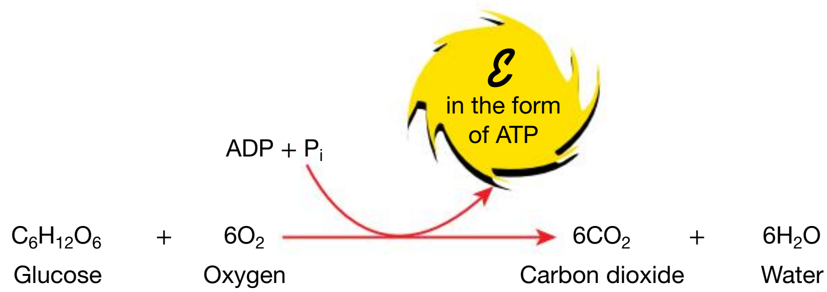
In the following sections, we will examine some aspects of these two types of cellular respiration.

**ODD FACT**

The term 'aerobic' was coined by Louis Pasteur in 1863 to mean 'living only in the presence of oxygen'. Its derivation comes from *aero-* = air and *bios* = life.

## Aerobic respiration: the big picture

Aerobic respiration can be shown by the simple equation:




For those organisms that depend on aerobic respiration for their energy for living, a severe shortage of oxygen is life-threatening and a complete lack of oxygen is fatal. This was seen, for example, in the tragic deaths that occurred in the manure pit on the dairy farm (refer back to page 132). Likewise, a scuba diver at depth whose oxygen supply runs out is in a life-threatening situation. An unacclimatised poorly prepared person attempting to climb Mt Everest (8900 metres) would experience a debilitating shortage of oxygen because the driving pressure for gas exchange in the lungs is reduced with increasing altitude. At sea level, the inspired oxygen pressure is almost 20 kPA, and it falls to 10 kPa at about 4000 metres above sea level, and to about 5 kPa at the summit of Mt Everest.

The critical importance of oxygen for aerobic respiration may also be seen in situations where oxygen is present in normal concentration in the air, but a person is either unable to access the oxygen for transport to the body cells or is unable to use the oxygen once it reaches the body cells.

- An inability to *access* oxygen occurs in severe untreated pneumonia when the alveolar sacs in the person's lungs become filled with fluid (see figure 4.32a). This condition greatly reduces the ability of red blood cells to take up oxygen from the lung alveoli, causing the oxygen supply to body cells to be inadequate.
- An inability to *use* oxygen occurs in people with **cyanide** poisoning. Such people can take oxygen into their lungs, transport it in their red blood cells and deliver the oxygen to body cells. However, cyanide is an inhibitor that stops oxygen from carrying out its essential role in aerobic respiration (as the terminal electron acceptor).

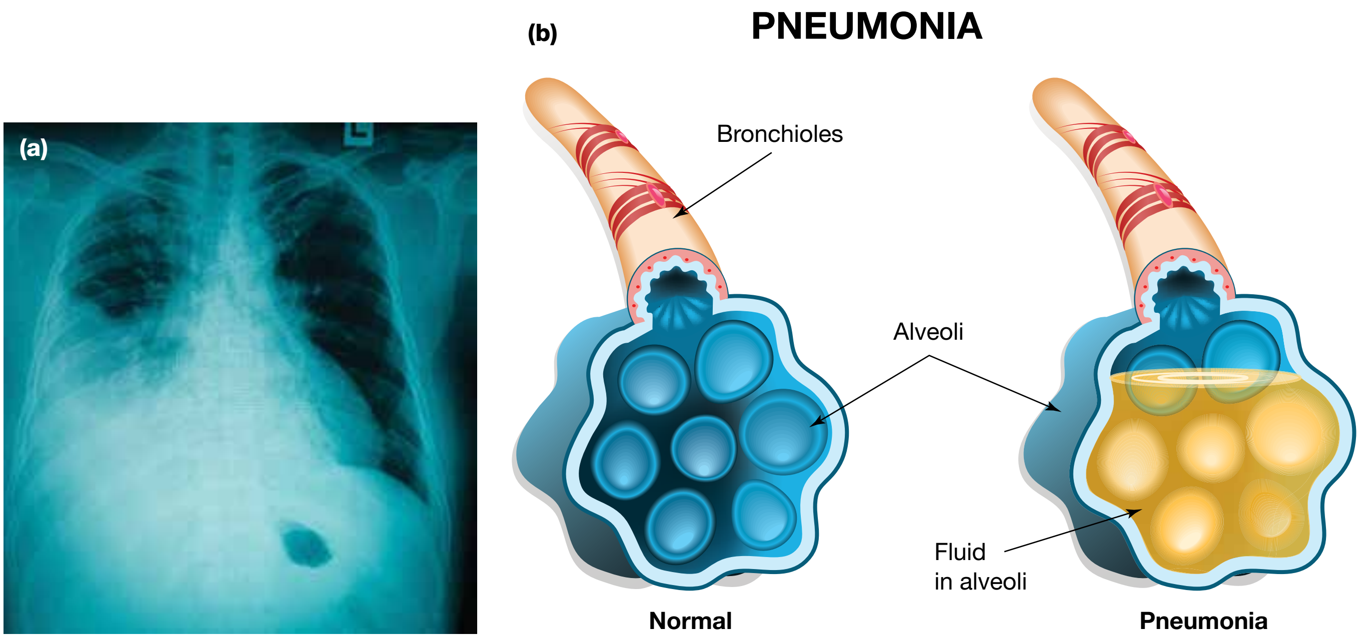



**FIGURE 4.32** **(a)** X-ray of human lungs. The dark areas in the lung are normal air-filled regions. The white areas in the lower lobes on both sides of the lung are caused by accumulated fluid resulting from a pneumonia infection. **(b)** Diagram showing (at left) a normal alveolar sac with air-filled alveoli and (at right) a fluid-filled alveolar sac. The presence of fluid in the sac adversely affects the supply of oxygen to the body cells and lungs, making breathing difficult.

## Aerobic respiration: inputs and outputs

Figure 4.33 shows the inputs and outputs of aerobic respiration.

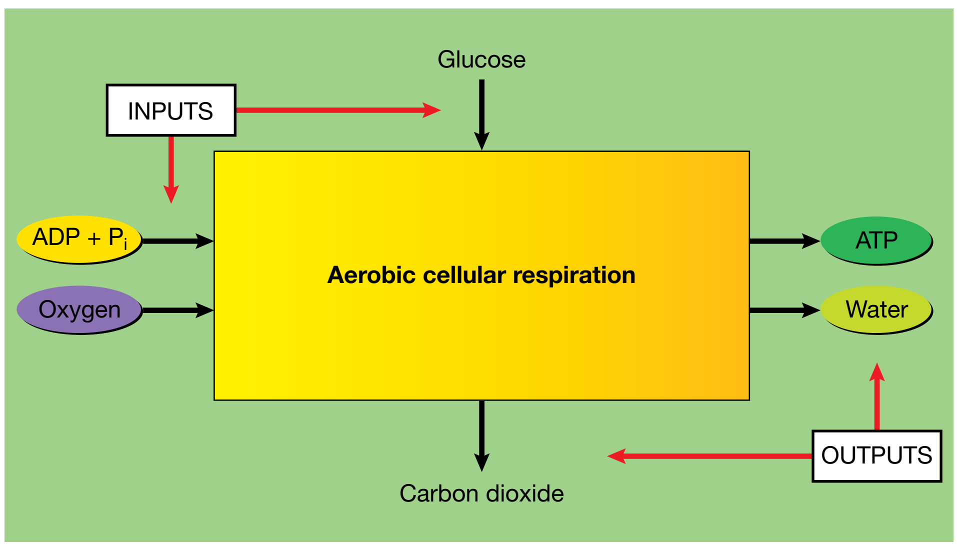



**FIGURE 4.33** Inputs and outputs of aerobic cellular respiration.

Heterotrophy and autotrophy are discussed on page 97 in *Nature of Biology 1 5th Edition*.

Let's look first at the inputs — glucose and oxygen.

- *What is the source of glucose?* People and all other **heterotrophic** organisms — all animals, all fungi and some microbes — obtain their glucose from food. In the case of animals, food is first ingested and then digested, while in the case of fungi, food is first digested and then ingested. For **autotrophic**

studyon

| | |
|---|---|
| Unit 3 | **Aerobic respiration: glycolysis** Summary screen and practice questions |
| AOS 1 | |
| Topic 7 | |
| Concept 2 | |

organisms, such as plants and algae, the source of their glucose is the output of their light-independent stage of photosynthesis.

- *Why is glucose needed?* Glucose is a high-energy sugar. It is the source of the energy that is ultimately transferred to ATP.
- *What is the source of the oxygen?* For terrestrial organisms, the source of oxygen is the air. In some cases, oxygen is taken in from the atmosphere through specialised organs such as the lungs of vertebrates, while, in other cases, oxygen is obtained by diffusion across moist outer surfaces, as in amphibians and small invertebrates. For aquatic animals, their source of oxygen is the dissolved oxygen in the water. Dissolved oxygen may be taken up in some cases through specialised organs, such as the gills of fish, while small invertebrates rely on diffusion for oxygen uptake.

What about plants? In daylight, plants and algae have their internal source of oxygen, which can be used for cellular respiration. This oxygen is a by-product of their photosynthetic activities. When this is not available, oxygen is taken up from their external environment.

The supply of oxygen in aquatic environments is more far variable than that in terrestrial environments. In aquatic environments, the concentration of dissolved oxygen changes markedly with changes in water temperature. In contrast, the concentration of oxygen in the air (about 21% by volume) is constant. Furthermore, the concentration of oxygen in air is much higher than in water, which is less than one per cent. See table 4.2, which shows how the oxygen content of water (in milligrams per litre) changes with water temperature.

**TABLE 4.2** Oxygen concentration in milligrams per litre of water at different temperatures. Would a fish population in a shallow lake be under greater oxygen stress in summer or in winter? Does this table give a clue as to why fluid-filled lungs caused by pneumonia result in a major restriction to the supply of oxygen to cells?

| Temperature (°C) | Oxygen (mg/L) | Temperature (°C) | Oxygen (mg/L) |
|---|---|---|---|
| 0 | 14.6 | 25 | 8.2 |
| 5 | 12.8 | 30 | 7.5 |
| 10 | 11.3 | 35 | 6.9 |
| 15 | 10.1 | 40 | 6.4 |
| 20 | 9.1 | 45 | 5.9 |

- *Why is oxygen needed?* Oxygen is the final electron acceptor at the end of an electron transport chain. In accepting electrons, oxygen also captures hydrogen ions, and the net resut is the production of water:

$$O_2 + 4e^- + 4H^+ \rightarrow 2H_2O$$

So, the oxygen that you take into your lungs with every breath is an essential electron acceptor of aerobic respiration. Without a continual supply of oxygen, the electron transport chain stops, effectively stopping aerobic respiration and its production of ATP. The limit on the time that a person can hold their breath is an indication of how long a person can manage without an input of oxygen — first they become unconscious as the brain cells are deprived of energy, then death follows.

Breathing is not a conscious activity as it is under control of the autonomic nervous system. But, stop, take a breath, and realise that those oxygen molecules will be transported by your circulatory system to various body cells where they will act as critical electron acceptors, enabling cells to produce the ATP, the energy source for staying alive.

- *How much oxygen does a person use?* The level of aerobic respiration in a person varies according to their level of physical activity. Aerobic respiration activity can most easily be quantified by measuring a person's rate of oxygen consumption, denoted $VO_2$. See table 4.3 for some values. Note that these oxygen consumption values refer to people, not specific tissues. Would you predict that various tissues within one person's body would have the same rates of oxygen consumption?

**TABLE 4.3** **Typical average values for rates of oxygen consumption ($VO_2$). The maximum oxygen consumption for a person is denoted by the symbol $VO_2$ max. The value of oxygen consumption of a person at rest can be denoted by the symbol $VO_2$ rest.**

| Subject | Average $VO_2$ (mL/kg/min) |
|---|---|
| young females (resting) | 29–33 |
| young females (maximal) | 38 |
| females (60 years) | 20–25 |
| young males (resting) | 36–42 |
| young males (maximal) | 45 |
| males (60 years) | 26–32 |
| elite cross-county skier (max) | 90 |
| thoroughbred horses (max) | 180 |
| husky sled dogs (max) | 240 |

- *If oxygen is being used all the time by all the cells of organisms that rely on aerobic respiration, why doesn't the supply of oxygen run out?*

Molecular oxygen ($O_2$) is an input to aerobic respiration, and it is constantly used, reappearing as part of the water output ($H_2O$). However, there is an offset. Plants and algae produce molecular oxygen as the output of the light-dependent stage of their photosynthesis. This oxygen is released into the external environment and so is available to other organisms for cellular respiration.

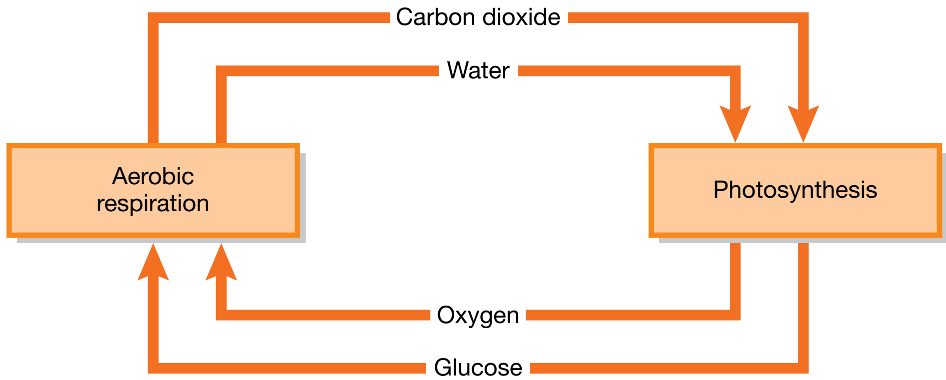
**Outputs of aerobic respiration:** The outputs of aerobic respiration are energy in the form of ATP, water and carbon dioxide.

ATP is the key output of aerobic respiration because, as we have seen, the purpose of cellular respiration is to transfer the chemical energy of glucose to the chemical energy of ATP, the versatile molecule that delivers chemical energy to cells.

The water produced in aerobic respiration is not a waste. Under normal conditions, about 300 mL of this 'metabolic' water is produced each day by cells, and it contributes to the water balance of the human body.

The carbon dioxide is a waste product that is excreted largely via the lungs into the air. However, carbon dioxide is an essential input to the light-independent stage of photosynthesis in plants and algae, and it is built into glucose via the Calvin cycle. The processes of aerobic respiration and photosynthesis can each make use of the outputs of the other process. Figure 4.34 shows the interrelationship between the inputs and outputs of aerobic respiration and photosynthesis.

FIGURE 4.34 Simplified diagram showing that the outputs of aerobic respiration are inputs to photosynthesis and that the outputs of photosynthesis can be used as inputs to aerobic respiration. In eukaryotes, photosynthesis occurs only in plants and algae while aerobic respiration occurs in plants, algae and animals.




The carbon atoms in the breath that you expire from your lungs can reappear as the starch in a kernel in a corn cob. When you eat a corn cob, you first digest the starch, then metabolise the glucose in your cells via aerobic respiration and then excrete the carbon atoms of the starch as gaseous carbon dioxide.

## Anaerobic respiration

**studyon**

Unit 3 | AOS 1 | Topic 7 | Concept 5

**Anaerobic respiration** Summary screen and practice questions

Some organisms, in particular many microbial species, live in environments in which oxygen is permanently absent. Such environments are said to be **anoxic** (*an* = not, *oxic* = ox(ygen) +ic). Examples of anoxic environments include marshes, hydrothermal vents, 'dead' zones of lakes and rivers, marine sediments and manure pits. As well as these larger scale anoxic environments, an anoxic environment may be as small as a deep puncture wound in a person's hand that can be happily occupied by *Clostridium tetani* bacteria, or the gut of a cow that is home to methanogenic microbes.

Organisms that live and reproduce in these environments are termed obligate anaerobes. They obtain the energy needed for living using **anaerobic respiration**.

**studyon**

Unit 3 | AOS 1 | Topic 7 | Concept 5

**Do more** Anaerobic respiration

**Anaerobic respiration** is a form of cellular respiration that does *not* require oxygen. (While aerobic respiration in all organisms has the same outputs of carbon dioxide, water and ATP, this is not the case for anaerobic respiration.) Anaerobic respiration in all microbes produces energy in the form of ATP, but different microbes do not all produce the same other outputs. In addition to ATP, the various outputs of anaerobic respiration include lactic acid, alcohol (ethanol) and carbon dioxide, acetic acid, acetone and methane.

Let's look in particular at:

- lactic acid fermentation
- alcoholic fermentation.

### Lactic acid fermentation in muscle cells

Mammals that normally rely on aerobic respiration can make use of anaerobic respiration in their muscle tissue for a short period only (measured in seconds). This can occur in skeletal muscle cells when the supply of oxygen to cells cannot keep up with their energy demands. Such a situation occurs, for example, during short intense bursts of muscle activity, such as a sprint over 100 metres or less, the sprint at the end of a stage in a bike road race or the lifting of a heavy weight by a competitive weightlifter (see figure 4.35). Because anaerobic respiration does not require oxygen, an athlete may not even take a breath when running 100 metres. A distinctive feature of highly strenuous physical activities is that they can be sustained for very short periods only, a maximum of just minutes. During periods of intense exercise, the muscle cells produce energy through lactic acid fermentation.

**FIGURE 4.35** Anaerobic respiration can provide ATP to power intense muscle activity, but only over a short time period, such as in this sprint to the finishing line. Would you expect that this type of respiration would be used to sustain a runner during a marathon race?

At the end of a sprint, athletes may be gasping for breath. The build-up of lactic acid in muscles has given them an oxygen debt. They breathe deeply and rapidly to restore the oxygen supply to their cells, including the supply of oxygen to oxidise lactic acid to pyruvic acid, and to re-oxygenate the myoglobin protein of skeletal muscle tissue. Their heart rates are elevated, increasing the blood flow as their circulatory systems remove the lactic acid from their muscles.

This type of anaerobic respiration is called **lactic acid fermentation**, and it can be shown as a simple equation:

**Glucose → Lactic acid + ATP**
**$C_6H_{12}O_6$ $\quad$ $2C_3H_6O_3$**

This same reaction is shown in figure 4.36, which shows the pyruvate and the involvement of the coenzyme, NADH. The 'loaded' acceptor molecules (NADH) that are produced in glycolysis are needed to drive the conversion of pyruvate to **lactate**.

Note that lactic acid fermentation has just a single input (glucose) and the end products are lactic acid and energy in the form of ATP. During strenuous exercise, lactic acid levels rise. Lactic acid dissociates, forming lactate and $H^+$ ions, which change the pH of muscle tissue, causing pain and fatigue.

The release of energy in lactic acid fermentation occurs more quickly than the release of ATP in aerobic respiration. However, anaerobic respiration *cannot* sustain life but, as identified above, it comes into play for very short periods when the energy needs of intense muscle activity cannot be met through aerobic respiration. It is energy for an emergency response. Handy for a gazelle being pursued by a cheetah, as well as for an Olympic sprinter, but also handy for the cheetah.

Note that the chemical output in lactic acid fermentation is lactic acid ($C_3H_6O_3$), not carbon dioxide ($CO_2$). Think about this. The end product of aerobic respiration is carbon dioxide, and this compound is completely oxidised and has no remaining energy that can be transferred to ATP. What about lactic acid? Does it appear to be completely oxidised? If not, what does this mean for the ATP yield from anaerobic respiration in comparison to that from aerobic respiration? (We will explore this issue later in this chapter; see page 145.)

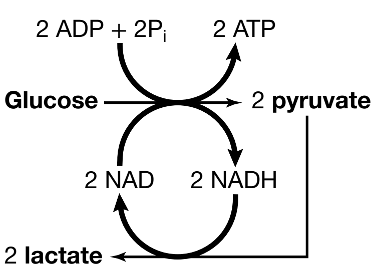



**FIGURE 4.36** Diagram showing the detail of lactic acid fermentation in human skeletal muscle cells. Is this an aerobic or an anaerobic process? Note that pyruvate is not an output — it is an intermediate product. What is the yield of ATP from this type of fermentation?

## Alcoholic fermentation in yeast cells

Yeasts are unicellular fungi. Some species of yeast are termed **facultative anaerobes** — this indicates that, when oxygen is present, these yeast cells produce ATP using aerobic respiration, but when oxygen is absent, they switch to anaerobic respiration for the energy to stay alive. In this event, the yeasts use a type of anaerobic respiration called **alcoholic fermentation**.

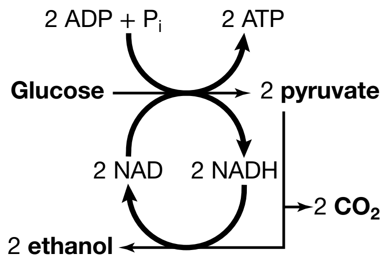



**FIGURE 4.37** Diagram showing detail of alcoholic fermentation in yeast cells. Is this an aerobic or an anaerobic process? Note that pyruvate is not an output — it is an intermediate product. What is the yield of ATP from this type of fermentation?

Alcoholic fermentation can be shown by a simple equation as follows:

$$\textbf{Glucose} \rightarrow \textbf{Ethanol} + \textbf{Carbon dioxide} + \textbf{ATP}$$
$$C_6H_{12}O_6 \quad 2C_2H_5OH \quad 2CO_2$$

This same reaction is shown in figure 4.37, which shows the pyruvate and the involvement of the coenzyme, NADH. The 'loaded' acceptor molecules (NADH) that are produced in glycolysis are needed to drive the conversion of pyruvate to ethanol.

Baker's yeast (*Saccharomyces cerevisae*) is used in bread making. Bread making has been part of human history for thousands of years, with much evidence of this practice found in ancient Egypt (see figure 4.38).

Breads can either be leavened or unleavened. Leavened refers to those breads made with a recipe that includes a 'raising' agent (also called a leavening agent) such as yeast, which causes the dough to expand and rise. Unleavened breads are made without any raising agent and are typically flatbreads such as Mexican tortillas.

**FIGURE 4.38** Wall painting from a tomb in Ancient Egypt showing bakers mixing and kneading dough. (Image courtesy of the Werner Forman Archive of the Bridgeman Art Library.)

The first breads baked by people were probably unleavened hard flatbreads. However, given that yeast spores are so widespread and float through the air, the first leavened breads may have been the result of accidental contamination of the wheat used to make bread.

The conditions in bread dough are anaerobic. If yeast is used as a leavening agent in bread making, the carbon dioxide gas produced during alcoholic fermentation by the yeast causes the dough to rise, and gives the baked bread its open texture and 'holes' (see figure 4.39). The ethanol is driven off during baking.

**FIGURE 4.39** Some breads: leavened (at left) and unleavened (at bottom right). Which type of bread depends on the alcoholic fermentation of yeast? (Image courtesy of Judith Kinnear.)

Wines and beers are also products of alcoholic fermentation by brewer's yeast (*Saccharomyces cerevisiae*). The starting point or input for wine production is the sugar in the juice of grapes. When the alcohol concentration in a wine vat (see figure 4.40) reaches about 11 per cent, the death of the yeast cells occurs and stops the fermentation process. The starting point of input for beer production is the carbohydrates of barley. In both cases, the alcohol output of the yeast fermentation is the desired product.

**FIGURE 4.40** Stainless steel wine vats in a winery are the sites of anaerobic alcoholic fermentation. What is the input to this process? What is the output? Wine vats like these are confined spaces. (Would a person needing to go into an empty wine vat to make an internal repair possibly be at risk. If so, why?)

**ODD FACT**

Baker's yeast and brewer's yeast are both members of the same species, but they are different strains that have been selected for different properties — the ability to make carbon dioxide in the case of baker's yeast and the ability to make alcohol in the case of brewer's yeast.

**study on**

| | |
|---|---|
| Unit 3 | **Factors affecting the rate of cellular respiration** |
| AOS 1 | Summary screen and practice questions |
| Topic 7 | |
| Concept 6 | |

## Comparing aerobic and anaerobic respiration

Table 4.4 shows a comparison of aerobic and anaerobic respiration.

**TABLE 4.4** Comparison of aerobic and anaerobic respiration. Complete breakdown of glucose refers to the oxidation of one molecule of glucose to six molecules of carbon dioxide.

| | Aerobic | Anaerobic (in animals) | Anaerobic (in yeasts, in plants) |
|---|---|---|---|
| Requires oxygen | YES | NO | NO |
| Number of steps in pathway | many | few | few |
| Complete breakdown of glucose | YES | NO, partial only | NO, partial only |
| Inputs | glucose & oxygen | glucose | glucose |
| Outputs | carbon dioxide & water | lactic acid | ethanol & carbon dioxide |
| Rate of ATP production | slower | faster | faster |
| ATP yield | high | low | low |
| Location in cell (in eukaryotes) | cytosol & mitochondria | cytosol | cytosol |

The above section has provided an overview of both aerobic and anaerobic respiration. In the following section, we will explore aerobic respiration in more detail, and meet the mitochondria that are the cell organelles where most of the ATP from aerobic respiration is generated.

## KEY IDEAS

- Cellular respiration occurs all the time in all living cells.
- The purpose of cellular respiration is to transfer the chemical energy of glucose to the chemical energy of ATP.
- Cellular respiration can be aerobic, with an essential requirement for oxygen, or it may be anaerobic, occurring in the absence of oxygen.
- The inputs of aerobic respiration are glucose, oxygen and ADP, and the outputs are ATP, carbon dioxide and water.
- Major oxygen deprivation or total lack of oxygen is a death sentence for organisms that rely on aerobic respiration.
- Examples of anaerobic respiration include lactic acid fermentation and alcoholic fermentation.
- The inputs and outputs of aerobic respiration are the same as the outputs and inputs of photosynthesis.

## QUICK CHECK

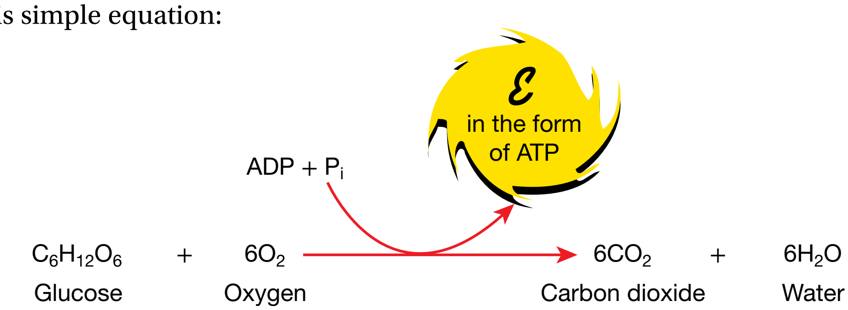
**20** Where and when might you expect to find the following:
   **a** lactic acid being produced in the human body
   **b** a facultative anaerobe producing carbon dioxide gas?

**21** Briefly explain the following observations:
   **a** In a liquor store, it is not possible to find a bottle of wine with an alcohol content of 20 per cent.
   **b** Production of ATP by aerobic respiration is much slower than by anaerobic fermentation.
   **c** Cyanide poisoning causes aerobic respiration to stop.
   **d** Leavened bread is raised and fluffy, and unleavened bread is flat and hard.

**22** Identify the following statements as true or false:
   **a** No form of cellular respiration can occur in the absence of oxygen.
   **b** If deprived of oxygen, a person could survive for several hours using anaerobic respiration.
   **c** The main purpose of aerobic respiration is different from that of anaerobic respiration.
   **d** The water output of aerobic respiration is derived from the oxygen input.
   **e** When competing, marathon runners gain their energy using aerobic respiration.

# Stages of aerobic respiration

In an earlier section of this chapter (see page 134) we looked at the major features of aerobic cellular respiration, seeing this process largely as shown in this simple equation:



While this equation shows the essential inputs and outputs of aerobic respiration, it conceals the complex nature of this process. The single reaction above is a simplification of what happens in aerobic respiration. The oxidation of glucose in cells does not occur in one step with the instantaneous release of a large amount of energy. In cells, the oxidation of glucose involves a controlled multistep reaction at low temperatures (37 °C in human cells). Several of these steps release small amounts of energy, which can be transferred to ATP.

To be oxidised outside cells, glucose must be heated to a temperature of more than 200 °C. This oxidation occurs in a single step, with most of the energy released as heat energy.

Aerobic respiration in cells involves many steps, and each step is catalysed by a specific enzyme. Coenzymes, in particular NAD (nicotinamide adenine dinucleotide) and FAD (flavin adenine dinucleotide), are also involved in many of these steps. NAD and FAD play important roles as **acceptor molecules**. In many steps of the aerobic respiration pathway, hydrogen atoms are removed from molecules, and these hydrogen atoms and, most importantly, their electrons, are accepted by NAD or FAD. In doing this, they become 'loaded' acceptor molecules, denoted as reduced NADH and $FADH_2$. Reactions that unload these loaded acceptor molecules release energy. This energy can be utilised to produce ATP. So, 'loaded' acceptor or carrier molecules are like vouchers that can later be turned into money.

We will look at a simplified account of aerobic respiration and focus on the energy transfer steps. Aerobic respiration can be divided into three stages: glycolysis, Krebs cycle and electron transport.

**studyon**

| | |
|---|---|
| Unit 3 | **Aerobic respiration: Krebs cycle and the electron transport chain** |
| AOS 1 | |
| Topic 7 | |
| Concept 4 | Summary screen and practice questions |

## Glycolysis: the first stage

The first stage of glucose breakdown, called glycolysis, consists of 10 steps. *Glycolysis comes from gluco* = 'sugar' *and lysis* = 'breakdown'.

### Where?

Glycolysis occurs in the cytosol of cells. The enzymes involved are in solution in the cytosol.

### What happens?

During glycolysis, one molecule of glucose is broken down into two molecules of a 3-C compound known as pyruvate (see figure 4.41). A net release of energy occurs and some of this energy is used directly to produce two molecules of ATP. Note also that two hydrogens are removed, resulting in the formation of two 'loaded' NADH acceptor molecules that will be used later in ATP production.

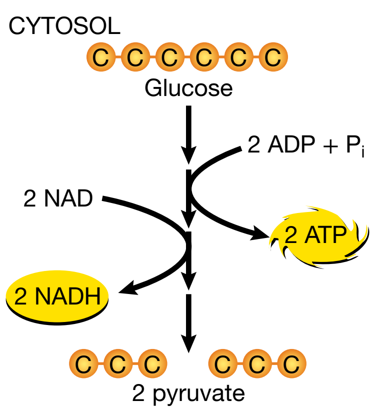



**FIGURE 4.41** During the process of glycolysis, each 6-carbon glucose molecule is broken down to form two 3-carbon pyruvate molecules. Energy is released in this process. What happens to this energy? Where in a cell does glycolysis occur? ($P_i$ = inorganic phosphate)

## Krebs cycle: the second stage

The second stage in aerobic respiration is known as the **Krebs cycle** (also known as the citric acid cycle), and its starting point is pyruvate.

### Where?

The reactions of the Krebs cycle occur in cell organelles, known as mitochondria, that form part of the cytoplasm of a cell. Because of their small size, mitochondria cannot be seen through a light microscope, but electron microscopy reveals their double-membrane structure. Each mitochondrion has an outer membrane and a highly folded inner membrane (refer to figure 4.42). The folds are termed cristae (singular: crista) and their presence increases the total surface area of the inner membrane as compared to the smooth outer membrane. The reactions of the Krebs cycle occur in the matrix that comprises the inner compartment of mitochondria.

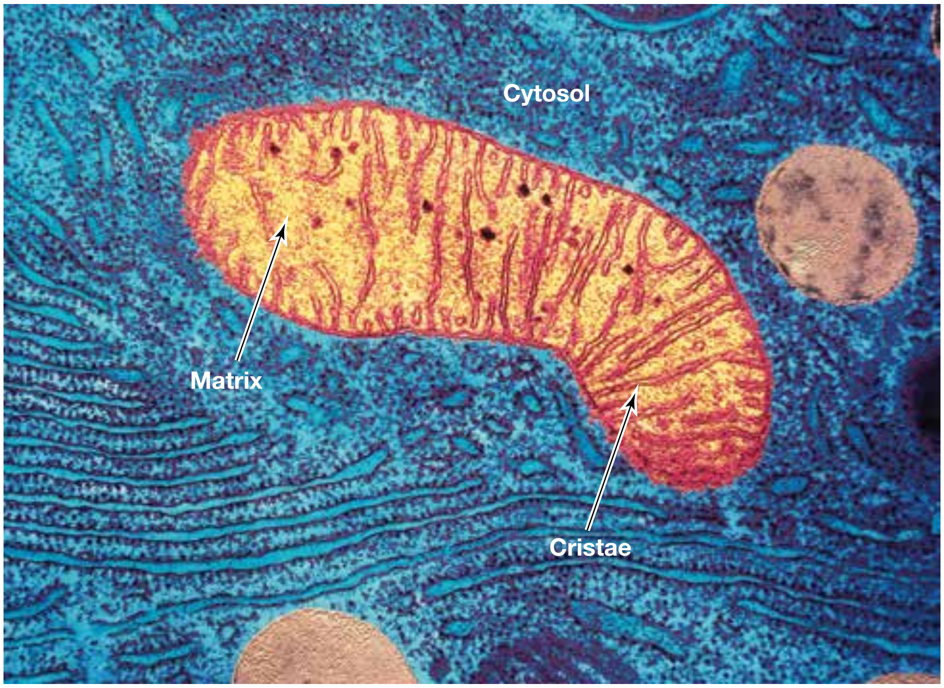



**FIGURE 4.42** Mitochondrion surrounded by cytosol and organelles. Enzymes of aerobic respiration gather in the matrix along cristae, which are extensive foldings of the internal membrane of the mitochondrion. The matrix also contains ribosomes and DNA molecules.

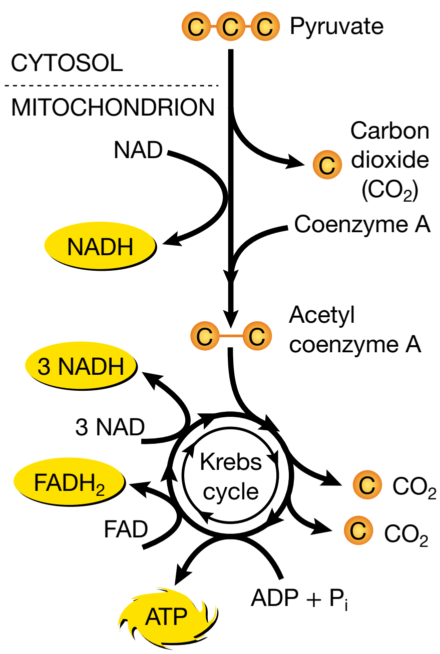



**FIGURE 4.43** Simplified version of the respiration of one molecule of pyruvate to form three molecules of carbon dioxide through the Krebs cycle. Are all carbon atoms accounted for? Energy released by this process is used to form ATP and 'loaded' acceptor molecules. How many of each would be formed from the breakdown of pyruvate from one molecule of glucose?

### What happens?

Pyruvate molecules leave the cytosol and enter the mitochondria where they react, forming an intermediate product called acetyl-CoA. In this step, one loaded NADH is formed. Acetyl-CoA now enters the Krebs cycle, which comprises nine steps.

For each molecule of acetyl-CoA that passes through the Krebs cycle, three molecules of carbon dioxide are formed (see figure 4.43). As the cycle proceeds, hydrogen atoms are gathered by acceptor molecules, with a total of four 'loaded' acceptor molecules being formed (three NADH and one $FADH_2$) for each acetyl-CoA molecule.

## Electron transport: the third stage

The final reactions in aerobic respiration are called electron transport reactions. Electron transport involves reactions of the 'loaded' acceptor molecules. Oxygen is essential for electron transport.

### Where?

Electron transport takes place on the inner membranes of mitochondria. The electron transport chain consists of a series of enzyme complexes embedded in the inner membrane of the mitochondria. The final electron acceptor, oxygen, is free within the matrix.

### What happens?

During electron transport, electrons from 'loaded' acceptors (NADH and $FADH_2$) are brought to the inner membranes of the mitochondria and accepted by the enzyme complexes of the electron transport chain. The electrons are successively transferred along the chain until they are finally accepted by oxygen. The enzyme complexes are reduced as they accept an electron and are then oxidised as they pass it on to the next carrier. These reversible reactions are reduction-oxidation or redox reactions. When oxygen combines with a pair of electrons, it becomes negatively charged ($O^{2-}$) and then reacts with hydrogen ions ($H^+$) to form water: $O_2 + 4e^- + 4H^+ \rightarrow 2H_2O$

Electron transport releases energy that is used for production of ATP. Theoretically, for each pair of electrons transferred down the electron transport chain, enough energy is released to drive the production of two or three molecules of ATP (see figure 4.44). (Some researchers propose that these numbers are not realised in practice.)

Check out the box at the end of this section to see how the energy released by the transport of electrons along the electron transport chain is used to produce ATP. Remember, the production of ATP for staying alive, using energy transferred from glucose, is the purpose of aerobic respiration.

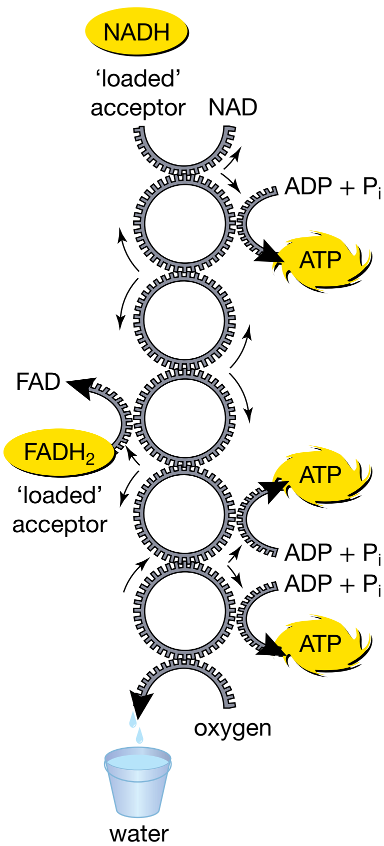



**FIGURE 4.44** A series of enzyme complexes that include cytochromes, located on the inner membranes of mitochondria, are involved in electron transport. This transport releases energy that can be used in ATP production. The electrons ($e^-$) and hydrogen ions ($H^+$) are finally accepted by oxygen to form water. Different 'loaded' acceptor molecules (NADH and $FADH_2$) start this process at different points.

## Outcome of the three stages

The aerobic respiration of one molecule of glucose to carbon dioxide releases energy that is coupled to the production of about 30 molecules of ATP. The question of the yield of ATP from aerobic respiration is discussed in the next section.

The process of aerobic respiration can be shown in several ways. Some examples are shown in figure 4.45 below.

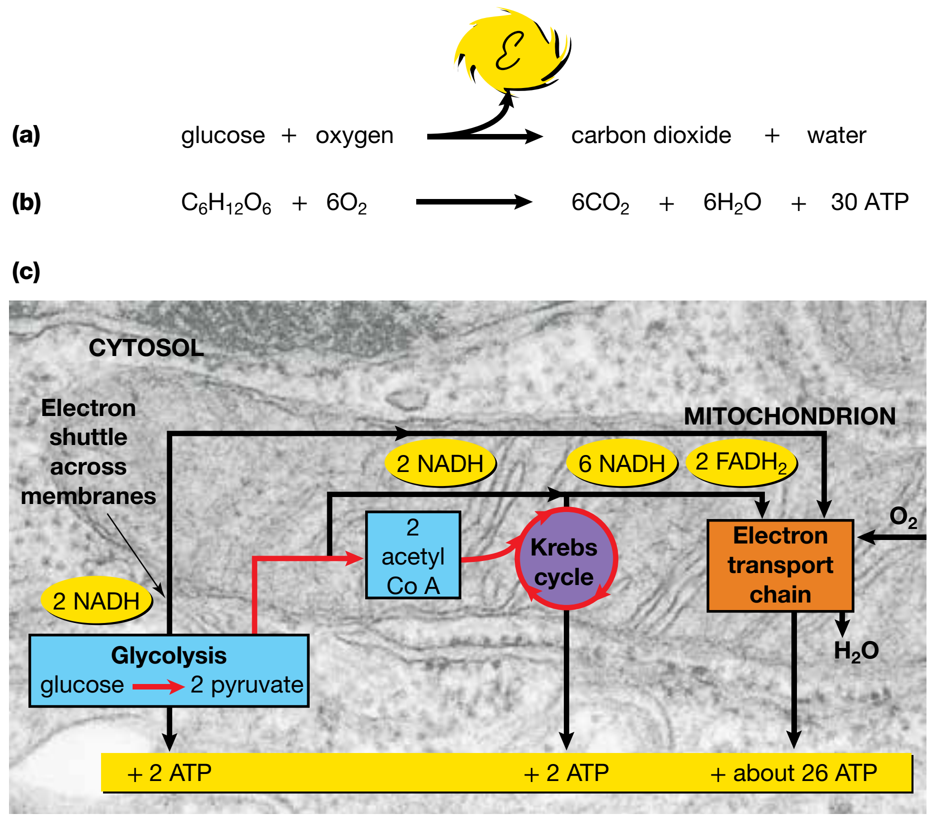



**FIGURE 4.45** The process of aerobic respiration can be shown in several ways. What is the difference between representations **(a)** and **(b)**, and between **(b)** and **(c)**?

### Yield of ATP from aerobic respiration

What is the yield of ATP from aerobic respiration? In other words, for every molecule of glucose that is broken down and oxidised to carbon dioxide through the process of aerobic respiration, how many molecules of ATP are produced? Pick up three different books and you may well see three different answers to this question.

The chemical reaction of glucose with oxygen under standard conditions (temperature of 298 kelvin, pressure of 1 atmosphere, pH of 7.0, and initial concentrations of reactants and products being equal) gives a theoretical yield of 36 ATP. However, the conditions in cells are not necessarily standard, so this is not what happens in cells.

Some researchers identify the *theoretical maximum* yield of 36 ATP but then qualify this value by stating that this yield is rarely achieved in cells. Others identify the *theoretical* yield of ATP from glucose as 30 or 32 moles of ATP per mole of glucose.

What is generally agreed is that, starting with one molecule of glucose, the outputs are:

| | | | |
|---|---|---|---|
| • yield from glycolysis: | 2 ATP | 2 NADH | |
| • yield from the Krebs cycle: | 2 ATP | 8 NADH | 2 $FADH_2$ |
| **TOTALS:** | **4 ATP** | **10 NADH** | **2 $FADH_2$** |

There is no disagreement about the numbers of NADH and $FADH_2$ produced, nor the number of ATPs produced outside the electron transport chain. The differences lie mainly in converting the energy released by the input of NADH and $FADH_2$ to ATP.

We know that the electrons carried by $FADH_2$ have less energy than the electrons carried by NADH. As a result, $FADH_2$ electrons pump fewer protons into the inter-membrane space than the electrons from NADH.

The following conversion values are favoured by some researchers:

1 NADH can transfer enough energy to produce 2.5 ATP.
1 $FADH_2$ can transfer enough for 1.5 ATP.

Using these values, this gives a yield of ATP as:

4 ATP + $(10 \times 2.5) + (2 \times 1.5) = 4 + 25 + 3 = 32$ ATP for prokaryotes.

**eBook*plus***

**Interactivity**
Photosynthesis or respiration
**int-3039**

Other conversion values have also been used, such as 3 for NADH and 2 for $FADH_2$. Using these values, the yield would of course increase to 38 ATP for prokaryotes (and 36 for eukaryotes).

Back to the 32 ATP for prokaryotes. In prokaryotes, all the reactions of aerobic respiration occur in the cytosol. In eukaryotic cells, however, only the initial reactions occur in the cytosol with the remainder occurring in the mitochondria. Because there is an energy cost in shuttling the NADH produced in glycolysis across the mitochondrial membranes, the value of 32 needs to discounted by that cost, namely 2 ATP.

So, for eukaryotes, a reasonable *estimate* of the theoretical maximal ATP yield of aerobic respiration is $32 - 2 = 30$ ATP. However, keep in mind that this is an estimate, not a definitive answer. One researcher has stated that '*about 30 molecules of ATP are formed when glucose is completely oxidized to $CO_2$;* this value supersedes the traditional estimate of 36 molecules of ATP.' (Source of quotation: Berg www.ncbi.nlm.nih.gov/books/NBK22448/)


A question whose answer depends on the assumptions made.

## Aerobic respiration: not just glucose!

Up to this stage, glucose has been identified as the only molecule that has its chemical energy transferred to ATP in the process of aerobic respiration. In fact, our dietary intake includes many other energy-rich organic compounds that can also be used in aerobic respiration and can transfer their energy to ATP (see table 4.5).

**TABLE 4.5** Energy-rich compounds of dietary intake. Along with glucose, note that fructose and galactose are simple sugars or monosaccharides.

| Dietary intake | After digestion |
|---|---|
| proteins | various amino acids |
| polysaccharides | |
| • starch | glucose |
| disaccharides | |
| • sucrose | glucose & fructose |
| • lactose | glucose & galactose |
| fats & oils (lipids) | glycerol & fatty acids |

All of these products of digestion, except for the amino acids, are channelled directly to a point where they can enter the aerobic respiration pathway. From there, they move through the steps of the pathway and, just as occurs with glucose, they are finally oxidised to carbon dioxide as their chemical energy is transferred to ATP.

In the case of the amino acids, they are first transported to the liver where their amino group ($-NH_3$) is stripped from those that are not needed for building the body's own proteins. This leaves a carbon 'skeleton' that is then channelled into the aerobic respiration pathway. Figure 4.46 shows these various entry points.

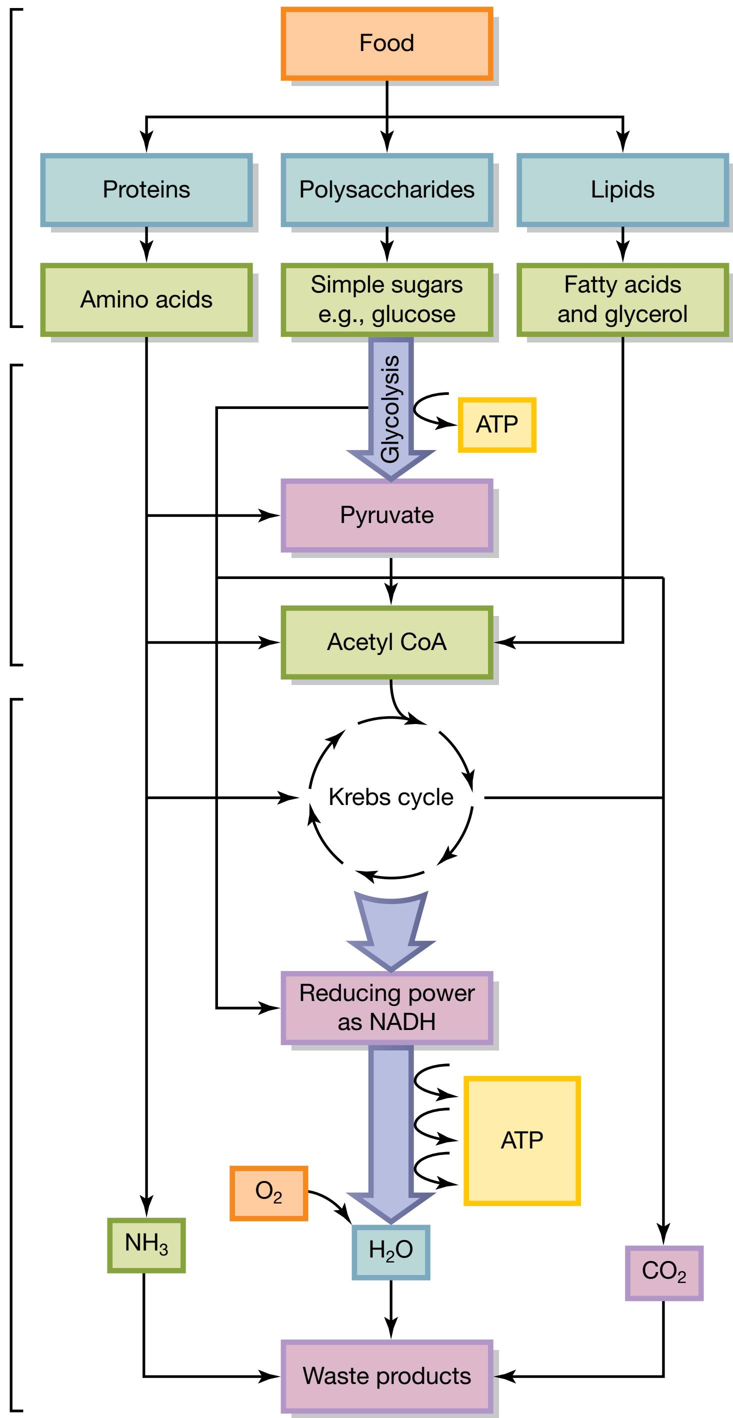



**FIGURE 4.46** Diagram showing the various entry points to the aerobic respiration pathway for products of digestion. Note that amino acids, because of their different R groups in their structure, can enter at several different points. Where do the monosaccharides (simple sugars) enter?

## USING ENERGY FROM ELECTRON TRANSPORT TO GENERATE ATP

Four enzyme complexes (I to IV) make up the electron transport chain. These enzyme complexes are embedded in the inner mitochondrial membrane. Electrons are transported along the electron transport chain from complex I to complex IV. Finally, the electrons are accepted by oxygen, and the reduced oxygen combines with hydrogen ions to form water by the following equation: $\frac{1}{2}O_2 + 2e^- + 2H^+ \rightarrow H_2O$ (see figure 4.47).

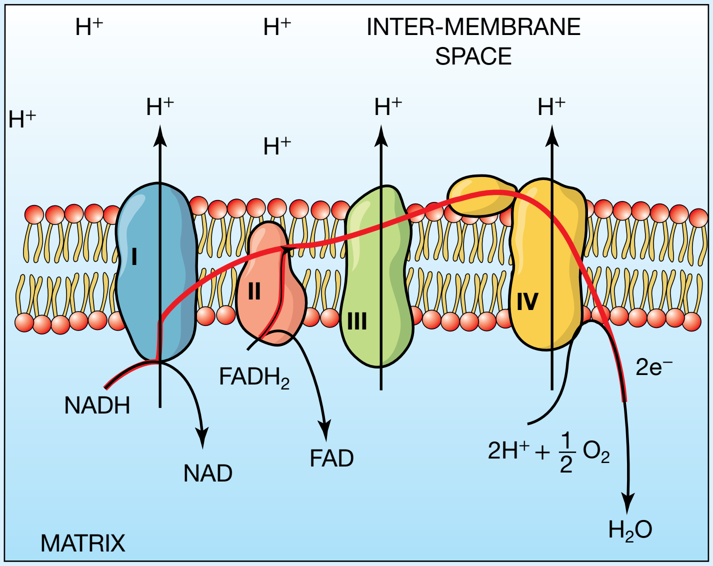



**FIGURE 4.47** Electron transport chain. Note the four enzyme complexes (I to IV). The transport route of the electrons is shown by the red line.

the inter-membrane space relative to the concentration in the matrix. After the gradient is established, protons diffuse down the electrochemical gradient. Their path back to the matrix takes the protons through an ion channel in a protein complex called ATP synthase (see figure 4.48).

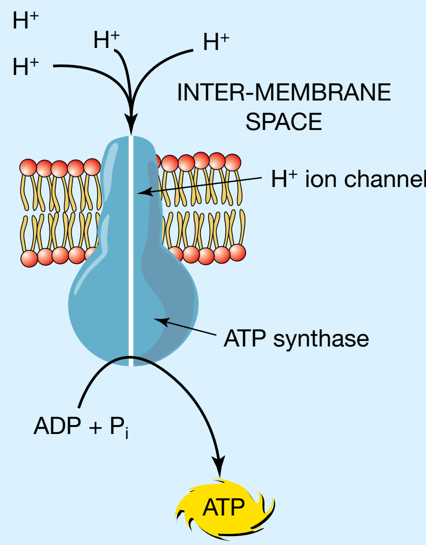



**FIGURE 4.48** ATP synthase is a protein complex, embedded in the inner mitochondrial matrix, that has both an enzyme function (catalysing the phosphorylation of ADP to form ATP) and a transport function (transporting hydrogen ions or protons across membranes).

Electron transport releases energy. However, this energy is *not* used directly to phosphorylate ADP to form ATP. What happens is that the energy is used to push hydrogen ions (protons) across the inner mitochondrial membrane into the inter-membrane space (refer back to figure 4.47).

The accumulation of protons in the inter-membrane space creates an electrochemical gradient, with a higher concentration of protons in the inter-membrane space relative to the concentration in the matrix.

ATP synthase traps the kinetic energy of the passage of the protons through the ion channel to generate ATP from ADP.

The generation of ATP by this means is termed **chemiosmosis**. In 1978, the Nobel Prize in Chemistry was awarded to the English biochemist Peter Mitchell (1920–1992) '*for his contribution to the understanding of biological energy transfer through the formulation of the chemiosmotic theory*'.

## KEY IDEAS

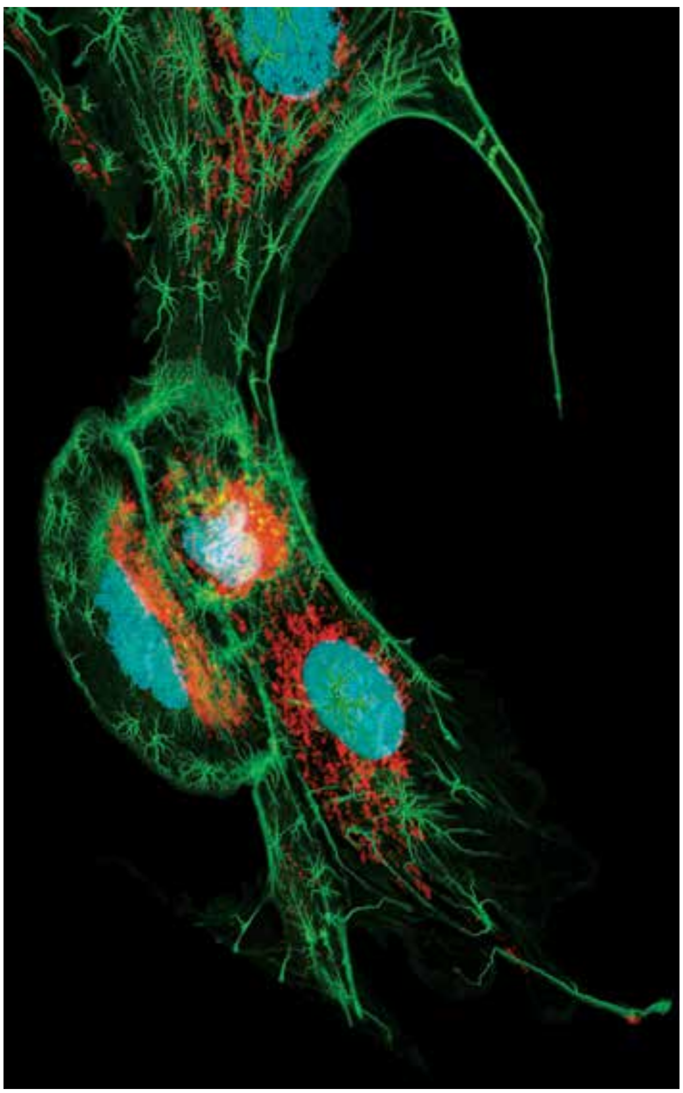
- Three stages are commonly recognised in aerobic respiration.
- Glycolysis occurs in the cell cytosol, while the Krebs cycle takes place in the mitochondrial matrix and the electron transport occurs on the surface of the cristae of the inner membrane.
- Each stage is a multistep pathway, only some steps directly yielding ATP.
- 'Loaded' coenzymes (reduced NADH and $FADH_2$) act as hydrogen and electron acceptors.
- In addition to glucose, other organic molecules can enter the aerobic respiration pathway and transfer their chemical energy to ATP.

## QUICK CHECK

**23** Identify the stage(s) of aerobic respiration that occur in:
- **a** the cytoplasm
- **b** the mitochondria
- **c** the mitochondrial matrix.

**24** Put a label on the following molecules or components that are part of aerobic respiration:
- **a** a reduced loaded acceptor
- **b** the end product of glycolysis
- **c** the final acceptor of the electron transport chain
- **d** the protein complex that produces ATP by chemiosmosis.

**25** Identify the following statements as true or false:
- **a** Electron transport precedes the Krebs cycle.
- **b** ATP synthase is needed to break down ATP to ADP.
- **c** Glucose transfers its chemical energy to ATP through the aerobic respiration pathway.
- **d** The stage of aerobic respiration that releases the greatest amount of energy for ATP production is the electron transfer stage.
- **e** It is reasonable to say that the yield from one molecule of glucose is about 30 molecules of ATP.



# Mitochondria: the powerhouses

We have seen that mitochondria are cell organelles present in the cytosol of all cells that rely on aerobic respiration for the supply of the ATP they need for living. Because they are not visible with conventional light microscopy, mitochondria are sometimes 'out-of-sight and out-of-mind'. However, fluorescence microscopy techniques have produced stunning images of the mitochondria in cells. Check out figure 4.49, which shows a fluorescence micrograph of a cell stained with the fluorescent dyes MitoTracker Red CMXRos and BODIPY FL phallacidin. These dyes target the intracellular mitochondrial network (red) and the cytoskeletal actin filaments (green) respectively.

Mitochondria are often called the 'powerhouses' of a cell. Why? The mitochondria, in particular the surfaces of their inner membrane (cristae), are the major site of ATP production in aerobic respiration. In addition to producing energy, mitochondria generate heat, and mediate cell death by apoptosis (see chapter 5, page 209).

**FIGURE 4.49** Fluorescence emission intensity from a culture of bovine pulmonary artery endothelial cells. Note the bright yellow-orange fluorescence exhibited by the tubular mitochondria and the green emission from actin filaments in the cytoplasm.

The number of mitochondria per cell varies, but, in general, cells that are most active either physically (such as cardiac muscle) or biochemically (such as liver cells) have the most dense populations of mitochondria. The number of mitochondria in heart (cardiac) muscle cells is estimated to be up to 5000 per cell, with mitochondria occupying about 40 per cent of the cell volume. The estimated number in liver cells is up to 2000 mitochondria per cell, and they occupy about 20 per cent of the cell volume.

In contrast, red blood cells have no mitochondria. As a result, they can only make ATP through the anaerobic lactic acid fermentation pathway. Still, they are not self-propelled — they just get carried passively in the bloodstream and they do little more than transport a bag of haemoglobin. When you compare the ATP outputs of aerobic respiration (30 ATP per molecule of glucose) with that of anaerobic fermentation (2 ATP per molecule of glucose), it is apparent that mitochondria are a great asset for cells. This is because cells with mitochondria can gain the energy for living very economically and efficiently. So, from where did these little powerhouses in our cells originate?

**studyon**

| | |
|---|---|
| Unit 3 | **Role of the mitochondria** Summary screen and practice questions |
| AOS 1 | |
| Topic 7 | |
| Concept 3 | |

## Origin of mitochondria

Earlier in this chapter (refer back to page 115), the endosymbiotic theory was introduced in relation to the origin of the plant chloroplasts. This theory proposes that chloroplasts were once free-living photoautotrophic bacteria that became endosymbionts in an early eukaryotic cell.

Mitochondria are also believed to have originated from free-living bacteria that gained their energy by aerobic respiration. Like the chloroplast ancestor, these aerobic bacteria became endosymbionts in another cell that was only capable of anaerobic respiration. This endosymbiotic relationship was mutually beneficial. The new intracellular tenant provided the host cell with a more efficient means of producing ATP, the source of energy for living. The host cell provided the tenant with a protective environment.

Evidence in support of the endosymbiotic origin of mitochondria comes from observations that highlight structural and biochemical similarities between mitochondria and bacteria:

- Mitochondria have lengths commonly in the range of 0.5 to 10 micrometres(μm). This falls within the size range of bacteria, most of which have lengths in the range of 2 to 8 μm.
- Mitochondria have their own DNA, separate from that of the host cell.
  - The mtDNA consists of 16 569 base pairs that encode the information for 37 genes, 2 ribosomal RNAs, 22 transfer RNAs and 13 polypeptide chains.
  - This genome would not support independent living outside its eukaryotic cell.
- The mitochondrial DNA is arranged as one circular molecule, such as occurs in bacterial cells (see figure 4.50a).
  (In contrast, the DNA of eukaryotic cells is organised into chromosomes.)
- The mtDNA replicates independently of the host DNA.
- Mitochondria have their own ribosomes with a 70S size, similar to those in bacterial cells.
  (In contrast, eukaryotic cells have larger 80S ribosomes.)
- Mitochondria divide by binary fission, such as occurs in bacterial cells (see figure 4.50b).
  (In contrast, eukaryotic cells divide by mitosis.)
- Mitochondria have a double membrane, similar to the double membranes of modern Gram-negative bacteria.
- The mitochondrial outer membrane contains porins that are transmembrane channel proteins; these are similar in structure to bacterial porins.

**eBook*plus***

**eLesson**
Binary fission
**eles-2464**

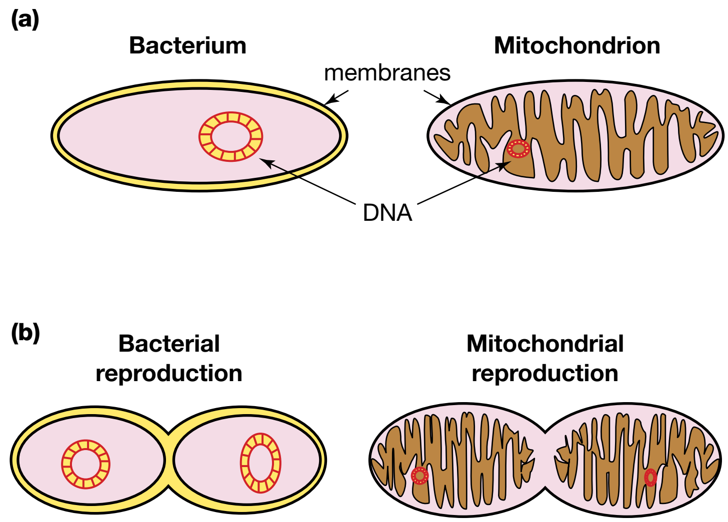



**FIGURE 4.50** Comparisons of a bacterium and a mitochondrion. **(a)** Both have DNA arranged as a circular molecule. **(b)** Both divide by a process of binary fission — 'splitting in two'!

Refer back to page 117 to check on some examples of endosymbiosis in today's living world.

*Which modern bacteria are most closely related to mitochondria?* Comparative studies of a ribosomal sub-unit from both mitochondrial RNA and bacterial RNA indicate that the free-living precursor of mitochondria was most likely a member of the proteobacteria group, possibly one of the rickettsias. One study published in the journal *Nature* (Sept 1998) details the results from researchers who analysed the complete genome of *Rickettsia prowazekii*, an obligate intracellular parasite, and concluded that this bacterial species '*is more closely related to mitochondria than is any other microbe studied so far*'.

## KEY IDEAS

- Mitochondria are the powerhouses of cells.
- Cells with higher levels of activity tend to have more dense populations of mitochondria.
- Strong evidence exists for the endosymbiotic origin of mitochondria.

## QUICK CHECK

**26** Identify the difference between how a eukaryotic cell divides and how its mitochondria divide.
**27** Briefly explain how red blood cells can survive without mitochondria.
**28** What is meant by the term 'endosymbiosis'?

## VARIOUS STYLES OF ENERGY CAPTURE IN THE LIVING WORLD

Organisms may be classified into various groups according to the forms of energy that they capture and the sources of the carbon atoms that they need for the processes of staying alive.

Which kinds of living organism belong to each group?

### 1. Eukaryotes

- Plants and algae are autotrophs that obtain radiant energy from sunlight and use it to build their own organic molecules from carbon dioxide.
- Animals and fungi are heterotrophs that obtain chemical energy and pre-formed organic molecules from the food that they ingest or absorb.
- The remaining eukaryotes are the protists that include both autotrophs and heterotrophs, such as the heterotrophic unicellular *Amoeba*, and autotrophic *Euglena*.

We can put this into a table — note that all the eukaryotes fit into two blocks: the photoautotrophs and the chemoautotrophs.

| Carbon source | Energy source: Light | Energy source: Chemical |
|---|---|---|
| Carbon dioxide | Plants, algae and some protists (e.g. *Euglena*)<br>Photoautotrophs | Chemoautotrophs |
| Organic molecules | Photoheterotrophs | Animals, fungi and some protists (e.g. *Amoeba*)<br>Chemoheterotrophs |

### 2. Prokaryotes

What about the bacteria and archaeans of the world of prokaryotic microbes?

***Some microbes are chemoheterotrophs like animals and fungi:***

Many bacteria are heterotrophs that gain their energy and carbon atoms from organic molecules. These are the pathogenic bacteria that cause diseases of living organisms and the decomposer bacteria that feed on dead organisms. Included among the many heterotrophic bacteria are *Mycobacterium tuberculosis*, which causes tuberculosis (TB) and *Streptococcus pyogenes,* one of the causes of toxic shock. These bacteria take in pre-formed organic molecules, which are a source of both energy and organic building blocks. We can add these microbes to the blue block in the table above, occupied by animals and fungi.

***Other microbes are photoautotrophs like plants and algae:***

Some bacterial species have pigments that allow them to capture sunlight energy and make glucose from carbon dioxide. These species include cyanobacteria, and the purple sulfur bacteria and the green sulfur bacteria. Cyanobacteria have chlorophyll and carry out photosynthesis like plants and algae.

Purple sulfur bacteria and green sulfur bacteria live in environments such as waterlogged soils and marshes where hydrogen sulfide ('rotten egg gas') accumulates. These bacteria possess a sunlight-trapping pigment called **bacteriochlorophyll** and, like plants, can build glucose from carbon dioxide. So they are photo-autotrophs. These bacteria use sunlight energy to split hydrogen sulfide ($H_2S$), rather than water ($H_2O$), to generate the hydrogen needed to reduce carbon dioxide. So, instead of producing oxygen, this type of photosynthesis produces sulfur:

***Purple sulfur bacteria:***

$$6\,CO_2 + 12\,H_2S \longrightarrow C_6H_{12}O_6 + 12S + 6H_2O$$

***Plants and cyanobacteria:***

$$6\,CO_2 + 12\,H_2O \longrightarrow C_6H_{12}O_6 + 6O_2 + 6H_2O$$

We can add these microbes to the green block in the table above, occupied by plants and algae.

***Some microbes are like heterotrophs,* but *they can also capture light energy:***

These microbes include the so-called AAP bacteria and the PR-containing bacteria and archaea that are found in freshwater and ocean habitats. Like heterotrophs, these microbes require pre-formed organic molecules that they obtain from dissolved organic matter in water. However, these microbes have pigment (proteorhodopsins) that can capture sunlight energy. They use sunlight energy to produce ATP directly from ADP. These microbes are photo-heterotrophs. We can add them to the yellow block in the table above.

***Some microbes are chemosynthetic autotrophs:***

Microbes found in extreme environments capture energy from chemical reactions, such as the oxidation of sulfides to sulfur or to sulfates, or the oxidation of nitrites to nitrates. They use this chemical energy from these oxidation reactions to build complex glucose from carbon dioxide. Examples of chemo-synthetic bacteria include *Nitrobacter*, a nitrifying bacteria, and *Sulfolobus*, a sulfur-oxidising archaea.

These microbes are chemosynthetic autotrophs. We can add them to the yellow block in the table.

This covers all living organisms, so we can now complete our table, which has been redrawn here:

| Carbon source | Energy source: Light | Energy source: Chemical |
|---|---|---|
| Carbon dioxide | Plants, algae and some protists, and cyanobacteria and green and purple sulfur bacteria<br>Photoautotrophs | Some bacteria and archaea (e.g. *Nitrobacter*, *Sulfolobus*)<br>Chemoautotrophs |
| Organic molecules | Some bacteria and archaea (e.g. AAP bacteria, and PR-containing bacteria and archaea)<br>Photoheterotrophs | Animals, fungi and some protists; decomposer and pathogenic bacteria<br>Chemoheterotrophs |

AAP = aerobic anoxygenic phototrophic; PR = proteorhodopsin (membrane proteins that operate as light-driven pumps)

Animals and plants each have just one lifestyle for capturing the energy and organic matter that they need for living. In contrast, the microbial world is remarkably diverse. This is not unexpected since microbes have been residents of planet Earth for billions of years and have evolved to exploit all sorts of habitats and energy sources.

# BIOCHALLENGE

## Aerobic respiration

1 Examine figure 4.51, which shows an overview of aerobic respiration.
  - a What is the other kind of cellular respiration?
  - b What are the stages of aerobic respiration?
  - c Where does each stage take place in a cell?
  - d By the end of which stage has a glucose molecule been completely broken down and fully oxidised to carbon dioxide?

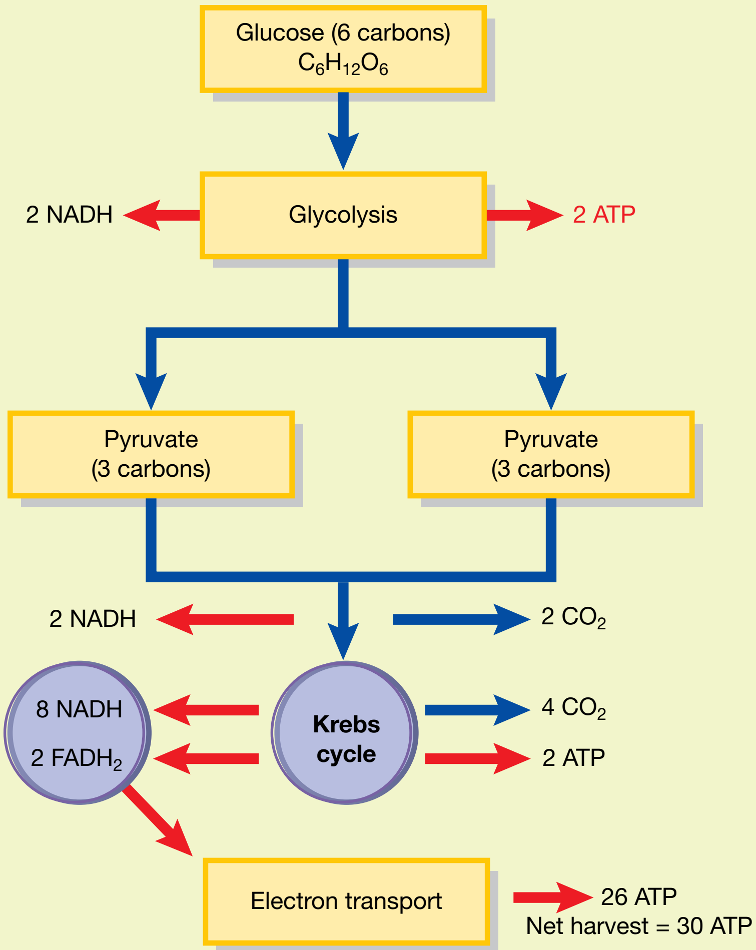



**FIGURE 4.51** Aerobic cellular respiration

2 Electron transport moves high-energy electrons produced in earlier stages of aerobic respiration and releases their energy, making it available for the production of ATP.
  - a How much ATP is theoretically produced as a result of electron transport?
  - b What percentage of the net total does this represent?
  - c What is the source of the high-energy electrons that are input to the electron transport chain?
  - d What is the final acceptor of electrons?
  - e What is produced when the electrons pass to the final acceptor?

3 The energy released from electron transport pushes $H^+$ (hydrogen ions or protons) from the matrix into the inter-membrane space. The protons diffuse back to the matrix through an ion channel in a protein complex called ATP synthase. The ATP synthase uses energy from this motion to generate ATP.
  - a What separates the matrix from the inter-membrane space?
  - b Write the equation that shows the formation of ATP from ADP.

4 Consider the information about the following compounds and identify the effect, if any on:
  - i the operation of the electron transport chain
  - ii the operation of ATP synthase.
  - a **Rotenone** is an insecticide that binds to complex I and prevents it from oxidising NADH to NAD.
  - b **Cyanide** binds to a cytochrome that is part of complex IV and prevents electron transfer.
  - c **Oligomycin** is an antibiotic that blocks the ion channel of ATP synthase.
  - d **DNP** (**dinitro phenol**) transports hydrogen ions/protons directly across mitochondrial membranes, bypassing the use of protein channels and reducing proton gradients.

# Chapter review

studyon
Unit 3 AOS 1 Topics 6&7
Photosynthesis
Cellular respiration
Practice questions

## Key words

acceptor molecules
ADP & $P_i$
aerobic respiration
alcoholic fermentation
anaerobic respiration
anoxic
ATP
autotrophic
bacteriochlorophyll
Calvin cycle
carbon fixation
chloroplast
cyanide
DNP (dinitro phenol)
ecosystem
endosymbiosis
facultative anaerobes
grana
granum
heterotrophic
Krebs cycle
lactate
lactic acid fermentation
limiting factor
MELiSSA (Micro-Ecological Life Support System Alternative)
$NADP^+$
NADPH
obligate aerobes
obligate anaerobes
oligomycin
organic compound
organic waste
photoautotroph
photosystem
producer
proton
rotenone
RuBisCo
stroma
thylakoid

## Questions

1 **Making connections between concepts** ➜ Use at least eight of the key words from this chapter to construct a concept map relating to photosynthesis and cellular respiration.

2 **Demonstrating knowledge and terminology** ➜ Examine the image below and identify the numbered structures.

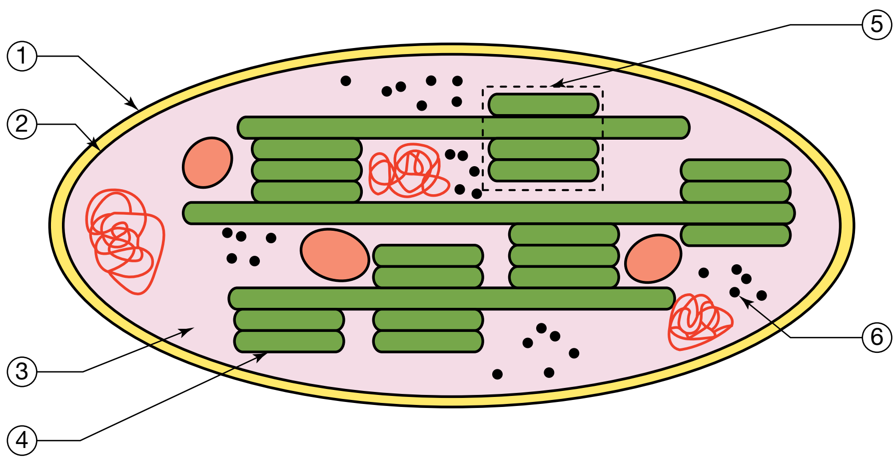



FIGURE 4.52

3 **Demonstrating knowledge and understanding** ➜ Put the following events in photosynthesis in order, from first to last:
- a generation of ATP
- b entry of carbon dioxide into leaf via stomata
- c trapping of sunlight energy
- d fixation of carbon
- e production of glucose
- f splitting of water molecules
- g release of 3-C molecules from the Calvin cycle.

4 **Analysing information and drawing conclusions** ➜ Live yeast was mixed with flour and water to make a dough. The dough was kept at a constant temperature of 30 °C and its volume measured every five minutes. The results are graphed in figure 4.53.
- a What was the increase in the volume of dough at the end of 30 minutes?
- b What was the percentage increase in the volume of dough during the 10-minute interval between the 30-minute and 40-minute stages of the experiment?
- c What kind of cellular respiration would be occurring in the yeast cells?
- d Explain the cause of the increase in volume of the dough.

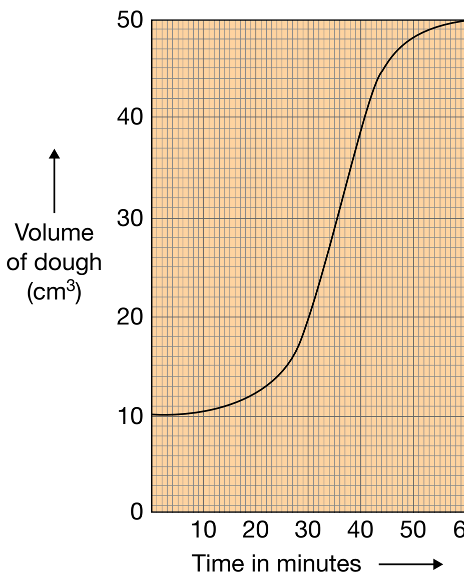



**FIGURE 4.53**

5 **Demonstrating understanding and communication** ➜ Student P said that the process of carbon fixation is 'the synthesis of ATP through the actions of the enzyme ATP synthase to produce useful energy.' Student Q said, 'No. It's the conversion of $CO_2$ to an organic compound.' Student R said, 'That's not right. I think that carbon fixation is the input to the Calvin cycle.'
  - a Which student(s) most accurately identified carbon fixation?
  - b Where you considered a student's description to be invalid, identify the inaccuracy involved.
  - c In which stage of photosynthesis does carbon fixation occur?
  - d Is the process of carbon fixation powered:
    (i) directly by light energy
    (ii) indirectly by light energy?
    Explain your decision.

6 **Demonstrating knowledge and understanding** ➜ Identify, as precisely as you can, the locations where the following processes or events occur in the chloroplast of a leaf cell:
  - a carbon fixation
  - b light trapping
  - c release of oxygen
  - d input of carbon dioxide
  - e reduction of carbon dioxide
  - f splitting of water.

7 **Demonstrating knowledge and understanding** ➜ You are given cell samples from five different sources and asked to match each to its correct source. You have access to light microscopes and electron microscopes to assist you.

The cells come from (1) unicellular green algae, (2) photosynthetic bacteria, (3) a human cell culture, (4) the spongy mesophyll of a hydrangea leaf, and (5) an obligate anaerobic bacterial species.

Your assistant views one of the cell samples using a light microscope and says, 'I'm certain I know what this one is!'
  - a Is your assistant likely to be correct? If so, suggest which cell sample was viewed and what was seen.
  - b Try to construct a dichotomous key that would allow you to distinguish with confidence between these five different cell types.

8 **Formulating and communicating valid biological explanations** ➜ Formulate a possible valid biological explanation for the following observations:
  - a Farmers managing crops growing in glasshouses increased their yields by extending the time period over which photosynthesis takes place. The farmers who chose to provide this light by burning paraffin oil lamps in their glasshouses obtained better results than the farmers who used standard electric-powered lighting. (Assume that the illuminance was the same in each situation.)
  - b The enzyme RuBisCo is regarded as the most abundant enzyme on this planet.
  - c A tomato grower increased the yield of his glasshouse-grown hydroponic tomatoes by using artificial lighting after daylight hours.
  - d Aquatic plants growing in an aquarium under bright lights were observed to produce bubbles.
  - e A cheetah chasing a gazelle has to stop its pursuit after a distance of about 100 metres.
  - f Workers entering a disused bulk oil tanker wear personal protection equipment.

9 **Interpreting data in graphical form and applying principles in a new context** ➜ Examine figure 4.54 below.

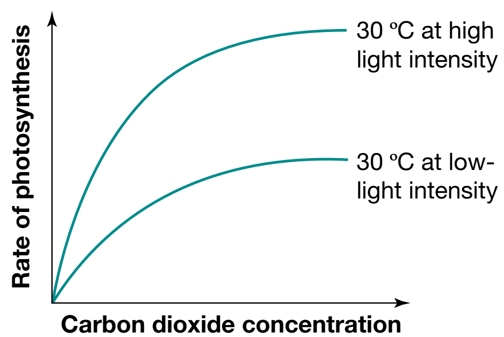



**FIGURE 4.54** Graph showing the rate of photosynthesis against carbon dioxide concentration in plant cells at two different light intensities.

a In which stage of photosynthesis is carbon dioxide used?
b What might have caused the rate of photosynthesis to plateau?
c Suggest why the rate of photosynthesis in the plant cells reaches higher values when illuminated by high-intensity light compared with that achieved under low-intensity lighting.

10 **Communicating knowledge and understanding** ➜ Consider a molecule of glucose ($C_6H_{12}O_6$) produced through photosynthesis in a leaf cell:
a What is the original source of the carbon atoms (C) in glucose?
b What is the original source of the hydrogen (H) atoms in glucose?
c How and where was this hydrogen produced?
d Which loaded carrier molecule transports hydrogen into the reactions involving reduction of glucose?
e Oxygen is present in carbon dioxide ($CO_2$) and oxygen is present in water ($H_2O$). From which source does the oxygen in glucose come? Explain.

11 **Applying principles in a new context** ➜ Consider a fish tank containing several fish and some aquatic plants.
Make a diagram showing one fish and one aquatic plant in a water-filled tank and showing the inputs and outputs related to cellular respiration and photosynthesis and any cycling that might occur:
a when the tank is brightly illuminated
b when the tank is in darkness.

12 **Demonstration analysis and communication** ➜ Consider two groups of plants growing under artificial light sources with identical illuminances of 5000 lux, but with different wavelengths:
Group 1. Wavelengths across the entire visible spectrum
Group 2. Wavelengths restricted to the red region of the spectrum.
a What prediction would you make about the rates of photosynthesis in each group? Explain your reasoning.
In the mitochondria of liver cells, the folds of the cristae increase the surface area of the inner membrane to about five times that of the outer membrane.
b What benefit, if any, might this increase in surface area bring?

13 **Evaluation of information and communicating understanding** ➜ Identify the following statements as true or false. Where you judge a statement to be false, rewrite it so that it would be judged to be true.
a Cellular respiration occurs all the time in animal cells but only in the dark in plant cells.
b Anaerobic respiration in mammals produces about half the amount of ATP compared to that produced in aerobic respiration.
c Carbon fixation is the breakdown of glucose to carbon dioxide.
d Endosymbiosis is a mutually beneficial relationship between two different organisms with one living inside the cells of the other.
e All light falling on a leaf is used in photosynthesis.
f The easiest way of measuring the rate of aerobic respiration in an animal would be to measure the rate of disappearance of glucose.

14 **Discussion question** ➜ Discuss with your group the validity or otherwise of the following statements:
a 'The purpose of lactic acid fermentation is to produce champion athletes.'
b 'The food we eat is a bit like packaged sunshine!'
c 'Methane is not a poisonous gas, so you cannot be hurt by exposure to it.'

CHAPTER

# 5 Cellular signals

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- recognise the various forms through which signals can be conveyed
- understand that chemical signals are the form through which cells most commonly communicate
- gain knowledge of cell-surface receptors and intracellular receptors
- understand signal transduction as it applies to hydrophilic and hydrophobic signals
- list examples of different kinds of chemical signals occurring in animal and plant cells
- become familiar with apoptosis as a process of programmed cell death, and its signal pathways.

**FIGURE 5.1** This zebra is not sniffing the wind but is displaying flehmen behaviour. Typically, the animal opens its mouth, curls its upper lip, holds its mouth open and draws air into its mouth. It is a way of detecting pheromones that perhaps signal a nearby male rival or a female ready to mate.

# Signals for communication

In our daily lives, communication is a constant component. Communication is variously defined as:

- the imparting of information by speaking, writing or some other medium (*Oxford English Dictionary*)
- the act or process of using words, sounds, signs, or behaviours to express or exchange information or to express your ideas, thoughts, feelings, etc., to someone else (Merriam-Webster).

We are constantly in communication, either receiving information from people and objects around us or imparting information to others through various channels.

Communication may involve a visual signal such as a red traffic light, an Instagram image, printed words in a textbook, an SMS message on an iPhone (see figure 5.2), a smile on a person's face, a hand gesture or the nod of a head. Communication may involve auditory signals, such as a spoken phrase: 'Don't do that!', a melody on an MP3 player, an announcement over a PA system or the screech of brakes. Communication may involve tactile signals, such as a tap on the shoulder, the heat of the handle of a saucepan on the stove, a hug, a handshake or the pain of a pinprick.

**FIGURE 5.2** A couple receiving a visual signal transmitted via an SMS message on a mobile phone. What might you infer about the visual signal from the signal given by the couple's facial expressions?

**ODD FACT**

The bitter taste receptor TAS2R38 can detect glucosinolates, a class of compounds with anti-thyroid activity. Genetically based differences in this receptor determine whether or not individuals perceive the bitter taste of foods (such as broccoli, watercress, kale and turnips) that contain glucosinolates.

Other communication can come via olfactory signals, such as the smell of food cooking, the smell of the smoke of a bushfire or the scent of a flower. Communication may also be transmitted in gustatory signals, such as the sweet taste of chocolate or, for *some* people, the bitter taste of broccoli (see Odd fact).

A signal has no effect unless it is received. Only after it is received can the signal be processed. Only after it is processed can a signal produce a response. Without the reception and the processing of a signal, no response to the signal can occur.

When we receive a signal, we process it and respond with a specific behaviour. The process by which a received signal is converted into a specific response is termed signal transduction. The driver of a car approaching an intersection who sees a red traffic light receives the visual signal and processes the information that the signal conveys as: 'I must stop!' The driver then responds by applying pressure to the foot brake, and the car comes to a stop just before the intersection. Note that it is *not* the signal that stops the car; it is the response triggered in the person who receives and processes this signal and takes relevant action.

The efficient functioning of cells depends on their ability to receive and respond to a variety of signals from the environment and from other cells. Cells are in constant communication, receiving signals from other cells and sending signals to other cells. This communication is sometimes called 'the constant chatter between cells.'

Just as we receive, process and respond to signals from our outside world, a similar process occurs in cells. The three components of cellular communication or 'cell chatter' may be summarised as:

1. **Reception** of a signal from a cell's external environment; such a signal may be a hormone or a neurotransmitter, a cytokine or a pheromone
2. **Transduction** of the signal into a form that can bring about a cellular response
3. Cellular **response** to the signal that comprises physical, biochemical or physiological changes brought about by effector proteins in the cell.

Figure 5.3 shows these processes as a stimulus-response model.

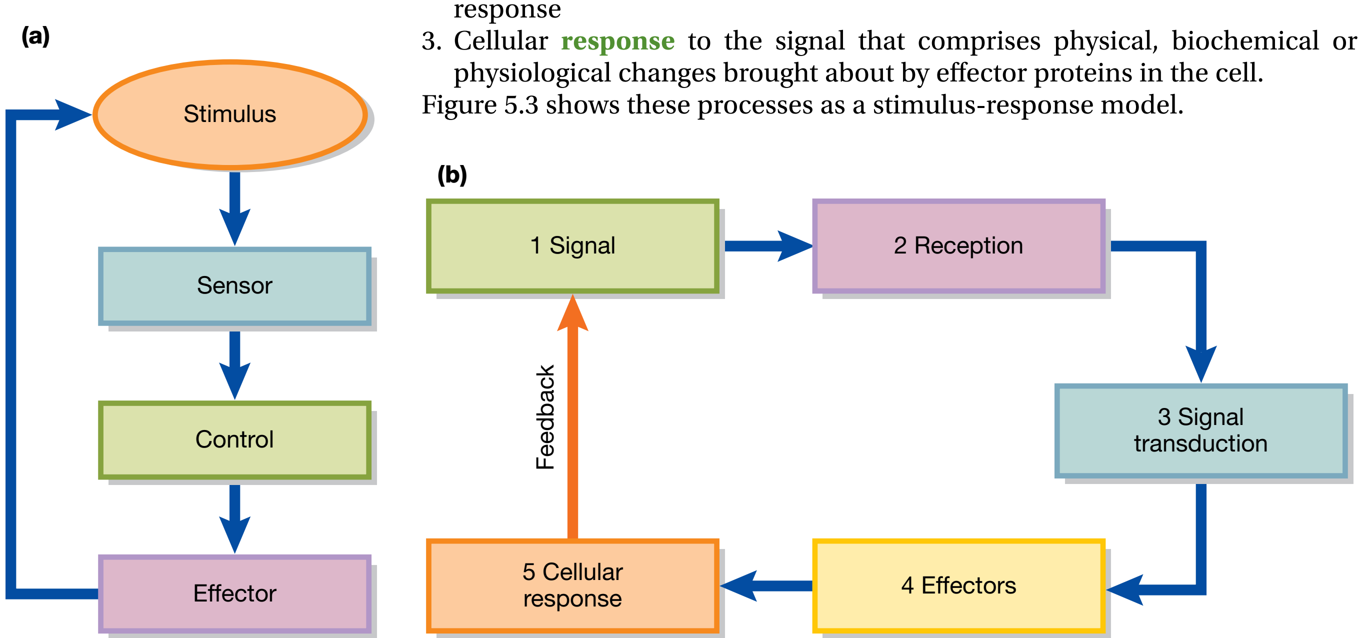



**FIGURE 5.3** **(a)** A generalised stimulus-response model. **(b)** Cellular communication shown as a stimulus-response model. The signal is like the stimulus in a stimulus-response model that is received and translated by effectors into a response. In the signalling process, a signal is received by a specific receptor at the cell surface and, within the cell, is converted via a signal transduction pathway into effector molecules that produce the cellular response. As in a stimulus-response model, feedback regulates the entire signalling process.

In this chapter, we will explore some of the features of this constant chatter between cells, with a focus on chemical signals, the most common type of signalling in cell communication.

## Communication between cells

**The most common signals that carry information between cells are chemical signals**. Other types of signals received by cells include mechanical signals, such as touch and pressure, and electromagnetic signals, such as light.

Signals can communicate different messages to cells, such as to stay alive and functioning, to self-destruct, to undergo cell division for growth or repair, to differentiate into a particular cell type, to eliminate a pathogen that has invaded the body, to activate a gene and produce the protein encoded by the gene, to silence a gene, or to produce an enzyme or a structural protein.

Let's first look at the various chemical signals involved in cellular communication.

## Signalling molecules

The most common signals are chemical molecules. Many different kinds of signalling molecules are involved in transmitting information between the cells of plants and between the cells of animals, and they come in a variety of shapes and sizes. Variation in signalling molecules includes:

- *differences in chemical structures*
  Molecules that act as chemical signals include amino acids, peptides, polypeptides, proteins and steroids
  - these differences mean that signalling molecules vary in size
  - these differences mean that signalling molecules differ in their affinity for water, some being hydrophilic and some hydrophobic.
- *differences in function*
  Signalling molecules belong to different functional groups: for example, animal hormones, plant hormones, neurotransmitters of nerve cells, antibodies of immune cells, cytokines and pheromones.
- *differences in scope*
  Some signalling molecules exert their effect over a distance while others only act locally.

More details of the various kinds of signal molecules are given throughout this chapter.

**Whatever their differences, all signal molecules share the same characteristic, namely: they bind to sites on specific receptors on their target cells. Receptors are proteins that receive the various signals, and a specific protein receptor exists for each different signalling molecule**. (Note that any molecule that binds to a specific target to form a biologically active complex is termed a **ligand**. In cells, common ligands are signalling molecules.)

## Signals move from source to target

A signalling molecule can only be effective if it reaches and binds to its specific receptor. The distance that a signalling molecule must travel to reach its targets can vary. Some signals are secreted by cells at a distance from their target cells, some are secreted by cells very close to their target cells, and some signals are produced and received by the same cells. The following sections identify an example of each situation.

1. **Long-distance travel to target cells:** Some signalling molecules must travel long distances to reach their target cells. Hormones secreted by various endocrine glands are the classic example of chemical signals that travel via the circulatory system over long distances to reach their target cells (see figure 5.4). For example, follicle stimulating hormone (FSH) that is secreted by cells in the pituitary gland travels via the blood to its target cells in the ovary.

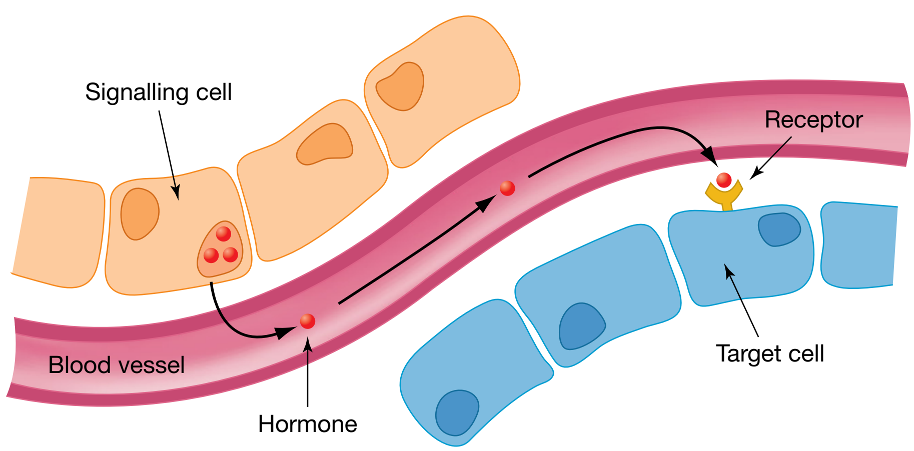



**FIGURE 5.4** Hormones (shown in red) are secreted by various endocrine glands. These chemical signalling molecules are carried by the bloodstream to target cells that are a distance away, where they are received by receptors (shown in yellow).

**ODD FACT**

Communication between two nerve cells is a situation in which a chemical signal released by one neuron needs to travel only a submicroscopic distance to reach its target, which is the next neuron. The distance involved is across a synaptic cleft, the gap that separates two nerve cells in a neural network. The width of a synaptic cleft is about 0.02 micrometre.

2. **Travel to nearby cells:** Some chemical signals are released by signalling cells and travel short distances to their nearby target cells. These signals move by diffusion through the interstitial fluid around cells (see figure 5.5). These signalling molecules are often called **local mediators**.

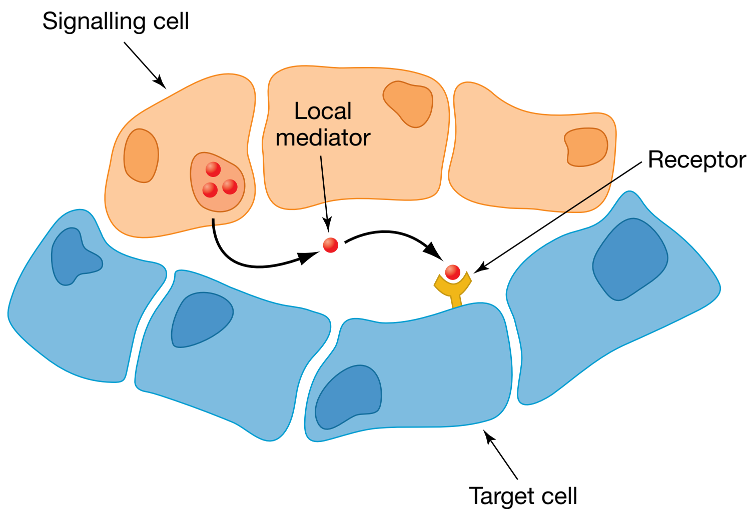
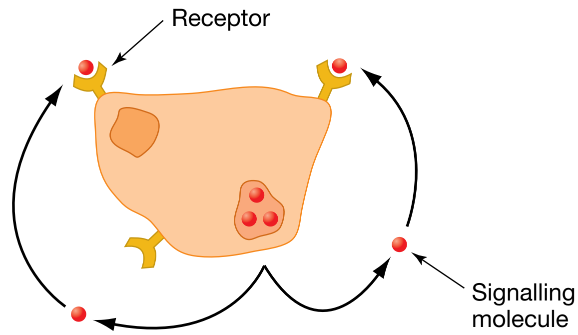



**FIGURE 5.5** Some signalling molecules (shown in red) diffuse over short distances from signalling cells to nearby local cells where the signal is received (receptors shown in yellow).

3. **One cell sends and receives a signal:** Chemical signals released by one cell may be received by the same cell (see figure 5.6). This can be seen in some immune cells, known as T cells. In the presence of a foreign antigen, one kind of T cell responds to the signal created by the foreign antigen by producing a growth factor (interleukin 2) that binds to receptors on the same type of cell and stimulates their replication. (This of course makes more T cells to attack the source of the foreign antigen.)

**FIGURE 5.6** In some cases, the signalling cell and the target cell are the same. Signalling molecules (shown in red) are produced and received by one type of cell that functions as both the signalling cell and the target cell.




4. **Direct cell-to-cell contact:** In some cases, a signal can move directly from the cytosol of one cell to that of another cell. This signalling involves direct cell-to-cell contact. Several structural features enable this direct contact between cells, as, for example:

- ***Gap junctions**, also known as communicating junctions, between adjacent cells in animal tissues*. In animal cells, gap junctions consist of protein-lined pores in the plasma membranes of adjacent cells (see figure 5.7a). Gap junctions allow various small molecules to pass between cells and also enable the transmission of electrical signals. For example, gap junctions between cells of your heart muscle enable the spread of an electrical impulse throughout the entire heart so that the cells 'beat as one'.

- ***Plasmodesmata** between adjacent cells in plant tissues*. Plant cells have relatively thick cell walls that lie outside their plasma membranes. Cell-to-cell communication in plants is achieved through gaps, known as plasmodesmata (singular: plasmodesma), through cell walls (see figure 5.7b). The plasma membranes of the cells are continuous and form the lining of the plasmosdesmata.

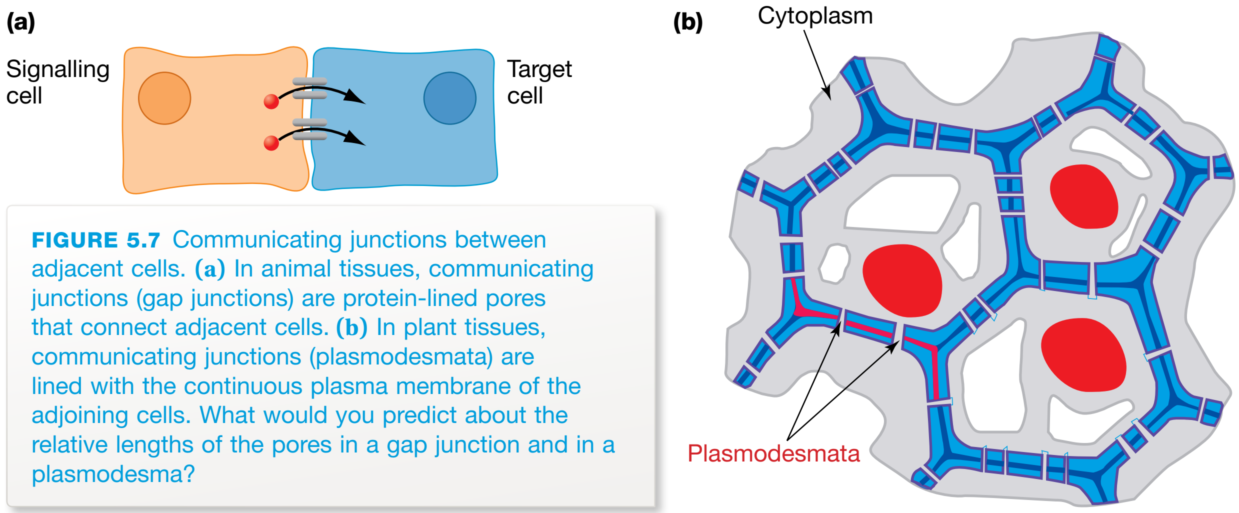



**FIGURE 5.7** Communicating junctions between adjacent cells. **(a)** In animal tissues, communicating junctions (gap junctions) are protein-lined pores that connect adjacent cells. **(b)** In plant tissues, communicating junctions (plasmodesmata) are lined with the continuous plasma membrane of the adjoining cells. What would you predict about the relative lengths of the pores in a gap junction and in a plasmodesma?

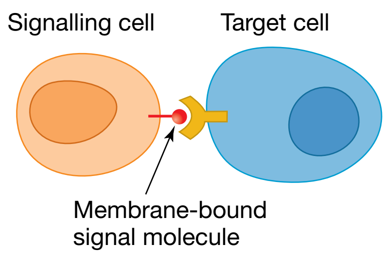



**FIGURE 5.8** Diagram showing a highly simplified version of cell-to-cell signalling where the signalling cell migrates to its target cell to deliver the signal.

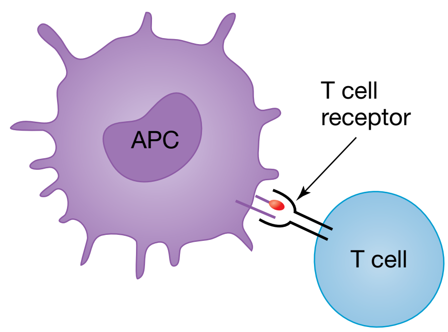
**Some signalling cells carry signals to target cells:** In some cases, signalling cells are highly mobile and deliver their signal to their target cells. Figure 5.8 shows a highly simplified version of this type of direct cell-to-cell signalling.

This type of cell-to-cell signalling occurs between cells of the body's immune system. Figure 5.9 shows a diagrammatic representation of cell signalling between two kinds of immune cells:

- The signalling cell (APC) carries, on its surface, a chemical signal (shown as a small red oval) that is a foreign antigen from part of a pathogen.
- The receiving cell is a T cell with a receptor (shown in black) for this signal.

The binding of the foreign antigen to the T cell receptor activates that cell so that it is ready to recognise and attack any cell it meets that carries the antigen.

**FIGURE 5.9** Diagram showing the cell-to-cell communication between a signalling cell (APC = antigen presenting cell) and its target T cell. Both of these cells are part of the human immune defence system. The signalling cell carries a signal molecule (shown in red), which is a foreign antigen. This signal molecule binds to the T cell receptor and activates that cell.




## KEY IDEAS

- Cells are in constant communication with other cells, both receiving and sending signals.
- Signals, most commonly, are chemical molecules, but they can be other environmental signals as well, such as light.
- Cellular communication comprises three stages: signal reception, signal transduction, and cellular response to the signal.

- Various kinds of signals may be received by cells, but the most common signals are chemical signals.
- The different chemical signals include animal hormones, plant hormones, cytokines, neurotransmitters, antibodies, and pheromones.
- Some chemical signals, such as mammalian hormones, exert their effect over a distance while others only act locally, yet others are transmitted by cell-to-cell contact.

## QUICK CHECK

1 What are the three stages of cellular communication?
2 Only certain cells respond to a chemical signal. What prevents all body cells from responding to a chemical signal?
3 Identify one key difference between:
   a a receptor cell and a target cell
   b a signalling molecule and a signalling cell
   c a gap junction and a plasmodesma.
4 By what main means do the following signals travel from their signalling cell to their target cell:
   a a mammalian hormone signal
   b a signal from a local mediator?

# Stages in 'cell chatter'

The three stages in cellular communication are signal reception, signal transduction and cellular response. Figure 5.10 shows a simple diagrammatic representation of this process. In the following sections we will explore each of these stages in turn.

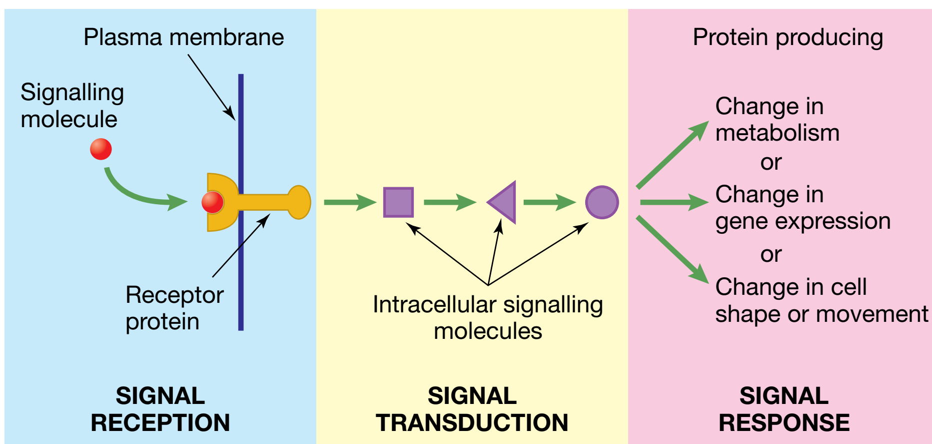



**FIGURE 5.10** Diagram showing a simple representation of the three stages in cellular communication.

**Cell signalling** stages:
1. **Signal reception**
2. **Signal transduction**
3. **Cellular response**

## Signal reception

**The first step of cellular communication is the reception of a signalling molecule (ligand) from a cell's external environment.** When signalling molecules (ligands) reach their target cells, they bind to a specific receptor.

Receptors are proteins and they are highly specific: each type of receptor protein binds to one signalling molecule with a specific shape. This specificity is achieved because the ligand and the binding site of its receptor have complementary shapes that fit together, just like a lock that will only fit a certain key. So, many different receptors exist.

Consider just the receptors concerned with the sense of smell, the so-called odorant receptors. A clue to the number of different odorant receptors came from the research of Richard Axel and Linda B. Buck, who asked the question: *How many different receptors do we have for the different odours, such as the perfume of a rose, the smell of rotting fish or burning gum leaves, that we can detect in our environment?*

The 2004 Nobel Prize in Medicine or Physiology was awarded to these researchers for their discovery of '*odorant receptors and the organization of the olfactory system*'. These researchers identified a large gene family comprising about 1000 different genes that encode an equivalent number of different olfactory receptors. These olfactory receptors are located on cells in the upper lining of the nose. Unexpectedly, it was discovered that each olfactory cell has just one kind of olfactory receptor on its surface.

**eBook*plus***

**Weblink**
Odorant receptors research

In most cases, the receptors involved in cell signalling are located on the plasma membrane of a specific target cell — such receptors are called **cell-surface receptors**. However, in other cases, the receptors are located either in the cytosol or in the nucleus of the target cell — such receptors are called **intracellular receptors**. Why are there two different locations for signal receptors?

**studyon**

| | |
|---|---|
| Unit 3 | **Receptors** |
| AOS 2 | Summary screen and practice questions |
| Topic 1 | |
| Concept 6 | |

We noted (refer back to page 162) that signalling molecules differ in their chemical structure and size. Some signalling molecules are small **polar** molecules, such epinephrine (adrenalin), other signalling molecules are larger **hydrophilic** molecules, such as the protein hormones, while yet others are **hydrophobic** molecules, such as the steroid hormones. Steroids are lipids, and they have four fused hydrocarbon rings in their chemical structures. The various steroid hormones are distinguished by the different functional groups attached to these rings (see figure 5.11), and these molecules are hydrophobic, and hence are lipophilic.

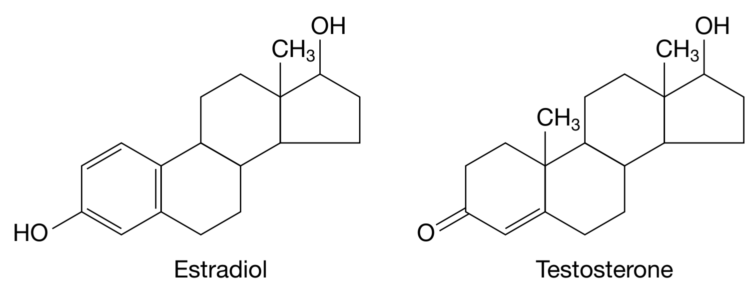



**FIGURE 5.11** The structures of two steroid hormones. Note the four hydrocarbon rings. How does estradiol differ from testosterone?

**Polar and hydrophilic signalling molecules cannot cross the lipid bilayer of the plasma membrane.** These signalling molecules can only bind to receptors at the cell surface. The binding site is that part of the receptor protein that is exposed to the extracellular environment (see figure 5.12a). All cell-surface receptors are either trans-membrane proteins on their own, or they operate as part of a complex that links the exterior and interior of the cell. One such complex receptor is the G protein coupled receptor.

In contrast, hydrophobic signalling molecules can readily cross the lipid bilayer of the plasma membrane. Hydrophobic signalling molecules, such as the steroid hormones, bind to receptors that are located within the target cell, either in the cytosol or in the nucleus (see figure 5.12b).

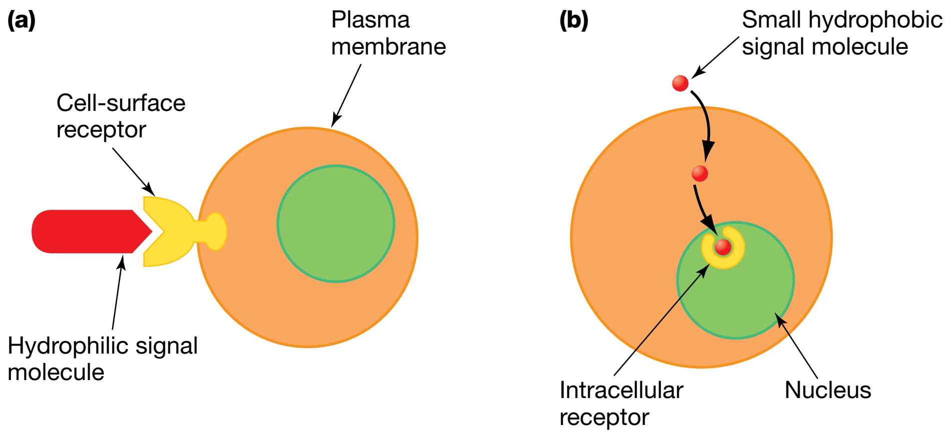



**FIGURE 5.12** **(a)** Hydrophilic signalling molecules can only bind to cell-surface receptors. **(b)** Hydrophobic signalling molecules that can readily pass across the lipid bilayer of the plasma membrane bind to intracellular receptors located either in the nucleus or in the cytosol. Can you name a hydrophobic signalling molecule?

The difference between signalling molecules in their affinity to water means that the receptors for hydrophilic signalling molecules are cell-surface receptors, while receptors for hydrophobic signalling molecules are intracellular receptors.

Look at the cell-surface receptor in figure 5.12a. Receptors of this type are trans-membrane proteins that are embedded in the plasma membrane. Each trans-membrane receptor has several regions or domains. One region is exposed to the outside of the cell — this is the **signal-binding domain**. Another region is in contact with the cytoplasm — this is the intracellular domain. (Trans-membrane receptors also have a hydrophobic domain. Where do you think the region of such a trans-membrane receptor might be located?)

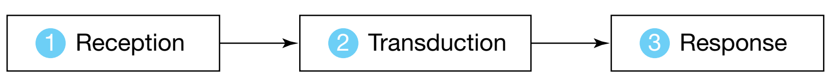
When a signalling molecule binds to its receptor, the 3D shape of the receptor is altered, activating the receptor and indicating 'Message received!' However, for cell-surface receptors, there is a problem: the message is at the cell surface. It's bit like an unread message on a smart phone: C U @ 8 2NITE:-) Unless the message is read and its contents acted upon, the communication has failed. The same applies to a signalling molecule at the cell surface. The control centre of the cell, the nucleus, is where the message must arrive. The nucleus is where the cellular response to the signal will be made, typically involving the gene activation and the production of effector proteins.

Getting the message from the cell surface to the nucleus is achieved through signal transduction, the second step in cell communication.

**Cell signalling** stages:
1. Signal reception
2. **Signal transduction**
3. Cellular response

## Signal transduction

The second stage of cell signalling is termed **signal transduction**, a process that converts a signal from outside a cell into a response within the cell. In signal transduction, a signal is received in one form, is changed to another mode or molecule, and is relayed to the appropriate target within the cell that responds through an effector protein. The signal transduction pathway links the signal reception to the cellular response to the signal:




The process of signal transduction starts after a signalling molecule binds to its specific receptor, changing its 3D shape and activating it.

We have seen in the previous section that some signalling molecules are hydrophilic (and hence lipophobic): for example, a peptide hormone such as insulin and a protein hormone such as growth hormone. Other signalling molecules are hydrophobic (and hence lipophilic): for example, the steroid hormones such as estrogen and testosterone. **The hydrophilic or the**

**hydrophobic nature of signalling molecules determines the location of the specific receptor where the signal is received.**

Let's look at signal transduction for each type of signalling molecule, starting with the transduction of a lipophilic signalling molecule such as a steroid hormone. This transduction process is less complex than that for a hydrophilic signalling molecule.

studyon

| | |
|---|---|
| Unit 3 | **Signal transduction** |
| AOS 2 | Summary screen and practice questions |
| Topic 1 | |
| Concept 7 | |

### Transduction of a hydrophobic signal:

Steroid hormones are hydrophobic and can readily cross the lipid bilayer of the plasma membrane by a process of diffusion. Once in the cell, the steroid hormone binds with its specific receptor. The receptor is a protein in solution either in the cytosol (as shown in figure 5.13) or in the fluid component of the nucleus.

The transduction pathway is as follows:

1. The binding of the steroid hormone to its specific receptor produces a change in the 3D shape of the receptor protein, exposing a region of the receptor that was previously within the molecule.
2. The hormone-receptor protein complex moves from the cytoplasm into the nucleus.
3. The exposed segment of the receptor protein attaches to a target DNA sequence next to certain genes, and activates those genes.

Figure 5.13 shows a diagram of this process. Note that steroid hormones directly regulate gene expression.

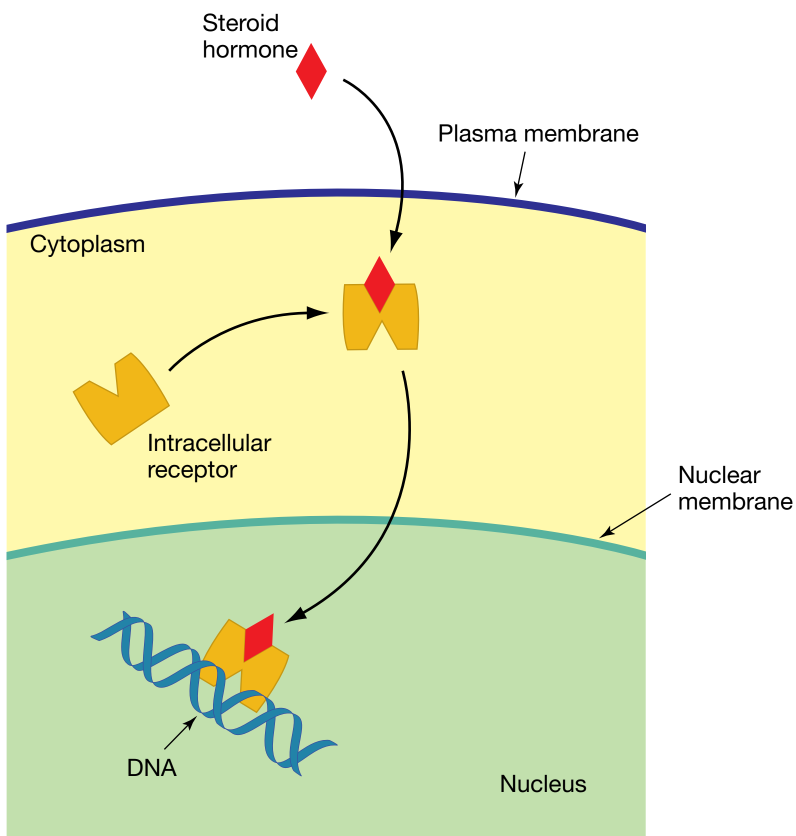



**FIGURE 5.13** Transduction of the signal from a hydrophobic signalling molecule, in this case a steroid hormone, is a very direct process.

### Transduction of a hydrophilic signal:

Now, let's look at the transduction of the signal from a peptide hormone that cannot cross the lipid bilayer of the plasma membrane. The process of signal

Proteins that can attach to DNA and influence levels of transcription are called **transcription factors**.

## ODD FACT

The second messengers for plant hormone signals are often calcium ions ($Ca^{++}$).

**study on**

Unit 3
AOS 2
Topic 1
Concept 7

**See more**
Signal transduction

transduction starts when the hormone binds to its specific receptor on the cell surface. The signalling molecule in this case is called the '**first messenger**' and, if its message remains on the cell surface, it will be ineffective. The message must reach the nucleus where the cellular response to the signal will occur. For this to happen, the signal must be transferred from the receptor to a '**second messenger**' within the cytoplasm.

**The transfer of the signal from receptor to nucleus occurs through a signal transduction pathway within the cytoplasm that involves production of large numbers of molecules of a second messenger**. This amplifies the original signal by a factor of up to 10 000. Second messengers are small molecules, for example, cAMP, and they are *not* enzymes, but they activate a key enzyme at the start of an enzyme relay.

The enzyme relay is not like a relay race in which a baton is passed from one person to the next. In the enzyme relay of signal transduction, the signal is amplified because, as each enzyme is activated, it activates multiple copies of the next enzyme in the relay, and so on. Molecules of the last enzyme in the relay to be activated finally carry the signal to the nucleus where the gene activation occurs.

Figure 5.14 provides a simplified version of the sequence of events in a signal transduction pathway.

1. The binding of the peptide hormone to its cell-surface receptor causes a change in the 3D shape of the receptor protein that activates it.
2. The receptor in turn activates an enzyme embedded in the inner surface of the plasma membrane.
3. This activated enzyme in turn catalyses the production of multiple copies of a **second messenger** (for example, cyclic AMP (cAMP)). This step amplifies the signal, because the binding of one external signalling molecule results in the production of multiple copies of second messenger molecules.
4. The second messenger molecules activate copies of a key enzyme (protein kinase) that, in turn, activate the next enzyme in the relay, and so on. This step further amplifies the signal.
5. This signal is finally relayed to the nucleus, where specific genes are activated and effector proteins are produced.

Note that signals from hydrophilic hormones indirectly bring about gene action, while signals from hydrophobic hormones directly activate genes.

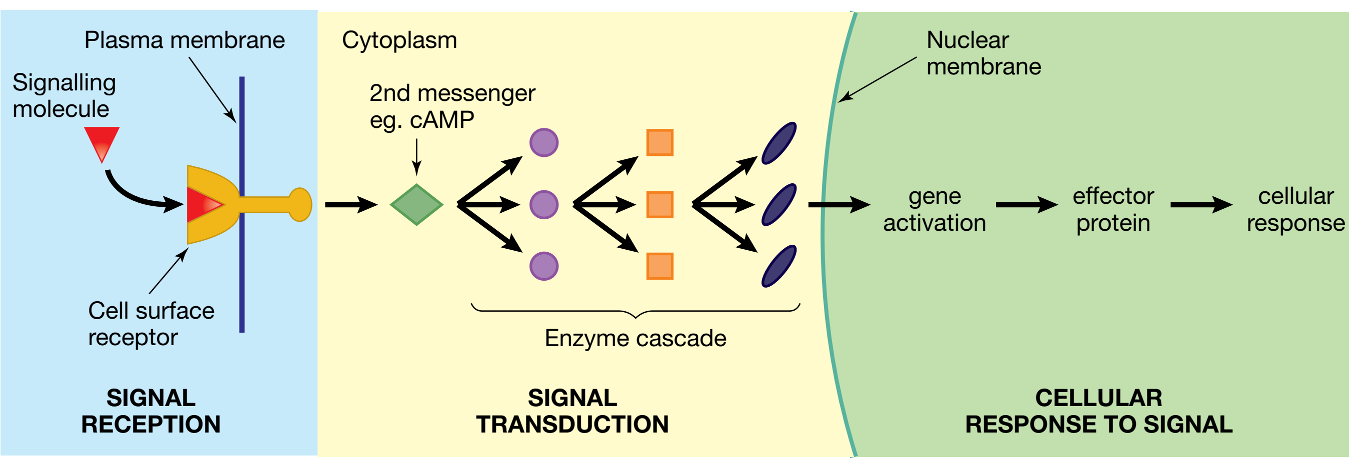



**FIGURE 5.14** Diagram showing a simplified version of the signal transduction pathway for a peptide hormone. The final cellular response differs, depending on the identity of the original signalling molecule.

Table 5.1 summarises the differences between the hydrophilic and the lipophilic hormones.

**TABLE 5.1** A summary of key differences between different groups of hormones, depending on their different affinities (love or hate) for water molecules.

| Steroid hormones | Peptide/protein hormones |
|---|---|
| hydrophobic: insoluble in water | hydrophilic: soluble in water |
| lipophilic: soluble in lipid solvent | lipophobic: insoluble in lipid solvents |
| transported in blood but only with the aid of carrier proteins | transported in solution in blood plasma |
| readily diffuse across the lipid bilayer of the plasma membrane | unable to cross the lipid bilayer of the plasma membranes |
| bind to intracellular receptors | bind to cell-surface receptors |
| directly regulate gene expression | indirectly act on genes |
| no second messenger involved | second messenger produced during signal transduction |
| longer lasting response | shorter period of response |

Figure 5.15 shows a comparison of signal transduction in lipid-soluble and water-soluble hormones.

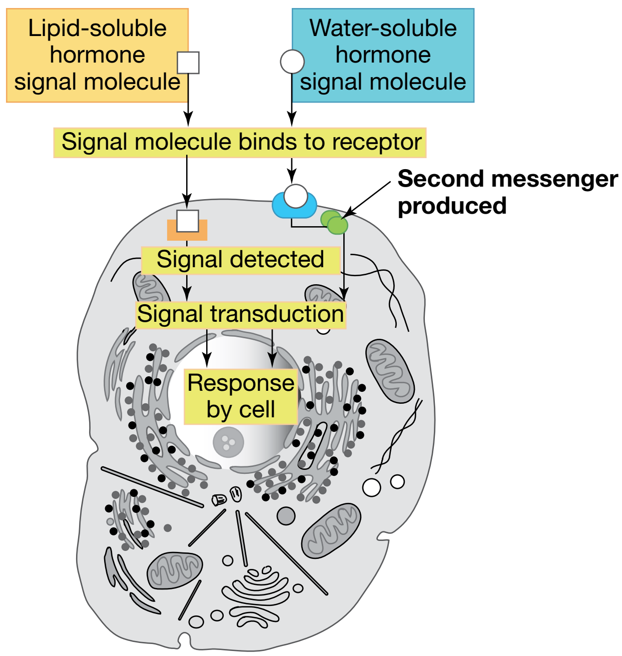



**FIGURE 5.15** A comparison of signal transduction in lipid-soluble and water-soluble hormones, showing the location of events within a cell.

## Cellular response to signal

The cellular response to the signal is the final stage in cellular signalling. Effector proteins are produced by gene activity stimulated directly by steroid hormones, or indirectly by peptide or protein hormones. These effector proteins produce the cellular responses to the original external signal molecule.

These cellular responses to different signals are extremely varied and may include:

- cell migration entailing the production of structural proteins, such as the microtubules of the cytoskeleton
- changes in metabolism that involve either the production of specific enzymes or a regulated increase in enzyme activity

**Cell signalling** stages:
1. Signal reception
2. Signal transduction
3. **Cellular response**

studyon

Unit 3 | AOS 2 | Topic 1 | Concept 8

**Cellular responses**
Summary screen and practice questions

- cell division involving DNA synthesis and the action of enzymes such as DNA polymerase
- apoptosis or programmed cell death involving the production of caspase enzymes.

Finally, cellular signalling pathways, once activated, do not stay switched 'on'. After a limited period of time, the system returns to its resting state, ready to receive and respond to the next signal.

**KEY IDEAS**

- The three stages in cellular communication are signal reception, signal transduction and cellular response.
- The nature of signalling molecules, either hydrophilic or hydrophobic, determines the location of the receptor where the signal is received.
- Hydrophilic signalling molecules are received by cell-surface receptors, while the receptors for hydrophobic signalling molecules are in the cytoplasm or the nucleus.
- Signal transduction of a hydrophilic signalling molecule involves a series of chemical reactions during which the signal is transferred and amplified before it reaches the cell nucleus.
- The cellular response involves changes in gene activity and the production of effector proteins that bring about the particular cellular response to the signal.

**QUICK CHECK**

**5** List the three stages involved in cellular communication.
**6** Identify the following statements as true or false:
  **a** Hormones are an example of signals that originate at a distance from the receiving cell.
  **b** Receptors for hydrophilic signalling molecules are located at the cell surface.
  **c** Steroid hormones are examples of hydrophobic signalling molecules.
  **d** Cellular response to signals involves effector proteins that bring about specific changes in the cell.
  **e** The receptor for a peptide hormone would be expected to be located on the surface of the receiving cell.
  **f** During signal transduction, the original signal is greatly amplified.
**7** Identify one key difference between:
  **a** a hydrophilic and a hydrophobic signalling molecule
  **b** the first messenger and the second messenger.
**8** Give an example of a cellular response to a signal.

# Chemical signals in animals

Earlier in this chapter, we met some chemicals that are signalling molecules in various animal cells. Among these were:

- hormones
- neurotransmitters.

In this section, we will explore these and other chemical signals that are involved in cellular communication in the animal world.

## Chemical signals: human hormones

Figure 5.16 shows the human endocrine system and the major hormones secreted by each endocrine gland.

The human hormones can be organised into three groups, as follows:

- **amino acid derivatives** — these hormones are derived from amino acids; they dissolve readily in water and therefore are hydrophilic. The names of hormones in this group typically end with the suffix *-ine*, as, for example, thyroxine and epinephrine.

- **lipid-derived hormones** — these hormones are derived from lipids; they are not water-soluble and therefore are hydrophobic. The names of hormones in this group typically end with the suffix *-ol* or *-one*, as, for example, testosterone and cortisol.
- **peptide and protein hormones** — some of these hormones are composed of shorter polypeptide chains; these are the peptide hormones, such as oxytocin and insulin. Other hormones in this group are made of long polypeptide chains — these are the protein hormones, such as growth hormone. These hormones are hydrophilic.

These groups are colour-coded in this figure. (Refer back to table 5.1 to check some of the differences between the lipid-derived hormones and the other groups of hormones.)

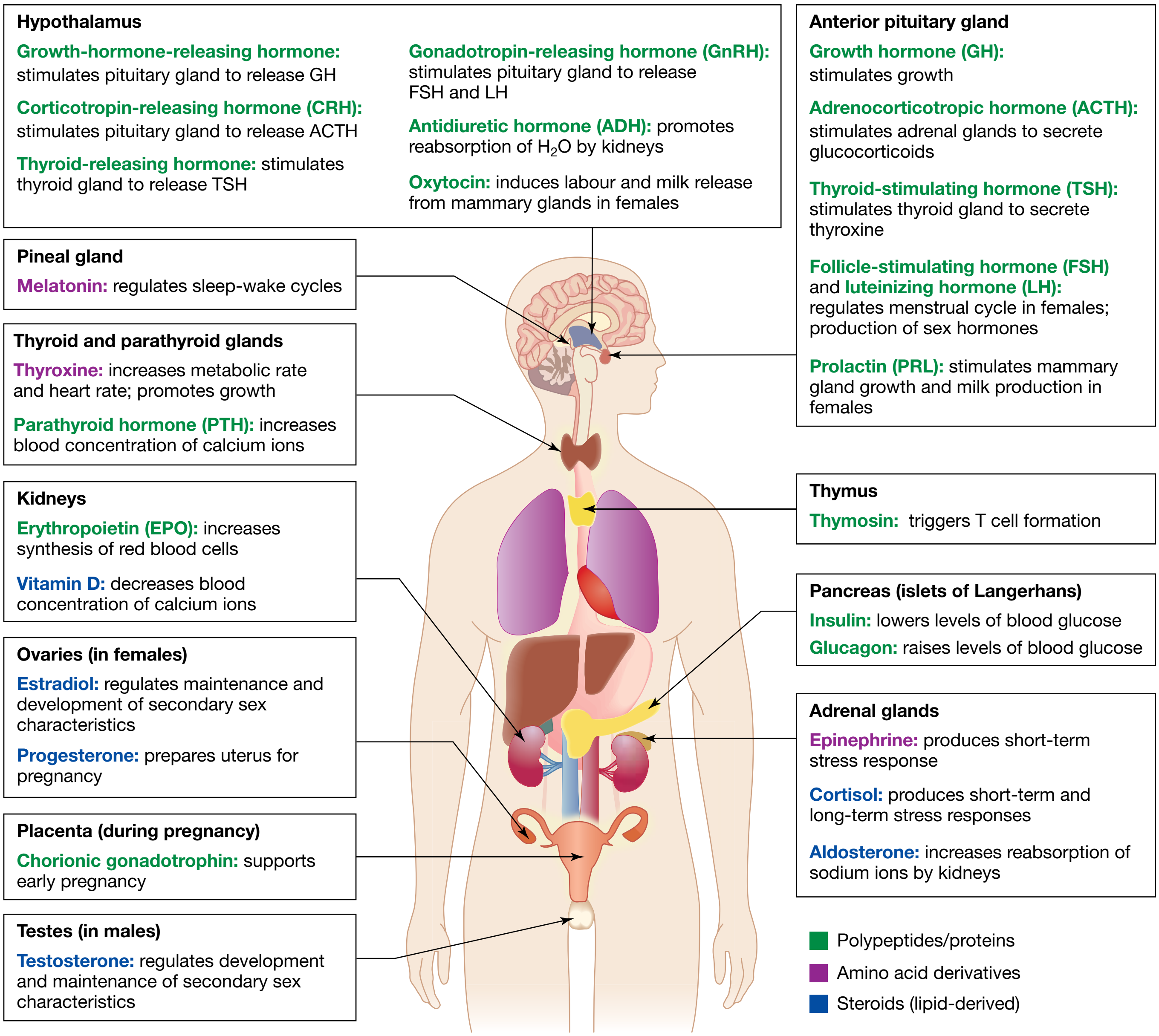



**FIGURE 5.16** The endocrine organs and the major hormone signalling molecules that they produce. Note the three major chemical classes into which the various hormones fall. Based on this figure, which class is the most common?

studyon

Unit 3
AOS 2
Topic 1
Concept 2

**See more**
How steroid hormones work

Hydrophilic hormones are transported in solution in the blood from their site of production to their specific targets, while hydrophobic hormones are also transported in the blood but they are attached to protein carriers.

Hormones act as signalling molecules that carry messages to their specific target cells:

- The hormone prolactin carries the message 'Produce milk' to its target cells in the mammary glands.
- The hormone insulin carries the message 'Take up glucose and convert it to glycogen' to target cells in the liver.
- The hormone vitamin D carries the message 'Absorb calcium' to its target cells in the small intestine.

In addition to the hormones produced by the endocrine glands, many other tissues of the human body also produce hormones. Some of these are listed in table 5.2 below.

studyon

Unit 3
AOS 2
Topic 1
Concept 2

**Hormones**
Summary screen and practice questions

**TABLE 5.2** Examples of hormones produced by organs and tissues other than the endocrine glands.

| Tissue/organ | Hormone | Major action |
|---|---|---|
| adipose tissue | leptin | suppresses appetite |
| skin | 1,25-dihydroxy vitamin D | stimulates uptake of $Ca^{2+}$ from small intestine |
| stomach | gastrin | stimulates acid secretion |
| heart | atrial natriuretic hormone | acts on kidney, promoting excretion of $Na^+$ ions |
| small intestine | cholecystokinin | inhibits gastric motility and stimulates secretion of bile and pancreatic juice |

The release of hormones operates within a negative feedback system. Figure 5.17 shows the signals that are exchanged between the hypothalamus, the anterior pituitary and the peripheral organs. Note the negative feedback loops, shown as red arrows, which ensure that the biological effects produced by the hormones are maintained within narrow limits. (This is the concept of homeostasis that is discussed in *Nature of Biology Book 1 Fifth Edition,* chapter 6).

Refer to this figure and follow the signal pathway that occurs if the body's metabolic rate falls to levels below normal limits:

1. The hypothalamus increases the release of thyroid releasing hormone (TRH).
2. TRH signals the anterior pituitary gland to increase the release of thyroid stimulating hormone (TSH).
3. TSH levels signal its target gland, the thyroid gland, to increase its release of the hormone thyroxine.
4. Increased thyroxine travels throughout the body, signalling changes in various target organs, and the resulting cellular responses produce an increase in the metabolic rate.

The increase in metabolic rate is kept within normal limits by the operation of two negative feedback loops as follows:

- The increased level of TSH gives negative feedback to the hypothalamus, stopping the release of TSH.
- The increased level of thyroxine gives negative feedback to both the anterior pituitary gland and the hypothalamus.

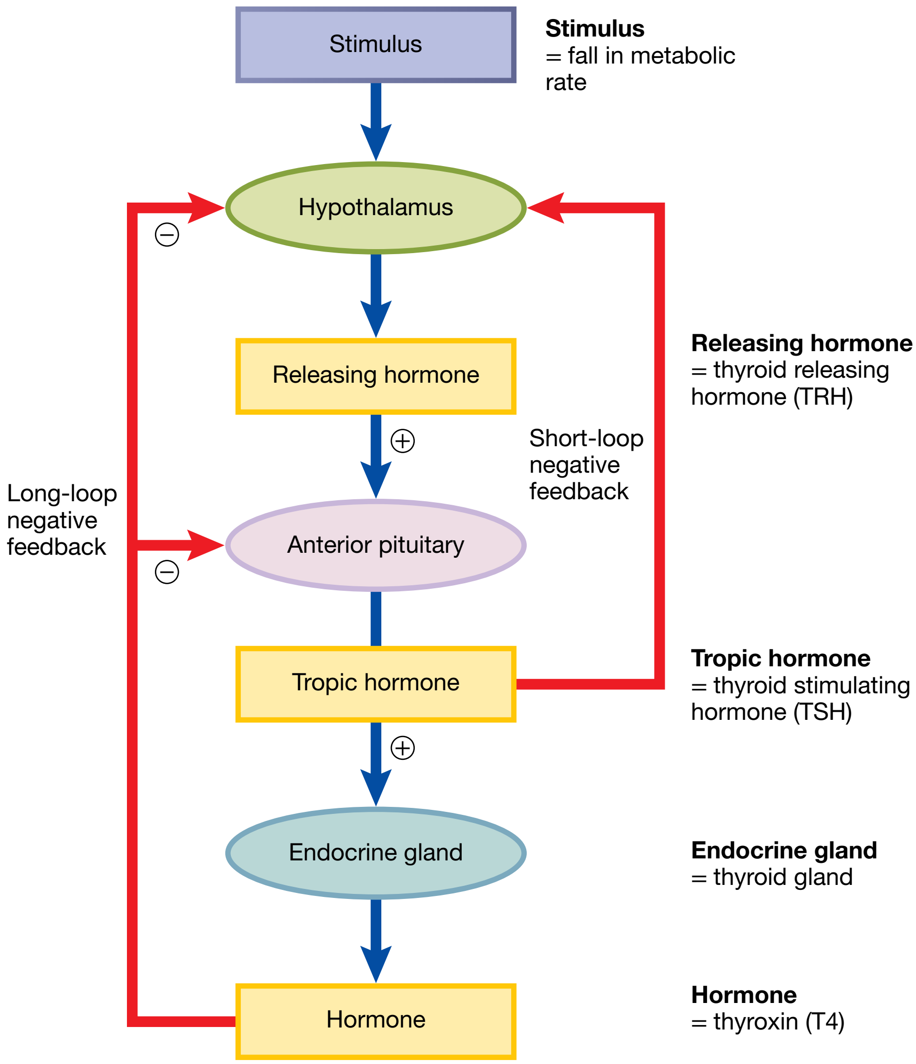



**FIGURE 5.17** The action of hormones typically involves negative feedback loops (shown as red arrows), ensuring that cellular responses to hormone signals are maintained within narrow physiological limits.

## KEY IDEAS

- **The hormones of the human endocrine system are a major group of signalling molecules.**
- **Hormone signals travel to their target cells via the bloodstream.**
- **In addition to the major hormones produced by the endocrine glands, other tissues and organs also produce hormones.**
- **The cellular responses to hormone signalling molecules are maintained within narrow limits principally through the operation of negative feedback loops.**
- **A hormone can communicate with cells only if the cells have receptors to receive the specific signal from that hormone.**

## QUICK CHECK

**9** Identify the target cells and organs of the following hormones:
- **a** insulin
- **b** prolactin.

**10** What is the cellular response to the hormonal signal of thyroxine?

**11** How do hormonal signals reach their targets?

studyon

Unit 3 | AOS 2 | Topic 1 | Concept 1

**Neurotransmitters**
Summary screen and practice questions

## Chemical signals: neurotransmitters

Neurotransmitters are the chemical signalling molecules of nerve cells (neurons).

Neurons use two types of signals:

1. Communication *within* a nerve cell depends on electrical signals, known as action potentials, that transmit a nerve impulse along the axon of a neuron. Action potentials are produced by the local movements of charged particles (sodium ions into the neuron and potassium ions out of the neuron). The electrical signal box at the end of this section provides a brief outline of how the electrical signal is produced.
2. Communication *between* nerve cells uses chemical signals, known as **neurotransmitters**, that diffuse across very small gaps, known as **synaptic clefts**, that separate one neuron from the next in a neural network.

Figure 5.18 shows these two types of signals — electrical signals that move a nerve impulse along the axon of a neuron, and chemical signals (neurotransmitters) that transmit the impulse from one neuron to the next.

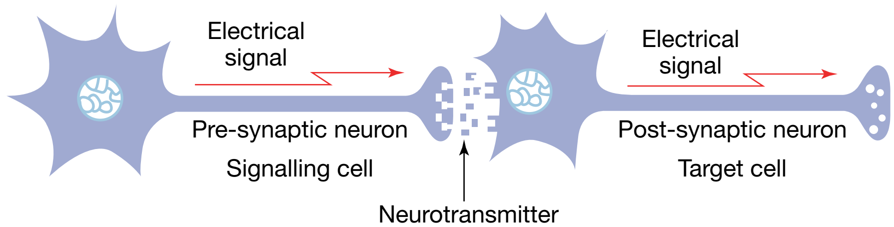



**FIGURE 5.18** The movement of a signal in nerves involves an electrical signal that travels *within* each neuron. A chemical signal, known as a neurotransmitter, transmits the signal across the synaptic cleft that separates the pre-synaptic neuron from the post-synaptic neuron.

In addition to carrying signals from one neuron to another, neurotransmitters transmit nerve impulses from neurons to muscle cells, stimulating their contraction, and to some glands.

Neurotransmitters include chemicals such as acetylcholine, epinephrine (adrenalin) and serotonin; these are excitatory neurotransmitters. Another neurotransmitter, gamma amino butyric acid (GABA), is an inhibitory neurotransmitter. Figure 5.19 shows one possible outcome of the action of excitatory and inhibitory neurotransmitters on target cells supplied by the branches of the axon of one neuron.

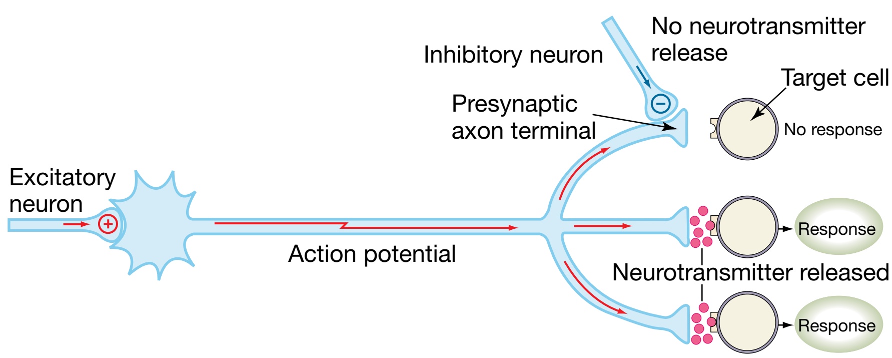



**FIGURE 5.19** One possible arrangement of excitatory and inhibitory inputs to neurons, and the result in three target cells. What neurotransmitter might be involved at the excitatory junction? At the inhibitory junction?

## ELECTRICAL SIGNAL WITHIN A NEURON

A resting nerve cell maintains a charge difference, known as the **resting membrane potential**, between the inside and the outside of the neuron. This difference is maintained in part by a channel in the plasma membrane of the neuron that is known as the $Na^+/K^+$ pump. The $Na^+/K^+$ pump actively transports sodium ions ($Na^+$) out of the cells and potassium ions ($K^+$) into the cells. From the energy supplied by one ATP molecule, three sodium ions are pumped out of the cell but only two potassium ions are pumped into it. This produces a net excess of positive charge outside the cell and a net negative charge of the cytosol inside the cell (see figure 5.20).

When neurotransmitter molecules bind to their receptors on the surface of the post-synaptic neuron, the situation changes. The receptors are activated, causing channels to open that allow large numbers of sodium ions ($Na^+$) to rush into the neuron (see figure 5.21a). The inflow of sodium ions causes a local **depolarisation** of the plasma membrane so that, at this point, the inside of the cell temporarily has a positive charge and the outside has a negative charge (see figure 5.21b). The region of depolarisation advances as an electrical signal along the axon carrying the nerve impulse. (A moving region of depolarisation is called an action potential.)

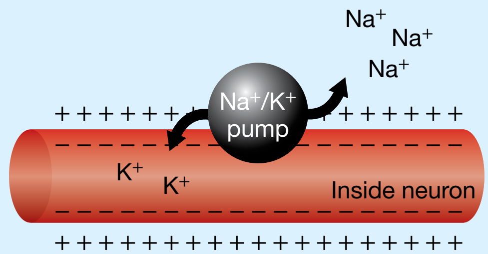



**FIGURE 5.20** Diagram showing part of a resting nerve cell. Note that the cytosol inside the cell has a negative charge relative to the outside of the cell, which has a positive charge. What maintains this charge difference in a resting neuron?

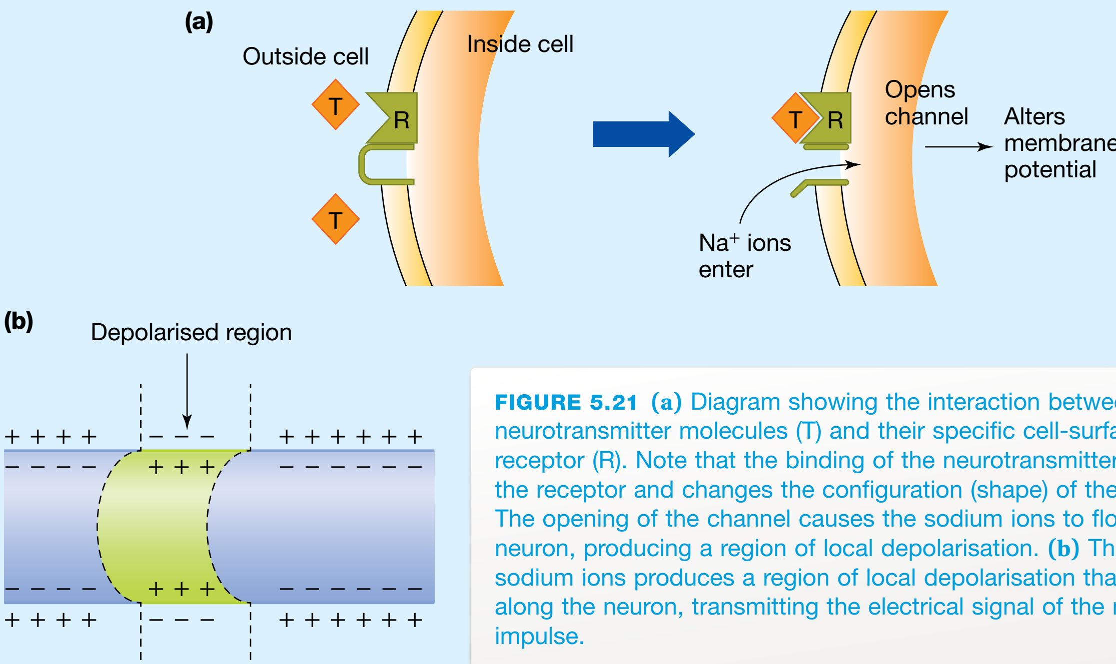



**FIGURE 5.21** **(a)** Diagram showing the interaction between neurotransmitter molecules (T) and their specific cell-surface receptor (R). Note that the binding of the neurotransmitter activates the receptor and changes the configuration (shape) of the channel. The opening of the channel causes the sodium ions to flow into the neuron, producing a region of local depolarisation. **(b)** The inflow of sodium ions produces a region of local depolarisation that advances along the neuron, transmitting the electrical signal of the nerve impulse.

### What happens at a junction between two neurons?

Look at figure 5.22, which shows a **synapse**, the term given to the junction of two neurons and the synaptic cleft that separates them. When a neuron is stimulated, it transmits a nerve impulse in the form of an electrical signal along its axon. At the end of the axon are packages or vesicles that contain neurotransmitters. The impulse is transmitted across the synaptic cleft by neurotransmitter molecules released from vesicles in the pre-synaptic neuron.

The neurotransmitter molecules diffuse across the synaptic cleft and bind with receptors on the surface of the post-synaptic neuron. Nerve impulses are transmitted in a one-way direction only across a synaptic cleft. Can you identify why the movement of the nerve impulse occurs in one direction only? Note the presence of an enzyme that destroys the neurotransmitter, ensuring that the nerve impulse has a limited duration.

studyon

Unit 3 | AOS 2 | Topic 1 | Concept 1

**Do More** Neurotransmitters and synapses

studyon

Unit 3 | AOS 2 | Topic 1 | Concept 1

**See More** Neurotransmitters

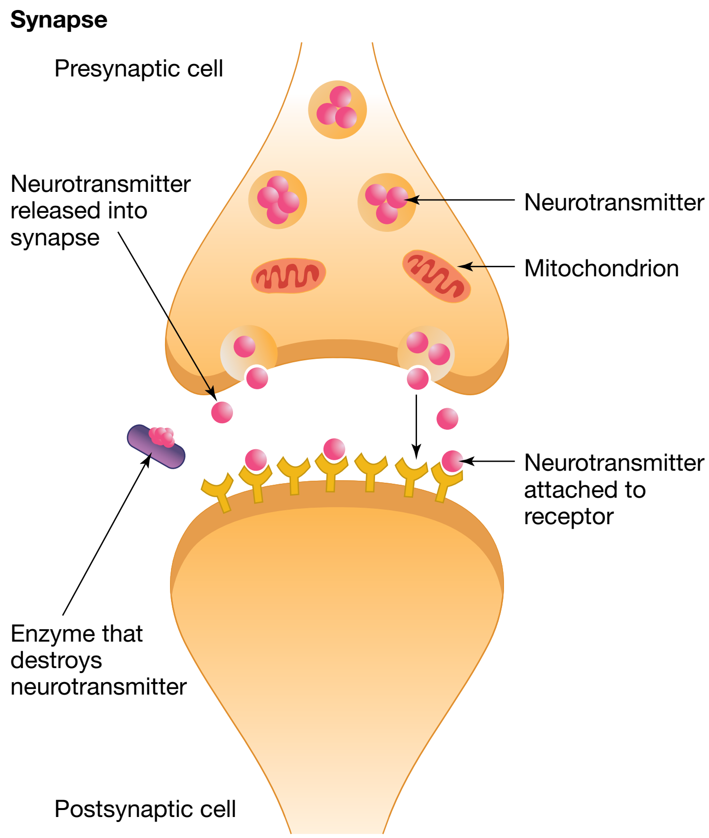



**FIGURE 5.22** Diagram showing the junction of two neurons, which forms a structure called a synapse. Note the vesicles that contain neurotransmitter molecules, the chemical signals that transmit the nerve impulse across the synaptic cleft.

Motor nerves also send impulses to muscles. Every involuntary muscle movement in your body, such as the peristaltic movement of your gut, involves nerve impulses to smooth muscle fibres in the gut wall. Every contraction of your skeletal muscles occurring when you turn the page of a book, take a step, turn your head, kick a football or throw a basketball requires that the muscles concerned receive stimulatory impulses from motor neurons.

At a neuro-muscular junction, a neurotransmitter is released from the motor end plates at the terminal ends of an axon. This neurotransmitter diffuses across the small gap between the axon and the muscle, and binds to receptor molecules on the membrane of the striated muscle (see figure 5.23). The cellular response contracts the muscle. What stops the action of transmitter substances on muscle? Muscle tissue produces an enzyme that inactivates the transmitter substance. Further action can occur only if more transmitter substance is released from the nerve motor end plate.

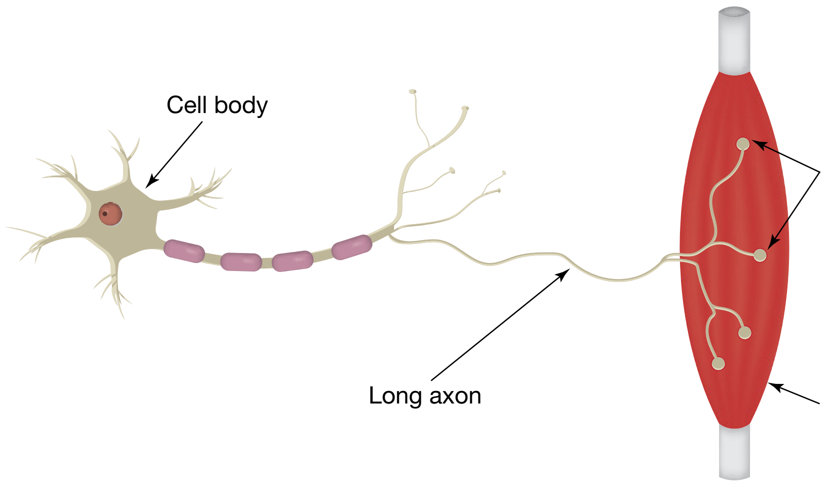



**FIGURE 5.23** Simplified diagram showing a motor neuron with a long axon that extends from the cell body and divides into many small branches at its terminus. At the end of each branch is a **motor end plate** that forms a junction with muscle cells. Neurotransmitter is released from vesicles in the motor end plate.

The critical importance of nerve impulses to voluntary movement was tragically apparent during past epidemics of poliomyelitis (polio). Polio is an infectious disease caused by the poliovirus. Major epidemics of polio occurred in Australia in the late 1930s, early 1940s and the 1950s, with the last epidemic occurring in 1956.

After gaining entry to the body, typically through the mouth, poliovirus may reach the central nervous system. If this occurs, the virus replicates in motor neurons, resulting in damage to and death of these cells. Motor neuron cells control the muscles of the trunk, legs and arms, and the muscles involved in swallowing and breathing. When breathing muscles were affected, a patient was placed in a tank respirator, better known as an iron lung, which did the mechanical work of breathing (see figure 5.24). Fortunately, mass vaccination has made poliomyelitis a disease of the past.

**ODD FACT**

In tissue culture, replication of poliovirus has been shown to produce 10 000 to 100 000 virus particles per cell in a period of about eight hours.

**FIGURE 5.24** Historical photograph showing a victim of polio inside an Emerson respirator, also known as an 'iron lung' machine. Patients stayed in the iron lung until they were able to breathe independently.

## KEY IDEAS

- Neurotransmitters are signalling molecules that enable communication between nerve cells.
- Communication from one neuron to another occurs when neurotransmitters are released from the pre-synaptic nerve cell and diffuse across the synaptic cleft to the post-synaptic nerve cell.
- Within a nerve cell, the signal moves as an electrical impulse.
- Communication between motor neurons and striated muscle cells is mediated by neurotransmitters released at motor end plates at the ends of the nerve axons.

## QUICK CHECK

**12** Identify the following as true or false:
   **a** Neurotransmitters can travel in either direction across a synaptic cleft.
   **b** Adjacent neurons are separated by a gap termed the synaptic cleft.
**13** Identify an example of a neurotransmitter.
**14** Where would you expect to find the following:
   **a** a synapse
   **b** a motor end plate
   **c** a neurotransmitter?
**15** What stops a muscle from continuing to contract after it receives a signal to contract?

### ODD FACT

The first pheromone was chemically identified in 1959. This was the chemical bombykol, a sex attractant that is released by female silkworm moths (*Bombyx mori*).

### ODD FACT

Ants use the chemical signal of a pheromone to mark a food trail. In contrast, bee species use the visual signal of a precise and elaborate dance for the same purpose.

FIGURE 5.25

## Chemical signals: pheromones

The dog sniffing around the lamppost, the two lemurs wiping secretions from their wrist glands onto their tails and waving their tails at each other, the ants forming a procession along a particular trail — these animals are receiving signals from and/or transmitting signals to other members of their species. The signal in each of these cases is a chemical signalling molecule called a pheromone.

**Pheromones** (*phero-* = to carry plus *-mone* from hor**mone**) are produced by a variety of animal species. As our understanding of pheromones has increased, so has the evidence that more and more animal species appear to communicate to each other using pheromones to convey particular messages.

Hormones and neurotransmitters are chemical signalling molecules that are involved in cellular communication within one organism. In contrast, pheromones are chemical signalling molecules released by one animal that can carry a signal to a second member of the *same* species (a conspecific) and, if this pheromone signal is received, the second animal produces a physiological or behavioural change.

The messages carried by different pheromones vary and include alarm signals ('Danger, predator approaching'), food trail signals ('Head this way to the food'), territorial signals ('This is my turf, so stay away!') and sexual attractant signals ('I'm ready to mate. Come find me!'). Communication by pheromones can be seen as being beneficial to the survival of individual animals and also to the survival of a population. For example, the individual worker ant who finds a source of food and lays a pheromone trail from the food source back to the ant colony provides a significant benefit to all members of the colony. The energy of the worker ants can then be concentrated on gathering food rather than on many non-productive searches for food.

### Detecting pheromones

Pheromones are secreted by an animal into the external environment. Highly volatile pheromones disperse rapidly, some travelling on air currents over great distances, such as insect sex pheromones. Others, such as the territorial

**FIGURE 5.26** The flehmen behaviour seen in many mammals involves drawing air into the mouth from where it can reach the vomeronasal organ.

pheromones of mammals that mark the perimeter of a territory, are less volatile and remain more localised.

A particular behaviour called the **flehmen** response may be seen in cats, dogs and ungulates, when they are investigating the urine, the faeces or the genital region of a conspecific, either up close or from a short distance. Figure 5.26 shows a horse displaying flehmen behaviour. Typically, the animal opens its mouth, curls its upper lip, holds its mouth open and draws air into its mouth. Flehmen behaviour in some animals may also involve licking and tongue flicking. A similar behaviour may be seen in dogs and cats. What is the purpose of this behaviour?

In addition to the olfactory system that detects smells and odours, mammals have a second system that detects pheromones. The detection of pheromone signals involves a different nerve pathway from the pathway for the detection of odours and smells. Can you identify a possible advantage of having a separate system for the detection of pheromones?

A structure called the **vomeronasal organ (VNO)** contains the detector cells of this second system. When an animal takes air into its mouth as part of the flehmen response, the animal is directing that air to the VNO. The VNO is located in the front portion of the nasal cavity. If cells detect pheromone molecules, a signal is relayed via several regions of the brain to the hypothalamus (H) (see figure 5.27). The hypothalamus is the brain centre where physiological and reproductive responses to pheromone signals are regulated.

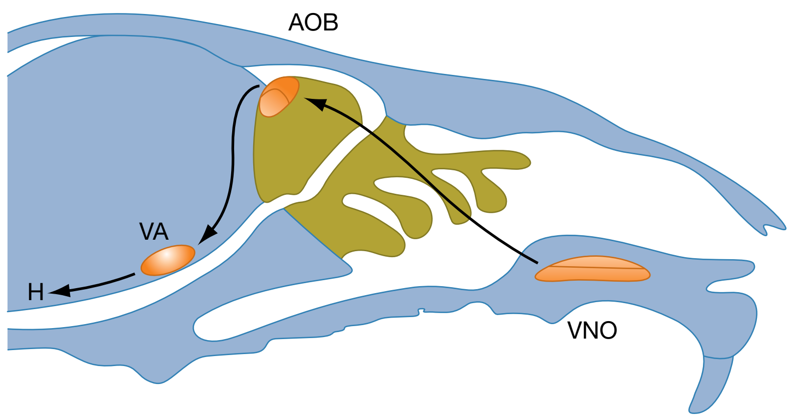



**FIGURE 5.27** Diagram showing a longitudinal section of a dog skull. Note the vomeronasal organ (VNO), where sensory neurons detect pheromone signals. The signal is first transferred by axons of VNO cells to a region of the brain (AOB = accessory olfactory bulb). From there, the signal is relayed in two steps to the hypothalamus (H) of the brain. The hypothalamus plays a major role in regulating physiological or reproductive responses to the pheromone signal. (It is not surprising that the signal travels to the hypothalamus — refer back to figure 5.17 to check on the range of hormones secreted by the hypothalamus.)

**studyon**

| | |
|---|---|
| Unit 3 | **Pheromones** |
| AOS 2 | Summary screen and practice questions |
| Topic 1 | |
| Concept 3 | |

**ODD FACT**

More than 4000 species of aphid have been identified, and many are serious pests in agriculture and forestry.

## Pheromones in the animal world

Some examples of pheromones in the animal world include:

**Trail pheromones:** Many insect species, in particular ant species, use pheromones to mark out a trail from a food source. While the food source remains, the trail continues to be renewed by pheromones released by more ants travelling to and from the food. The trail continues to be marked out as long as the food source exists. Would you predict that the pheromone involved would be a chemical that remained in place for an extended period?

**FIGURE 5.28** Wingless aphids on a plant stem. Note the paired cornicles near the rear of the abdomen.

**Alarm pheromones:** When some insect species are attacked by predators, they release alarm signals. Included among these species are aphids. Look at figure 5.28 and note the paired tubular structures, called cornicles, near the end of the aphid abdomen. Cornicles are present on most aphid species and it is from the cornicles that aphids release their alarm pheromones when attacked.

**Territorial pheromones:** Many mammals, including lions, tigers, dogs and wolves, define the perimeters of their territories by spraying urine on particular landmarks. Present in the urine are pheromones that proclaim the 'owner' of the territory to a passing member of the same species. Other mammals, including species of wildebeest and deer, mark their territories using strong-smelling secretions from scent glands on various parts of the animal's body. Ringtail possums (*Pseudocheirus peregrinus*), an Australian native marsupial species, also mark their territory with scent.

**Sex pheromones:** Females of many insect species, in particular species of moth and butterfly, release pheromones signalling that the females are ready to mate. These pheromones drift downwind in the air and can attract males from up to ten kilometres away. A female moth emits pheromones from brush-like scent organs that protrude from the distal end of her abdomen (see figure 5.29a). Male moths (see figure 5.29b) have large feathered antennae with specialised hair-like sensory receptors to detect pheromones released by a conspecific female moth. The sensory cells on the antennae of male moths can distinguish minute traces of pheromones, even a few molecules, in the presence of many odours. Would you predict that sex attractant pheromones are species specific?

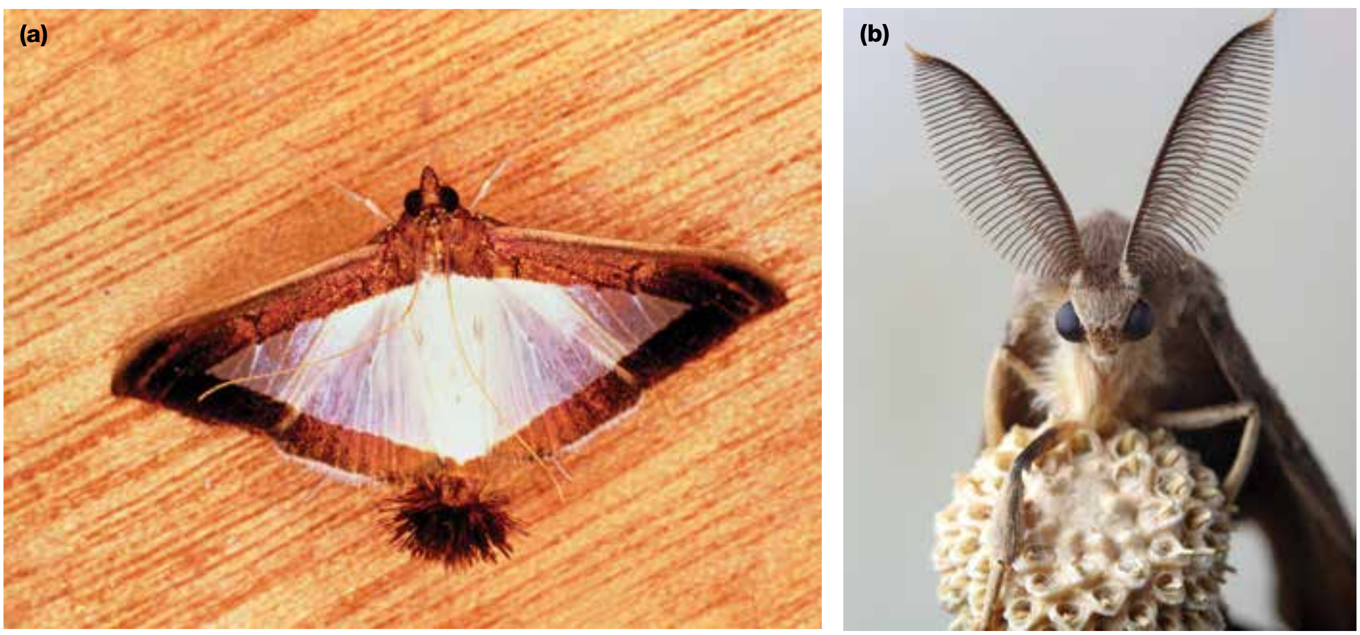



**FIGURE 5.29** **(a)** A female moth (*Estigmene acraea*) releases pheromones from brush-like scent organs at the distal tip of her abdomen. **(b)** A male gypsy moth (*Lymantria dispar*) with his feathered antennae that he uses to detect sex pheromones released by a female gypsy moth.

## PHEROMONES IN AGRICULTURE

### Making use of pheromones

Pheromones are used to reduce insect pests in both glasshouse crops and field crops. Artificial chemicals that mimic particular pest-insect pheromones are used. Insect traps, baited with artificial hormones, attract male pest insects into the traps from which they cannot escape. In a confined space such as a glasshouse (see figure 5.30a), such a strategy allows informed release of other insects for biological control. The insects chosen for the biological control prey on the pest insect but otherwise have no impact on the crop.

In a field crop or for orchards, traps can be used as early warning signals (see figure 5.30b). They indicate that adult insects are emerging and give some idea about the size of the pest population. This enables insecticide to be sprayed only when and where it is needed.


**FIGURE 5.30** **(a)** Pheromone traps used in glasshouses detect the presence of pest insects. An appropriate 'non-pest' or 'crop-friendly' species can be used in biological control against the pest insect. **(b)** A pheromone-baited trap in an orchard. The artificial pheromone used here mimics that of the codling moth, *Carpocapsa pomonella*, an important pest of apple and pear orchards.

## KEY IDEAS

- Pheromones are chemical signalling molecules for communication between members of one animal species.
- Pheromones convey many different kinds of messages, including alarm signals and territorial signals.
- Pheromones are used in various ways to reduce pest infestation of plant crops.
- Artificial molecules can be made that mimic the effects of natural pheromones.

## QUICK CHECK

16 Identify two functionally different kinds of pheromone.
17 What is the difference between a hormone and a pheromone?
18 By what means do male moths detect the pheromone signals produced by females of the same species?
19 What is the flehmen response?

## Chemical signals: cytokines

**Cytokines** are a diverse category of signalling molecules — more than 200 have been identified. Many cytokines are peptides, but some are proteins and some are glycoproteins. The major cells that secrete cytokines are the immune cells (white blood cells) of mammals, but they are also produced by other cell types. Cytokines are '*cell signalling molecules that aid cell-to-cell communication in immune responses and stimulate the movement of cells towards sites of inflammation, infection and trauma*' (Source: News Medical).

**Cytokines act as messengers between cells of the immune system**, just like the hormones of the endocrine system communicate with various cells throughout the body. Figure 5.31 shows a diagrammatic representation of cytokine-mediated communication between immune cells.

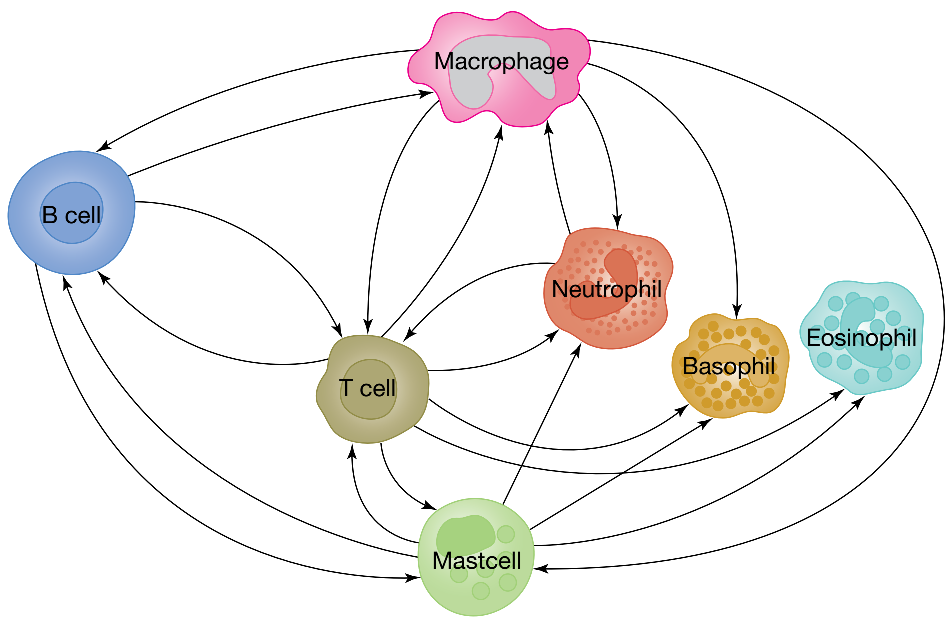



**FIGURE 5.31** Diagram showing some of the communication between immune cells that is mediated by cytokine signalling molecules.

Cytokines have a large number of effects — stimulatory, inhibitory and regulatory — on surrounding cells, with the specific effect depending on the cytokine involved and the target cell. Cytokines are examples of chemical **mediators**, that is, intracellular proteins that enhance and activate the functions of other proteins. So, we can say: 'Cytokines mediate a number of cell processes including immunity, inflammation, and blood cell formation.' Cytokines are also described as **modulators**, that is, molecules that directly influence the effects of other molecules. So, we can say: 'Some cytokines are immune-modulators', meaning that cytokines directly influence immune responses, such as inflammation.

Table 5.3 identifies the major actions of the major subgroups of cytokines. These subgroups are named according to a major function of the subgroup, as, for example, interleukins (*inter* = between; *leuko* = white) are produced by white blood cells and act mainly on other white blood cells.

Cytokines are secreted *by cells* in response to various stimuli, and they usually act briefly, locally, and at very low concentrations. However, cells that produce cytokines do *not* hold stores of pre-formed cytokines. Instead, cytokines must be produced anew in response to a stimulus, such as the presence of a pathogen. Because of this need to synthesise new cytokines, cellular responses to cytokine signals are slow, and may take hours to appear.

**ODD FACT**

Peptide hormones are important for cell-to-cell communication in plants. One researcher has recently proposed that these peptide hormones in plants be called plant cytokines or phytocytokines

In chapter 7, you will read about some of these various cytokines in action in innate immune responses, including inflammation.

**TABLE 5.3** Major actions of various subgroups of cytokines.

| Subgroup | Major actions |
|---|---|
| interleukins (IL-1 to IL-35) | • stimulate growth and differentiation of white blood cells<br>• promote inflammation (pro-inflammatory) |
| interferons (IFRs) | • act as antiviral agents<br>• interfere with replication of viruses |
| colony-stimulating factors (CSFs) | • signal stem cells in bone marrow to differentiate into various kinds of blood cells (see the box on CSFs in chapter 6) |
| tumour necrosis factors (TNFs) | • promote inflammation<br>• activate some immune cells<br>• high levels of TNF are seen in cancer patients with wasting disease (cachexia)<br>• induce apoptosis of tumour cells |

Once cytokines are produced and released from the signalling cell, the cytokines diffuse to nearby target cells where each binds to a specific cell-surface receptor. The binding of a cytokine to its specific receptor triggers a signal transduction pathway in the cytoplasm. This process involves the production of second messenger molecules (such as cAMP) and enzyme activation. The signal is finally transduced to a specific cell response that is brought about by changes in gene transcription, in which cytokine-related genes are switched 'on' or 'off'.

The cellular responses produced in response to cytokine signalling include promotion of cell growth and differentiation, cell proliferation, cell migration, cell activation, apoptosis, as well as immune responses such as inflammation and phagocytosis.

One cytokine can provide different signals to various cells. Figure 5.32 shows the effects of interleukin-2 (IL-2), a cytokine that is released by one kind of immune cell. IL-2 acts as a signal to other immune cells, but the cellular response to the cytokine signal varies in different kinds of cells.

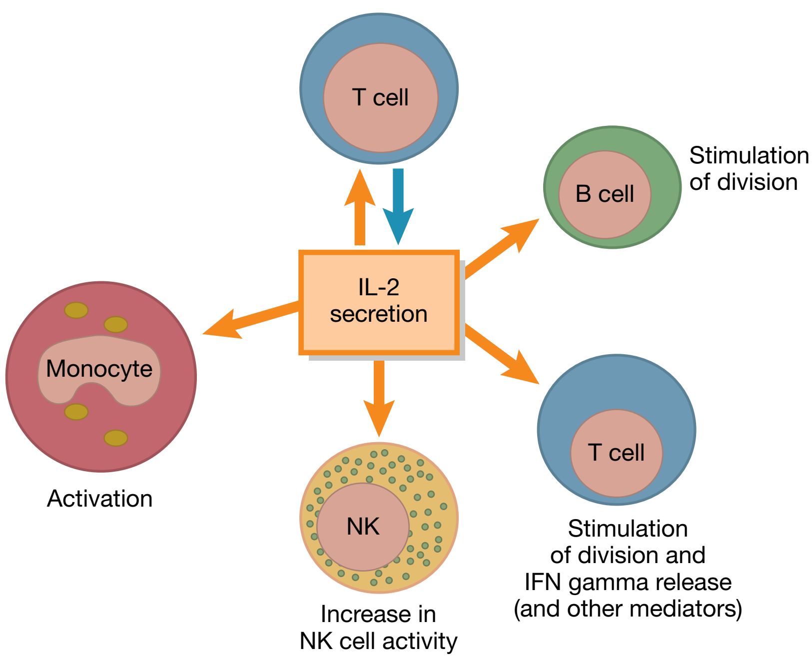



**FIGURE 5.32** The same cytokine, interleukin 2 (IL-2), produces different cellular responses in different types of immune cells. Note that one of the actions includes the release of IFN (interferon). What is the cellular response of B cells to the IL-2 signal? (You will meet these and other immune cells in chapters 7 and 8.)

### Cytokines in human disease

**Several cytokines act as mediators in the normal inflammatory response to infection**. For example, one cytokine, tumour necrosis factor (TNF), regulates the production of several other pro-inflammatory molecules, including several interleukins. However, overproduction or overstimulation of pro-inflammatory cytokines can produce undesirable outcomes. For example, pro-inflammatory cytokines appear to be directly linked to the disease process of rheumatoid arthritis and play a role in the joint destruction seen in this disease.

An acute condition, popularly called a 'cytokine storm' is life-threatening. Cytokine-stimulated inflammation is a normal and healthy *localised* response to infection. A cytokine storm refers to the excessive or uncontrolled release of pro-inflammatory cytokines at a *system-wide* level. This involves a positive feedback loop with cytokines stimulating the production of more white blood cells and these cells, in turn, produce more cytokines that attract more cells, and so on.

**ODD FACT**

The disproportionately high death toll among young people during the 1918 influenza pandemic is thought to have been a result of a cytokine storm in those people.

**KEY IDEAS**

- Cytokines are important signalling molecules that are involved in communication between immune cells.
- Cytokines are a large and diverse group of proteins and protein derivatives, and their subgroups include interferons and interleukins.
- Cellular responses to cytokine signals vary, including cell growth and differentiation, cell migration, cell apoptosis, as well as immune responses such as inflammation and phagocytosis.
- Cytokines are involved in a number of inflammatory diseases.

**QUICK CHECK**

**20** Briefly describe the function of the following cytokines:
  **a** interferon
  **b** colony-stimulating factors.
**21** In what way are cytokines like hormones?
**22** In what way are cytokines different from hormones?

# Chemical signalling in plants

Plant cells, like animal cells, engage in a 'constant chatter' that is carried on through signalling molecules. Intercellular signalling is essential for the coordinated growth, development and reproduction of plants. These signals regulate the development of the undifferentiated mass of cells of a plant embryo into an adult plant with its differentiated organs, including leaves, stems and buds above the ground, and roots below ground. Signalling molecules include messages relating to timing, such as: when seeds should germinate, when dormant buds should sprout, when flowers should bloom, when fruit should ripen, when leaves should drop (see figure 5.33).

Fruits appear on flowering plants and ripen at certain seasons; so, for example, peaches and cherries typically appear for sale in the summer months that follow the ripening season for these fruit. Fruit growth and maturation are just two of many changes that occur during the life cycle of a flowering plant. Changes during the life cycle — from the time of its formation from a germinating seed to the time that the plant matures and itself produces seeds — involve both growth and development. Growth (increase in size) and development (change in form) are seen in many events, such as formation of buds, lengthening of stems, downward growth of roots, expansion of leaves, appearance of flowers, development and ripening of fruits. For each plant species in a particular environment, the events of growth and development occur in a predictable sequence.

Plant growth and development are influenced by both internal and external factors and by interactions between these factors.

- Internal factors include chemical substances known as plant hormones. These hormones are the major signalling molecules that influence germination, growth, development and metabolism in plants.
- External factors include environmental factors such as light intensity, light wavelength, temperature, and day length. For example:
  - Increases in day length and temperature may act as signals to plants to shift resources to sexual reproduction and produce growth in flowers and then in seeds
  - Low light intensity may act as a signal to plants for the stem to elongate (see figure 5.33).

In this chapter, we will explore plant hormones and identify their role as signalling molecules that regulate many aspects of plant life.

studyon

Unit 3 | AOS 2 | Topic 1 | Concept 4

**See more**
Plants and hormones

# Chemical signals: plant hormones

Plant hormones are chemical messengers or signalling molecules that are produced in one part of a plant and have a physiological effect on cells in a specific target tissue. This target tissue can be both the cells that produce the hormone as well as cells in other parts of the plant.

Plant hormones have many of the same general characteristics as mammalian hormones. Plant hormones are produced by plant cells in small amounts and are active at low concentrations. Plant hormones act as signalling molecules that target various cells in plant tissues and produce specific effects. Both mammalian hormones and plant hormones act as signalling molecules. Hormone signalling in plants is similar to that occurring in animals. The hormone signal binds to a specific receptor, and once received, the signal is transmitted via a signal transduction pathway to the nucleus where a specific cellular response is generated:

1 Reception → 2 Transduction → 3 Response

studyon

Unit 3 | AOS 2 | Topic 1 | Concept 4

**Plant hormones**
Summary screen and practice questions

Responses of the plant vary according to the particular signalling molecule involved, its concentration, the region of the plant where the signal is received, and the stage of development. Responses include biochemical and physiological responses ranging from seed germination, shoot and root development, initiation of flowering, ripening of fruit and the regulation of photosynthesis.

Some differences between plant and animal hormones:

Mammalian hormones are produced by discrete endocrine glands. Plant hormones are not produced in glands but are produced mainly in the cells of growing regions (meristems) of shoots and roots, in young leaves, in germinating seeds and in developing fruits. Mammalian tissues are typically transported in the bloodstream, while transport of plant hormones is mainly in the phloem tissue.

The classical plant hormones are:

1. auxins, which are responsible for cell division and growth in cell size
2. cytokinins, which are responsible for an increase in cell division
3. gibberellins, which are responsible for cell division and growth in cell size
4. abscisic acid, which is responsible for dormancy
5. ethylene, which is responsible for ageing (senescence).

Figure 5.33 shows some of the actions of these classical hormones in a plant.

studyon

Unit 3 | AOS 2 | Topic 1 | Concept 4

**Do more**
Hormone production in plants

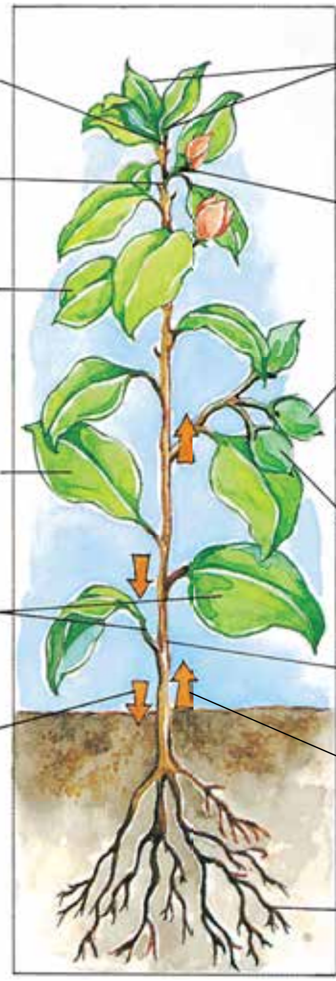
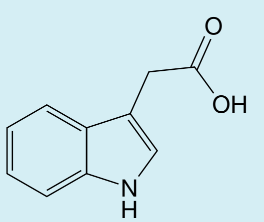



**FIGURE 5.33** Plant hormones regulate many aspects of growth and development in flowering plants. Some typical hormone actions are shown in this figure.

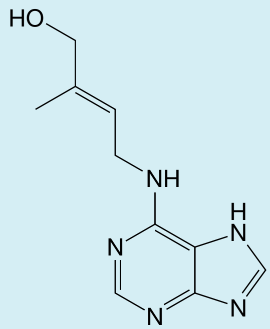
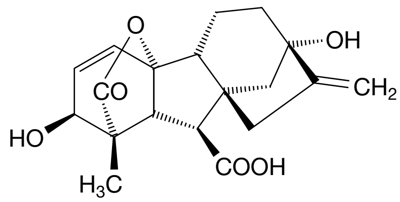
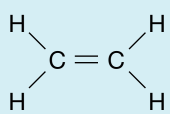
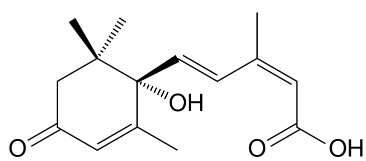
Table 5.4 summarises the main features of the classical plant hormones.

**TABLE 5.4** A summary of major actions of the classical plant hormones. In column 2, the chemical structure of a major hormone or the only hormone in a group is shown.

| Hormone | Structure | Where found | Major actions include |
|---|---|---|---|
| AUXINS include<br>IAA<br>a growth hormone | <br>indole acetic acid (IAA) | in meristems at:<br>apical buds (growing tips)<br>in developing seeds<br>in leaf primordia | regulate embryonic growth and form<br>promote flowering and fruit development<br>promote stem and coleoptile elongation<br>produce trophic responses to light and gravity<br>induce formation of xylem and phloem<br>induce formation of adventitious roots<br>prevent leaf drop (abscission) |

*(continued)*

**TABLE 5.4** *(continued)*

| Hormone | Structure | Where found | Major actions include |
|---|---|---|---|
| CYTOKININS<br>growth hormones | <br>zeatin | found in actively dividing meristematic regions of stems, roots and leaves<br>believed to be synthesised in roots | promote cell division<br>regulate growth and form (with auxin)<br>promote root growth<br>delay senescence (ageing)<br>stimulate leaf expansion<br>stimulate growth of lateral buds<br>regulate development and activity of chloroplasts |
| GIBBERELLINS<br>growth hormones | <br>gibberellic acid | produced by roots, young leaves and seeds | promote rapid stem elongation ('bolting')<br>stimulate shoot growth<br>break dormancy of some seeds and promote germination<br>mobilise food stores in seeds<br>promote flowering, fruit and leaf growth<br>stimulate production of seedless fruits |
| ETHYLENE<br>the gaseous hormone | <br>ethylene | produced by most parts of a plant, with highest concentrations during senescence, leaf abscission and fruit ripening | stimulates shoot and root growth<br>promotes flower opening and fruit ripening<br>promotes abscission of leaves and fruit<br>stimulates leaf and flower senescence |
| ABSCISIC ACID (ABA)<br>the 'stress' hormone | <br>abscisic acid | produced in roots and terminal buds | inhibits growth of shoots<br>blocks seed germination<br>promotes dormancy<br>promotes synthesis of storage proteins in seeds<br>promotes stomatal closure (in times of water stress) |

It should be noted that:

- The same hormone may induce different responses in different tissues of a plant, such as promoting growth in one region (such as the stem) while inhibiting growth in another region (such as the roots).
- The action of a hormone may depend on the concentration of the hormone, with different concentrations producing different responses by a plant.

Further, plant hormones do not act in isolation — one hormone may influence the action of another hormone.

Let's look at the classical plant hormones.

*'Adventitious root' refers to any root that develops in an unusual place, such as roots that grow from stems. 'Lateral root' refers to a root that branches out from a main or primary root.*

## Plant hormones: auxins

The label *auxins* (from the Greek *auxeun* = to grow or increase) refers to several hormones, such as indole acetic acid (IAA), the main naturally occurring auxin in plants. Auxins are essential plant hormones whose signalling

## ODD FACT

Bamboos are tropical grasses that grow at rapid rates, with some species growing at rates of up to four centimetres per hour. This growth is due mainly to cell elongation. What class of plant hormone is involved in the control of this growth?

influences almost every aspect of plant growth and development. The major roles of auxins are to promote cell growth and cell elongation (expansion), and they are important in root development by inducing the growth of lateral roots and adventitious roots. Auxins also promote the growth of flowers and fruits, and influence the differentiation of unspecialised cells into vascular tissue (xylem and phloem).

### Where are auxins produced?

Auxins are produced by meristem cells in the growing tips of plant shoots, and move unidirectionally from shoots to roots in the phloem. The site of production of auxins was first identified in germinating grass seeds from the result of the following experiment:

*Experiment 1*: **The coleoptile is the outer protective sheath that covers the embryonic shoot of a new grass seedling as it emerges from below ground.** Seeds were allowed to germinate and the tips of the coleoptiles of some seedlings were removed, while other seedlings were left with an intact coleoptile (see Figure 5.34). It was observed that only those seedlings with an intact coleoptile continued to grow, while the seedlings with topless coleoptiles stopped growing. From this observation, it was concluded that the tip of a coleoptile was the source of a growth hormone, which we now know to be the auxin, indole acetic acid (IAA).

Another experiment provided some information about the nature of auxins.

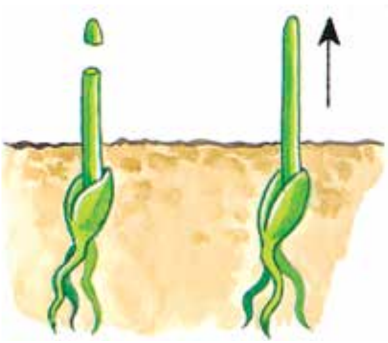



**FIGURE 5.34** Experiment demonstrating that the coleoptile tip is the source of a growth factor, later identified as indole acetic acid (IAA), which is an auxin.

*Experiment 2*: Tips cut from coleoptiles were placed on agar blocks for several hours. It was observed that, when one of these agar blocks was placed onto a topless coleoptile, normal growth recommenced (see figure 5.35).

From this observation, it was concluded that the substance causing growth was a water-soluble chemical that could diffuse from the coleoptile tip into the agar block and then diffuse downwards from the agar block into the cut end of a topless coleoptile. What is the control in the experiment shown in figure 5.35?

*Agar is a firm jelly-like (gelatinous) substance.*

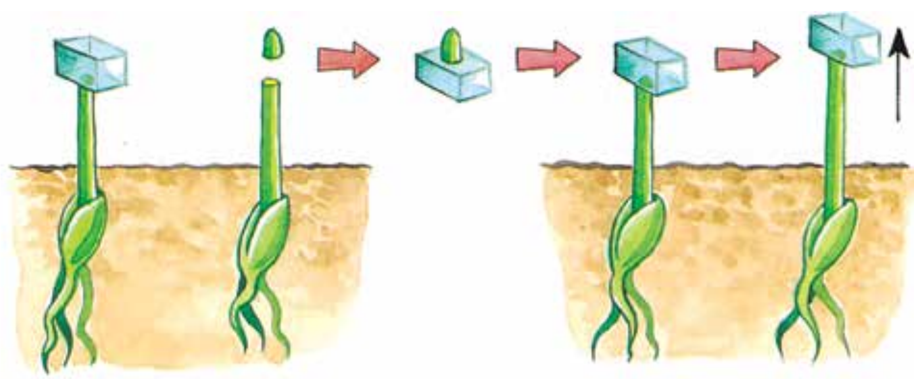



**FIGURE 5.35** Experiment demonstrating that auxins are water-soluble chemicals produced in growing coleoptile tips and that auxins promote the elongation of cells in the region below the tip.

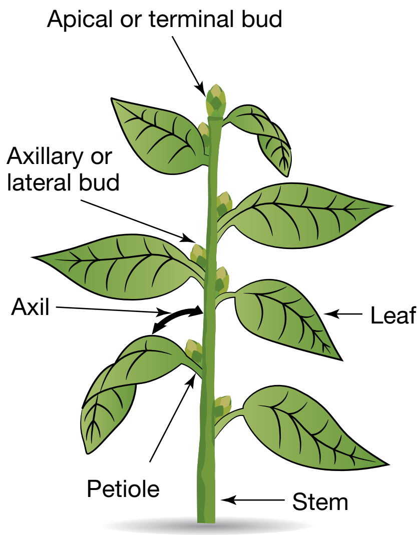



**FIGURE 5.36** Diagram showing the types of buds (apical or terminal, and axillary or lateral) and their relative positions on a shoot.

**ODD FACT**

Apical dominance was one of the first developmental events in plants shown to be regulated by plant hormones. The experimental evidence is described in a classic paper by Thimann and Skoog (1934): *Proc R Soc London Ser B* volume **114**: pages 317–339.

## Auxins in action

Let's now look at some specific actions of auxins in:

- apical dominance
- phototropism
- promotion of cell growth:
  - strawberries
  - adventitious roots.

### Auxins and apical dominance:

**Apical dominance** is the control exerted by the shoot tip on the outgrowth of axillary buds. Figure 5.36 shows the positions and kinds of buds on a shoot. Apical dominance can be seen when lateral (axillary) buds on a stem close to the apex of a plant do not develop because of the presence of an actively growing terminal (apical) bud at the apex of a plant.

The principal auxin, indole acetic acid (IAA), is responsible for apical dominance. The IAA produced by cells of the terminal bud moves down the stem through the phloem and sends a signal to cells of the lateral buds that inhibits their growth. The auxin signal travels to target cells in the lateral buds. Receipt of this signal initiates a signal transduction pathway that produces a cellular response.

When apical dominance is strong, plants develop few, if any, lateral branches. What would be expected to happen to a plant when the terminal (apical) bud at the apex is removed (see figure 5.37)?

When the terminal bud at the apex is snipped off, the source of IAA is removed and lateral (axillary) buds lower down on the stem begin to develop. Why does this happen? The answer lies in the relative balance between signals from auxins and signals from another group of plant hormones, the cytokinins.

The removal of the terminal bud removes the major source of auxin. This means that the auxin signals ('Stay asleep') that have dominated the chemical communication to the lateral buds stop. Instead, signals from another group of plant hormones, the cytokinins, can now dominate. The cytokinin signals ('Start growing') stimulate the growth of lateral buds. By selectively pruning or removing growing tips with their terminal buds, gardeners can produce more bushy plants because more lateral buds develop into leafy shoots.


**FIGURE 5.37**

### Auxins and trophic movements:

External factors, such as light, gravity and touch, exert an influence on plant growth and development. The growth of a plant in response to a stimulus such as light or water is called a **tropism**. When a plant grows towards a stimulus, the term *positive tropism* is used. Growing away from the stimulus is called *negative tropism*. Let's look at an example of the signalling that occurs when a plant shoot responds positively to the stimulus of light, an example of positive phototrophism.

In an early investigation of growing seedlings, Charles Darwin (1809–1882) and his son Francis found that they could prevent grass coleoptiles from bending towards the light by covering their growing tips (see figure 5.38). Darwin concluded that the tip of the seedling produced an 'influence' that passed from the tip to the region below the tip where bending occurred.

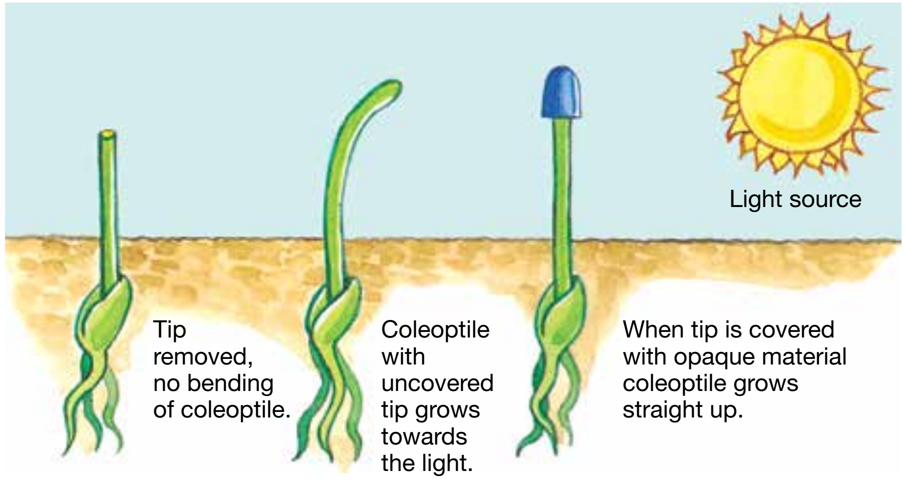



**FIGURE 5.38** Growing coleoptiles bend towards a light source. Bending does not occur if the tip is shielded from the light. What conclusion can you draw from the observation that a coleoptile without its tip fails to respond to the light?

We now know that auxin, a plant hormone, is produced in the tip of a coleoptile and causes elongation of cells in the coleoptile. The tip is the site of reception of the light stimulus. Cells of the growing region below the tip are the effector or the target cells.

What causes a plant to bend towards a light? If a coleoptile is evenly illuminated, auxin is evenly distributed throughout the tip and the coleoptile grows straight up.

However, if light is concentrated to one side of a coleoptile, **the auxin moves away from the light source to the darker side and becomes more concentrated in cells in that region**. The increased concentration of auxin in the cells on the darker side causes these cells to elongate to a greater degree than cells nearer the light (see figure 5.39). The uneven growth of cells leads to bending of the coleoptile.

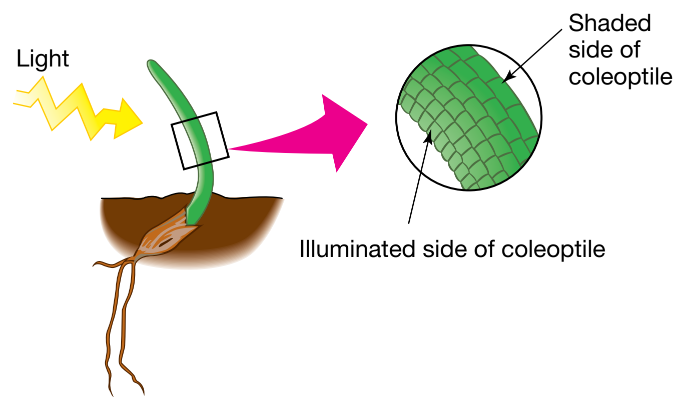



**FIGURE 5.39** The distribution of auxin in cells controls the bending of a coleoptile. If light shines from one side, auxin moves to the shaded side and concentrates there. The higher concentration of auxin causes greater elongation of cells in that area (as compared to cells on the lighted side). The different rates of increase in cell size on each side of the shoot result in the bending of the coleoptile towards the source of light.

**ODD FACT**

Auxin activates an enzyme that loosens the structure of the cellulose in the cell wall. The reduced rigidity of the cell walls enables cells to bend more easily.

**Auxins and cell growth:** *1. Growth of strawberries.*
Figure 5.40 shows strawberries (at left) and a scanning electron micrograph of the surface of a strawberry (at right). Note the large ovoid shields that are commonly referred to as 'seeds'. The flesh of the familiar strawberry that we eat is an enlarged receptacle. The receptacle is part of the strawberry flower on which the 'seeds' form (see figure 5.40c). However, the 'seeds' on a strawberry are not true seeds. Each is a dry fruit, with the botanical name 'achene', and inside each achene is a tiny seed.

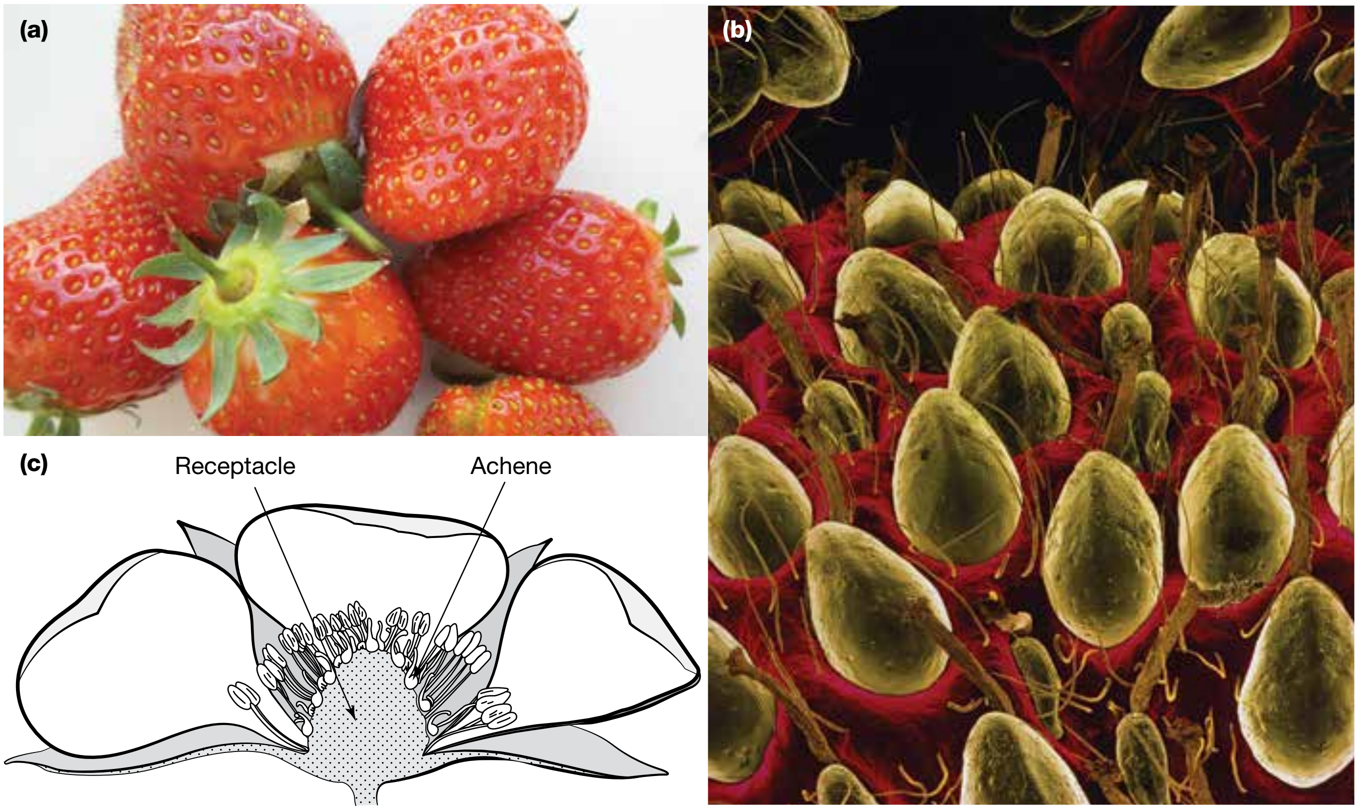



**FIGURE 5.40** **(a)** Strawberries. (Image courtesy of Judith Kinnear.) **(b)** Scanning electron micrograph of the part of the surface of a strawberry showing the red flesh of the strawberry and the prominent ovoid-shaped 'seeds' or achenes. The brown tubular structures between the achenes are the styles and stigmas of the female reproductive organs. **(c)** Longitudinal section of a strawberry flower showing the receptacle with developing achenes on its surface.

**ODD FACT**

One medium-sized strawberry of most varieties has about 200 achenes. So, eat one strawberry and you are eating the strawberry receptacle along with several hundred strawberry fruits.

The achenes are a source of auxins. Refer to figure 5.41 to see the effect of auxin from the achenes on the development of the strawberry receptacle. The receptacle enlarges **only if the achenes are present on its surface and they release auxin signals for cell growth**. The size of a strawberry is directly proportional to the number of achenes on the receptacle.


**FIGURE 5.41** Diagram showing the effect of auxin signals on the enlargement of the strawberry. What happens if 'seeds' are removed from the receptacle after flowering? What is the effect if auxins are applied directly to a small 'de-seeded' receptacle?

**Auxins and cell growth:** *2. Growth of adventitious roots*

Auxins at particular concentrations stimulate the growth of adventitious roots. Commercial rooting powders or solutions are used by gardeners to stimulate root growth in cuttings (see figure 5.42). These products are applied at concentrations of 500–1500 ppm for herbaceous and softwood cuttings, and at higher concentrations (1000–3000 ppm) for cuttings that have more woody tissues.

ppm = parts per million

**FIGURE 5.42** A gardener dipping the stems of *Cistus* sp. (rockrose) cuttings into rooting powder prior to planting. Rooting powder contains auxin hormones that promote adventitious root development.

## Plant hormones: cytokinins

**Cytokinins** are another class of growth-promoting hormones that act on shoots, roots and fruits. A common action of this class of hormone is to **promote cell reproduction**. Not surprisingly, high concentrations of cytokinins are found in plant tissues where cell reproduction is occurring at a rapid rate, such as in young, growing fruits.

Along with auxins, cytokinins are a major class of growth-promoting hormones that are important in the organised growth and development of plants. One of the most common cytokinins found in plants is called zeatin. (Zeatin was the first cytokinin to be isolated from plants, and was originally isolated from corn (*Zea mays*) in 1961.) A common action of this class of hormone is to promote cell division. Cytokinins are produced in roots and travel upwards to the shoots. Not surprisingly, high concentrations of cytokinins are found in plant tissues where cell division is occurring at a rapid rate, such as in the meristematic regions of roots, and in young, growing fruits.

Careful! Do not confuse plant cytokinins with mammalian cytokines. Cytokines are discussed earlier in this chapter (see page 183).

**Cytokinins and adventitious shoots:**

Cytokinins promote the production of adventitious shoots. The term 'adventitious shoot' refers to any shoot that develops in an unusual place, such as shoots that grow from leaves. Some plants, such as *Bryophyllum* sp., have meristematic tissue at notches along the edges of their leaves. This meristematic tissue can divide and give rise to adventitious shoots (see figure 5.43). The signalling molecules that stimulate the production of adventitious shoots are cytokinins that are present in cells at the leaf margins.

Meristematic tissue is located at the tips of stems and roots (= apical meristem), in the sides of stems and roots (= lateral meristem) and at the base of leaves and at internodes (= intercalary meristem).

**FIGURE 5.43** A *Bryophyllum* sp. leaf showing the formation of adventitious shoots along its leaf margin, in response to signals from cytokinins.

Evidence of the role of cytokinins in stimulating shoot production came from a study of genetically engineered cytokinin-deficient plants. These plants developed stunted shoots, and their leaf production was just 3–4 per cent of that of normal plants.

**Interactions between cytokinins and auxins:** Cytokinins typically interact with auxins in plant growth and development. The effects of these two hormones on plant growth is complementary. One effect of cytokinins is to induce shoot formation, while that of auxins is to induce root formation. For example, the small adventitious shoots on the *Bryophyllum* leaf margin are formed in response to cytokinin signals. When these small shoots fall to the soil below, root formation occurs in response to auxin signals.

Another example of the interaction between cytokinins and auxins may be seen in the spread of blackberry bushes (*Rubus fruticosus*). Examine figure 5.44. Note that blackberry plants have long stems, called canes, that grow from the crown of the plant. When a blackberry cane forms an arch that touches the ground, auxins accumulate in the horizontal portion of the cane, and the signals from the auxin promote root development. In turn, the developing roots produce lots of cytokinins, and their signals induce shoot development — hey presto, a new blackberry plant. Daughter plants can be formed at the ends of canes up to 6 m from the crown.

**studyon**

| | |
|---|---|
| Unit 3 | **Cytokines** |
| AOS 2 | Summary screen and practice questions |
| Topic 1 | |
| Concept 5 | |

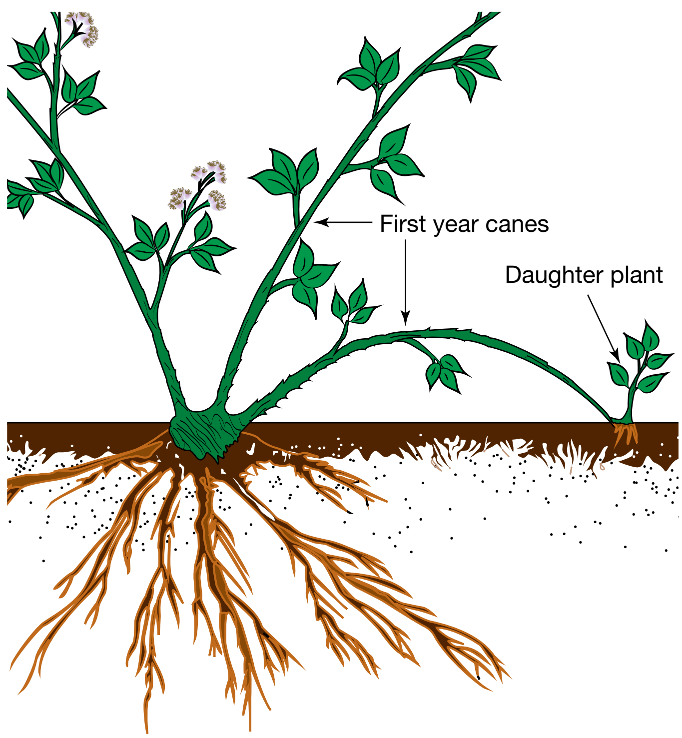



**FIGURE 5.44** A blackberry plant can spread when its arching shoots touch the ground, each forming a new daughter plant that is a clone of the parent plant.

**ODD FACT**

Blackberry is a declared noxious weed in several Australian states, including Victoria.

## COMMERCIAL USES OF PLANT HORMONES

### Herbicides

Synthetic auxins can be produced in very large quantities in the laboratory. One common auxin, known as 2,4-dichlorophenoxyacetic acid (2,4-D), is used as a selective **herbicide** or weedkiller. This is possible because 2,4-D kills some plant species but not others. Among the susceptible plant species are many of the dicotyledonous noxious weeds. In contrast, monocotyledonous cereals and other grasses show little or no effect from 2,4-D. Crops or pastures where weeds are growing are sprayed, and the herbicide kills the weeds by disrupting normal growth.

Some weeds now have strains that are resistant to 2,4-D, so use of the herbicide in the future may become less effective. In addition, questions are being raised about the possible effects on people and other animals of the long-term use of such herbicides.

### Hormones and tissue culture

Whole plants can grow from small cuttings. This is often done by keen home gardeners who take cuttings from favourite plants and propagate them.

When large numbers of plants with a particular genetic make-up or of particular economic importance are required, growth from cuttings or even from a small group of cells is carried out in the laboratory using special techniques. The technique of **tissue culture** or **cloning** (see 'Cloning plants from tissue culture' in *Nature of Biology Book 1 Fourth Edition*) can be used to obtain large numbers of plants in a relatively short time. Cloned plants are genetically identical to the plant from which the original cells were taken and so they have the same valuable genetically determined features present in the parent plant.

When small groups of unspecialised cells are used, they are sterilised and grown on agar in a test tube or other container (see figure 5.45). Each group is called a callus. The hormone cytokinin is added. High levels of cytokinin combined with relatively low levels of auxin result in the growth of shoots. The shoots on each callus are then treated with auxin, leading to a relatively high level of auxin and a relatively low level of cytokinin compared to the level previously. This results in the formation of roots.

Each callus, which started as a small group of cells, gives rise to a complete new plant. By this technique, many genetically identical plants can be quickly produced from the one parent plant.

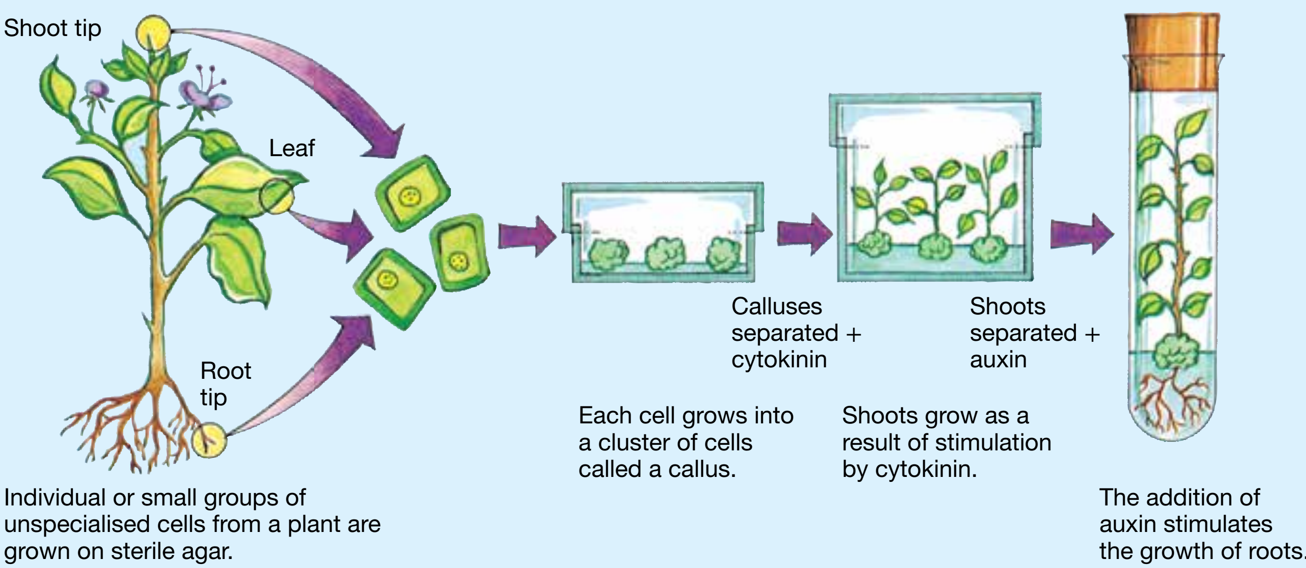



**FIGURE 5.45** Plant hormones are used to promote growth when new plants are cloned from unspecialised cells.

## Plant hormones: gibberellins

In 1926, some plants in rice crops growing in Japan were observed to be diseased. They were thinner and taller than normal and, instead of standing erect, these rice plants tended to fall over. This disease was called the 'foolish seedling' or bakanae disease. These abnormally tall plants had a fungal infection, and the fungus involved was the species *Gibberella fujikuroi*.

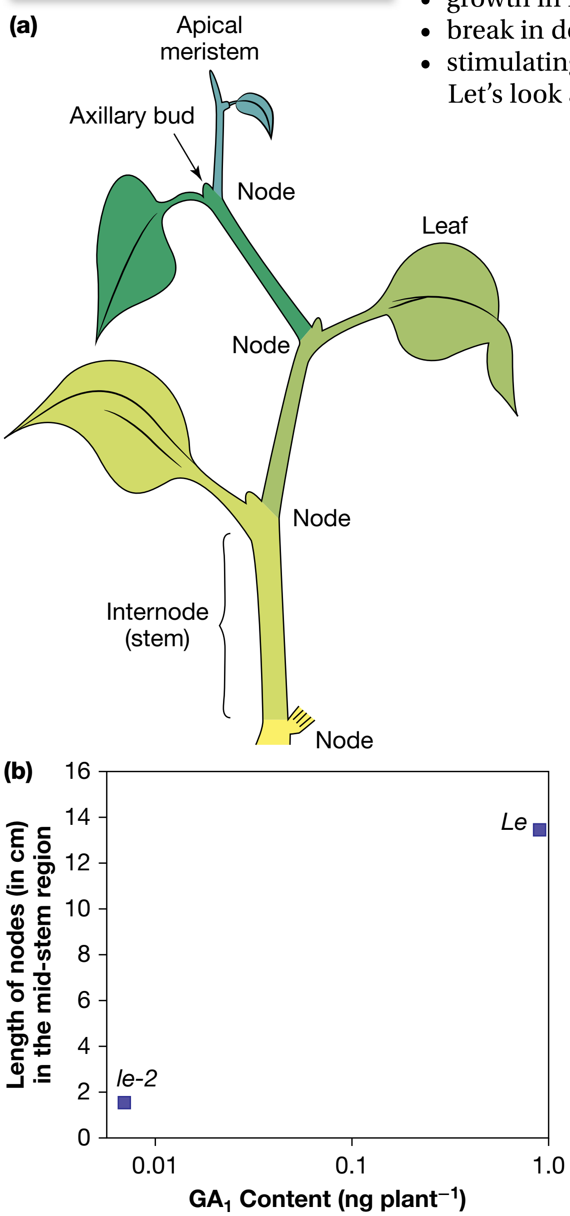
A Japanese scientist extracted many compounds from this fungus and identified one particular compound as the cause of the abnormal growth of the rice plants that made them tall and thin. The compound responsible for the increased height of the rice was named **gibberellin**, after the fungus from which it was isolated. Although the first gibberellin hormone was extracted from a fungus, many naturally occurring gibberellins have now been extracted from many plant species. The best known gibberellin is called **gibberellic acid (GA)**.

*Gibberellins* are growth hormones, produced by roots and young leaves, that act as signalling molecules for cellular responses including:

- stem growth
- growth in fruit size (in conjunction with auxins)
- break in dormancy of some seeds, promoting seed germination
- stimulating the development of seedless fruits.

Let's look at the results of some of these gibberellin signals.

**FIGURE 5.46** **(a)** Diagram showing the nodes along a stem. Nodes are regions from where leaves and axillary buds develop. An internode is the region between two nodes. **(b)** Graph showing internode lengths in the mid-stem region plotted against total GA content of a tall plant (Le) and a dwarf plant (le-2). By what multiple does the GA content of the tall plant differ from that of the dwarf plant?




### Gibberellins and stem growth:

1. *Mendel's peas:*

Gregor Mendel (1822–1884), an Augustinian monk in a monastery at Brno in the present day Czech Republic, is famous for his experiments with pea plants (*Pisum sativum*). One of the traits studied by Mendel was plant height, with pea plants being either tall or dwarf. Pea plants of the tall variety grow to heights of about 180 to 200 centimetres, while plants of the dwarf variety grow to a height of 30 centimetres.

While Mendel solved the mystery of how these traits were inherited, it was not until 1997 that the underlying cause of dwarfness in pea plants was identified. In 1997, scientists demonstrated that the dwarf condition in pea plants was due to a defect in a particular gene. This gene encodes the enzyme needed to convert an inactive precursor of gibberellic acid into active gibberellic acid (GA). GA is the signalling molecule that carries the messages to stem cells in the internode region to divide and to elongate. In the dwarf plant, this enzyme is missing, resulting in the lack of GA signalling to the cells in their stem. The role of GA in stem elongation was confirmed when it was observed that spraying dwarf plants with GA caused them to grow to normal height.

Both tall and dwarf plants have the same number of nodes (and hence internodes) in their stems (see figure 5.46a). The difference between them is that the internode distance (distance between nodes) is much greater in tall plants than in dwarf plants (see figure 5.46b). The increase in internode distance is a result of GA signals to stem cells. These signals carry messages for both cell division and cell elongation of stem cells in the internode region.

### Gibberellins and fruit size:

Gibberellins have several uses in agriculture. One example is the treatment of developing clusters of seedless grapes with gibberellic acid (GA). Treatment with GA typically involves spraying when the grape diameter is 3 to 5 mm. This treatment increases the size of individual grapes by a factor of two or three (see figure 5.47) by stimulating the division and the enlargement of individual cells, producing an overall increase in grape size. Interestingly, GA has little or no effect on seed-containing grapes.

**FIGURE 5.47** Clusters of seedless crimson table grapes. The cluster at the left was sprayed with gibberellic acid (GA) while the cluster at the right was not treated. Note the difference between the sizes of individual grapes in the two bunches. (Image courtesy of Stephen Vasquez and Matthew Fidelibus, UC Cooperative.)

**ODD FACT**

Botanically speaking, a grape is a berry. Breeding programs have selected grape plants that naturally produce large berries. This means that many newer cultivars require less GA to attain optimal berry size as compared to older cultivars.

## Gibberellins and leaf shape:

Some species of plants, including *Eucalyptus* species, have two leaf forms — juvenile leaves and adult leaves. These two types of leaf have distinctly different shapes (see figure 5.48).

A gibberellin signal promotes the development of branches with juvenile-type leaves; these branches cannot produce flowers. In the absence of gibberellins, branches with adult leaves form, and these branches are able to produce flowers. What would you predict would happen if buds on an adult branch were sprayed with gibberellic acid?

**FIGURE 5.48** Distinctly different adult (right-hand side) and juvenile leaves (left-hand side) on the same species of *Eucalyptus*. The presence of gibberellins promotes the development of branches with juvenile-type leaves. Branches with adult leaves form on the same tree in the absence of gibberellins. (Image courtesy of Judith Kinnear.)

## Gibberellins in seed germination:

Germination is the growth of a plant that is enclosed within a seed. Enclosed within the seed coat are the plant embryo and its energy store, the endosperm (see figure 5.49). In the seeds of cereals, such as wheat, oats and rice, this energy store is predominantly in form of starch — for example, starch comprises about 60–70 per cent of the whole grain. In addition to starch, the seeds of some plants also have significant oil stores in their endosperm.

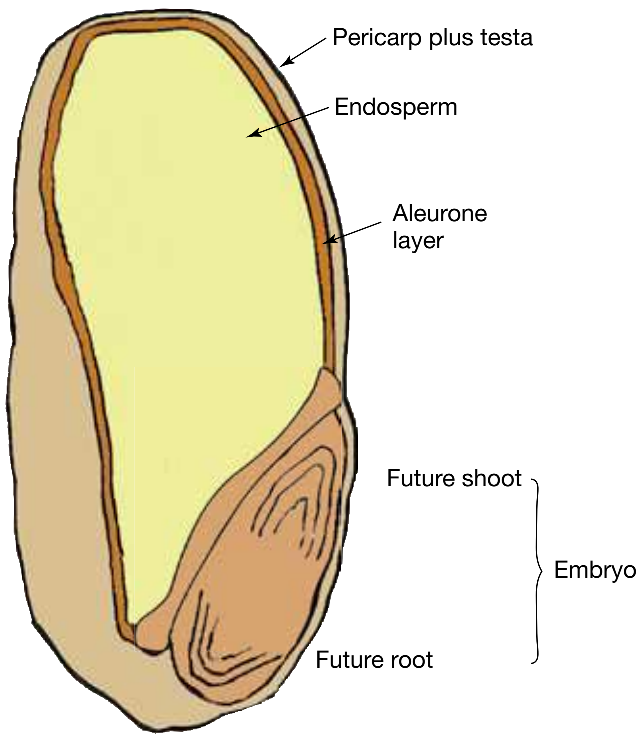



**FIGURE 5.49** Diagram showing a longitudinal section through a grain of wheat (*Triticum aestivum*). Note the embryo in the lower right-hand corner and the endosperm, which is an energy reserve mainly in the form of starch. The outer seed coats form wheat bran, and the embryo is wheat germ.

**ODD FACT**

The oil content of small grains, such as wheat seeds, is around one per cent, while that of oil seeds ranges from about 20 per cent for soybeans to more than 40 per cent for the seeds of sunflower and canola.

At the cellular level, plant hormones produce signal transduction pathways in much the same way as that occurring in animal cells. For example, consider the germination of a grain of wheat (figure 5.50).

- When water enters a seed, cells of the plant embryo release gibberellic acid (GA), which diffuses to the aleurone, the layer of cells that surrounds the starchy endosperm.
- The GA binds to a specific receptor on the plasma membrane of aleurone cells; this is a G protein coupled receptor.
- The binding of GA to the receptor starts the signal transduction pathway.
- The end result of the signal transduction pathway is the activation of a gene that encodes the instructions for the enzyme α-amylase. α-amylase is an enzyme that breaks down starch.
- α-amylase is secreted from the aleurone cells into the endosperm where it catalyses the breakdown of starch molecules into their glucose sub-units. The glucose molecules then diffuse to the embryo and provide it with an energy source for the initial stages of growth of a new plant.

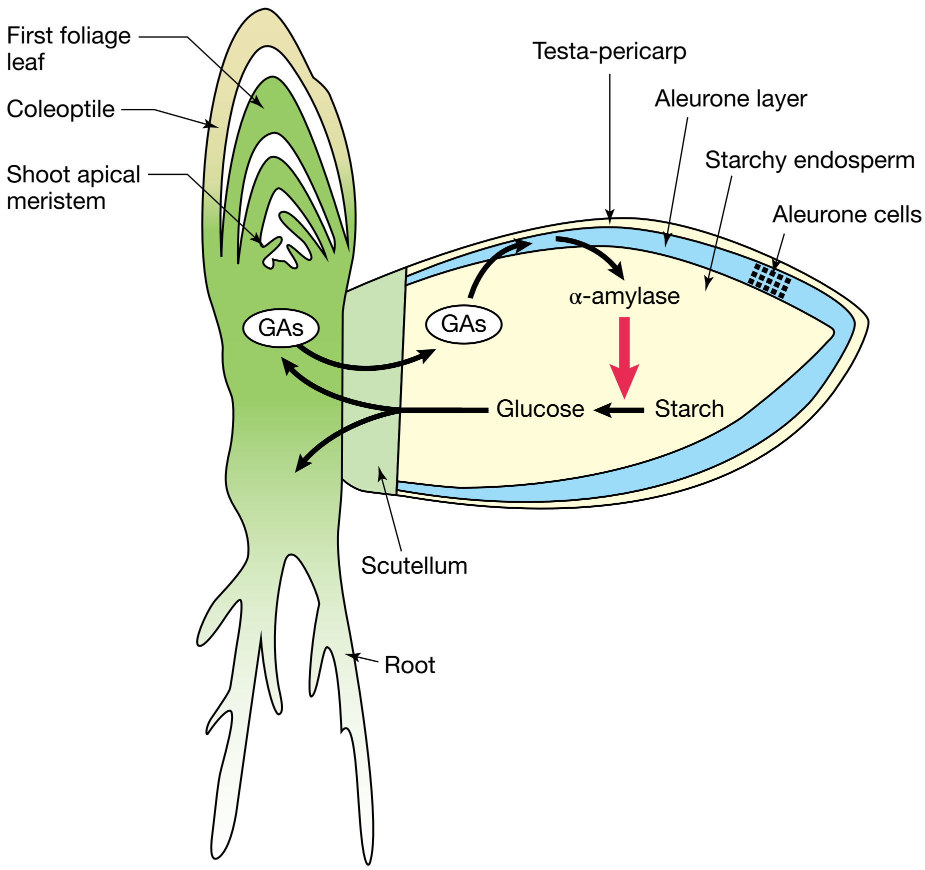



**FIGURE 5.50** Diagram of a longitudinal section (LS) through a grain of wheat (at right-hand side) with an LS of an enlarged embryo (at left-hand side). Normally the embryo occupies the lower portion of the seed as shown by the light-green shading. The stimulus to start the release of gibberellins (GAs) from embryonic tissue is the entry of water into the seed.

**ODD FACT**

Ethylene is produced by many fruits, including apples, pears, bananas, avocados, and tomatoes as they ripen, but not by cherries, grapes or citrus fruit.


**FIGURE 5.51**

## Plant hormones: ethylene

In the mid-nineteenth century, unusual changes were observed in plants growing near street gas lamps. These changes, such as the thickening and twisting of stems and premature leaf loss, were not seen in the same kind of plants growing in areas away from the gas lamps. Further investigation showed that the gas lamps concerned were leaking gas, and this suggested that something in the gas was causing the changes in the plants. In 1901, the active component of the gas was identified as ethylene.

However, it was not until 1934 that the first evidence was produced that plants could synthesise ethylene. It is now known that ethylene is produced from the amino acid methionine in almost all the tissues of higher plants.

Ethylene interacts with other hormones to influence special events in the life cycle of a plant, including the ripening of fruit and the autumn drop of leaves. Let's look at these two events.

### Ethylene and fruit ripening

Young bananas are green. If left on the plant, they eventually ripen and turn yellow. The ripening of bananas is stimulated by a signal from ethylene that **increases the rate of respiration** of the banana cells (see figure 5.51).

Under natural conditions, bananas in the same bunch can be at different stages of ripening. However, if some are over-ripe while others are still unripe, this reduces their commercial value.

A hand of bananas comprises a few bananas joined together. A bunch of bananas consists of many hands of bananas.

For marketing purposes, the ripening of bananas is controlled and synchronised. Bananas are picked while still green but sufficiently mature to ripen after they have been picked. If picked too young, bananas remain green and are generally unpleasant to taste. After picking, mature green bananas are kept under conditions of controlled humidity and temperature and in an atmosphere to which ethylene gas is added at a concentration of 100 μL/L. Under these conditions, uniform ripening of the fruit occurs and bananas in the same bunch ripen synchronously (at the same time).

After they are picked, bananas sometimes need to be kept for some time. How should they be stored? For maximum storage life, bananas are vacuum packed in plastic bags containing material that absorbs ethylene. This method of storage extends the life of the bananas because any ethylene gas produced by the bananas is absorbed and is unavailable to cause ripening of the fruit. In addition, the plastic used for these bags has a low permeability to oxygen. As a result, oxygen is not freely available for the living cells of the fruit and the cellular respiration rates are low.

Through controlled atmosphere storage methods, some fruits that are naturally available only during certain seasons can be made available over longer periods of the year. Are you aware of other fruits that are stored under controlled atmosphere conditions? Can you suggest what the label on a box of 'Controlled atmosphere apples' means?

### Ethylene and leaf drop

In September 2008, the scientific journal *PNAS* included a paper titled: 'Regulation of floral organ abscission in *Arabidopsis thaliana*'. Not an attention-grabbing title! However, a summary of this research appeared in a prominent UK newspaper, *The Telegraph*, under the snappier headline: *'Why leaves fall off trees is discovered.'* This research paper was of great interest because it correctly identified ethylene as the plant hormone with a major role in regulating the fall of leaves (see figure 5.52). Previously, this role had been assigned to the plant hormone abscisic acid.

**FIGURE 5.52** The fall of leaves from deciduous trees in autumn is regulated by signals from the plant hormone, ethylene.


**FIGURE 5.53** Isaac Newton assigned the fall of the apple to gravity. Gravity, yes, but with a little assistance from the plant hormone ethylene.

The separation of a part of a plant, such as a leaf or a flower or fruit from its parent plant, is called **abscission**. Autumn is a time of the year when deciduous trees lose all of their leaves. Before a leaf falls, a special zone called the **abscission zone** forms at the base of the leaf stalk (petiole). **Fruits and flowers have similar abscission zones** in their petioles (see figure 5.53).

Figure 5.54 below shows the abscission zone in a leaf petiole. Note that the abscission zone contains two layers of cells:

1. The separation layer: this multi-layer of cells across the base of the leaf stalk (petiole) marks the end point of the leaf that will fall, either when blown off by the wind or under the influence of gravity. The cells of the separation layer have thin cell walls weakened by the action of cellulase, an enzyme produced in response to an ethylene signal. A split in the separation layer spreads until finally the leaf is either blown away or falls.
2. The protective layer: these cells form a seal that protects the plant against the entry of microbes and against water loss. The cell walls in the protective layer contain suberin, a waxy waterproof substance. Expansion of the protective layer produces a tear in the weakened cells of the separation zone so that this layer splits and separates from the plant.

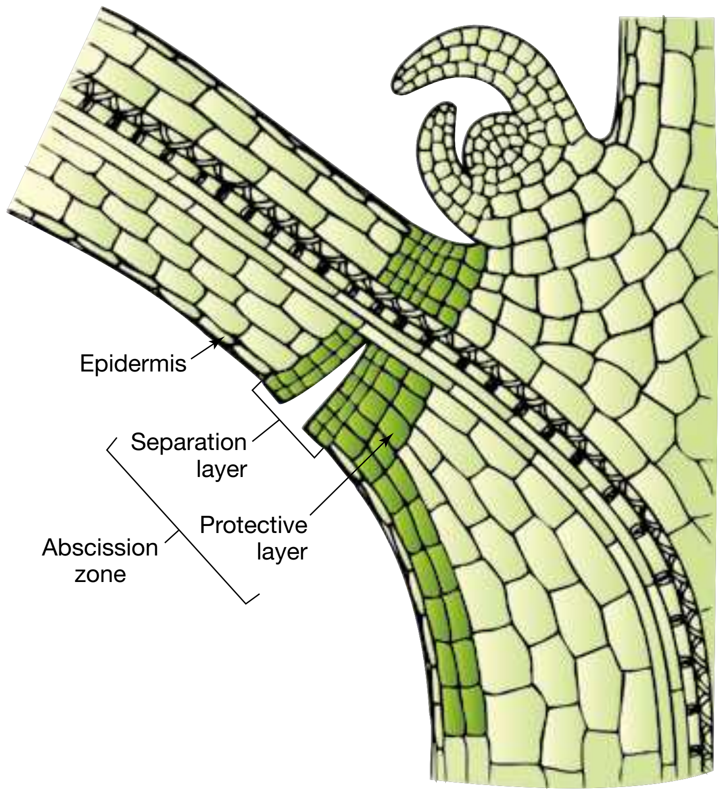



**FIGURE 5.54** Abscission zone in a leaf stalk (petiole). Separation of a leaf from the plant occurs across the abscission zone at the base of the leaf stalk or petiole. What is the function of the protective layer? Can you identify the curved structure above the abscission zone?

Abscission is important for plants because it provides a means by which a plant can discard non-functional organs, such as leaves in autumn, or infected organs. In addition, abscission of fruit assists in the dispersal of the seeds, which are the next generation of the plant.

### Ethylene and flower senescence

A flower on a plant may be dying at the same time as another flower is forming on the same plant. A flower dies because the petals die. Studies on picked carnations show that petals die because there is an increase in ethylene produced by the plant (see figure 5.55). What is the level of ethylene production by a carnation flower three days after it is cut? What is the level six days after it is cut?

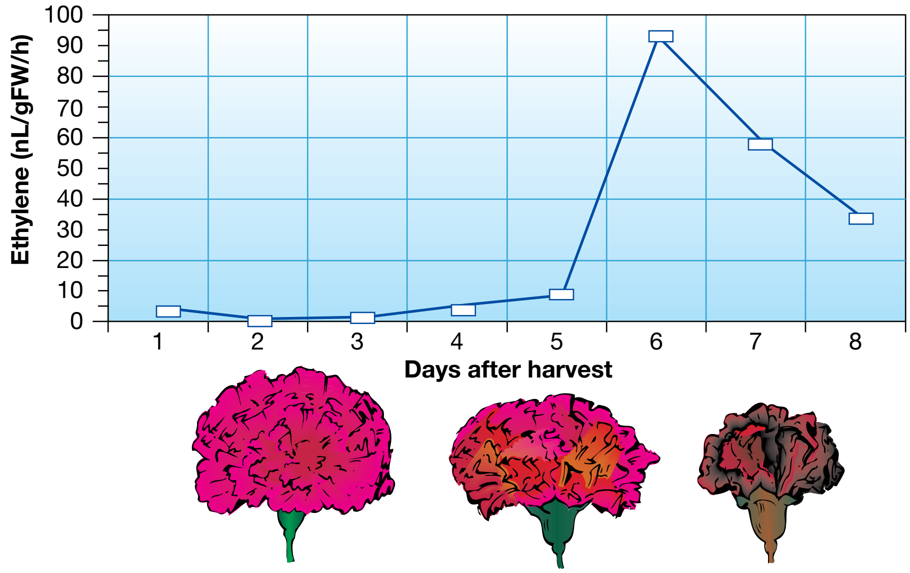



**FIGURE 5.55** Ethylene production by harvested carnation flowers. Initially, a cut carnation flower produces very little ethylene. As the amount of ethylene produced increases, the petals begin to curl and die. By about the eighth day after being cut, the flower is in an advanced stage of death.

The less ethylene produced by a carnation, the longer the cut flower lives. Genetic engineering experiments have produced a carnation in which the production of ethylene is blocked. This delays the death of flower petals and increases the life of the cut flower. This reduces the need to use environmentally hostile chemicals to extend the life of cut flowers.

## Plant hormones: abscisic acid (ABA)

Abscisic acid (ABA) is sometimes called the stress hormone. When conditions become unfavourable, as in freezing conditions or in drought, animals may migrate to more favourable conditions. This is not an option for plants, which, instead, must depend on other strategies to survive. For plants, the hormone ABA is a key mediator in plant responses to unfavourable conditions. Let's look at two examples of ABA in action.

**ODD FACT**

Abscisic acid (ABA) has this name because scientists originally thought that ABA was the signal for the abscission (fall) of leaves and fruit. It is now known that the hormone ethylene is the main signalling molecule for abscission, but the abscisic acid name was retained.

### ABA: Inducing dormancy

Signals from abscisic acid (ABA) inhibit growth and induce dormancy in leaf buds and seeds during periods of short day length and low temperatures. Abscisic acid signals are important in preventing the growth of leaves and the germination of seeds during periods when the delicate tissues of new leaves and seedlings would be at risk of being killed by freezing conditions. When days lengthen and temperatures rise, ABA concentrations fall and ABA signals decrease. At the same time, gibberellic acid concentrations rise and GA signals increase. As a result, dormant buds sprout, producing new leaves, and dormant seeds begin to germinate.

### ABA: Closing stomata

Plants lose most of the water taken up by their roots through their **stomata** (singular = stoma). Each stoma consists of two guard cells. If the guard cells are flaccid, the stomata are closed; if the guard cells are turgid and swollen, the stomata are open and water will be lost from a plant (see figure 5.56).

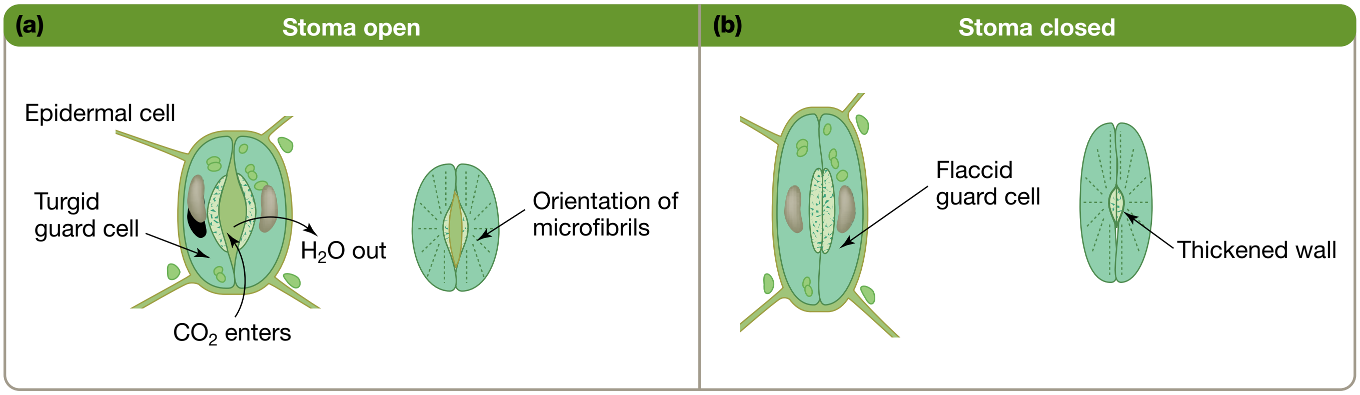



**FIGURE 5.56** Diagram showing stomata. **(a)** Open stoma with turgid guard cells that allow water loss through the central pore. **(b)** Closed stoma with flaccid guard cells prevent water loss.

In drought conditions, survival of a plant depends on stopping this water loss. Under drought conditions, plants produce and accumulate abscisic acid (ABA) in their leaves. Signals from ABA lead to the closing of stomata, so preventing the major source of water loss from a plant.

The signal pathway by which ABA leads to closing of stomata may be summarised in simple terms as follows:

ABA molecules bind to receptors at the surface of the plasma membrane of the guard cells.

This sets up a signal transduction pathway that leads to a net loss of ions from the guard cells.

The loss of these solutes from the cytosol of the guard cells reduces the osmotic pressure inside those cells, with the result that the cytosol no longer exerts an outward pressure on the cell walls of the guard cells because the guard cells are no longer turgid (swollen) and the stomata close.

## Other plant hormones

The first plant hormones to be recognised were the auxins. In addition to the classical plant hormones, other plant hormones have been identified, including some plant hormones that act as signalling molecules when plants come under attack by herbivores. Herbivore attack may come from sap-sucking or leaf-eating insects at their larval or adult stages, or from the leaf-eating activities of grazing or browsing mammals (see figure 5.57).



**FIGURE 5.57** **(a)** Damage to leaves by the feeding activities of insect larvae can be extensive. Here we see the damage to a leaf from the eating activities of caterpillars. **(b)** A gerenuk or Waller's gazelle (*Litocranius walleri*) feeding on leaves of an acacia tree in shrubland in Kenya.

## Jasmonates: defence against herbivores

One group of plant hormones known as **jasmonates (JAs)** plays a key role in plant defence against herbivores. JAs, such as methyl jasmonate, are produced by a plant in response to wounding or attack by herbivorous animals and build up in damaged areas of the plant. When JA molecules bind to their specific receptors, this starts a signal transduction pathway leading to the expression of response genes that encode production of a range of defensive chemicals.

These defensive chemicals include enzymes that interfere with leaf digestion. For example, tomato plants that come under insect attack produce a proteinase inhibitor that interferes with the insects' digestive processes. This interference presumably discourages these insects from eating that type of plant again.

Other examples include defence chemicals that attract predators or parasites of the attacking insects, as for example:

- Lima bean plants are often attacked by spider mites. When under attack from spider mites, lima bean plants produce a chemical that attracts a predator that preys on spider mites.
- Corn plants, when attacked by caterpillars, produce a volatile chemical that attracts a parasitoid wasp that lays its eggs in the caterpillar.

Some JAs are volatile compounds that readily evaporate into vapour form. Remarkably, these volatile compounds act as signalling molecules *not only* for the plant that produces them *but also* for other plants of the same species. JA vapours can spread downwind through the air from insect-damaged plants to neighbouring unharmed plants. The JA vapours enter leaf cells of neighbouring plants, either through the stomata or by diffusion. JAs enable communication between plants so that unharmed plants are primed in advance for defence against a herbivore attack.

### KEY IDEAS

- Plant hormones are signalling molecules involved in all processes of plant growth and development.
- Signals from plant hormones are transmitted through signal transduction pathways to the nucleus.
- The classic plant hormones are auxins, cyokinins, gibberellins, ethylene and abscisic acid.
- Plant hormones result in signal transduction pathways in plant cells in much the same way as mammalian hormones in mammalian cells.
- Each plant hormone has different target cells and induces one or more specific cellular responses.

### QUICK CHECK

**23** Identify the plant hormone that is a major signalling molecule in:
- **a** the fall of leaves
- **b** the growth of adventitious roots
- **c** the closing of leaf stomata in drought conditions
- **d** the growth of lateral buds
- **e** apical dominance
- **f** bending response to light.

**24** Check back on the signal transduction pathway in wheat seeds (figure 5.50).
- **a** What is the signalling molecule?
- **b** Where is the receptor for this signalling molecule located?
- **c** Name the protein produced by the cellular response to this pathway.
- **d** What is the function of this protein?

25 What is a key difference between the members of the following pairs:
   **a** the separation layer and the protection layer in the abscission zone of a leaf
   **b** an apical and an axillary bud?
26 What plant hormone signal is absent in dwarf pea plants?
27 List the cascade of signal transduction events when gibberellins are activated in a wheat seed.

## Apoptosis: programmed cell death

**ODD FACT**

An adult loses an estimated 30 000 epidermal skin cells every minute. Don't panic. These lost cells are replaced at the same rate by new skin cells.

**ODD FACT**

The term *apoptosis* was introduced in 1972 and comes from the Greek *apo* = from, and *ptosis* = falling, which refers to the process of leaves falling from trees or petals from flowers.

At any time, the cells in many organs and tissues of our bodies are in a state of turnover — old cells are dying and new cells are being formed. This turnover is happening, for example, in cells that form the lining of the gut, in cells of the epidermis of the skin, in cells of hair follicles, in cells of bone and in cells of the blood. In adults, the rates of the two processes are closely related so that, normally, a balance exists between the rate of new cell production and the rate of cell loss; that is: **rate of cell renewal = rate of cell death**. This balance keeps the number of cells relatively constant in adults. What would you predict about the situation in a growing child?

The process known as **apoptosis** or **programmed cell death** is a genetically controlled and highly regulated process of cell self-destruction. Apoptosis plays an essential role in ensuring that a balance exists between cell production and cell loss. It may seem strange at first, but programmed cell death is a normal healthy process in cells.

The rate of cell loss and replacement varies between different tissues. For example, the epithelial lining of the small intestine consists of a single layer of cells that is completely replaced every four to five days. Over this period, an entire layer of cells progressively dies and is replaced by new cells. The replacement of the intestinal lining involves two processes:

(i) the production of new cells deep in infoldings (crypts) of the intestine (see figure 5.58a). Production of these new gut lining cells is by mitotic divisions of stem cells located in the crypts.

(ii) the death of cells at the tips of the intestinal villi by apoptosis or programmed cell death (see figure 5.58b).

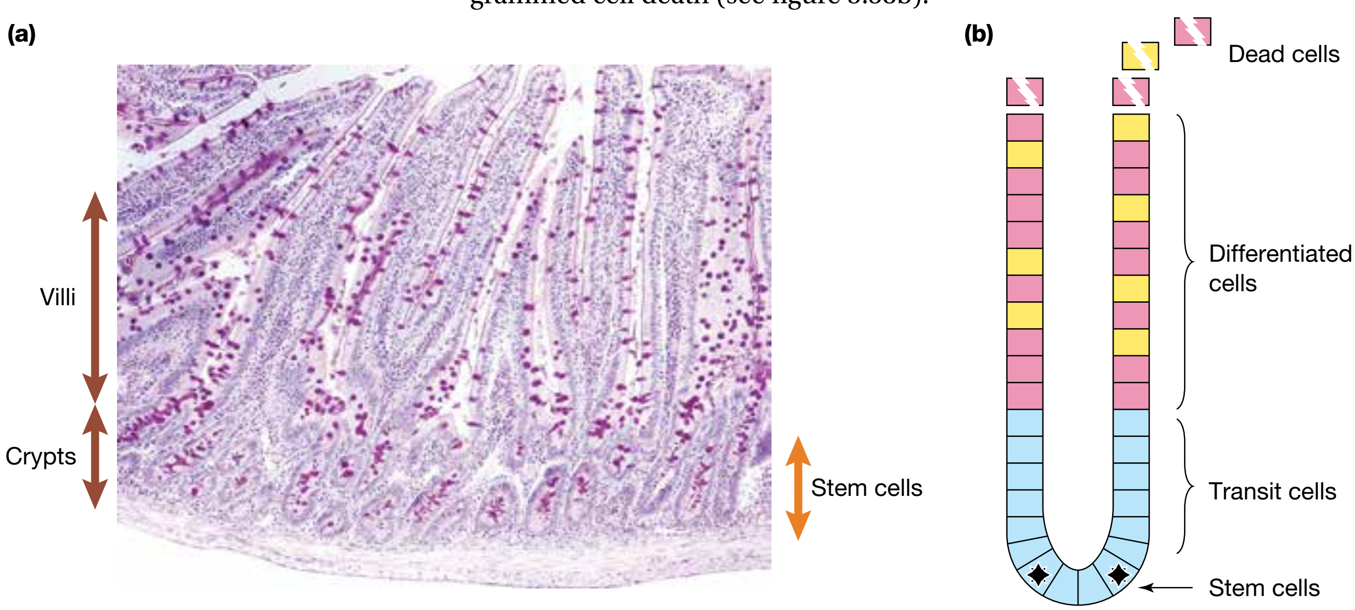



**FIGURE 5.58** **(a)** Longitudinal section through the small intestine showing the deep infoldings (crypts) and the upward projecting villi (singular = villus). **(b)** Diagram showing the progression of cells from their production by stem cell mitosis in the crypts, to their death by apoptosis at the tips of the villi.

studyon

Unit 3
AOS 2
Topic 1
Concept 9

See more
Apoptosis

studyon

Unit 3
AOS 2
Topic 1
Concept 9

Apoptosis
Summary screen and practice questions

In apoptosis, the death of cells occurs in a highly regulated sequence in response to specific signals received by the 'condemned' cell. The signal pathway of apoptosis brings about the planned death of various cells including:

- *Cells at the end of their natural life* (such as gut lining cells, as described above, and skin cells)
- *Dysfunctional, damaged or diseased cells*
  - Cells infected with a virus may be eliminated by apoptosis. The programmed death of these cells prevents the virus from replication and spreading to other cells (see chapter 7, page 323).
- *Excessive cells*
  The programmed removal of excessive cells by apoptosis is important in many contexts:
  - Every day, the bone marrow produces millions of new immune cells (white blood cells) and red blood cells. Over the same period, this production of new cells must be balanced by the loss of a similar number of cells. This is achieved in an orderly manner through apoptosis.
  - During embryonic development, the final shaping of organs depends on the programmed cell death of excess cells. This process is seen, for example, in the reshaping of the embryonic limb buds of mammals into digits. Figure 5.59 shows the role of apoptosis in digit formation in a human embryo. This process starts at about week four of embryonic development with the formation of a paddle-shaped limb bud. This is followed by the appearance of five ridges on the limb bud. The cells in the valleys between these ridges are progressively removed by programmed cell death, making the valleys deeper. By week eight, a structure with five distinct digits is clearly visible, the 'webbing' between the digits having been removed by apoptosis.
  - During the body's response to infection, immune cells are produced in larger numbers than required. These excess cells are removed by apoptosis.
  - During the development of the normal brain, enormous numbers of cells are formed, but only those that make the best neural connection are retained. Those cells not selected are culled by apoptosis.

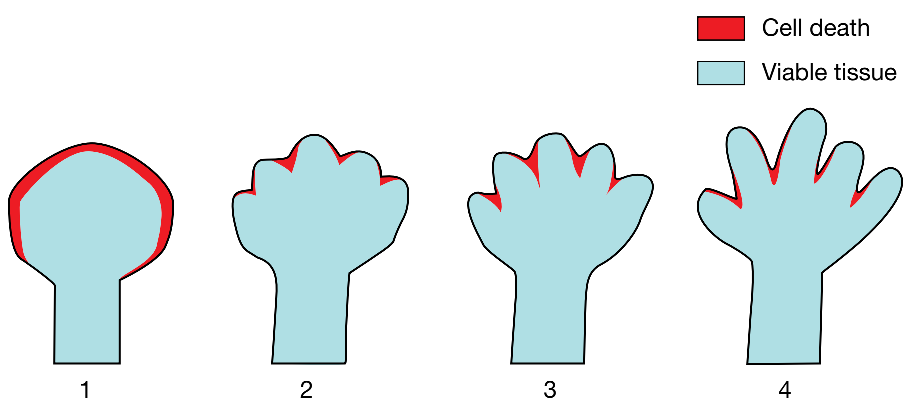



**FIGURE 5.59** Diagram showing the development of digits in a human embryo. The process starts with (1) a paddle-like limb bud and (2) the appearance of five ridges on the outer edge of the limb bud. (3) Apoptosis or programmed cell death continues in cells in the valleys (shown in red) so that the valleys become deeper. (4) By the eighth week of embryonic development, the five digits are well separated.

As well as programmed cell death (apoptosis), body cells may die by a process known as necrosis. **Necrosis** is unplanned cell death such as that occurring when cells suffer trauma and die prematurely. Necrotic cell death occurs when cells suffer mechanical damage or chemical trauma, thermal burns, frostbite (such as that occurring when the fluid in cells freezes and crystallises, rupturing their plasma membranes) and oxygen deprivation (such as that occurring in cardiac muscle cells when the flow of oxygen-rich blood is blocked during a heart attack). In uncontrolled necrotic cell death, the dying cells swell, their plasma membranes rupture and their cell contents pour out into the surrounding tissue, causing local inflammation. If an extensive volume of cells undergoes necrotic cell death and the blood supply is cut off, a condition known as gangrene can develop.

Check out figure 5.60, which shows a foot that has suffered frostbite. The blood supply to the ends of the toes has been destroyed and the cells have undergone necrotic cell death, releasing their contents and causing cell damage. Gangrene has developed, as shown by the blackened tissue. This tissue is permanently damaged and will not recover. In cases of severe frostbite, the affected areas may need to be amputated.

**FIGURE 5.60** Frostbite of the toes. The blackened tissue has been deprived of its blood supply, and the cells have undergone necrotic cell death.

See figure 5.61 below for a diagrammatic comparison showing the differences in cell death between apoptosis and necrosis.

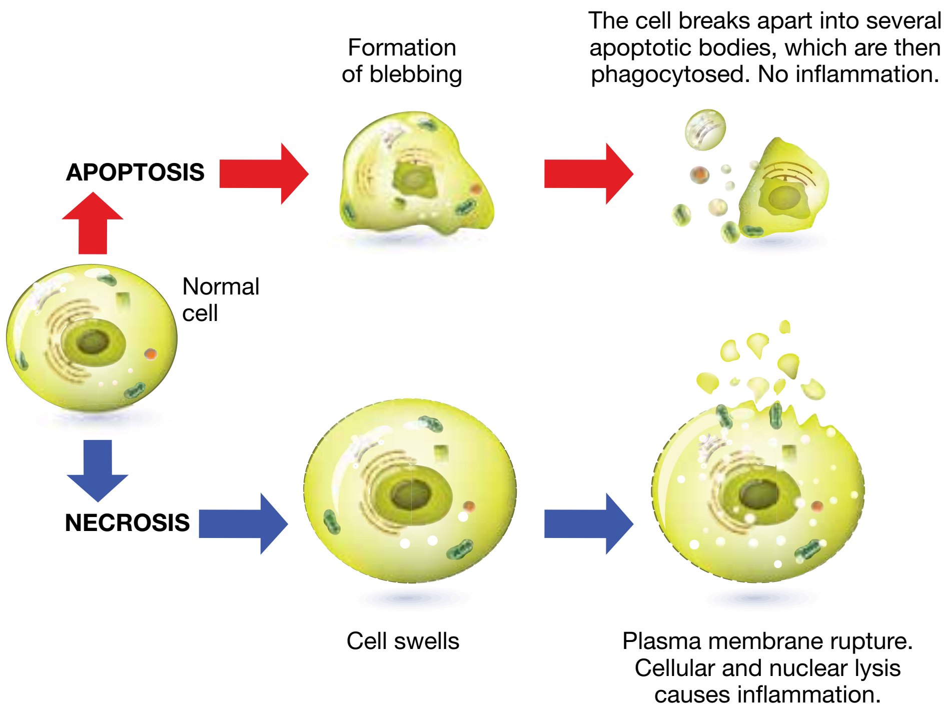



**FIGURE 5.61** Diagram showing differences in cell death by apoptosis and by necrosis. Death by apoptosis is an orderly process in which the cell is dismantled into apoptotic bodies that are engulfed by phagocytic cells. In necrosis, the plasma membrane bursts and the cell contents are spilled to the exterior, resulting in inflammation.

## Cell changes in apoptosis

The process of phagocytosis is detailed in chapter 7, page 291.

An animal cell undergoing apoptosis shows characteristic physical changes: the cell shrinks, its plasma membrane forms bubbles or 'blebs', the nucleus and cytoplasmic organelles condense and fragment, and the cell breaks up and its contents form into parcels called **apoptotic bodies** (see figure 5.62). Apoptotic bodies contain remnants of nuclear and cytoplasmic materials and they are engulfed by nearby phagocytic cells. This process prevents damage to surrounding cells and also enables efficient recycling of macromolecules.

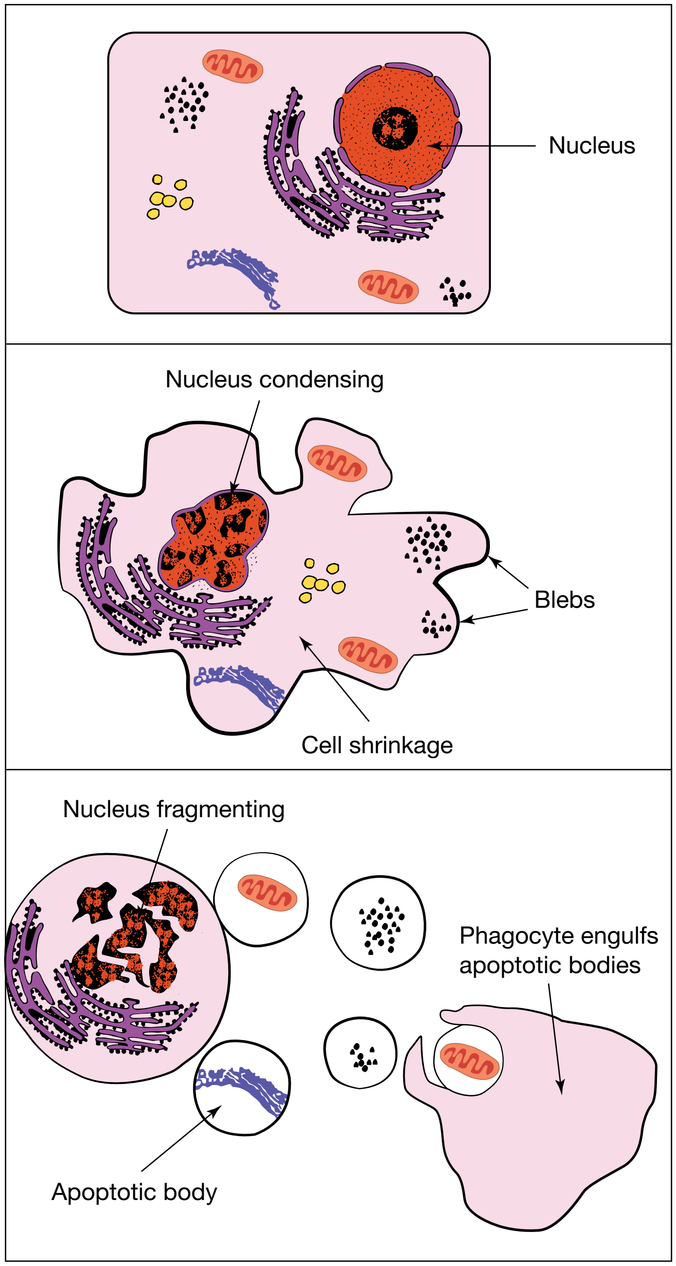



**FIGURE 5.62** Programmed cell death or apoptotic cell death. Note how the various cell components are broken down and packaged into apoptotic bodies that are engulfed by phagocytic cells. The removal of dying cells by phagocytes occurs in an orderly manner without producing the inflammatory response that occurs in necrotic cell death.

Figure 5.63 is a scanning electron micrograph (SEM) of lymphocytes (white blood cells) undergoing apoptosis. Note the blebs on the plasma membrane, which are one of the distinctive physical changes that can be seen in cells undergoing apoptosis.

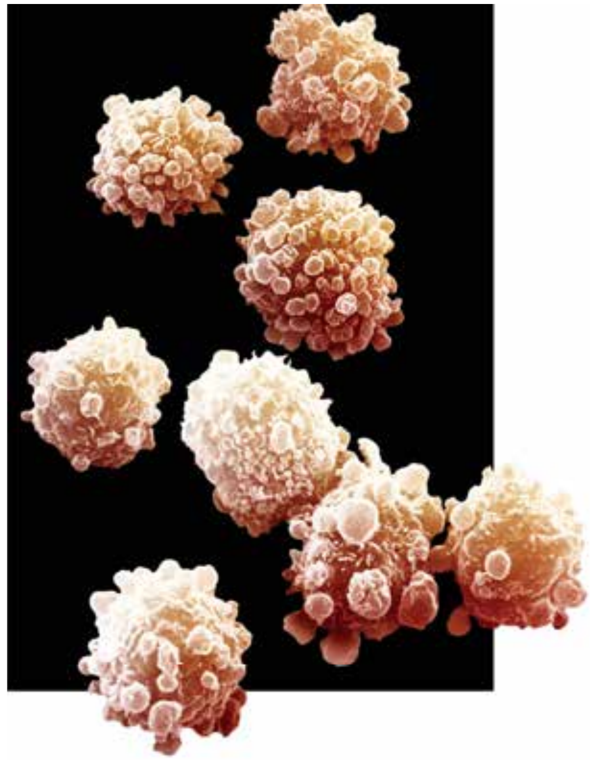


**FIGURE 5.63** White blood cells or lymphocytes undergoing apoptosis. The clue that these cells are undergoing programmed cell death is the presence of many blebs on their plasma membranes.

## Apoptosis pathways

Apoptosis can be initiated by one of two signal pathways:

1. the intrinsic pathway or mitochondrial pathway, initiated from within a cell
2. the extrinsic (outside-the-cell) pathway or death receptor pathway, initiated by factors external to the cell.

In both the intrinsic and the extrinsic pathways, a family of protease enzymes, known as **caspases**, drives the apoptotic death process. Caspases are present in cells in an inactive form, called procaspases. When activated by the removal of a small part of their molecular structure, procaspases are converted to caspases that are functional enzymes. Once activation of the first caspases starts, this sets in train a cascade of reactions that produces more caspases. This cascade can neither be stopped nor reversed.

Caspases are a family of cys protease enzymes that can cleave proteins at a particular amino acid residue (asp) near their active sites. The name, caspase, comes from **c**ys and **asp**art**ase**.

### The intrinsic apoptosis pathway (or mitochondrial pathway):

The intrinsic apoptosis pathway is also known as the mitochondrial pathway because it depends on factors released from the mitochondria. The mitochondria are important in determining how and when damaged and stressed cells recognise that it's time to die.

On occasion, cells come under stress resulting from factors such as DNA damage, viral infection, hypoxia, ultraviolet radiation, and deprivation of growth factors. Cellular stresses cause the cells concerned to activate the intrinsic (mitochondrial) pathway of apoptosis and kill themselves. For the organism concerned, the decision by its stressed cells to commit suicide by apoptosis is beneficial. Virus-infected cells, for example, remove themselves before the infection worsens and spreads, and this is done without damage to neighbouring cells such as would occur if stressed cells underwent necrotic cell death and released the virus particles.

When the intrinsic apoptosis pathway is activated, pores form in the outer membranes of the mitochondria of the cell concerned. These pores allow the release of cytochrome c from the mitochondria into the cytoplasm. (You may recall from chapter 4 that cytochrome c in mitochondria is an electron acceptor in cellular respiration.) In the cytoplasm, however, cytochrome c forms aggregates with other compounds and this activates caspase-9, the initiator caspase. (Activation chops a fragment from inactive pro-caspase-9 molecules, converting them to active caspase-9 enzymes.)

Active caspase-9 then activates other caspases (caspase-3, -6 and -7) that are the executioners. These executioner caspases then systematically dismantle the cell by fragmenting its DNA, cleaving the proteins of its nuclear envelope, cleaving the proteins of its cytoskeleton and Golgi complex. (What might cause the blebs to appear?) By the end, the executioner caspases have cleaved hundreds of proteins, and the cell contents have been dismantled into packages (apoptotic bodies) that are removed by phagocytic cells.

Figure 5.64 shows a highly simplified diagram of the intrinsic (mitochondrial) apoptosis pathway.

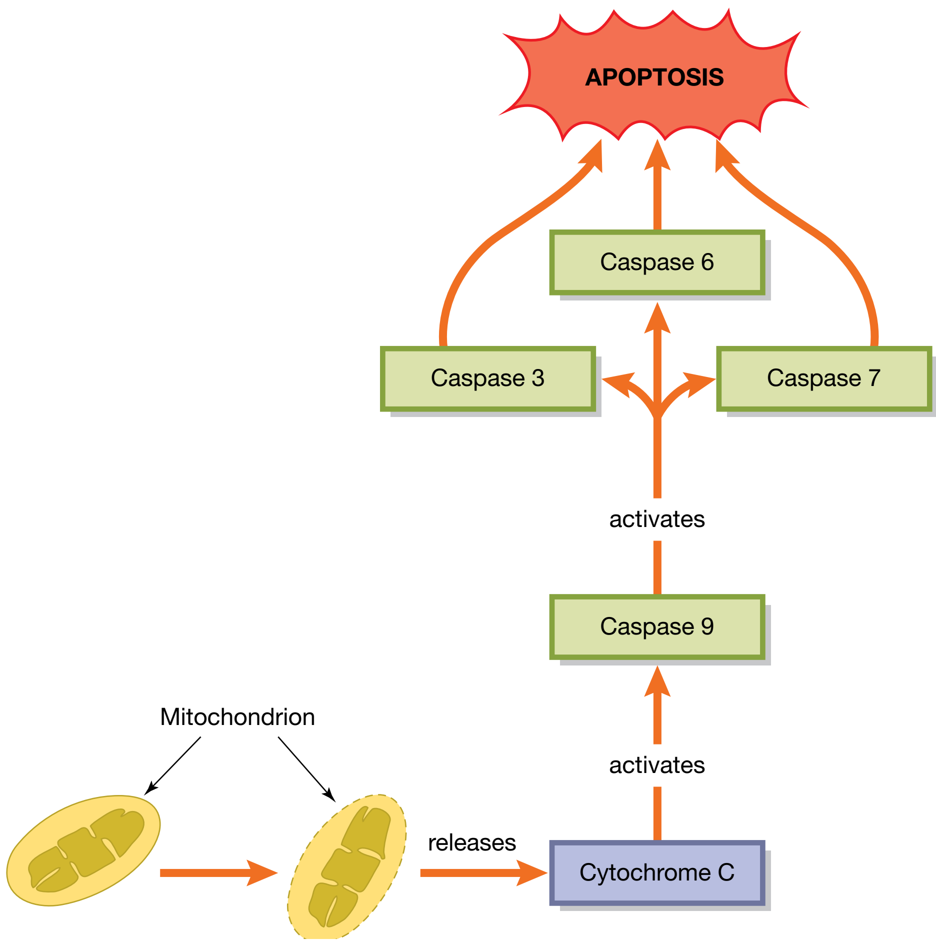



**FIGURE 5.64** Diagram showing a simplified representation of the intrinsic or mitochondrial apoptosis pathway. In reality, the process is far more complicated, but this diagram shows the key elements in this process.

## The extrinsic pathway (or death receptor pathway) of apoptosis:

The extrinsic or **death receptor pathway** is activated when molecules from *outside* a cell bind to death receptors on the plasma membrane of a cell and signal the cell: 'It's time to die'. Every nucleated cell carries within it the molecules for its own self-destruction. These are the inactive pro-caspases, but they become active caspases when the signal 'It's time to die' is received.

The cells that undergo programmed cell death by the extrinsic pathway are not stressed cells (as is the case for the mitochondrial pathway). Why send an external signal to healthy cells to self-destruct? In some cases, these cells are no longer needed, or they are being replaced by new cells, or they are in excess of current needs. The death of these cells is necessary to maintain the balance between cell production and cell loss.

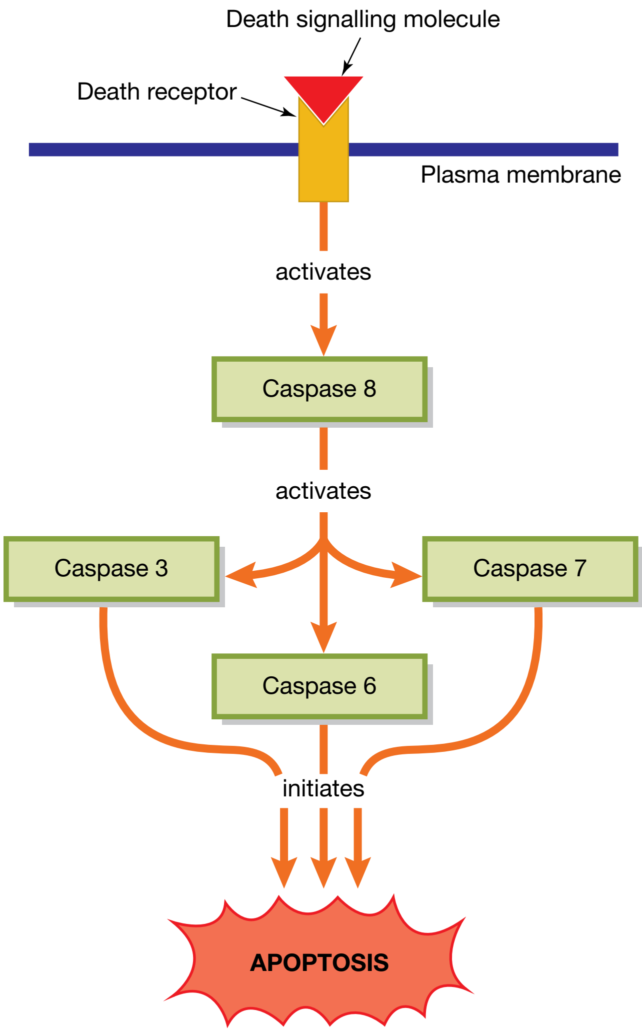



**FIGURE 5.65** Diagram showing a simplified version of the extrinsic or death receptor pathway of apoptosis. Note that caspase 8 is the initiate caspase that activates caspases 3, 6 and 7.

The binding of external signalling molecules (ligands) to the death receptors activates these receptors. This starts the extrinsic pathway of apoptosis. The first step is to cleave (cut) inactive pro-caspase-8 molecules in the cell and convert them to active caspase-8, the initiator. Caspase-8 then cleaves and activates the executioner caspases (caspase-3, -6 and -7). The executioner caspases dismantle the cell in an orderly manner and package its various components into apoptotic bodies (see figure 5.65).

In chapter 7 (see page 323), you will meet some immune cells called cytotoxic T cells. On the outer surface of the plasma membranes of these cells is a signalling molecule (ligand) called FasL. The signalling molecules bind to death receptors on target cells and this activates the extrinsic pathway of apoptosis.

### What about plants?

So far, the discussion of apoptosis in this chapter has related to mammals. So, what about plants? Since programmed cell death is a normal part of the life of multicellular organisms, it would be expected that apoptosis would play a role in the life of plants by enabling an orderly process of removal of unwanted or damaged cells.

Researchers have shown that plant cells go through an orderly process of cell death when treated with chemicals known to induce apoptosis in animals. Further investigations have revealed that the treatment of plant cells with ethylene induces apoptotic cell death. Further, the treatment of cells with inhibitors of ethylene biosynthesis block this chemically induced cell death. These experimental findings indicate that ethylene is a mediator of programmed cell death in some plant species at least.

## Apoptosis and human disease

The caspase proteases in a cell generate benefits when their activity is highly regulated. However, the complex process of apoptosis can malfunction. Many human diseases are now known to be the direct or indirect results of malfunctions of the apoptotic pathway. If the normal operation of apoptosis is disrupted so that apoptosis is reduced to inadequate levels, too little cell loss will occur and, instead, cells will live beyond their 'use-by-date' and accumulate abnormally. On the other hand, if apoptosis increases above normal levels, the excessive cell loss that results can cause disease.

Examples of diseases associated with too little apoptosis and an abnormal accumulation of cells include the following:

- Many cases of cancer appear to be the result of too little apoptosis. Mutations present in cancer cells enable these cells to grow in an uncontrolled fashion, crowding out normal healthy cells, and forming tumours. This situation arises because the cancer cells do not respond to the normal apoptotic signal to self-destruct.
- Autoimmune diseases appear to involve disruption of apoptosis; for example, rheumatoid arthritis is characterised by an excessive proliferation of cells in the synovial tissue of the joints that leads to joint destruction (see figure 5.66). These cells do not respond to the signals of the extrinsic apoptosis pathway so that a situation of too little apoptosis exists.


**studyon**

| | |
|---|---|
| Unit 3 | **Malfunction of apoptosis** |
| AOS 2 | Summary screen and practice questions |
| Topic 1 | |
| Concept 10 | |

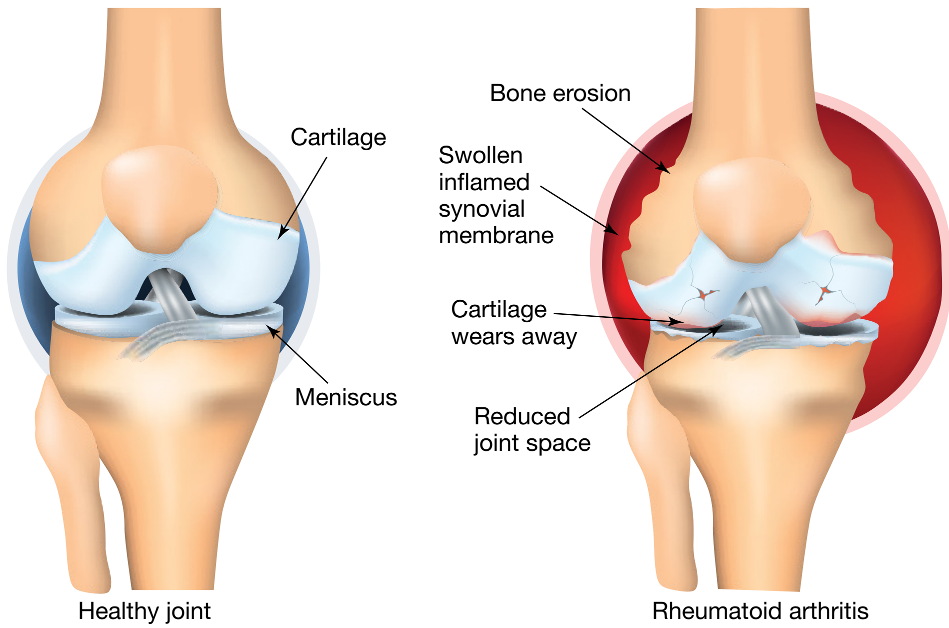



**FIGURE 5.66** Rheumatoid arthritis is a chronic inflammatory disease in which the synovial lining becomes inflamed and the adjacent bone and cartilage are destroyed. Alterations in the apoptosis of synovial cells is associated with the pathogenesis of rheumatoid arthritis.

Examples of diseases associated with too much apoptosis and an abnormal loss of cells include the following:

- In neurodegenerative disorders, such as Parkinson's disease and Alzheimer disease, abnormal progressive loss of nerve cells occurs in particular regions of the brain. The loss of nerve cells is associated with unregulated activation of caspases and apoptosis that produces excessive cell death.
- In untreated acquired immune deficiency syndrome (AIDS), excessive unregulated apoptosis leads to the loss of a particular group of immune cells, known as T-helper cells, also known as CD4+ cells.

AIDS is explored further in chapter 8, page 358.

## KEY IDEAS

- Apoptosis is a process of programmed cell death.
- Apoptosis can occur by one of two pathways, the mitochondrial pathway and the death receptor pathway.
- In an adult, apoptosis ensures that a balance exists between the rates of cell renewal and cell death.
- Uncontrolled apoptosis appears to be associated directly or indirectly with a number of human disorders.

## QUICK CHECK

**28** Briefly identify the important balance that is maintained by apoptosis.

**29** Identify a key difference between the members of the following pairs:
- **a** apoptosis and necrosis
- **b** mitochondrial pathway and death receptor pathway of apoptosis
- **c** a caspase and a pro-caspase.

**30** Which pathway of apoptosis will be undertaken by a cell that is stressed by a viral infection?

**31** Give an example of the following:
- **a** normal apoptosis in human embryonic development
- **b** normal apoptosis in an adult human
- **c** unregulated apoptosis associated with a human disorder or disease.

# BIOCHALLENGE

1 **Epinephrine signalling pathway:** When the blood glucose levels drop below the normal range, several hormones are released that raise the blood glucose level. These hormones include epinephrine (also known as adrenalin), cortisol, growth hormone and glucagon. Epinephrine is secreted by cells of the adrenal glands and its targets are liver cells.

The pathway from epinephrine to a cellular response in the form of an increase in blood glucose levels is a signal transduction pathway. In this case, the cellular response is the breakdown of glycogen into its glucose sub-units and their release into the blood, returning the blood glucose levels to within the normal range.

Figure 5.67 shows a diagrammatic representation of this signal transduction pathway that involves several steps, but only a few are shown here. (Note that you do not need to know the specific details, just the 'big picture'.)

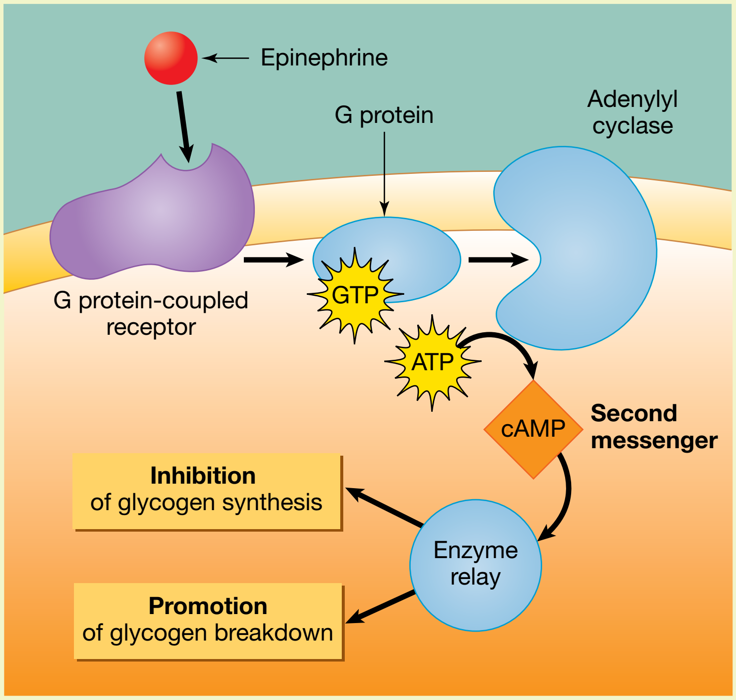



**FIGURE 5.67** Diagram showing *some* features of the signal transduction pathway that occurs after an epinephrine signal is received by a G protein linked receptor on the surface of a liver cell.

- a Identify the signalling molecule or first messenger in this pathway.
- b Where are the receptors that can receive this signal?
- c Identify the second messenger in this pathway.
- d Why is a second messenger necessary — why not just use the first messenger to continue to carry the message?
- e Would you expect that this signal transduction pathway would be continual or would you expect it to stop after a limited time? Explain your choice.
- f What is the cellular response to the transduced signal?
- g Would you predict that epinephrine is hydrophilic in nature? Explain.

2 An important part of a signal transduction pathway is the amplification of the signal. This is needed because signalling molecules (first messengers) are typically present in very low concentrations. Amplification is achieved through an enzyme relay that is a bit like a chain reaction. The amplification means that a small number of signalling molecules can produce a large cellular response (see figure 5.68).

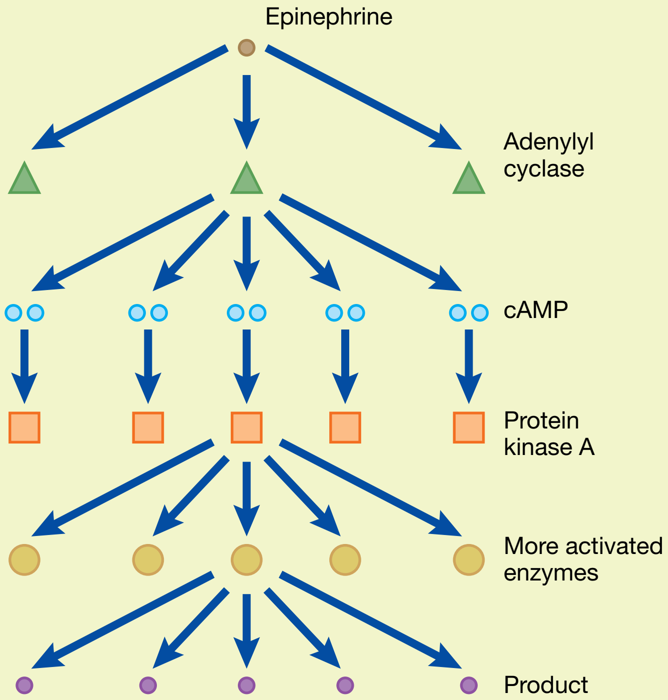



**FIGURE 5.68**

Amplification occurs at various steps in the signal transduction pathway.

- a Reception of one epinephrine molecule sets off a signal transduction pathway that produces a large number of cAMP second messenger molecules. Is this an example of amplification?
- b Each cAMP molecule activates one protein kinase A molecule. Is this a further amplification of the signal?

Each protein kinase A enzyme molecule then activates a large number of the enzymes that are the next members of the relay.

- c Explain why the action by the various enzymes in the relay can result in amplification of the signal.
- d The final product of this pathway is an enzyme called glycogen phosphorylase. Explain how this enzyme produces the cellular response.

# Chapter review

## Key words

abscission
abscission zone
apical dominance
apoptosis
apoptotic bodies
caspases
cell-surface receptors
cloning
cytokine
cytokinin
death receptor pathway
depolarisation
endocrine-disrupting chemicals (EDCs)
first messenger
flehmen
gibberellic acid (GA)
gibberellin
herbicide
hydrophilic
hydrophobic
intracellular receptors
jasmonates (JAs)
ligand
local mediators
mediator
modulator
motor end plate
necrosis
neurotransmitters
pheromone
polar
programmed cell death
rate of cell renewal = rate of cell death
reception
response
resting membrane potential
second messenger
signal transduction
signal-binding domain
stomata
synapse
synaptic clefts
tissue culture
transcription factors
transduction
tropism
vomeronasal organ (VNO)

## Questions

1 **Making connections between concepts →** Use at least eight of the key words from this chapter to construct a concept map relating to cellular communication.

2 **Communicating knowledge and understanding →** Examine figure 5.69, which shows a cell with two different receptors, S and R.

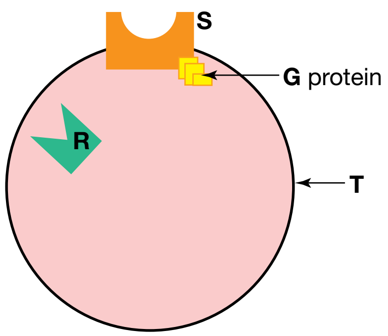



**FIGURE 5.69**

a Identify structure T.
b Identify two characteristics of the signalling molecule that has receptor R as its target.
c Identify two characteristics of the signalling molecule that has receptor S as its target.
d What name is given to the kind of receptor that is physically associated with a G protein?

3 **Applying principles in a new context →** Cholera is a bacterial disease caused by *Vibrio cholera*. These bacteria produce an exotoxin that enters particular gut cells and activates a signal transduction pathway, causing it to be to be switched 'on' permanently. This results in the continual activation of the CFTR protein channel that pumps chloride ions (Cl–) out of the gut cells. These negatively charged chloride ions are accompanied by the loss of an equivalent number of positively charged sodium ions ($Na^+$). This movement of NaCl into the gut drags with it large amounts of water from gut cells.

Suggest the likely effects of this particular malfunction of a signal transduction pathway caused by the cholera toxin.

4 **Developing and communicating explanations →** Examine figure 5.70 and suggest a carefully argued explanation for the cause of this condition of fused digits, which is known as syndactyly.

**FIGURE 5.70** Photo of the hand of a young child showing fusion of digits 3 and 4.

5 **Demonstrating knowledge and understanding** ➔ Consider each of the following statements and identify whether each is true or false. Where you judge a statement to be false, (i) give a reason for your decision and (ii) rewrite the statement so that it would be judged to be true.

- **a** One type of receptor can receive signals from many different kinds of signalling molecule.
- **b** A signalling molecule must cross the plasma membrane in order to be received.
- **c** Only some body cells respond to a given signalling molecule.
- **d** Communication between two neurons can occur in both directions across a synapse.
- **e** A given signalling molecule that is received by several different kinds of cell will produce the same cellular response in those cells.
- **f** Signalling molecules must be small enough so that they can reach the cell interior.

6 **Communicating knowledge and understanding** ➔ Examine figure 5.71 and tell the story that it communicates.

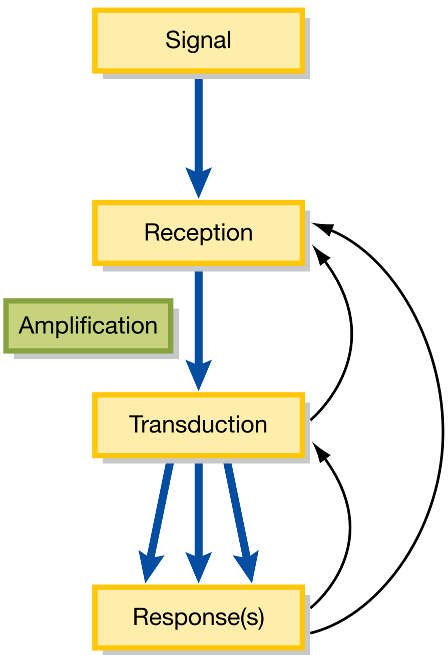



FIGURE 5.71

7 **Developing and communicating explanations** ➔ Suggest an explanation in biological terms for each of the following observations:

- **a** Caspase-9-deficient mice suffer from large brain outgrowths characterised by excessive neurons.
- **b** Pruning the shoots of a newly planted tree slows root regeneration.
- **c** The direction of travel of a nerve impulse along a nerve pathway is always one-way.
- **d** Hormones reach all cells of the body, but only particular cells respond to their signals.
- **e** Cytokines are sometimes called the 'messengers' of the immune system.

8 **Demonstrating understanding of biological terminology** ➔

- **a** Identify an essential difference between the members of the following pairs:
  - i pheromone and hormone
  - ii hydrophilic and hydrophobic signalling molecules
  - iii signal reception and signal transduction
  - iv an interferon and an interleukin
  - v a hormone and a local mediator
  - vi the intrinsic apoptosis pathway and the extrinsic apoptosis pathway.
- **b** For each pair, identify one feature that they have in common.

9 **Developing valid biological explanations** ➔ Give an explanation in biological terms for each of the following observations:

- **a** Grapes ripen more quickly if they are placed near bananas.
- **b** Some receptors for signalling molecules are located on cell surfaces while other receptors are intracellular.
- **c** Cytokine production as part of a localised response to infection is normal, but a system-wide production can be fatal.
- **d** Loss of nerve tissue occurs in neurodegenerative diseases such as Parkinson's disease.

10 **Applying understanding to familiar and new contexts** ➔ The drug tamoxifen is an anti-estrogen compound and is an example of a hormone blocker. This drug is used in the treatment of forms of breast cancer that have high levels of estrogen receptors. These cancers are said to be ER-positive, and their growth is enhanced by estrogen.

- **a** Where would these receptors be located in the breast cancer cells? Give a reason for your answer.
- **b** Identify a possible rationale for the use of tamoxifen in the treatment of ER-positive breast cancers?

11 **Making predictions** ➔ In 1926, the Dutch biologist Fritz Went (1903–1990) carried out experiments in total darkness in which he placed blocks of agar containing auxin in offset positions on the cut tops of the coleoptiles of grass seedlings (see figure 5.72). He also prepared some control coleoptiles.

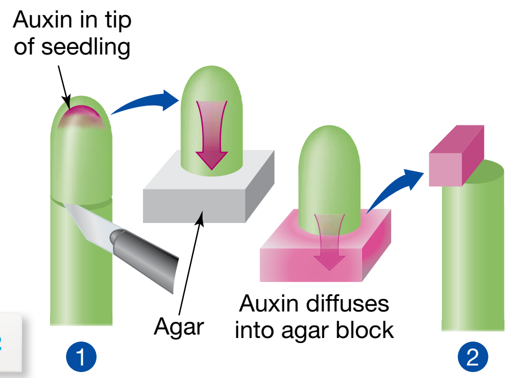



FIGURE 5.72

a Predict what would be expected to happen to the coleoptile (2) and give an explanation in biological terms for your decision.
b What control would be expected to be part of this experiment?
c What valid conclusion may be drawn from the results of this experiment?

12 **Applying knowledge in a new situation** → Patients with rheumatoid arthritis (RA) were treated with an anti-tumour necrosis factor (anti-TNF).
a To which group of signalling molecules does tumour necrosis factor (TNF) belong?
b Identify three characteristics of this group of molecules.
c Identify a possible rationale for use of anti-TNF in treatment of patients with RA.
d What effect, if any, would you predict that treatment with anti-TNF might have on patients with RA?

13 **Communicating information and ideas effectively** → Explain how pheromones could be useful:
a in reducing the use of insecticides in crop farming
b as control agents of insect pests in glasshouses.

14 **Interpreting data** → Figure 5.73 shows how the various responses to the plant hormones auxin and cytokinin depend on the relative concentrations of these hormones. Note that the term 'callus' refers to a mass of unorganised plant tissue consisting of cells of one kind.

Small shoots of a particular plant species were treated with solutions as follows:

Solution A: High concentration of auxins and low concentration of cytokinins

Solution B: High concentration of cytokinins and low concentration of auxins.
a Identify the growth response that would be expected on shoots treated with solution A.
b Identify the growth response that would be expected on shoots treated with solution B.
c At what relative concentrations of the two hormones did callus formation occur?
d What solution would have been used in this experiment to serve as the control?

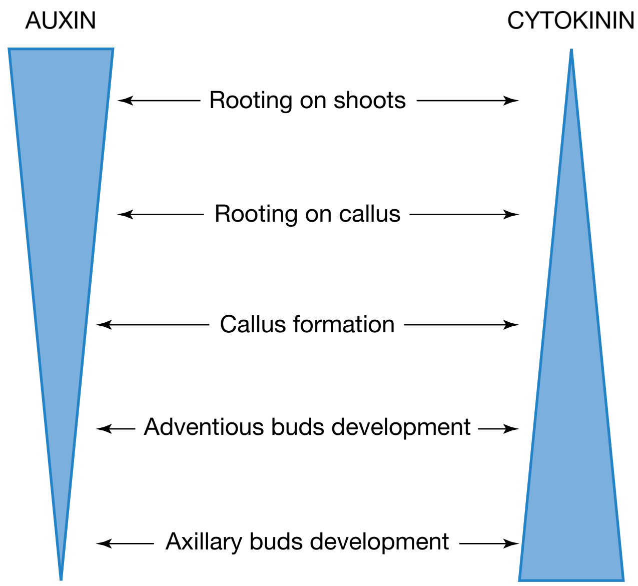



FIGURE 5.73 Various plant responses to different relative concentrations of the plant hormones auxin and cytokinin.

15 **Discussion question** → Some synthetic chemicals can bind to the receptors that are normally sites where natural hormones bind. These synthetic chemicals include DDT (dichloro-diphenyl-trichloroethane), bisphenols used in the manufacture of plastics, dioxin, and some PCBs (polychlorinated biphenyls). In the main, these synthetic compounds are lipophilic (hydrophobic) and persist in the environment. Some natural compounds including phytoestrogens from fungi and plants can also bind to hormone receptors on mammalian cells.

These chemicals, termed **endocrine-disrupting chemicals (EDCs)**, can interfere with the function of the normal hormones. Any system in the body controlled by hormones can be affected by EDCs. In a 2015 publication, the Endocrine Society identified evidence from several sources, including animal experimentation, for an association between endocrine disruption and conditions such as obesity, diabetes, abnormal puberty, female reproductive impairments, testicular dysgenesis, hormone-sensitive cancers in females, prostate cancer, and disorders of the thyroid, as well as neurodevelopment and neuroendocrine systems.
a Do a search on the web and find an example of an animal population in the wild that shows some reproductive abnormalities associated with EDCs.
b Suggest how an EDC might produce a particular disorder, such as precocious puberty in girls.

CHAPTER

# 6 Pathogens all around us


**KEY KNOWLEDGE**

This chapter is designed to enable students to:

- understand that pathogens are organisms or agents capable of producing a disease in a host they infect
- gain knowledge of the germ theory of disease
- recognise how causal relationships between infectious disease and microbes are established
- distinguish between cellular and non-cellular pathogens
- become familiar with the concept of antigens
- recognise the difference between self and non-self antigens
- recognise the difference between infection and disease.

**FIGURE 6.1** A laboratory technician working at a negative-pressure cabinet at the Uganda Virus Research Institute, handling viral-infected samples from the Ebola outbreak in July–August 2012. Note the personal protective equipment (PPE) that covers all body surfaces, including a filtered breathing system to prevent viruses entering the technician's airways. In this chapter we will explore human infectious disease and meet some disease-causing microbes and agents. (Credit: Centers for Disease Control and Prevention.)

# Bubonic plague or black death

Imagine you are living in Milan, Italy, in 1630, during an outbreak of the plague. Plague is spreading through the city and many people around you are dying. Your neighbour has fallen ill. A 'plague doctor', dressed in his distinctive garb, comes to the house next door to check your neighbour. Does your neighbour have the dreaded plague?

The doctor has a startling appearance. He wears a dark neck-to-toe overcoat, breeches, boots, gloves and a wide-brimmed hat, all made either of waxed canvas or leather, and he carries a long cane. What is most striking is the bird-like mask that the doctor wears over his face and through which he breathes (see figure 6.2 ). A closer look at the mask reveals that it has eyeholes and that the beak of the mask is filled with aromatic herbs or dried flowers as a protection against the bad-smelling air (**miasma**) that is believed to be the cause of the plague.


**FIGURE 6.2** Engraving, dated 1721, of a 'plague doctor' in 1656 showing his distinctive mode of dressing. Note the bird-like mask that was thought to protect against the plague. At that time, what was believed to be the cause of the plague? What was the nature of the 'protection' that the mask was believed to give?

In examining the ill man, the doctor does not touch the patient but, instead, uses his cane to shift the patient's clothing in order to check for the presence of very tender and swollen lymph nodes (**buboes**; singular = bubo; from the Greek *boubon* = groin) in the neck, groin and armpits of the patient. Their presence is an early symptom of this disease, known as **bubonic plague** or black death (see figure 6.3).

**FIGURE 6.3** Miniature from the *Toggenberg Bible*, published in Switzerland in 1411. This image, believed to illustrate people with the plague, shows numerous swellings or buboes on their bodies. In reality, swollen lymph nodes are typically restricted to the groin, neck and armpits. Note, in the background, the person throwing what are probably aromatic herbs or dried flowers in a futile attempt to dispel the foul or bad air that was then believed to cause plague.

**ODD FACT**

DNA analysis has shown the presence of the plague bacteria in human skeletons from locations across Europe. The oldest bones date back almost 5000 years ago to the Bronze Age.

The first recorded major outbreak of plague occurred in Constantinople (now Ankara, Turkey) during the sixth century AD (recurring in the seventh century AD) and is described in the writings of the historian Procopius, who lived between 500 and 560 AD. Procopius wrote that *'a bubonic swelling developed there in the groin of the body, which is below the abdomen, but also in the armpit, and also behind the ear and at different places on the thighs . . . the bubonic swelling became worse and the sufferer, no longer able to endure the pain, died . . . In some cases death came immediately, in others, after many days; and with some the body broke out with black pustules about as large as a lentil and these did not survive even one day, but all succumbed immediately. Vomiting of blood ensued in many, without visible cause, and immediately brought death.'*

The second major outbreak of plague originated in central Asia in the fourteenth century and spread from there to Europe. Perhaps the plague entered Europe at the port of Messina in Italy when a black rat (*Rattus rattus*) carrying plague-infected fleas (*Xenopsylla cheopis*) escaped via a mooring rope from the ship to the shore. In the period from 1347 to 1351, the plague spread across Europe, causing the deaths of an estimated 30 per cent to 50 per cent of the total population. Sporadic outbreaks of plague then continued over Europe until the mid-1700s. These outbreaks included the Italian Plague of 1629–31 and the Great Plague of London, with outbreaks occurring in 1603, 1625, 1636 and 1665 (see figure 6.4).

**ODD FACT**

During the Great Plague of London, when a person fell ill with the plague, that person and all family members were locked in their house for forty days, the door was marked with a red cross, and entry was restricted to ‘nurses’ who were local women paid to bring food.

**eBook*plus***

**Weblink**
Plague

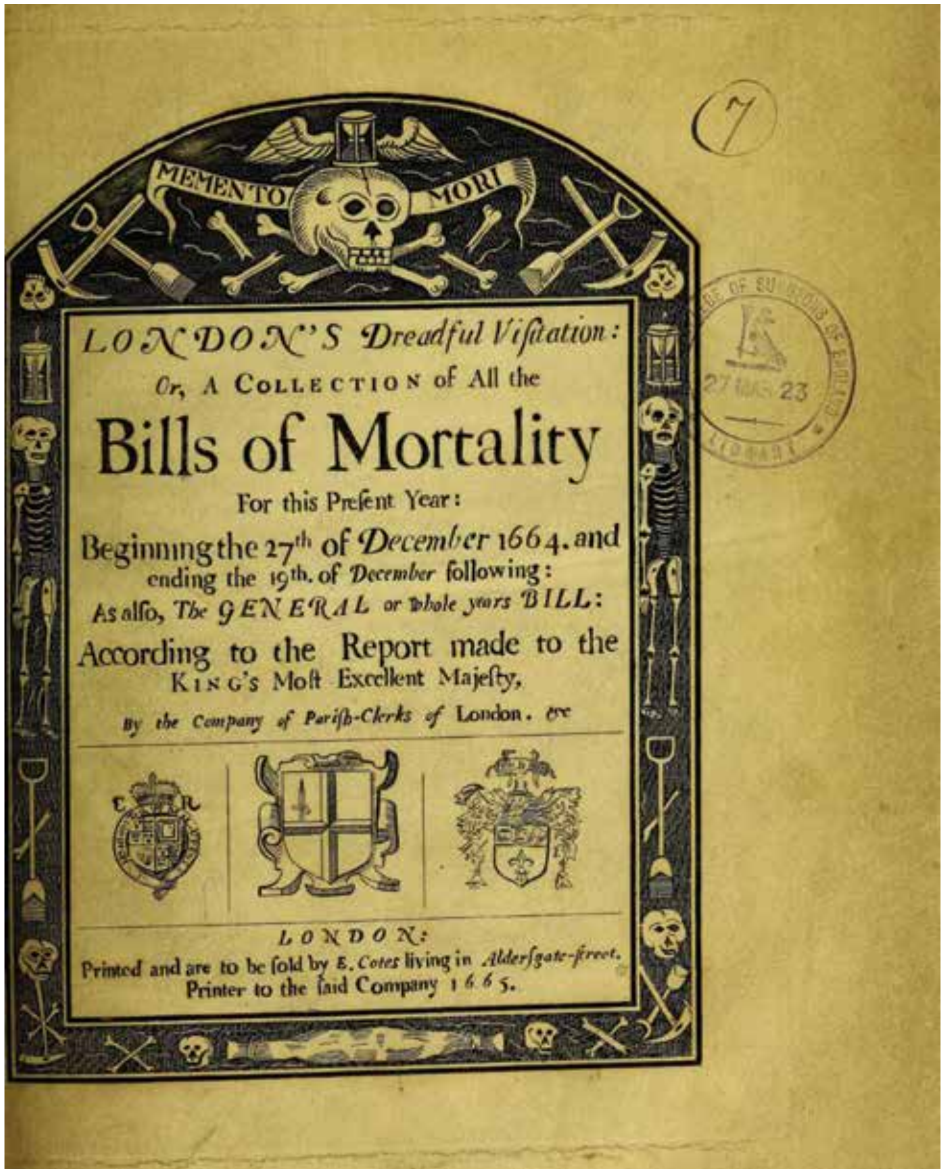



**FIGURE 6.4** Front page of the *Bills of Mortality* listing deaths by cause for the city of London for the period beginning 27 December 1664 and ending on 19 December 1665. (The symbol that looks like **f** is in fact **s**.) This publication reveals that the major cause of death during the week of 15–22 August was plague. The number of burials for that week was 1289 more than in the previous week. What does this suggest? (Image courtesy of The Wellcome Library.)

## Cause of the plague: views from the Middle Ages

We now know that plague is a disease resulting from an infection by a particular bacterial species, *Yestinia pestis,* the plague bacteria. Three forms of plague are recognised:

- Bubonic plague results when plague bacteria become established in the lymphatic system. Bubonic plague is most commonly transmitted to people by the bites of fleas that are infected with plague bacteria (*Yersinia pestis*) (see figure 6.5).
- Pneumonic plague involves the lungs and is spread when a plague-infected person coughs, producing airborne droplets containing plague bacteria that are breathed in by a person in close contact.
- Septicemic plague results when the blood becomes infected with plague bacteria, as may occur when plague bacteria gain entry to the body via a cut in the skin.

**ODD FACT**

The last major outbreak of plague originated in Asia in the mid-eighteenth century and spread to other countries, including Australia, in the period 1895 to 1930. Twelve outbreaks of plague were recorded in Australia between 1900 and 1925.

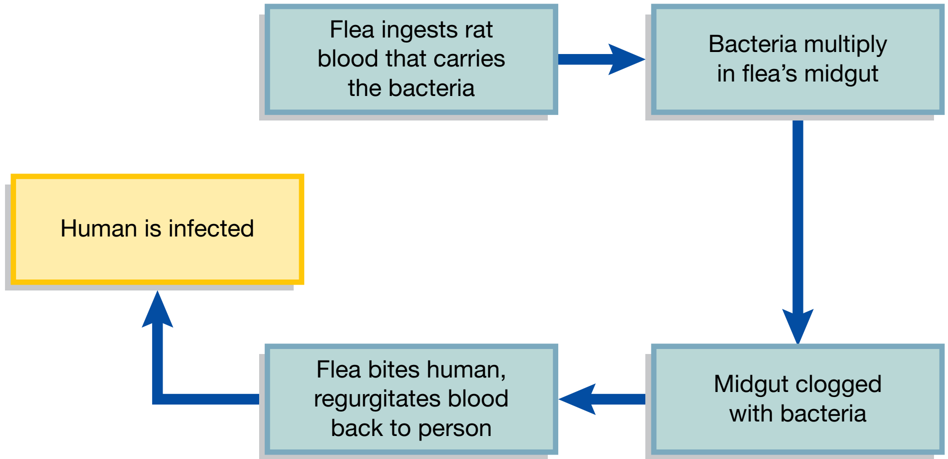



**FIGURE 6.5** Diagram showing the transmission of the plague bacteria *Yersinia pestis* to people. Rats (and other rodents) are the natural host of plague bacteria and act as reservoirs of the plague bacteria. Rat fleas become infected when they feed from an infected rat. The bacteria colonise the midgut of the flea and multiply there. Plague bacteria are most commonly transmitted to people through bites by these infected rat fleas. When an infected flea bites a person, blood from the person enters the flea's gut and mixes with bacteria. When the flea stops feeding, traces of this bacteria-infested blood are returned to the person bitten.

During the Middle Ages, the concept of disease-causing microbes was non-existent. Various causes of the plague were postulated. One cause drew on religion, with the plague being identified by church authorities as God's punishment for the sins of mankind. Another cause of the plague drew on astronomy, with medical authorities reporting to King Philip VI of France that the plague was the result of the alignment of the planets Mars, Jupiter and Saturn on 20 March 1345. Another cause was based on medieval medicine, with the plague due to an imbalance between the four so-called humors of which the human body was believed to be composed, namely blood, phlegm, yellow bile and black bile. (A balance between these four humors was thought to be essential for a person's mental and physical health.)

However, a commonly accepted view was that poisonous vapours (**miasmas**) identified by their foul smell were the cause of infectious diseases. Disease-producing vapours were said to originate from sources as diverse as earthquakes, swamps, rotting rubbish and decaying corpses. So, a commonly accepted means of preventing plague was to ward off the foul-smelling air or miasma through the use of sweet-smelling agents. This may seem strange to us, but, after all, which idea would have seemed more plausible to people of medieval times: the idea that diseases were caused by noxious vapours from rotting matter or that diseases were caused by invisible living organisms floating in the air? Noxious vapours were obvious to all because of their foul smell, and the rotting matter from which they emanated could be seen. Microbes were totally invisible to the unaided eye, so their existence was not suspected.

In order to protect against the plague, people placed bunches of dried flowers at the doors and the windows of their homes and scattered dried flowers and herbs around their houses to keep the air smelling sweet. When outside their homes, peasants wrapped dried flowers or herbs in a cloth and held it to their noses in order to avoid foul smells in the belief that this would protect them from the plague. For the same purpose, more wealthy people used ornate pomanders filled with aromatic herbs, such as rosemary, lavender,

**ODD FACT**

The 1594 publication *Present Remedies against the Plague* proposed that leaving a few peeled onions around a neighbourhood would absorb miasma (bad air) and so prevent the plague in that area. In 1636, the College of Physician in London identified pomanders as one means of preventing plague.

sage, mint, nutmeg or cinnamon (see figure 6.6a). Pomanders were held in the hand or worn on neck chains, or could be attached to a chain around a person's waist.

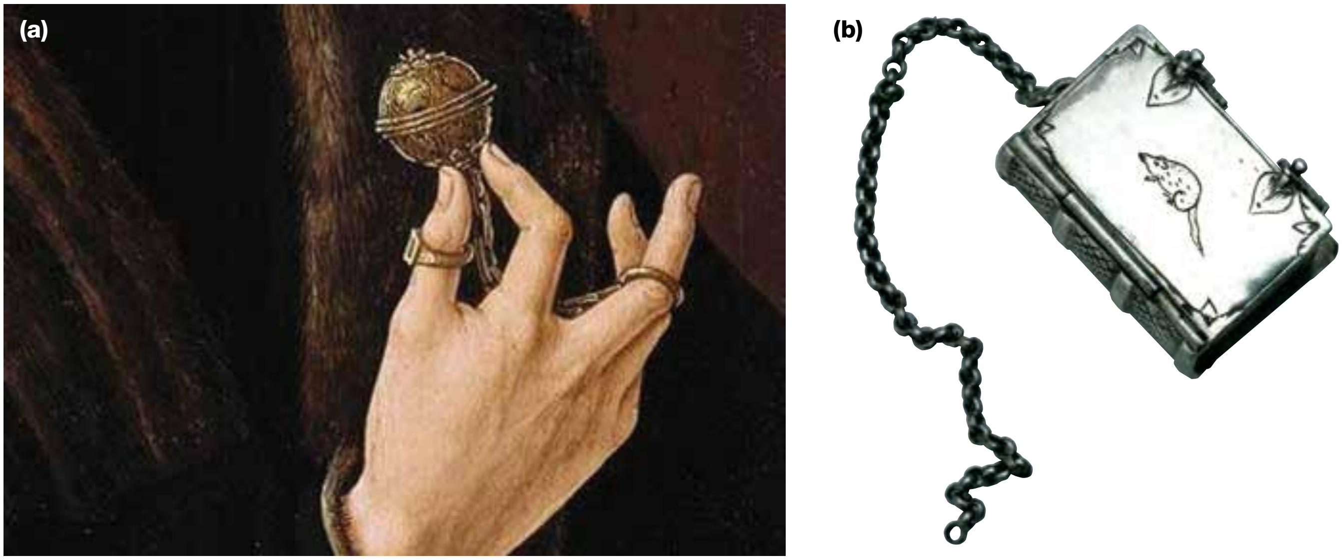



**FIGURE 6.6** **(a)** The first pomanders (from the French *pomme* = apple, and *d'ambre* = of amber) appeared in the thirteenth century. These pomanders were apple shaped and contained herbs mixed with a waxy substance, such as ambergris. This image shows a sixteenth-century pomander. Elaborately decorated pomanders also served as items of jewellery. **(b)** Pomanders from the seventeenth century appeared in other shapes, including skulls and animals. This seventeenth-century pomander is in the shape of a hinged book with a rat engraved on the top surface. Inside are several internal compartments to contain different aromatic herbs. The chain allowed the owner to attach the pomander to a waistband.

## The germ theory of disease

Our modern understanding of microbes as the cause of infectious diseases began to emerge in the mid-nineteenth century as a result, in particular, of laboratory investigations by the French chemist Louis Pasteur (1822–1895) and, later, by the German doctor Robert Koch (1843–1910) (see figure 6.7 a & b). In the 1880s, their research, supported by improvements in microscopes, led to the formulation of the **germ theory of disease**. **The germ theory of disease states that specific microbes cause specific infectious diseases**. This theory replaced the so-called **miasma theory of disease**, which identified foul-smelling vapours as the cause of infectious disease.

The concept of 'germs' had been proposed before Pasteur and Koch, but no experimental proof of the role of microbes as agents of disease had been demonstrated. **Remember that biological theory is not founded on untested proposals, but is built on testable hypotheses that are subject to revision and replacement**. Testing of hypotheses includes both observation and experimentation, and the results may either confirm a hypothesis or disprove it.

The investigations of Louis Pasteur were key to the development of the germ theory of disease. In the 1850s, the wine industry in France was at risk because of contamination. Pasteur demonstrated that this was the result of microbial action, and he introduced a process of heating wines to a temperature in the range of 50 to 56 °C to kill these microbes. This process of heating is now known as **pasteurisation** and, while not used in wine production, it is still used today in the treatment of many foods, in particular, milk, and in beer production. The observation that heating stopped the contamination supported the conclusion that the contaminating agents were living organisms.

**ODD FACT**

In Tudor England, pomanders were popular fashion accessories as well as safeguards against plague. Henry VIII, who ruled England from 1509 to 1547, is reported to have owned at least 16 ornate pomanders.

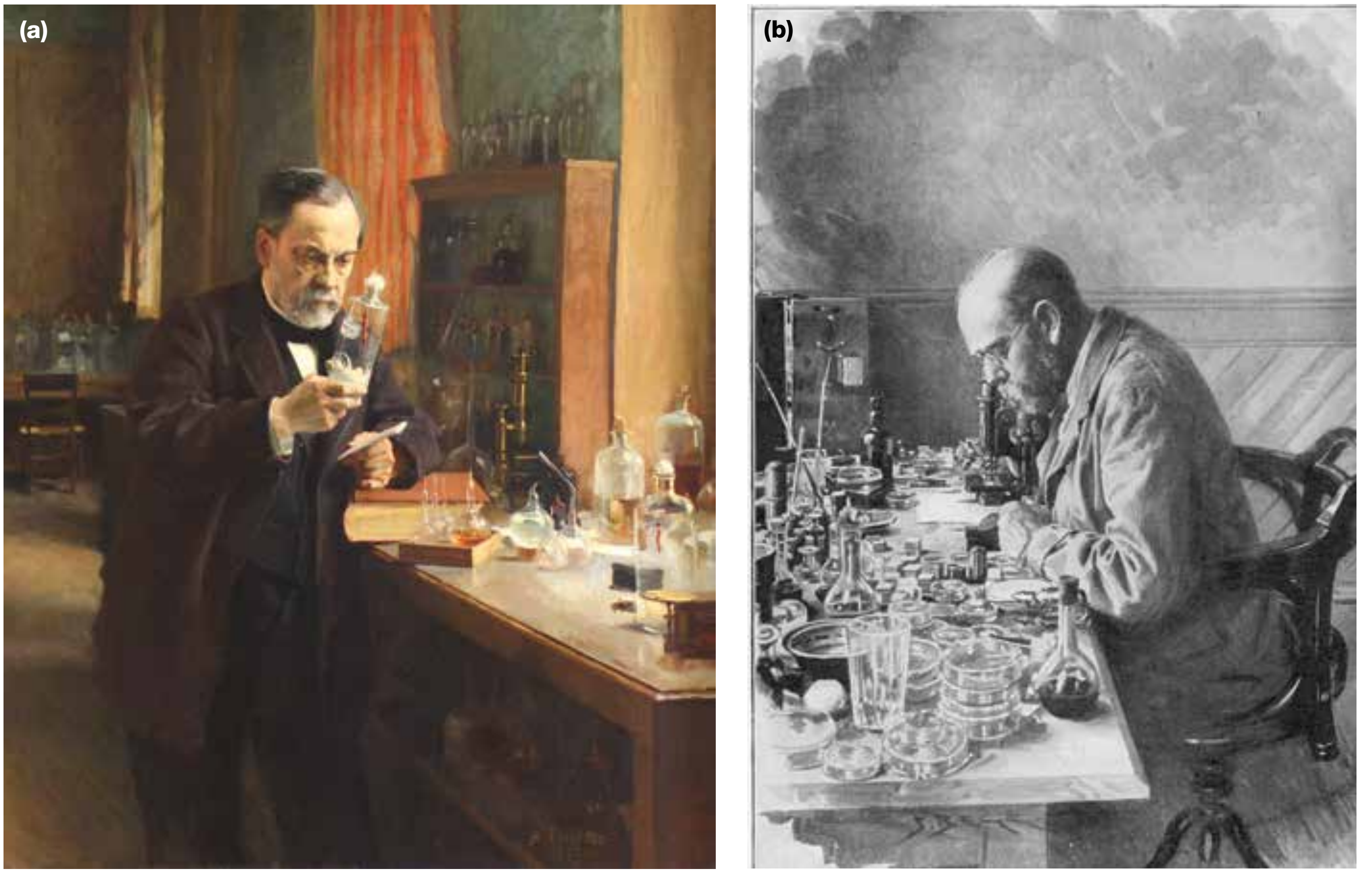



**FIGURE 6.7** **(a)** 1885 Painting of Louis Pasteur in his laboratory. **(b)** Robert Koch in his laboratory. Koch was an unknown country doctor in Germany until his wife gave him a microscope for his 29th birthday, which enabled him to explore the world of bacteria. In 1905, Koch was awarded the Nobel Prize in Medicine or Physiology for his contribution to the study of bacteria and the understanding of disease.

**eBook*plus***

**Weblink**
Read more about Robert Koch

In 1876, Robert Koch reported that anthrax was a disease resulting from infection by a specific bacterial species, *Bacillus anthracis*. Later, Koch showed that other diseases — tuberculosis (in 1882) and cholera (in 1883) — were also the result of infection by specific microbes. How might you show that a particular microbe is the cause of a specific disease rather than the result of the disease?

The approach of Robert Koch to this question was rigorous and definitive (see figure 6.8). For example, Koch's conclusion that anthrax was the result of a bacterial infection was based on the following careful observations and experimental results:

- Koch examined the blood of sheep that had died of anthrax and identified small rods that he believed were living microbes.
- He confirmed that these rods were not present in the blood of healthy sheep.
- Koch hypothesised that these bacteria were the cause of anthrax and he carried out experiments to test this hypothesis:
  1. He extracted bacteria from sheep that had died of anthrax.
  2. He grew these bacteria in culture, and injected them into mice. On every occasion, he found that the mice died of anthrax.
  3. He extracted the bacteria from the dead mice.
  4. He grew them in culture and injected them into healthy mice. Again, these mice died of anthrax.

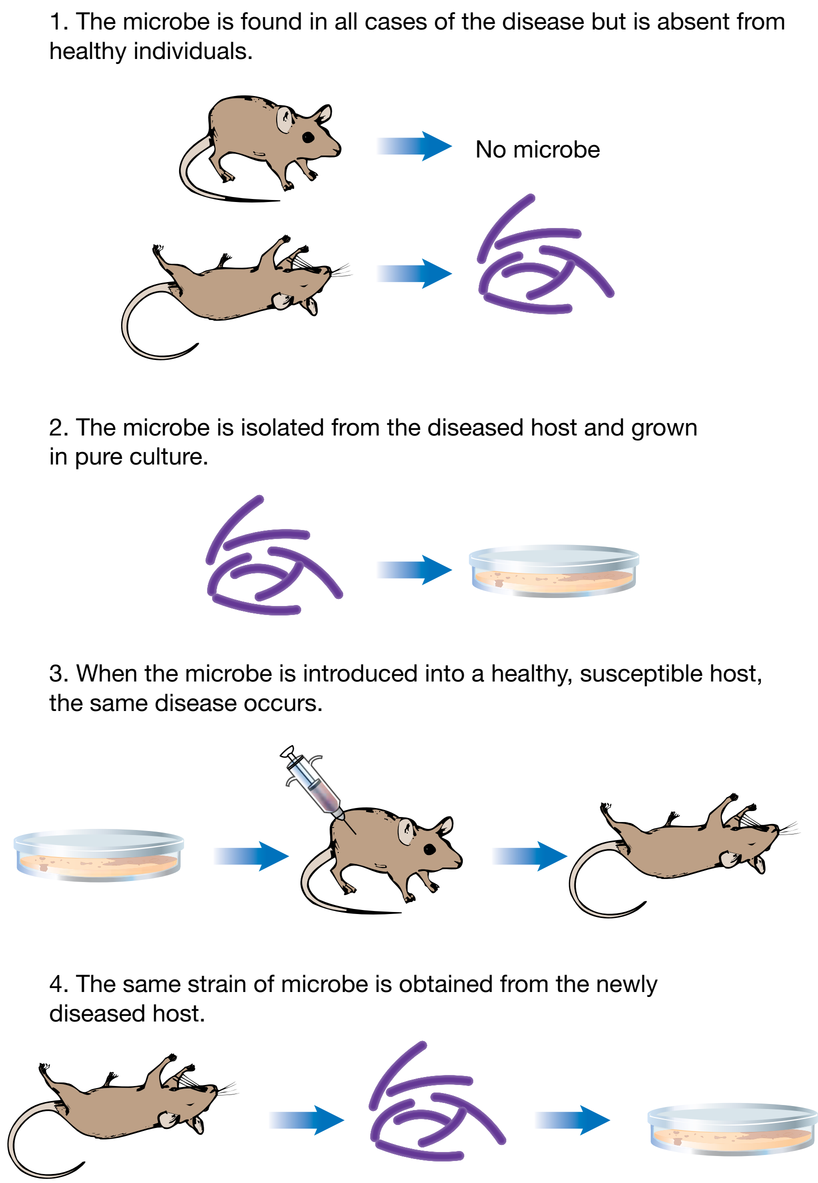



**FIGURE 6.8** The postulates developed by Robert Koch were used to establish that a particular microbe was the cause of a specific infectious disease.

What about the cause of bubonic plague? We now know that the cause of bubonic plague is a bacterium, *Yersinia pestis* (see figure 6.5). This bacterium was first discovered by a French-born Swiss bacteriologist named Alexander Yersin when he was researching a plague outbreak in Hong Kong in 1894.

## KEY IDEAS

- Until the late nineteenth century, infectious diseases were commonly believed to be caused by foul-smelling vapours, giving rise to the so-called miasma theory of disease.
- Pomanders were used in medieval times as one means of 'protection' against the bubonic plague.
- Major outbreaks of bubonic plague or black death have occurred in human history.
- In the mid-to-late nineteenth century, the germ theory of disease was developed, in particular, from the work of Louis Pasteur and Robert Koch.
- Koch's postulates provided a means of identifying a particular microbe as the cause of a specific disease.

**QUICK CHECK**

1 Who was the first person to provide experimental proof that specific microbes caused particular diseases?
2 Identify the following statements as true or false:
 a Black death is caused by foul vapours.
 b One symptom of bubonic plague is the appearance of swollen lymph nodes at particular sites on the body.
 c Bubonic plague is a bacterial infection involving the lymphatic system.
 d The germ theory of disease replaced the miasma theory of disease.
3 Briefly explain why pomanders were regarded as protection against the plague.
4 Identify a probable means by which the black death pathogen first reached Europe from Asia.

# Microbes are all around us

The term '**microbe**' collectively refers to microscopic organisms, including bacteria, archaea, protists and fungi with a cellular structure, and also includes viruses that have a non-cellular structure.

Microbes are present in the air that we breathe, the soil and water around us, decaying organic matter, the food we eat and the water we drink. Because they are not visible to an unaided eye, we are far less aware of the presence of microbes than the presence of macroscopic organisms. However, the **biomass** (total amount of living matter) of microbes is greater than that of all the animals and plants on Earth! In fact, we are part of a living world that, in terms of numbers, consists principally of microbes, but these are only made visible through use of microscopes.

In terms of bacteria, their concentrations in our surroundings have been estimated as follows:

- Air from the lower atmosphere (troposphere) contains up to 10 000 viable bacteria per cubic metre (see figure 6.9).
- Upper levels of ocean water contain an average of 500 000 bacterial cells per millilitre.
- Agricultural and forest soils have bacterial cells or spores with concentrations ranging from 100 million to 1000 million per gram — that's a teaspoon of soil!

**FIGURE 6.9** Colonies of various bacterial species growing on a nutrient-enriched agar plate. These bacteria were present in air in a public hall. Each discrete colony consists of millions of bacterial cells produced by multiplication of one or a few bacteria that settled onto the plate when the plate was briefly exposed to the air. Various species can be distinguished by features, including colour, form, opacity and edge (margin) of the colony.

**ODD FACT**

Microbial cells on and in the human body outnumber our human cells by 10 to 1. However, because of their much smaller size, this microbial population makes up just one to three per cent of the mass of a human body.

**ODD FACT**

When added to milk, bacteria, such as *Lactobacillus bulgaricus* and *Streptococcus thermophiles*, ferment the milk sugar (lactose), producing lactic acid, which converts the liquid milk into gel-like yoghurt. Evidence suggests that people have been making yoghurt since at least 500 BCE.

In addition, all plants and animals possess their own microbial communities, known as **microbiota**, on and in their bodies. The first microbiota community is acquired during a baby's passage through the mother's birth canal, and further microbes are continually acquired throughout life. The human microbiota is most concentrated in the gut but is also present on the skin and on the mucous membranes that line the respiratory and the alimentary tracts of the body:

- The human gut is home to an estimated 100 trillion microbes, mainly bacteria. In the colon of the gut, where most bacteria are found, the numbers are estimated at $3.2 \times 10^{11}$ per gram wet weight.
- On most of the human skin, the microbial density is about 1000 to 10 000 microbes per square centimetre, while on the skin of armpits and groin, the number is estimated at 1 000 000 per square centimetre.

The microbial community of the human body is normally confined to the surface of the skin, and surfaces of the mucous membranes that line the mouth, nose, throat and trachea of the respiratory system, and that line the genital tract and the gut. **All the body cells and the various body fluids, namely the plasma of the blood and the interstitial fluid, are sterile (microbe-free).** Maintaining this sterility is critical to good health and even survival, and the defences of our immune system are critical in ensuring this.

We are constantly exposed to the world of microbes. The majority of these microbes are harmless, and many microbes are beneficial, as, for example, our normal gut microbes that play a key role in both our digestive health and our general health, the decomposer microbes in soil and water that play a key role in the cycling of matter within ecosystems, and the nitrogen-fixing bacteria that convert nitrogen in the air to nitrogenous compounds that are essential for plants. In addition, many microbes are important in the food industry, such as the bacteria involved in producing sauerkraut, pickles, olives, cheese and yoghurt (see figure 6.10), and the yeasts involved in producing bread and wine.

**FIGURE 6.10** Yoghurt is essentially milk in which sugar (lactose) has been converted to lactic acid in the process of fermentation by bacteria such as *Lactobacillus bulgaricus* and *Streptococcus thermophiles*.

## Disease-causing microbes and agents

A small number of the microbes — less than one per cent — are harmful and can cause diseases. These microbes are said to be **pathogenic** (*pathos* = disease; *gen* = producing). In addition, non-cellular agents, such as viruses and prions, are also known pathogens. Despite their small number, some pathogens produce disease and death annually on a large scale; for example, estimates by the World Health Organization (WHO) indicate that each year in developing countries more than 200 000 newborns die from tetanus caused by the bacterial species *Clostridium tentani*.

Check out table 6.1 to see the approximate numbers of different pathogens that are responsible for human diseases. **Note that infection by a pathogen typically causes a disease in just one specific host or a few specific hosts.** The same microbes in other hosts may have few, if any, negative effects; for example, the human immunodeficiency virus (HIV) causes the serious disease AIDS (acquired immunodeficiency syndrome) in people, but, in various monkey hosts, the same virus does not cause the same serious disease progression seen in people with untreated AIDS.

**TABLE 6.1** Different kinds and numbers of human pathogens (disease-causing microbes and agents). Note that human pathogens include both (i) cellular agents, mainly bacteria, fungi, protists and some multicellular parasitic helminth worms, and (ii) non-cellular agents, mainly viruses, that lack a cellular structure. (It is interesting to note that no evidence has yet been found of a human disease caused by an archaeal species.) (Ref: Aminov, R.I., 'Role of Archaea in Human Disease,' *Frontiers in Cellular and Infection Microbiology*, 3 (2013): 42.)

| Human pathogens | No. of disease-causing species |
|---|---|
| ***CELLULAR:*** | |
| • bacteria | 541 |
| • fungi | 325 |
| • protists | 57 |
| • parasitic helminth worms | 285 |
| ***NON-CELLULAR:*** | |
| • DNA viruses | 36 |
| • RNA viruses | 153 |
| • prions | 2 |

## Infection is not disease

If a pathogen can gain entry to the human body and reach the target cells, the pathogen may multiply rapidly and produce an infection. If the body's immune system cannot overcome an infection, the infection will develop into a **disease**. So, an infection is not equivalent to a disease, but infection is a necessary pre-condition for an infectious disease.

**An infection occurs when pathogenic bacteria, viruses or other microbes gain entry to the human body and, once within the body, begin to multiply.** In some cases, the infection is eliminated by the body's defence (immune) system, so not all infections will develop into a disease.

**A disease occurs only after body cells or organs are damaged by the infectious microbes and the signs and symptoms of the specific infectious disease are visible.** These various signs and symptoms result from damage to the host caused by the increasing numbers of pathogens.

Before an infection can develop into a disease, several events must occur. The pathogen, whether bacterial, fungal or viral must first:

- gain entry to the body and reach target site(s) in the body; typically, the portals of entry are the skin, the mucous membranes of the respiratory, uro-genital, and digestive systems of the body
- overcome the defence mechanisms of the body
- become established at one or more sites
- multiply rapidly, causing harm to the host and producing the symptoms of the disease.

Harm to the host may result from cell damage and cell death locally at the site of infection, or from the actions of toxins that spread through the body and cause systemic cell damage and death. Which would be expected to produce more harm?

The period after infection and before the first symptoms of a disease appear is called the **incubation period**. **An incubation period is the interval between a person's exposure to a pathogen and the onset of disease symptoms in that person.** During the incubation period, the disease-causing agent multiplies to concentrations that are sufficient to produce the symptoms of the disease.

Incubation periods can vary. For many pathogens, the incubation period is short; for example, the symptoms of acute gastroenteritis caused by a *Salmonella typhimurium* bacterial infection typically appear within 12 to 72 hours after a person is infected. For a few pathogens, however, the incubation period can be months or even years. For example, the serious symptoms of AIDS may appear several years after a person is infected with the human immunodeficiency virus (HIV). Likewise, variant Creutzfeldt-Jakob disease (vCJD) has an incubation period that can exceed 30 years! This means that a person can be infected with the prion that causes vCJD, but the disease symptoms may not appear until decades later.

Table 6.2 shows the incubation periods for several infectious diseases.

**TABLE 6.2** Incubation periods for several diseases. Can you suggest how long a person who has been exposed to the Ebola virus should be kept in quarantine before that person can be said to be free of Ebola virus disease?

| Disease | Pathogen# | Incubation period## |
|---|---|---|
| Salmonellosis | B | 6 to 72 hours |
| influenza | V | 1 to 3 days |
| common cold | V | 1 to 3 days |
| diphtheria | B | 2 to 5 days |
| bubonic plague | B | 2 to 6 days |
| Ebola virus disease | V | 2 to 21 days |
| tetanus | B | 4 to 21 days |
| meningitis | B | 7 to 21 days |
| chickenpox (varicella) | V | 7 to 21 days |
| typhoid | B | 8 to 14 days |
| measles | V | 10 to 21 days |
| rubella | V | 14 to 21 days |
| pertussis (whooping cough) | B | up to 21 days |
| tuberculosis | B | 2 to 12 weeks |
| HIV | V | weeks to months |
| kuru | prion | 10 to 13 years |

# B = bacteria; V = virus

##Values given as usual or typical incubation periods.

Some pathogens can be transmitted from an infected person to a healthy person only after the infected person shows visible symptoms of the disease. These pathogens include those responsible for diseases such as Ebola virus disease, typhoid and pertussis. In contrast, other pathogens, including those causing diphtheria, rubella, influenza, measles and tuberculosis, can be transmitted, not only by people showing obvious disease symptoms, but also by infected people during the incubation period when they show no symptoms of the disease.

In addition, some people can be infected by a pathogen but be in good health and never show any signs or symptoms of the disease concerned. Such people are said to be **asymptomatic carriers** of the pathogen concerned and they can be a source of infection of people with whom they come in contact. Perhaps the most famous asymptomatic carrier of an infectious disease was Mary Mallon (1869–1938), a domestic cook, who was infected with the *Salmonella typhi* bacteria, which cause typhoid fever. The following box tells the story of this asymptomatic carrier of typhoid.

**eBook*plus***

**Weblink**
Mary Mallon

## MARY MALLON, AN ASYMPTOMATIC CARRIER

Mary Mallon (1869–1938) worked as a cook for several wealthy families in New York City in the early 1900s. In each household where she was employed, family members came down with typhoid fever. Health authorities recognised the connection between Mary, the cook in each affected household, and the outbreak of typhoid fever in the family members in that household. This was the first time that an asymptomatic carrier of a pathogen had been identified, in this case, an asymptomatic carrier of the bacteria *Salmonella typhi*, the causative agent of typhoid fever. Asymptomatic carriers are perfectly healthy and show no symptoms of the disease, but they can transmit the bacteria to other people.

Mary is presumed to have transmitted the typhoid bacteria to about 47 people, three of whom died from typhoid fever. Typhoid bacteria are typically present in the urine and faeces of an asymptomatic carrier, and these bacteria can be spread to others if the carriers do not take care to wash their hands after using the toilet and before handling food. Because typhoid bacteria are generally killed at temperatures used to cook food, it is likely that Mary spread the bacteria when she prepared one of the uncooked desserts that was greatly favoured by her clients — ice cream containing cut-up pieces of raw peaches.

In 1907, Mary was arrested and forcibly placed in quarantine. Her first quarantine period was on a small island in the East River where she remained for more than two years, from 1907 to early 1910. The label 'Typhoid Mary', by which she is commonly known, first appeared in an article published in an issue of the 1908 *Journal of the American Medical Association*. Some newspapers described her as 'crawling with typhoid bugs'.

Mary argued for her release from quarantine. In one letter she stated that she was '*kept as a prisoner without being sick nor needing medical treatment*' and felt that she was '*a peep show for everybody*'.

Mary was released from this quarantine on condition that she never again work as a cook. On release, Mary first worked in a poorly paid job in a laundry but eventually returned to the better-paid position of cook. She also changed her name — not surprisingly, since she carried the derogatory and widely publicised label of 'Typhoid Mary'. An outbreak of typhoid fever in a hospital in 1915 was traced to the hospital cook, a Mrs Brown, who was, in fact, Mary Mallon. Once Mary was recognised, she was immediately placed in quarantine again. This second period of quarantine lasted from 1915 until Mary's death in 1938.


**FIGURE 6.11** An article in the New York newspaper *New York American* on 20 June 1909. The headline continues, '. . . *a Prisoner on New York's Quarantine Hospital Island, Not Because She Is Sick, But Because She Breeds Typhoid Germs and Scatters Them Wherever She Goes.*'

Mary Mallon's story raises the issue of setting an appropriate balance between protecting a person's civil liberties and protecting the heath of the community.

## Diseases: transmissible or not?

So far, the discussion in this chapter has centred on those diseases that result from an infection of the human body by pathogenic agents, such as bacteria, viruses, protists, fungi, parasitic helminth worms or prions. The majority of these **infectious diseases** are also known as **transmissible diseases** or **communicable diseases**.

An infectious disease can be defined as '*an illness caused by a specific infectious agent or its toxic product that results from transmission of that agent or its products from an infected person, animal, or reservoir to a susceptible host, either directly or indirectly through an intermediate plant or animal host, vector or inanimate environment*' (Source: Last, J.M., ed., *A dictionary of epidemiology*, 4th ed. (New York: Oxford University Press, 1988).)

However, it is important to recognise the existence of **non-infectious** or **non-communicable diseases**. These diseases cannot spread from affected people to healthy people via the environment.

Non-communicable diseases have several causes and include:

- **environmental diseases** caused by agents in the environment, such as lung cancer caused by exposure to smoke carcinogens, and mesothelioma caused by exposure to asbestos fibres
- **nutritional deficiency diseases** resulting from diets that are deficient in a particular essential component, as, for example, scurvy, which results from a vitamin C deficiency, and pellagra, which results from a deficiency of vitamin B3 (niacin)
- **inherited diseases** that result from mutant alleles or chromosomal aberrations inherited by children from one or both parents via eggs and/or sperm.

Can you think of other examples of non-communicable diseases?

Infectious diseases may spread from infected people to healthy people by various means, as, for example:

by **direct transmission**, such as

- by person-to-person contact, such as may occur through kissing or sexual contact

by **indirect transmission**, such as

- by airborne droplets or particles, as, for example, an uncovered sneeze or cough
- by contact with contaminated objects, such as bedding, cups and dishes, or medical instruments
- by ingestion of contaminated food or water
- by biological vehicles, such as contaminated blood, sputum, or faeces
- by **vectors** that carry pathogenic agents and spread them to people through bites from infected ticks, mites, fleas or mosquitoes, or through contaminated particles that they leave on material, such as fly droppings on food. In many cases, these infected vectors transmit pathogens from their natural hosts (birds, bats, rodents or monkeys) to people. Diseases of this type are called **zoonotic diseases** or **zoonoses** (*zoo-* = animal; *nosos* = disease). Examples of zoonotic diseases include bubonic plague, psittacosis, anthrax, West Nile virus disease and rabies.

Figure 6.12 shows some common forms of transmission of disease-causing agents.

If sewage treatments are inadequate, water may be contaminated with pathogens from human faeces. Fish that swim in contaminated water and shellfish that filter feed in contaminated water can carry pathogens. If contaminated fish or shellfish are eaten, these pathogens can infect people. Food may also be contaminated by infected food handlers, such as cooks and waiters, who do not carry out adequate procedures such as hand washing.

**ODD FACT**

Until the twentieth century, infectious diseases were the greatest health concern worldwide. Since then, in developed countries, non-infectious and degenerative diseases have emerged as health issues of major concern. Can you suggest factors that may have caused this change?

**zoonotic disease**: disease that normally exists in vertebrate animals but that can infect humans either by direct transmission or via an intermediate vector.

**ODD FACT**

In 1893, Theobald Smith and F. L. Kilbourne were the first to demonstrate that pathogens could be transmitted to people through the bite of an infected organism. They showed that *Babesia microti*, the cause of babesiosis, a malaria-like disease, was transmitted by the bite of an infected tick. More than 150 zoonoses have now been identified.

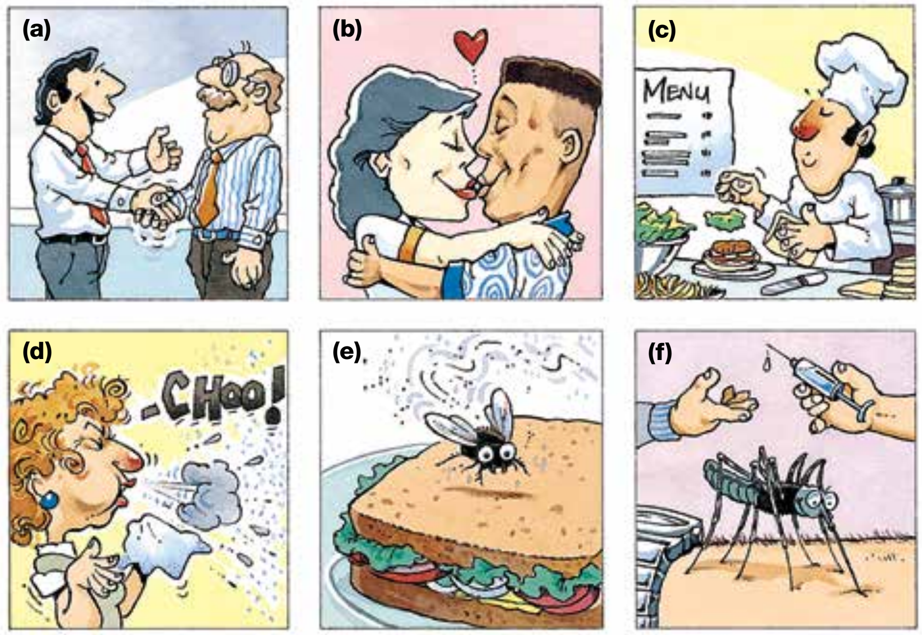



**FIGURE 6.12** The most common means of transmission of an infecting agent from one host to another are **(a)** body contact, **(b)** sexual contact, **(c)** food and water, **(d)** droplet or airborne, **(e)** carrying vector and **(f)** injecting vector.

**FIGURE 6.13** A female mosquito on human skin in the process of biting. The mosquito drives two thin parallel tubes through the skin to find a blood vessel. One tube forces saliva into the tissue (to stop blood vessels contracting and to prevent blood from clotting) while the other tube pumps blood from under the skin. If the mosquito is infected, its saliva can introduce pathogens into the person.

## Entry points of pathogens to the human body

Pathogens *may* gain entry to the human body via:

- the nose and throat (naso-pharynx) into the ducts and tubes of the respiratory system
- the mouth into the digestive tract
- the urethra or via genital openings into the genital tract.

Note that the cavities (lumens) within these ducts, tubes and tracts are still outside the body. To gain entry to the body requires that the pathogens cross the cellular boundary to the internal environment of the body.

Cuts, punctures, burns or insect bites (see figure 6.13) that penetrate the skin give pathogens direct entry to the internal environment of the body.

## KEY IDEAS

- Not all diseases are infectious diseases; other kinds of diseases exist.
- Infection differs from disease and precedes a disease.
- The incubation period of a disease is the interval between a person's exposure to a pathogen and the onset of disease symptoms in that person.
- Asymptomatic carriers of a disease show no symptoms of a disease but can spread it.
- Pathogens can enter the human body via various entry points.
- Infectious diseases may be spread directly (by person-to-person contact) or they may be spread indirectly.

**studyon**

| | |
|---|---|
| Unit 3 | **Infectious diseases and pathogens** |
| AOS 2 | Summary screen and practice questions |
| Topic 2 | |
| Concept 2 | |

**QUICK CHECK**

5 What is the difference between an infection and a disease?
6 Identify the following as true or false:
  a Diseases can only be spread by a person who is showing the symptoms of the disease.
  b Only some infections develop into diseases.
  c An environmental disease is an example of a communicable disease.
  d An example of indirect transmission of disease is from airborne droplets that contain pathogens.
  e The incubation period of all diseases varies from days to weeks.
  f Bubonic plague is an example of a zoonotic disease.
7 Give an example of the following:
  a a step that must occur before an infection can develop into a disease
  b the incubation period of a disease
  c indirect transmission of a disease
  d a non-infectious disease
  e an entry point for a pathogen to the body.

## Pathogens: cellular and non-cellular

Infectious diseases of people are caused either by pathogens with a cellular structure, such as bacteria or fungi, or by non-cellular agents, such as viruses or prions. Table 6.3 shows some examples of human diseases caused by these various pathogens. Many human diseases have a long history of occurrence in human populations. However, new human diseases continue to be identified, the majority being the result of a transfer of pathogens from their natural non-human animal hosts to people.

**TABLE 6.3** Various kinds of pathogens, the infectious diseases that they cause, and the various means of transmission. Archaea are *not* included in this table as no evidence has been found to date that archaea are responsible for any human diseases.

| Causative agent | Disease | Example of how it spreads |
|---|---|---|
| **BACTERIA** | | |
| *Neisseria meningitidis* | bacterial meningitis; septicemia | most commonly, via person-to-person through saliva or mucus; more rarely, from a person's own throat bacteria that cross the mucous membrane |
| *Streptococcus pneumoniae* | pneumonia | via person-to-person through inhalation of contaminated droplets, such as from the sneeze of an infected person |
| *Corynebacterium diphtheria* | diphtheria | via person-to-person from close respiratory contact |
| *Mycobacterium tuberculosis* | tuberculosis | via person-to-person through inhalation of contaminated droplets that reach the lung alveoli (see figure 6.14 (b) below) |
| *Bacillus anthracis* | anthrax | via inhalation or ingestion of spores or via break in skin that enables entry of spores |
| *Legionella pneumophila* | legionnaire's disease | via inhalation of water droplets from a contaminated air conditioner |
| *Yersinia pestis* | Black Death (bubonic plague) | via the bite from infected fleas, from an infected rodent |
| *Vibrio cholerae* | cholera | via ingestion of faecal-contaminated water or food |
| *Streptococcus pyogenes* | toxic shock syndrome; strep throat & others | via upper respiratory tract or skin lesion |
| *Clostridium tetani* | tetanus | via deep puncture wound |
| *Treponema pallidum* | syphilis | via person-to-person through direct sexual contact with syphilis sore of an affected person |

*(continued)*

| Causative agent | Disease | Example of how it spreads |
|---|---|---|
| *Helicobacter pylori* | stomach ulcers (see figure 6.14 (e) below) | via person-to-person contact through saliva of infected person |
| *Salmonella typhi* | typhoid fever | via ingestion of faecal-contaminated water or food; can also be transmitted from person-to-person |
| *Chlamydia trachomatis* | pelvic inflammatory disease (PID)# | via sexual contact with an infected person |
| *Clostridium botulinum* | botulism, a form of food poisoning | via ingestion of improperly canned food that exposes person to the bacterial exotoxin that has not been heat-inactivated |
| *Salmonella typhimurium* | Salmonellosis, a form of food poisoning | via ingestion of faecal-contaminated foods such as raw eggs, and improperly cooked chicken |
| *Rickettsia australia* | Queensland tick typhus (QTT) | via saliva from bite of infected paralysis tick, *Ixodes holocyclus* |
| **VIRUSES** | | |
| Influenza viruses | influenza | via inhalation of airborne droplets from sneeze or cough of infected person |
| Herpes virus | chickenpox (varicella) | via nose or mouth from direct contact with infected person |
| Rubella virus | rubella (German measles) | via naso-pharynx from airborne droplets in sneeze/cough of infected person |
| Hepatitis B virus | viral hepatitis type B | via contact with body fluids of infected person |
| Human immuno-deficiency virus (HIV) | Acquired immuno-deficiency disease (AIDS) | via sexual contact with infected person or from use of contaminated needles |
| *Poliovirus* | poliomyelitis (polio) | via mouth from faeces or sneeze droplets of infected person |
| *Rhinovirus* | common cold# | via inhalation of droplets in air from infected person or via hand- to-hand contact with infected person |
| SARS coronavirus (SARS CoV) | Severe acute respiratory syndrome | via inhalation of infected droplets or by touching contaminated surfaces and then touching one's mouth or nose |
| *Ebolavirus* | Ebola virus disease (see figure 6.14 (a) below) | via direct contact with body fluids of an infected person (e.g. blood, urine, saliva, sweat, faeces, vomit, breast milk, and semen) or contact with contaminated needles or syringes. |
| **FUNGI** | | |
| *Aspergillus fumigatus* | pulmonary aspergillosis | via inhalation of airborne spores |
| *Candida albicans* | thrush (oral candidiasis) | via mouth-to-mouth contact with infected person |
| *Trichophyton* spp. | athlete's foot (tinea) | via contact with infected person or with contaminated surfaces |
| **PROTOZOA** | | |
| *Giardia intestinalisasis* | giardiasis | via ingestion of water or food contaminated with *Giardia* cysts |
| *Entamoeba histolytica* | amebiasis | via ingestion of water contaminated with *Entamoeba* cysts |
| *Pneumocystis jirovecii* | pneumocystis pneumonia (pcp) | via inhalation of spores through nose or throat into lungs |
| *Plasmodium falciparum* | malaria | via bite of infected mosquito |
| **PRIONS** | Creutzfeldt-Jakob disease (CJD) | via tissue grafts from infected cadavers, including corneas, brain membrane or via growth hormone isolated from infected pituitary glands of cadavers |
| | variant Creutzfeldt-Jakob disease (vCJD) | via ingestion of beef from cattle infected with bovine spongiform encephalopathy (BSE), commonly known as 'mad cow disease' |
| | kuru | via past practice of ceremonial ingestion of brain tissue of deceased infected person. |

# also caused by other bacteria

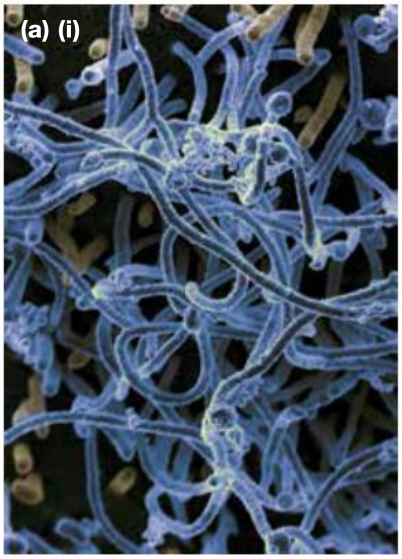
**ODD FACT**

Eggs of parasitic helminths have been found in mummified human faeces from thousands of year ago.

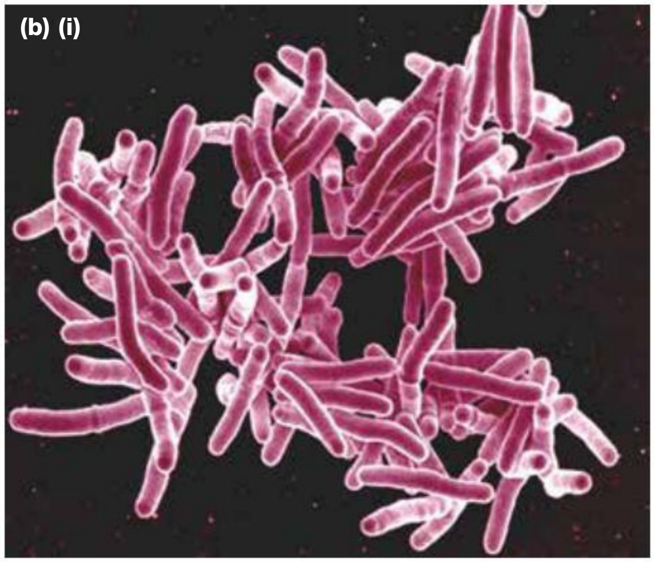
Another group of disease-causing organisms are parasitic helminth worms, including roundworms, hookworms and whipworms. Helminth infections are transmitted via soil contaminated with human faeces that contain eggs of these worms. However, many helminths do not fall into the category of microbes as they are visible to the unaided eye.

Figure 6.14 shows some pathogenic agents and the infectious diseases that they cause.

(a) (i)



(b) (i)



(a) (ii)

(b) (ii)

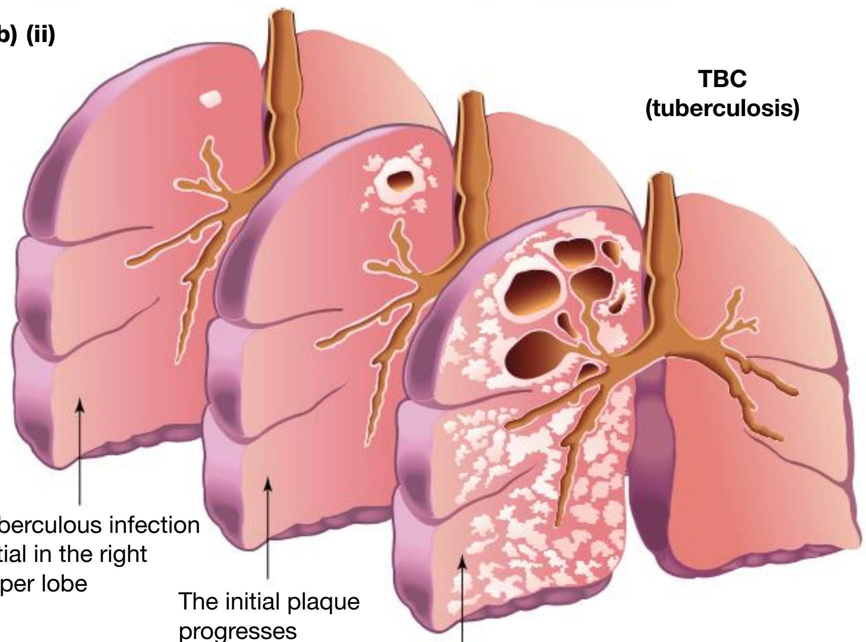



**FIGURE 6.14** **(a) (i)** Ebola virus: Colourised scanning electron micrograph from the 2014 outbreak Mali isolate of the Ebola virus shedding from the surface of Vero cells, six days post-infection. Note the mass of filamentous Ebola viral particles (virions) (blue) budding from an infected cell (25 000 × magnification). (Image courtesy of the Rocky Mountain Laboratories/NIAID/NIH.) **(ii)** One of the early symptoms of an Ebola virus disease is the appearance of a haemorrhagic rash over the body. **(b) (i)** SEM image of *Mycobacterium tuberculosis* bacteria, the cause of tuberculosis. (Credit: NIAID) **(ii)** Progressive damage to lungs from untreated tuberculosis results in lesions and loss of lung tissue.

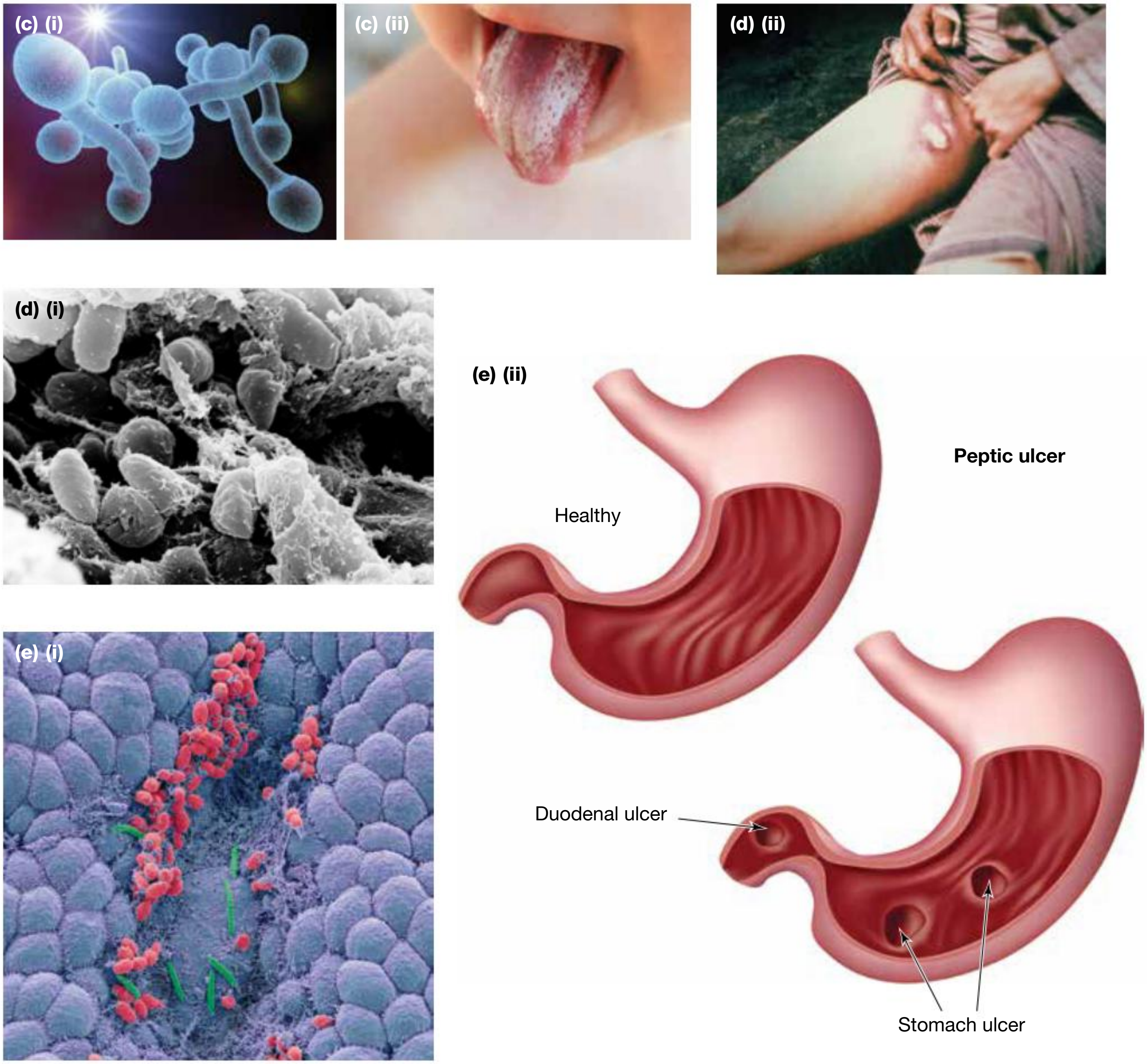



**FIGURE 6.14** *(continued)* **(c) (i)** SEM image of the yeast *Candida albicans*. These yeasts are a species of fungi that are normal inhabitants of the skin and mucous membranes. However, if these yeasts grow excessively, they produce the diseases termed candidiasis. This disease has various common names depending on the location of the overgrowth, such as ‘thrush’ (in the mouth or throat) and ‘yeast infection’ or ‘vaginitis’ (in the vagina). **(ii)** Image of a young child with thrush, a yeast infection of the mouth, which is most commonly caused by the fungus *Candida albicans*. **(d) (i)** SEM image depicting a mass of rod-shaped *Yersinia pestis* bacteria, the cause of bubonic plague (black death) in the foregut of the rat flea (Credit: NIAID). Bites by infected fleas (*Xenopsylla cheopis*), whose main host is the black rat, can transmit plague bacteria to people. Refer back to figure 6.3 to see the swollen lymph glands or buboes that are one symptom of bubonic plague. **(ii)** Image showing swollen lymph gland (buboes) in the groin of a person with bubonic plague. (Credit: Centers for Disease Control and Prevention.) **(e) (i)** False-coloured SEM image of *Helicobacter pylori* bacteria (green) and yeast cells (red) in the mucous lining of the stomach (blue). *H. pylori* bacteria are the most common cause of stomach ulcers. **(ii)** Diagram showing the lesions caused by a *Helicobacter pylori* infection of the lining of the stomach and the duodenum, resulting in peptic ulcers, which appear as lesions of the lining.

studyon

Unit 3
AOS 2
Topic 2
Concept 2

See more
Bacteria

studyon

Unit 3
AOS 2
Topic 2
Concept 2

See more
Malaria

**KEY IDEAS**

- Diseases may be caused by a variety of cellular pathogens, including bacteria, and by non-cellular agents such as viruses and prions.
- Different pathogens show varied modes of transmission.

**QUICK CHECK**

8 Give an example of a disease that is:
  a caused by bacteria and spread by faecal-contaminated water or food
  b caused by a virus and may be spread by the use of infected needles
  c caused by a fungus and spread by mouth-to-mouth contact with an infected person.
9 Give an example of a pathogen that is the cause of:
  a a disease that produces a severe form of food poisoning
  b the disease spread by infected droplets from air conditioners
  c a disease that is transmitted by the bite of an infected mosquito.
10 What is the difference between a cellular and a non-cellular disease-causing agent?

# Meet some pathogens!

In the following sections, we will identify the features of some disease-causing agents:

- bacteria
- viruses
- prions.

# Bacteria and human diseases

**Bacteria** are prokaryotic microbes, and their genetic material is double-stranded DNA. As identified earlier in this chapter, only a very small percentage of bacteria are human pathogens that are responsible for a number of diseases. Figure 6.15 shows the locations in the human body — extracellular and/or intracellular — where bacterial pathogens can live, reproduce and cause disease if they can gain entry to the human body.

Note that:

- Some bacterial pathogens can live and reproduce *only* on the external surfaces of human cells or in body fluids, such as plasma, lymph or interstitial fluid. These bacteria are 'exclusive extracellular pathogens', as, for example, *Vibrio cholera*, which causes the uncontrolled diarrhoea of cholera.
- Other bacteria can live and reproduce either inside or outside the cells of their host, as, for example, *Neisseria meningitidis*, the cause of bacterial meningitis, and *Salmonella typhi*, the cause of typhus disease. These bacteria are 'facultative intracellular pathogens'.
- Some other bacterial pathogens can only survive and reproduce within host cells. These bacteria are 'exclusive intracellular pathogens'. Examples include *Chlamydia trachomatis*, the cause of a sexually transmitted infection that, if left untreated, can develop into pelvic inflammatory disease and infertility. Another example are bacteria of the genus *Rickettsia* that cause a group of serious diseases known as spotted fevers, and *Mycobacterium leprae*, the cause of leprosy.

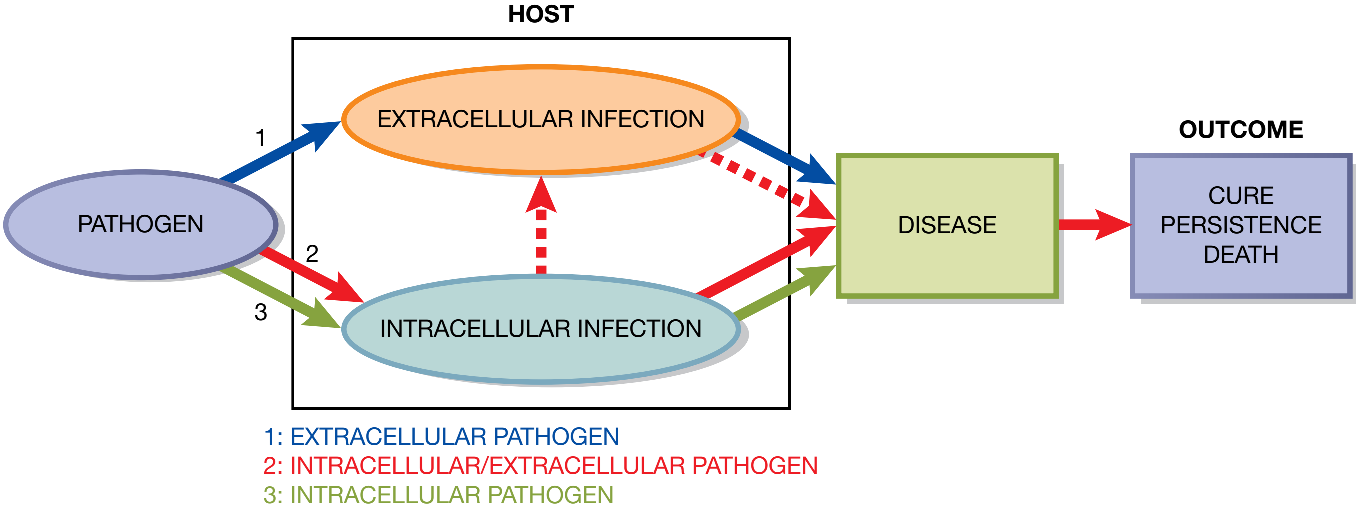



**FIGURE 6.15** Diagram showing possible outcomes of an infection by various pathogens once they gain entry to the body. Would you classify pathogen 2 as an exclusive intracellular pathogen? Note that an infection does not always lead to a disease. In many cases, the body's immune system eliminates the infection and prevents the development of a disease.

Under favourable conditions, bacteria can, on average, reproduce every 20 to 30 minutes so that, over a period of 12 hours, enormous numbers of bacterial cells can potentially be produced. If pathogenic bacteria can cross the body's defences, such an infection has the potential to develop rapidly into a disease.

## Features of bacteria

Bacterial pathogens can be distinguished in several ways, including:

1. shape
2. response to Gram stain
3. ability to produce toxins
4. presence/absence of a **capsule**
5. oxygen requirements
6. nutritional needs.

These features are useful in identifying bacteria. In addition, some features influence the **virulence,** that is, the ability of bacteria to produce their harmful effects.

### Bacteria come in various shapes

Figure 6.16 shows various shapes of bacterial species. Cocci are spherical with diameters of 0.5 to 2.0 micrometres (μm), while bacilli are rod-shaped with typical lengths of 1 to 10 μm and diameters ranging from 0.25 to 1.0 μm. Other shapes of bacterial species include spirilla (spiral shape), vibrios (comma-shaped), and spirochaetes (corkscrew-shaped). On average, bacterial cells are about one-tenth the size of eukaryotic cells. Bacterial cells can show different arrangements and may exist, for example, as single cells, as groups of two (pairs), as groups of four (tetrads) or larger, or as chains.

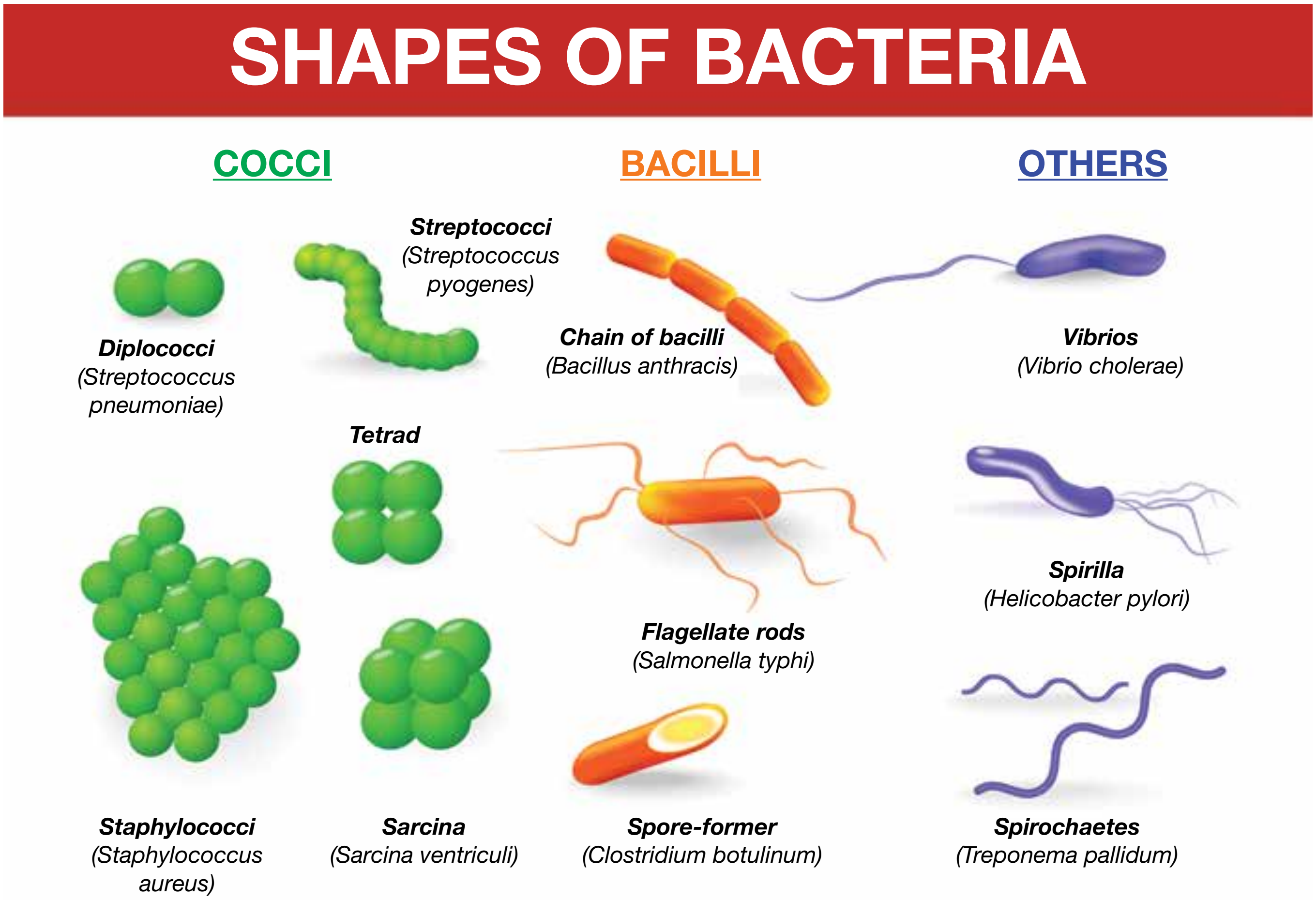



**FIGURE 6.16** Diagram showing some common bacterial shapes. Note that some bacterial species have one or more flagella. What disease is caused by Vibrio cholerae?

## Bacteria: responses to Gram stain

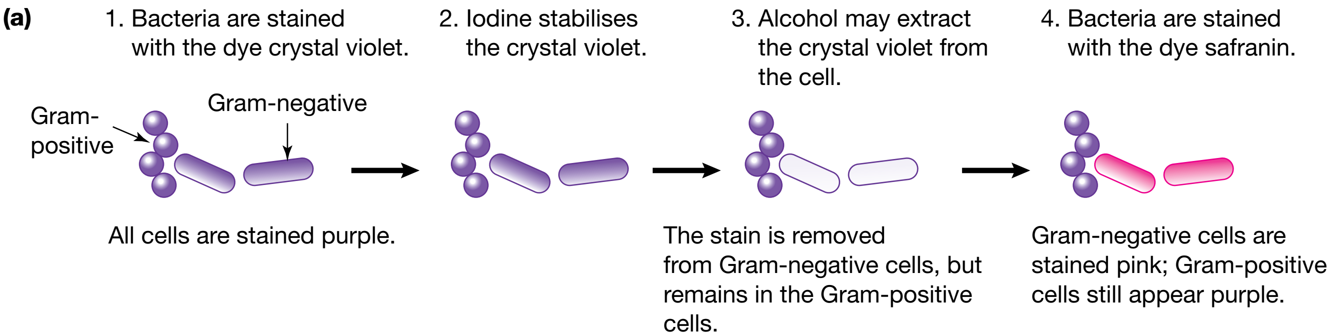
A special stain, known as Gram stain, is commonly used for the general identification of bacteria. This stain separates bacteria into two main groups, **Gram-positive** and **Gram-negative**, depending on the structure of their cell wall. This staining technique involves the use of two dyes as outlined in figure 6.17a. Some bacteria retain the first dye (crystal violet) and appear dark blue — these bacteria are said to be Gram-positive. In contrast, other bacteria do not retain the crystal violet dye and instead are stained by a second dye (safranin) so that they appear red or pink — these bacteria are said to be Gram-negative (see figure 6.17b).

The contrasting responses of bacteria to Gram staining are due to differences in the structure of the external cell wall that lies outside the plasma membrane of bacterial cells. The cell wall of Gram-positive bacteria contains a thick layer of **peptidoglycan** (a complex of amino acids (*peptid-*) and sugars (*-glyc* = sweet)), which retains the crystal violet dye used in the Gram staining technique. The cell wall of Gram-negative bacteria has a very thin layer of peptidoglycan that does not retain the violet dye (see step 3 in figure 6.17a).

**ODD FACT**

Gram staining was developed by Danish bacteriologist Joachim Gram (1853–1938).

(a)




(b)

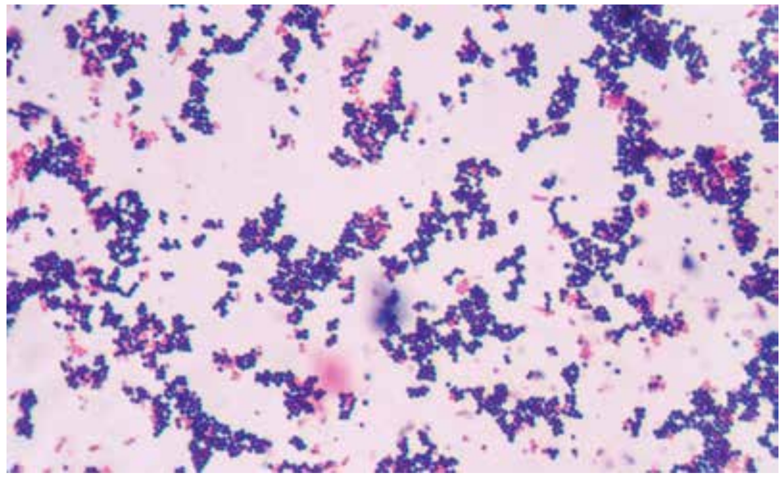


**FIGURE 6.17** **(a)** Steps in the Gram staining technique. **(b)** Light microscope (LM) image showing the results of Gram staining of two kinds of bacteria. The majority of the bacteria are Gram-positive cocci stained purple, while the other bacteria are Gram-negative and are stained pink.

Gram-negative bacteria have an outer membrane made of lipopolysaccharide (LPS) (see figure 6.18). This LPS outer membrane enables these bacteria to expel or exclude certain drugs and antibiotics. As a result, Gram-negative bacteria are generally far more resistant to antibiotic treatment than are Gram-positive bacteria. Gram-positive bacteria are normally susceptible to antibiotics, such as penicillin and the sulfonamide drugs, but these antibiotics are not effective against Gram-negative bacteria. So, the result of Gram staining can give an indication of the type of antibiotic treatment that could usefully be given to a person with a bacterial disease.

If a bacterial infection is suspected, Gram staining is carried out on either body fluids or cell samples. This procedure is a quicker method of identifying the presence of bacteria than that of culturing them (growing bacteria on a medium such as nutrient agar). What is a possible advantage of this faster method? Table 6.4 identifies examples of Gram-positive and Gram-negative bacteria.

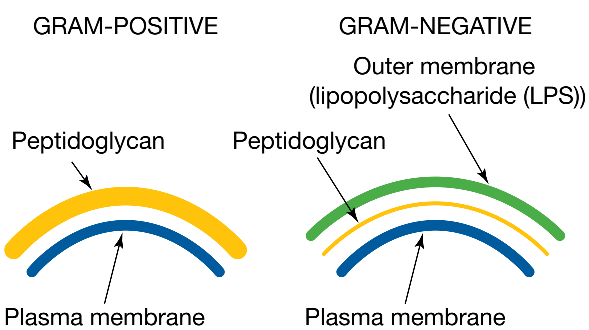



**FIGURE 6.18** Cell wall structure of bacteria. At the left is a simplified representation of the cell wall of Gram-positive bacteria, which consists of a thick layer of peptidoglycan lying outside the plasma membrane. At the right is a simplified representation of the more complex cell wall of Gram-negative bacteria. Note the much thinner peptidoglycan layer and the presence of a lipopolysaccharide (LPS) outer membrane.

## Some bacteria produce toxins

A major factor that increases the virulence of some bacteria is their ability to produce specific toxins. **Bacterial toxins are substances produced by some bacteria that damage particular tissues of their host organism and cause disease.** Bacterial toxins are of two types:

- exotoxins
- endotoxins.

## ODD FACT

Exotoxins that damage or kill cells of any kind are called cytotoxins, while exotoxins that attack nerve cells are termed neurotoxins, and those that damage gut cells are termed enterotoxins.

**TABLE 6.4** Examples of Gram-positive and Gram-negative bacteria. Check table 6.3 to see some of the human diseases caused by these bacteria.

| Gram-positive bacteria | Gram-negative bacteria |
|---|---|
| *Streptococcus pyogenes* | *Escherichia coli* |
| *Staphylococcus pneumonia* | *Vibrio cholerae* |
| *Staphylococcus aureus* | *Yersinia pestis* |
| *Clostridium botulinum* | *Neisseria meningitidis* |
| *Clostridium tetani* | *Helicobacter pylori* |
| *Bacillus anthracis* | *Legionella pneumophila* |
| *Corynebacterium diphtheria* | *Treponema pallidum* |

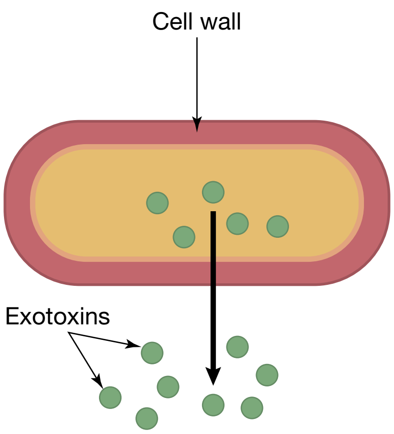



**FIGURE 6.19** Exotoxins are highly toxic soluble proteins produced by living bacteria. Exotoxins are far more poisonous than chemicals such as strychnine and arsenic.

**What is an exotoxin?** **Exotoxins are highly toxic soluble proteins that are produced by living bacterial pathogens as part of their metabolism and are released into their surroundings** (see figure 6.19). Because they are soluble proteins, exotoxins do not remain at their site of production. Instead, exotoxins can spread throughout the body and cause system-wide damage.

Several bacteria, mainly Gram-positive bacteria, produce exotoxins that can damage or kill cells of all kinds, while other bacteria produce exotoxins that can damage cells of particular kinds, such as nerve cells. Exotoxins act in different ways on host tissues; some exotoxins damage plasma membranes and disrupt transport of compounds across these membranes, some inhibit protein synthesis, and some block normal nerve function.

Exotoxin-producing bacteria include the following:

- *Clostridium tetani* releases a neurotoxin that blocks muscle relaxation, resulting in tetanus, a disease that is characterised by painful muscle spasms and lockjaw and that can result in respiratory failure.
- *Vibrio cholerae* releases a toxin that damages the cells of the gut lining, leading to the uncontrolled production of watery diarrhoea that is seen in the disease cholera.
- *Streptococcus pyogenes* releases an exotoxin that kills cells and can lead to major organ failure as seen in streptococcal toxic shock syndrome.

Here is Matthew's story and his rehabilitation from a *Streptococcus pyogenes* infection that developed into streptococcal toxic shock syndrome.

## MATTHEW'S STORY

The bacterium *Streptococcus pyogenes* is also known as Group A *Streptococcus* or streptococcal A. Most commonly, people associate a streptococcal A infection with the relatively minor illness of a sore throat or 'strep throat'. Streptococcal infections of the throat and tonsils typically respond to a short course of antibiotic treatment. In fact, many people carry strep A in their throats or on their skin and show no symptoms.

On very rare occasions, however, streptococcal A may cross cell barriers and enter the bloodstream after a person has suffered a minor injury. Once there, the bacteria can rapidly divide by binary fission, perhaps every 15 to 20 minutes, producing a period of exponential growth in the numbers of bacteria, which can develop into a life-threatening situation.

In June 2012, Matthew Ames, 39 years old, devoted husband of Diane and proud father of their four young children, had a sore throat. What started as a sore throat then became a more serious condition as his muscles and joints became affected. As a result, Matthew was admitted to hospital, but within 12 hours, he had been placed in an induced coma and was on life support. In spite of the treatments administered, the streptococcal A infection had become established in Matthew's bloodstream and increasing numbers of bacteria were producing exotoxin that threatened widespread organ damage. Forty-eight hours later, in an attempt to slow the progress of this infection, surgeons took the drastic step of amputating Matthew's left arm, which was the initial site of the infection.

This measure failed to slow the streptococcal A infection and the bacterial exotoxin produced a condition known as toxic shock syndrome, which affected organs throughout Matthew's body. As a result, his systems began to shut down and, deprived of blood flow, his extremities began to blacken as the cells of muscle, skin and underlying tissue died. Matthew was in an extremely critical condition and his family members were told that his death was imminent. The only hope for possible survival lay in the amputation of Matthew's remaining limbs, a procedure that would remove the necrotic tissue that was an increasing source of further toxins. Matthew's family decided to give him this chance for life.

Happily, Matthew survived and he then began a long journey of rehabilitation. His progress is remarkable and his story is inspiring and one of great courage (see figure 6.20).

You can read about Matthew's progress as, with the support of his family, he adjusts to life without his limbs and gains capability through the use of prosthetic devices. Matthew and Diane's own account is in their book, *Will to Live*, published by Penguin Books in 2014.


**FIGURE 6.20** Matthew undergoing rehabilitation. **(a)** In the pool **(b)** Getting around. (Images courtesy of Kate Ames.)

**Exotoxins damage cells, even in the absence of the bacteria that produce them.** On Monday 14 August 1922, a group of people were enjoying a picnic near Loch Maree in Scotland. Among the refreshments were sandwiches containing duck paste that had been prepared at the hotel where the picnickers were staying. That evening, two of the picnickers became ill, and they died the following morning. Other picnickers became ill and two died on Wednesday. In all, within a week, eight people were dead, all having eaten the duck paste sandwiches. What had caused these deaths?

The inquest determined that the deaths were a result of food poisoning caused by the bacterium *Clostridium botulinum.* The duck paste was contaminated with spores of this anaerobic, exotoxin-producing bacterium. Before being sealed into a jar, the duck paste was insufficiently heated, so the spores remained viable. In the anaerobic conditions of the sealed jar, the bacterial spores germinated and produced the deadly botulinum toxin, which is colourless, odourless and tasteless. This exotoxin is the most lethal toxin known, with a lethal dose by ingestion for a 70 kg adult being just 70 micrograms. (To put this amount into context, you might note that one grain of table salt weighs about 0.5 milligram, that is, about 500 micrograms!) Eating even very small quantities of the duck paste sandwiches proved to be fatal. The exotoxin blocks signal transmission by nerve cells and causes death by respiratory paralysis.

**What is an endotoxin? Endotoxins are toxins derived from the lipid portion of the lipopolysaccharide (LPS) of the outer membrane of Gram-negative bacteria.** These endotoxins are produced by some Gram-negative pathogenic bacteria. Endotoxins are released only after these bacteria die and their outer membrane breaks down (see figure 6.21).

**ODD FACT**

Botulinum toxin can be inactivated by heating at 85 °C or higher for 5 minutes. This is the botox used for cosmetic purposes.

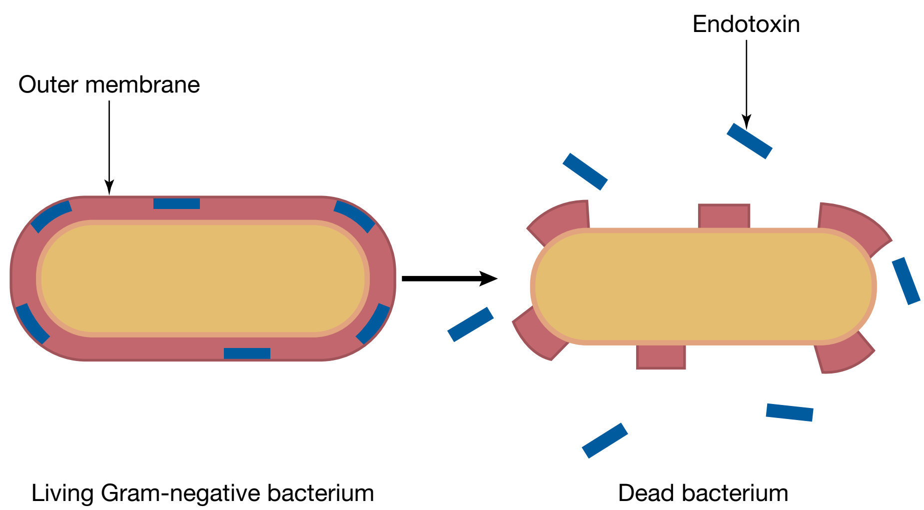



**FIGURE 6.21** Endotoxins are the lipid portion of the lipopolysaccharide (LPS) that forms the outer membrane of Gram-negative bacteria. Endotoxins are released only when these bacteria die and their outer membranes are broken down.

Release of endotoxins from large numbers of dead Gram-negative bacteria produce similar symptoms. The symptoms include chills, fever, weakness, blood clotting, tissue death and endotoxic septic shock. Examples of endotoxin-producing bacteria include *Salmonella typhi,* the cause of typhoid fever, and *Neisseria meningitidis*, the cause of meningitis.

Table 6.5 summarises some differences between bacterial exotoxins and endotoxins.

**TABLE 6.5** Summary of differences between bacterial exotoxins and endotoxins.

| Exotoxins | Endotoxins |
|---|---|
| • soluble proteins | • lipopolysaccharide |
| • produced by some Gram-positive and Gram-negative bacteria | • produced only by Gram-negative bacteria |
| • released by living bacterial cells | • part of outer membrane of bacterial cell wall and released when bacteria die |
| • many different types, each producing specific damage | • one type only, producing a diverse range of effects |
| • typically destroyed by heat | • heat stable |

## Some bacteria have a capsule

Many pathogenic bacteria, both Gram-positive and Gram-negative, have a gelatinous layer of polysaccharide that lies outside their cell walls (see figure 6.22). This outermost layer, known as a capsule, protects these bacteria from one of the defence mechanism of the human body — being engulfed by defence cells of the body known as macrophages (see pages 292 and 293, in chapter 7). As a result, **the presence of a capsule increases the virulence of pathogenic bacteria; that is, it increases the likelihood of disease resulting from an infection by these bacteria.**

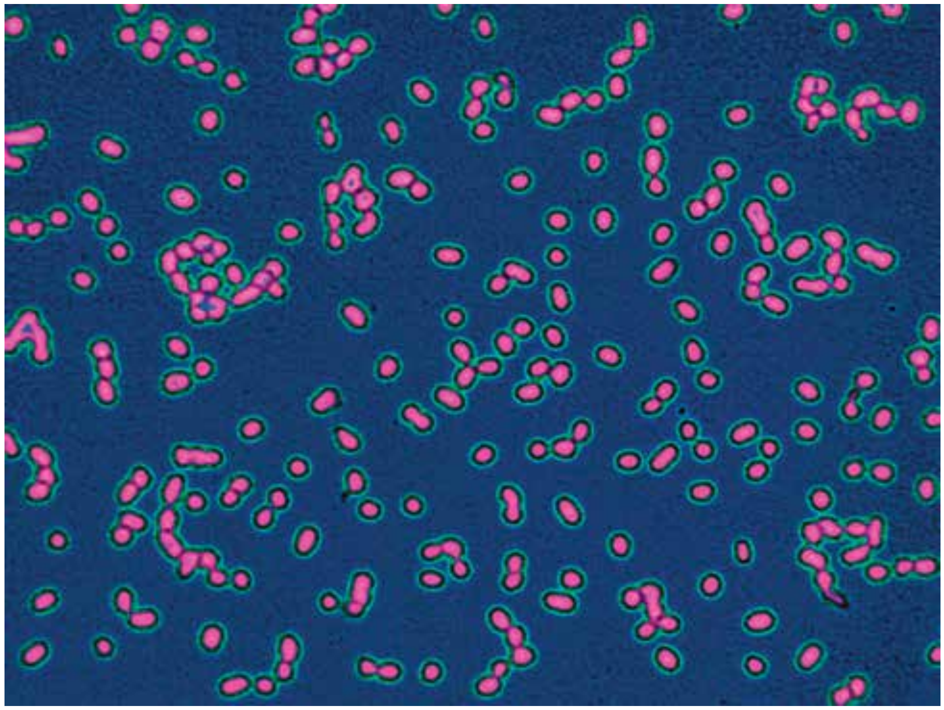


**FIGURE 6.22** Bacterial cells, each surrounded by a gelatinous capsule that appears in this photomicrograph as a greenish halo. Capsules enable these bacteria to avoid some of the immune defence mechanisms of the human body, such as being engulfed by phagocytes, special cells of the body's immune system (see pages 291–294, in chapter 7).

## Bacteria differ in their oxygen requirements

Table 6.6 shows the variations in oxygen requirements of bacteria for energy production (and life).

Many bacteria are obligate aerobes. This means they can produce energy by aerobic respiration only, and so can only live and reproduce in the presence of oxygen. Oxygen absent: no energy production, no life!

Other bacteria are **obligate anaerobes**. This means they can only produce energy in the absence of oxygen, and oxygen is toxic for them. Oxygen in high concentration: no growth, no life! These bacteria are commonly found in the gut, in the reproductive tract and in deep wounds. However, some bacteria that are obligate anaerobes survive in the presence of high oxygen concentrations by forming dormant spores.

Other bacteria can generate energy through both aerobic respiration and anaerobic respiration, and so can live and reproduce with or without oxygen. These bacteria are said to be **facultative anaerobes** as they produce energy by aerobic respiration when oxygen is present and by anaerobic respiration (fermentation) when oxygen is absent.

**ODD FACT**

In a healthy person, the lower gut (colon) has very low concentrations of oxygen, and more than 90 per cent of the microbes in the colon are obligate anaerobes.

**TABLE 6.6** Summary of oxygen requirements of various bacteria.

| Group | Conditions for growth | Examples |
|---|---|---|
| obligate aerobes | require oxygen for producing energy by aerobic respiration | *Mycobacterium tuberculosis* *Pseudomonas aerugenosa* |
| faculative anaerobes | produce energy by either aerobic or anaerobic respiration; that is, with or without oxygen | *Streptococcus pyogenes* *Escherichia coli* *Staphylococcus* spp. |
| obligate anaerobes | produce energy only in the absence of oxygen; poisoned by oxygen | *Clostridium botulinum* *Clostridium tetani* *Clostridium perfringens* |

If conditions become unfavourable, some bacterial species have the ability to form special structures called **spores**. The tough outer coats of a spore make bacteria highly resistant to harsh conditions and enable them to remain inactive for very long periods. When temperature, moisture and other conditions become favourable again, spores revert to the normal active cellular form of the bacteria. Spore-forming bacteria include members of the genus *Bacillus* and the genus *Clostridium*.

### Bacteria: nutritional needs

Different bacteria have different energy and carbon needs. The wide range of ways in which they meet these needs has to be taken into account when disease-causing bacteria are being isolated and classified.

When specimens from infected patients are cultured in a microbiology laboratory, several different kinds of culture media may be used (see figure 6.23). The way in which the bacteria grow on different media helps to determine exactly what bacterium is causing the problem. Some species will grow on a sugar medium and produce gas; another might only produce acid; still another might produce acid and gas. Such differences allow bacteria to be correctly identified. This information is useful in determining the most appropriate treatment for a patient.

**FIGURE 6.23** Different bacteria grow on different kinds of media. The photograph on the right shows five dishes containing different kinds of agar, with different bacteria growing on them. These kinds of media are also called selective media because if a mixed culture of bacteria is plated onto the surface of each, only some species, or a selection of the bacteria, will grow on each. A species of bacteria will only grow on the special medium if the medium provides an appropriate source of energy and carbon.

## How do bacteria cause disease?

Table 6.7 shows various factors that are involved in the production of some human diseases by bacteria.

**TABLE 6.7** Factors involved in disease-production by bacteria.

| Disease | Bacteria responsible | Factor involved in disease | Mode of action |
|---|---|---|---|
| botulism (one form of food poisoning) | *Clostridium botulinum* | exotoxins that are neurotoxins | paralysis |
| salmonellosis (most common form of food poisoning) | several species of *Salmonella* | invasive properties | bacteria invade tissue lining intestine |
| cholera (severe gastroenteritis) | *Vibrio cholerae* | exotoxin | alters intestinal permeability |
| diphtheria (sore throat and fever) | *Corynebacterium diphtheriae* | exotoxin | prevents protein synthesis |
| scarlet fever (fever and rash) | *Streptococcus pyogenes* | exotoxin | bacteria invade tissue and damage blood vessels |
| tuberculosis (lesions in lung and other tissues) | *Mycobacterium tuberculosis* | invasive properties | bacteria invade tissue |

## KEY IDEAS

- **Various bacterial species show differences in structural, biochemical and physiological features.**
- **The structure of the bacterial cell wall determines whether bacteria are Gram-positive or Gram-negative.**
- **Some bacteria produce an external capsule that makes them more virulent.**
- **Various bacteria differ in their requirement for oxygen, ranging from obligate aerobes to obligate anaerobes.**
- **Many bacteria cause disease because of the toxins that they produce.**

## QUICK CHECK

**11** Identify a key difference between the members of the following pairs:
- **a** exotoxin and endotoxin
- **b** obligate aerobic bacteria and obligate anaerobic bacteria
- **c** Gram-negative bacteria and Gram-positive bacteria
- **d** a spirochete and a coccus.

**12** Identify bacteria that cause disease because of:
- **a** an exotoxin that they produce
- **b** their invasive properties.

**13** Which group of bacteria is susceptible to antibiotics such as penicillin?

# Viruses: non-cellular pathogens

Let's meet the viruses — interesting entities because they span the boundary between non-living and living.

- Non-living features: viruses are non-cellular in structure, have no metabolic activity, and they cannot reproduce independently of a host cell.
- Features shared with living organisms: viruses have genetic material, either DNA or RNA, that encodes specific proteins and, *once within a host cell*, this genetic material is self-replicating.

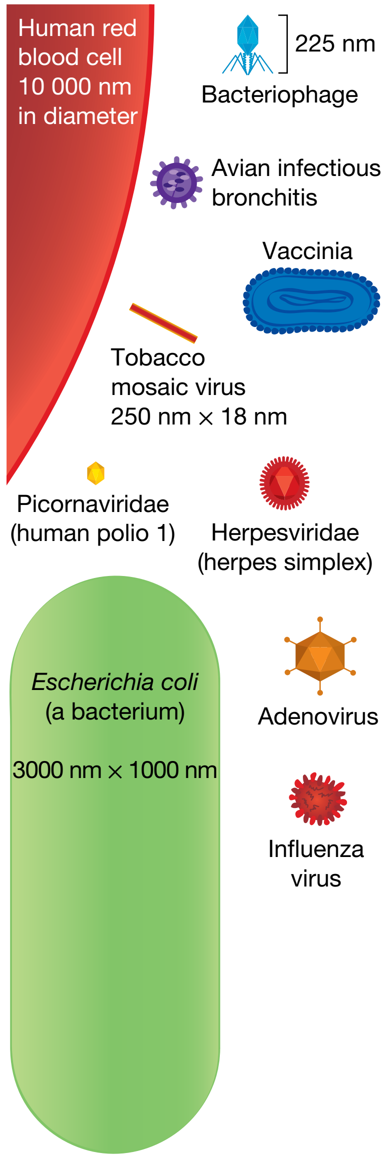



**FIGURE 6.24** Diagram showing the size of several viruses in comparison to a bacterial cell and a human red blood cell. Measurements are in nanometres (nm); note that 1000 nm = 1 micrometre (μm) = 0.001 mm.

Viruses may be defined as non-cellular agents consisting of genetic material, either DNA or RNA, that can replicate only within a host cell — they are obligate intracellular parasites.

Outside a host cell, viruses exist in a form called a **virion** or viral particle. A virion that is the extracellular form of a virus is the means by which a virus can transfer from one host cell to another.

Viruses are very much smaller than bacteria (see figure 6.24). Their extremely small size means that they can only be seen using electron microscopy.

On average, a viral particle or virion is about 1/100th of the size of a typical bacterial cell. Various viral particles have sizes ranging from about 20 to 500 nanometres (nm), that is, 0.02 to 0.5 micrometres (μm).

## Viruses and human disease

Human diseases caused by viruses include rabies, smallpox, influenza, polio and measles. These diseases have caused countless deaths in human populations over thousands of years. These diseases were well known and, in many cases, greatly feared, but their cause was a mystery — until the twentieth century.

**Viruses have preferred target cells in which they can replicate**; for example:

- Rhinoviruses enter the body via the nasopharynx and preferentially infect cells of the upper respiratory tract where the conditions of temperature and other factors are most suitable for multiplication of this virus.
- Hepatitis B virus can enter the body via a skin puncture and travel in the blood to its preferred target cells in the liver. Receptors on the surface of liver cells facilitate the entry of these viruses into the liver cells where they multiply.
- Human immunodeficiency virus (HIV) can be transmitted by various means, including the exchange of body fluids. The preferred target cells of this virus are immune cells called T-helper cells (also known as CD4+ cells).

**New viral diseases affecting people continue to appear**. Since the mid-twentieth century, new viral diseases affecting people have appeared. These new viral diseases include acquired immunodeficiency syndrome or AIDS (first known case in the African Congo in 1959), Ebola virus disease (first recorded case in Sudan in 1976), Hendra virus disease (first recorded case in Queensland in 1994), and severe acute respiratory syndrome or SARS (first recorded case in Guangdong Province in southern China in 2002).

These new human diseases are due to people being infected by viruses whose normal hosts are other animals, such as birds, bats, pigs, rodents and monkeys. New opportunities for viruses to transfer from their natural animal hosts to people have occurred because of the great increases in human populations in some parts of the world and the consequent spread of human occupation into natural forests and woodlands, creating new contacts between wild animals and people.

**ODD FACT**

The first recorded description of rabies comes from ancient Babylon in the year 2300 BCE, more than 4000 years ago.

HIV is discussed further in chapter 8, pages 358–362.

ODD FACT

The first connection between a virus and cancer was identified in 1964 when EBV was shown to be linked with cancers, including Burkitt lymphoma and Hodgkin lymphoma.

## Viruses: how do they cause disease?

Viral diseases are due to the disturbances caused when viruses infect their target cells in the human body. Virus-infected cells show morphological and biochemical changes that lead to disease symptoms. For example, rotaviruses that infect cells lining the human gut change the permeability of their plasma membranes, resulting in severe diarrhoea that causes the deaths of thousands of young children in developing countries annually. Some DNA viruses can alter the DNA of their host cells in a way that leads to the development of cancers. Viruses that can cause human cancers include Epstein-Barr Virus (EBV), hepatitis B virus (HBV) and human papillomavirus (HPV).

Direct damage to host cells from viral infections also occurs because viruses 'take over' the host cells so that they cannot perform their normal functions. Instead, the host cells are dedicated to the production of the genetic material and proteins of the infecting viruses, rather than making the macromolecules that they need for their own use in energy generation and for repair. For example, in an influenza virus-infected cell, only virus-specific RNA binds to ribosomes and the host-cell mRNA stops binding. Virus-infected cells eventually die because the viruses stop cells from performing their normal functions.

Damage to host cells also occurs as a result of the body's immune responses to the viral infection; for example, the symptoms of the common cold from a rhinovirus infection are the result of the inflammatory responses of the body's immune system to the virus.

Viruses replicate within their host cells. Multiple copies of the virus are released from an infected cell either following cell lysis (for naked viruses) or by gradual release by **budding** (for enveloped viruses). In lysis, the infected cell 'explodes,' releasing large numbers of viral particles that can infect other target cells, compounding the damage from the viral infection. A clue to the scale of viral replication and release comes from the observation that one gram of faeces of a child infected with rotavirus may contain trillions of rotavirus particles.

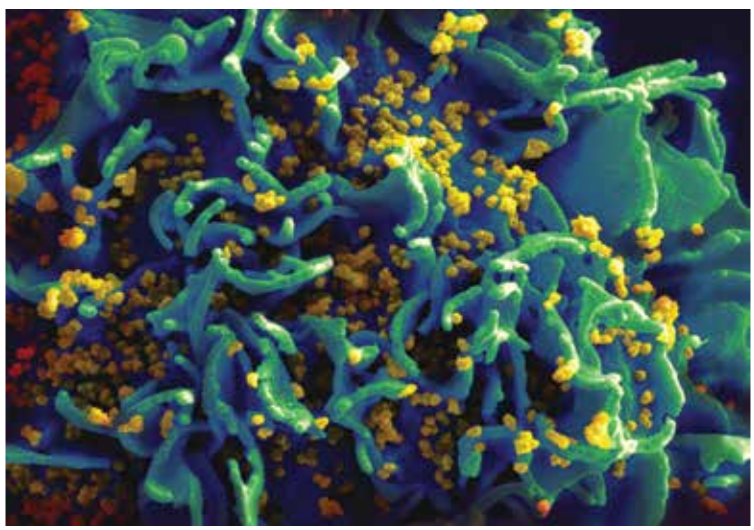


**FIGURE 6.25** Viruses on a T cell. (Credit: NIAID)

## Discovery of viruses

The first clue to the existence of viruses came in 1892 when Dmitry Ivanovsky, a Russian scientist (1864–1920), examined a particular disease of tobacco plants. Ivanovsky made an extract of sap from the diseased plants and passed the extract through a special type of filter that completely removed all the bacterial cells. He was surprised to find that the filtered extract could still infect healthy plants. This meant that the infectious agent was something smaller than the extremely small pores of the special filter. The question remained: what was this infectious agent? Ivanovsky thought that it might be a toxin but carried out no further investigations.

ODD FACT

Some viruses can infect bacteria. Viruses that infect and reproduce in bacterial cells are known as **bacteriophages** (= bacteria eaters).

In 1898, Martinus Beijerinck, a Dutch microbiologist (1851–1931), found that the infectious agent could multiply in plant cells (and so was not a toxin).

Further research showed that this 'invisible' infectious agent was a particle that could only reproduce in the living cells of specific host organisms. Further, these particles, termed viruses, were recognised as causing diseases, not only in plants, but also in animals and people. Only after the invention of the electron microscope in the 1930s were the first viruses able to be seen.

In the following section, we will explore how viruses differ in various ways, including the nature of their genetic material and whether or not they have an external envelope.

### Viruses: enveloped or not?

All viral particles have a simple non-cellular structure that includes:

(1) genetic material, either DNA or RNA, organised as a single molecule or as several molecules

(2) a protein shell, known as a **capsid**, that surrounds the genetic material.

The combination of genetic material plus its surrounding protein coat or capsid = **nucleocapsid**.

For some viruses, the nucelocapsid comprises their total structure; these viruses are said to be **naked** or non-enveloped. Examples of naked viruses include those that cause polio, hepatitis A, and human papillomavirus (HPV).

Other viruses have an additional outer envelope that surrounds the nucelocapsid. Typically, this envelope is a segment of the plasma membrane that is 'captured' when the virus buds from the infected host cell (see figure 6.26). The budding process does not destroy the plasma membrane so that the host cell is not killed. As a result, some virus-infected cells may continue to shed viruses for some time so that the viral infection can persist until the cell eventually dies. Examples of these **enveloped viruses** include those that cause chickenpox, influenza, rubella, rabies, AIDS and mumps.

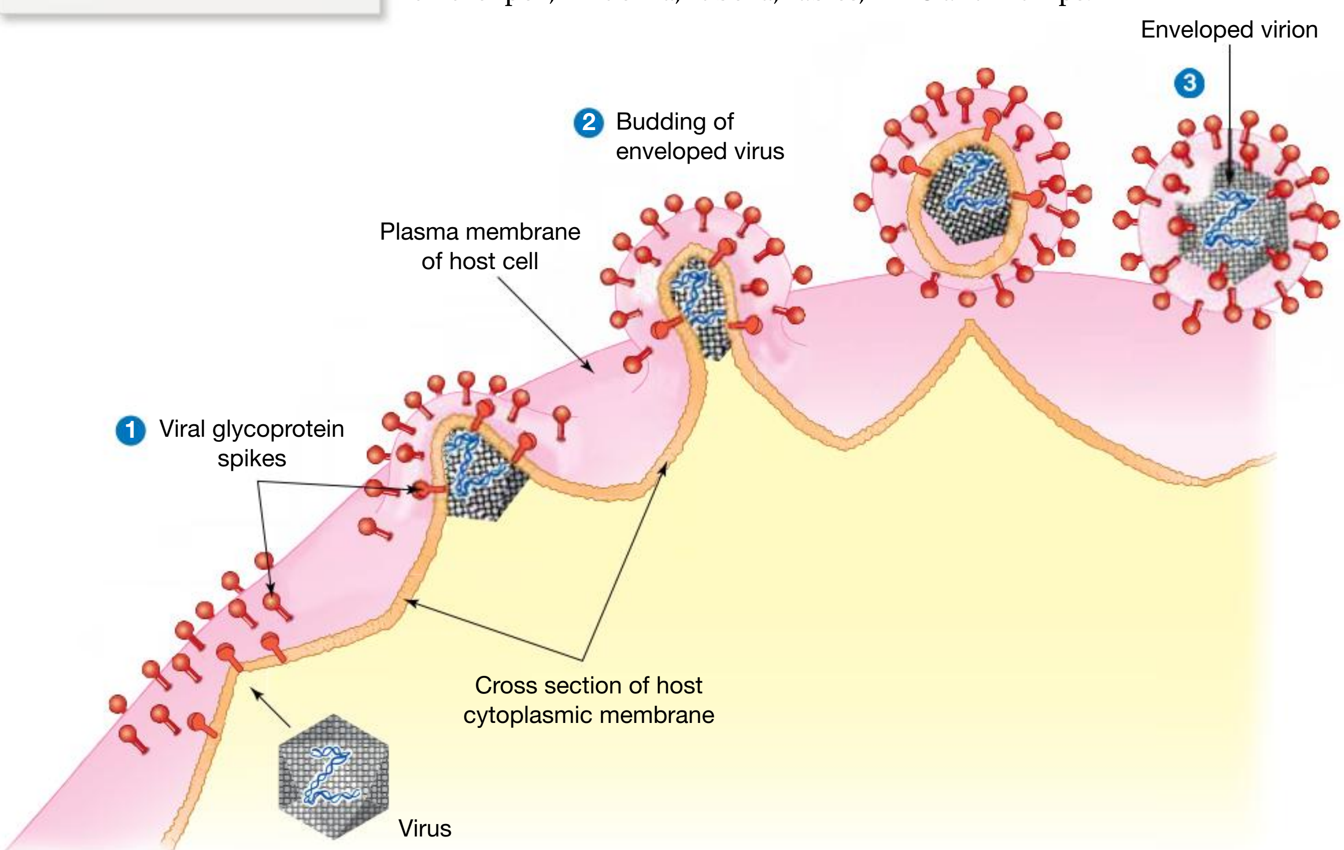



**FIGURE 6.26** Enveloped viruses capture a segment of the plasma membrane of their host cell as each new virus particle buds out. Note that the virus embeds some of its own proteins into this envelope.

Figure 6.27 shows examples of these two kinds of viruses. The outer envelope has viral components (spikes) studded over its outer surface. **These spikes are glycoproteins (proteins that are combined with sugars) and they are specific to each kind of virus**. For example, the influenza A virus has two glycoproteins on its envelope, hemagglutinin (HA) and neuraminidase (NA ). (We will meet HA and NA again in chapter 16.) The human immunodeficiency virus (HIV) has two glycoproteins, GP120 and GP160, on its outer envelope.

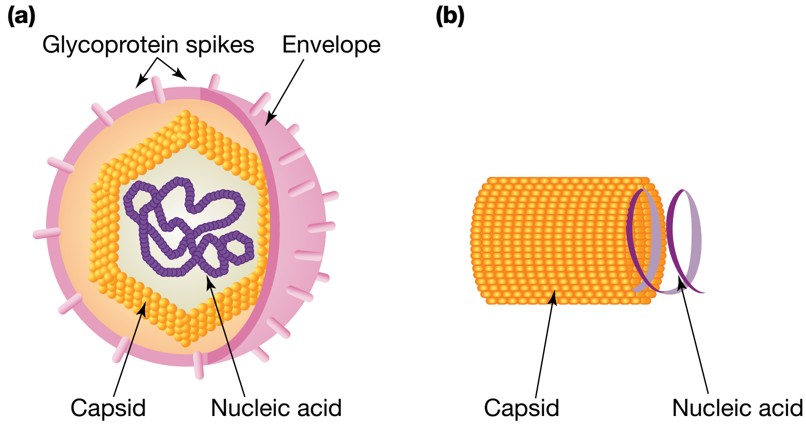



**FIGURE 6.27** **(a)** Diagram of an enveloped virus. The 'spikes' are viral glycoproteins and are specific to different viruses embedded in its outer envelope.**(b)** Diagram showing a naked virus that is composed of nucleic acid and a capsid made of protein.

Viruses with an outer envelope are more sensitive to heat, drying, acid, and detergent treatment than naked viruses. As a result, enveloped viruses can be destroyed more easily by sterilisation. Because enveloped viruses do not last long outside their host cells, they are typically transmitted directly from host to host in body fluids or in airborne particles. In contrast, naked viruses can be transmitted from contaminated surfaces as well as by direct host-to-host transmission.

## Viruses: genetic material

The genetic material of all eukaryotic organisms (plants, animals and fungi) and all prokaryotic organisms (bacteria and archaea) is double-stranded DNA (dsDNA).

In contrast, the genetic material of viruses is much more diverse. **The genomes of different viruses are composed of either DNA or RNA, and this DNA and RNA may be either double-stranded (ds) or single-stranded (ss)**. For some viruses, their genetic material is present as a single nucleic acid molecule. In other viruses, the genetic material consists of several segments of nucleic acid. For example, the genome of the influenza A virus consists of eight segments of ssRNA, the HIV genome consists of two segments of ssRNA, the polio virus genome is a single segment of ssRNA, and the genome of the smallpox virus is a single segment of dsDNA.

The genomes of viruses are relatively tiny and contain only the genetic instructions to make viral genetic material and its small number of viral proteins. The ssRNA of the Ebola virus (EBOV) is about 19 000 nucleotides long and encodes information for just seven proteins. (Contrast this with the much larger genomes of bacteria, such as the four million or so base pairs in the DNA of *Escherichia coli*, a common resident of the human gut, with a genome that codes for several thousand proteins.) What is the size of the human genome?

The rate of mutation of RNA viruses is much higher than that of DNA viruses. As a result, RNA viruses, such as the influenza A virus, are constantly changing and can produce new strains. For this reason, vaccines against influenza that are effective in one year can be ineffective if a new strain of influenza emerges.

Table 6.8 shows the nature of the genetic material of several viruses responsible for human diseases and also identifies whether or not they have an envelope.

**TABLE 6.8** Some viruses responsible for human diseases. The structure of these viruses, either enveloped or naked, and the nature of the genetic material that comprises the viral genomes are identified. (The identifier 'positive' indicates that this RNA can act directly as the mRNA template for the translation of viral proteins. The identifier 'negative' indicates that this RNA must first be copied into its complementary RNA strand, which then acts as mRNA.)

| Virus | Disease | Structure | Genetic material |
|---|---|---|---|
| herpes viruses | chickenpox (varicella) | enveloped | dsDNA |
| smallpox virus | smallpox | enveloped | dsDNA |
| parvovirus B19 | erythema infectiosum | naked | ssDNA |
| rotavirus | gastroenteritis | naked | dsRNA |
| Ebola virus | Ebolavirus disease | enveloped | ssRNA (negative) |
| influenza viruses | influenza | enveloped | ssRNA (negative) |
| measles virus | measles | enveloped | ssRNA (negative) |
| rabies virus | rabies | enveloped | ssRNA (negative) |
| SARS coronavirus | severe acute respiratory syndrome | enveloped | ssRNA (positive) |
| poliovirus | poliomyelitis | naked | ssRNA (positive) |
| rubella virus | German measles | enveloped | ssRNA (positive) |
| rhinovirus | common cold | naked | ssRNA (positive) |
| hepatitis A virus | viral hepatitis type A | naked | ssRNA (positive) |
| HIV | AIDS | enveloped | ssRNA-retro |

ss = single-stranded; ds = double-stranded; HIV = human immunodeficiency virus;
AIDS = acquired immune deficiency syndrome; retro = retrovirus

The very small genomes of viruses reflect the fact that viruses have no genes for energy production or any metabolic processes. Viruses are totally dependent on their host cells for their replication and the energy needed for this process. Outside its host cell, a viral particle is metabolically inert. The box at the end of this section gives further details of the mode of replication of viruses.

## Virus release by lysis or budding

A host cell infected by a virus becomes a factory for the production of multiple copies of the virus. The mode of release of viral particles from an infected cell may be by budding or by cell lysis.

- Enveloped viruses are released from an infected cell by a process of budding, in which virions are released until the infected cell finally dies. Figure 6.28 shows viral particles that have budded from a virus-infected eukaryotic cell.
- Naked virus particles are commonly released from the host cell in a process called **lysis**. In this process, the infected host cell 'explodes' as its plasma membrane disintegrates and viral particles are released into the extracellular fluid from where they can infect other cells.

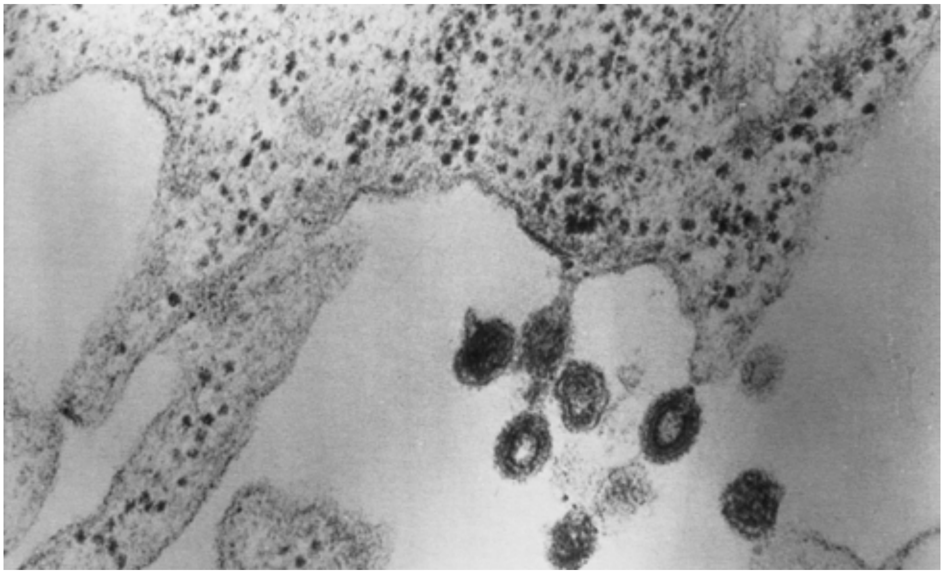


**FIGURE 6.28** Electron micrograph showing viral particles that have budded from the plasma membrane of a virus-infected cell. As each particle buds out, it captures part of the plasma membrane of the host cell that forms the viral envelope. Refer also to figure 6.26.

A clue to the scale of viral replication and release comes from the observation that one gram of faeces of a child infected with rotavirus may contain trillions of rotavirus particles that have been shed from infected gut cells.

## KEY IDEAS

- **Viruses are responsible for many diseases in humans, animals and plants.**
- **Viruses are non-cellular agents consisting of genetic material, enclosed within a protein capsid.**
- **Viruses can only reproduce within the living target cells of their particular hosts.**
- **The genetic material of viruses is either DNA or RNA, and the nucleic acid is either single-stranded or double-stranded.**
- **Some viruses have an outer envelope while other viruses are naked.**
- **Viral glycoproteins on the outer envelope of different viruses are specific to each kind of virus.**
- **The mode of release of viral particles from an infected host cell may be by budding or by cell lysis.**

## QUICK CHECK

**14** List three human diseases that are caused by viruses.

**15** Identify the following statements as true or false:

- **a** Viruses are smaller than eukaryotic cells but are about the same size as bacterial cells.
- **b** Viruses are totally dependent on their host cells for reproduction.
- **c** Viruses are obligate extracellular parasites.
- **d** Multiple copies of naked viruses are released by lysis of the infected host cell.
- **e** Viral genomes are composed of double-stranded DNA.

**16** Consider enveloped viruses P, Q and R that cause three different human diseases.

Identify the following statement as either true or false:

- **a** Viruses P, Q and R would be expected to have distinctively different spikes.
- **b** The information encoded in the genetic material of virus P and R would be identical.
- **c** The genetic material of virus Q would be expected to encode tens of different proteins.
- **d** Copies of virus R would be expected to be released by budding from its host cells.

## VIRUSES AND THEIR REPLICATION

The genetic material of different viruses may be:

- DNA, either double-stranded DNA (dsDNA) or single-stranded DNA (ssDNA).
- RNA, either single-stranded RNA (ssRNA) or double-stranded RNA (dsRNA).

The single-stranded RNA may be either a positive (+) or a negative (−) sense strand.

A positive (+) sense ssRNA molecule can act directly as a messenger RNA (mRNA) template from which viral proteins can be translated. A negative (−) sense ssRNA molecule must first be transcribed into its complementary (positive sense) strand before it can be used as the mRNA template.

Events involved in the replication of an RNA-positive virus include the following steps:

1. **Attachment** of the viral particle to specific receptors on the plasma membrane of the host cell.
2. **Entry** of the viral particle into the host cell, usually by endocytosis. Once inside the cell, the viral RNA is released.
3. **Translation** of the viral RNA to produce viral proteins. This process occurs in the cytoplasm of the host cell and uses the ribosomes of the host. The translated proteins include the structural proteins of the viral capsid. (In the case of an enveloped virus, the proteins that form the 'spikes' of the viral envelope are also translated.)
4. One of the key viral proteins transcribed is **RNA polymerase**, an enzyme needed for **transcription** of the viral RNA.
5. **Assembly** of the new viral particles comprising the viral RNA enclosed within the protein capsid — this assembly occurs on the endoplasmic reticulum of the host cell.
6. **Release** of mature virions by budding. (In the case of an enveloped virus, as the virus buds from the host cell, it captures part of the plasma membrane of the host cell that forms the envelope.)

Figure 6.29 shows a highly simplified version of the replication of a naked ssRNA (positive) virus in its host cell.

The time required for the reproductive cycle of viruses varies from several hours to more than three days (72 hours). In terms of the number of viral particles produced, one of the highest is for the poliovirus, which can produce more than 100 000 viral particles per host cell (Ref: Baron, S., ed., *Medical Microbiology*, 4th ed. (Galveston TX: University of Texas Medical Branch at Galveston, 1996).

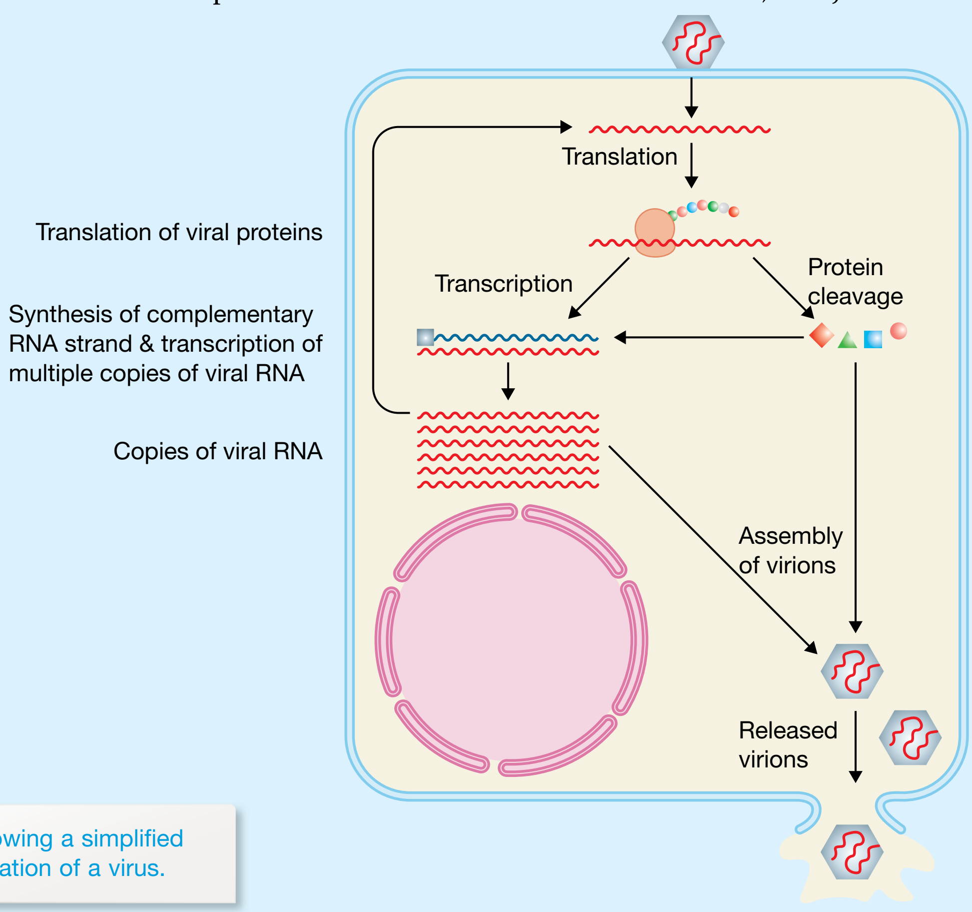



**FIGURE 6.29** Diagram showing a simplified representation of the replication of a virus.

# Prions: a new kind of pathogen

**ODD FACT**

Scrapie has been known since the eighteenth century. Spanish shepherds noted that diseased sheep displayed a strange behaviour — they persistently scraped against fences, often removing the wool from their skin. This behaviour led to the disease being called scrapie!

In April 1982, a research paper by Stanley Prusiner (1942–) appeared in a scientific journal (*Science* vol 216, pp. 136–144, 9 April 1982). The title of this paper was *Novel proteinaceous infectious particles cause scrapie*. Scrapie is a degenerative disease affecting the nervous system of sheep, which is found in sheep flocks worldwide, except in Australia and New Zealand. The disease develops slowly in sheep but, after the first symptoms of scrapie appear, they progressively worsen and inevitably cause the death of the sheep.

**In his 1982 paper, Prusiner proposed that the agent that caused scrapie was a small infectious protein**. He introduced the term 'prion' to label this pathogen, creating this label from two words: '***pr**oteinaceous*' and '*infect**ion***'. Prusiner's proposal caused quite a stir in the scientific community because prions created a completely new concept of pathogens — they were simply misfolded versions of a normal cellular protein — nothing like bacteria or viruses. What was most controversial was Prusiner's view that prions could self-propagate.

studyon

Unit 3 | AOS 2 | Topic 2 | Concept 2

**See more** Pathogenic agents: prions

Not surprisingly, many people were skeptical that misfolded proteins could cause an infectious disease. Why? Infectious diseases involve the multiplication of invading pathogens. All other pathogens rely on the information encoded in their genetic material for replication. Yet prions have no genetic material. Without any genetic material, how could proteins multiply and cause a disease? This seemed to defy the rules of biology. Prusiner's critics said that the causative agent of scrapie must be a 'slow virus'. However, experiments showed that the infectivity of the causative agent of scrapie:

(i) was reduced by procedures that destroyed protein
(ii) was not reduced by procedures that affected nucleic acids.

These results supported the view that the causative agent of scrapie was made of protein.

Since then, more evidence in support of the existence of prions has been identified. **Prions are now defined as infectious particles made of protein and lacking nucleic acids.** More diseases have been identified as caused by prions (see below). In 1997, Stanley Prusiner's pioneering research on prions was acknowledged at the highest level when he was awarded the Nobel Prize in Physiology or Medicine for his discovery of 'prions — a new biological principle of infection'.

## Comparing normal and harmful prion protein

Prion proteins are most abundant in nerve cells. The prion protein can exist in two forms:

- a normal harmless cellular form of the protein that is denoted **$PrP^{C}$**
- a harmful infectious prion form of the protein that is denoted **$PrP^{Sc}$**.

(Note: Pr = prion; P = protein; C = cellular; Sc = scrapie.)

Note that the amino acid sequence, that is, the primary structure of the two forms, is identical. **The difference between the harmless and the disease-causing form is the secondary structure of the protein involved**.

The prion protein is in fact a glycoprotein, that is, a protein molecule joined to a sugar.

Figure 6.30 shows the secondary structures of the normal protein and the misfolded protein of a prion. Note that the structure of the normal protein is mainly an alpha helix while that of the harmful prion protein is mainly a beta sheet.

Table 6.9 shows various properties of the normal **$PrP^{C}$** protein and the infectious **$PrP^{Sc}$** prion.

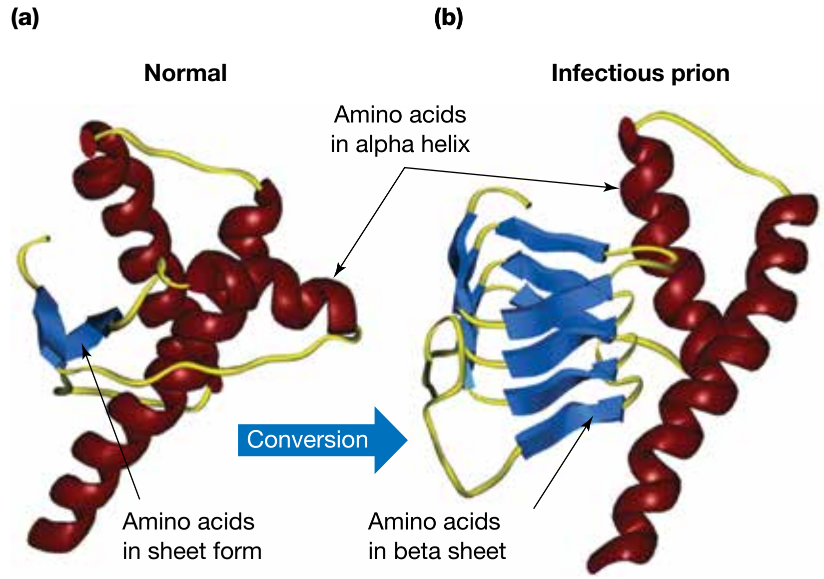



**FIGURE 6.30** Secondary structure of the normal $PrP^{C}$ protein and that of the infectious $PrP^{Sc}$ prion protein. The folding is altered when the normal protein is changed into the harmful prion. This refolding results in part of the alpha helix of the normal protein being converted to a beta-sheet structure. (Image courtesy of Dr Fred Cohen, University of California, San Francisco.)

**TABLE 6.9** Comparison of normal cellular protein and its infectious prion form. Prions are extremely small; individual prions cannot be seen even with an electron microscope, but aggregates of prions can.

| Properties | $PrP^{C}$ | $PrP^{Sc}$ |
|---|---|---|
| **prion form** | normal, harmless | pathogenic |
| **infectivity** | no | yes |
| **amino acids** | same sequence but different shape | same sequence but different shape |
| **secondary structure** | mainly alpha helix: 40% | mainly beta sheets: 45% |
| **cellular location** | on plasma membrane of nerve cells | in vesicles in cytoplasm of nerve cells |
| **organisation** | exist as monomers | aggregate into multimers (prion rods) |
| **solubility** | soluble | insoluble (as rods) |
| **heat treatment** | sensitive | resistant |
| **UV treatment** | sensitive | resistant |

Prions are extremely small, smaller than viruses, and even through an electron microscope only aggregations (clusters), not individual prions, can be seen.

## How do prions 'reproduce'?

The normal form of the prion protein is found mainly in nerve cells. **The normal $PrP^{C}$ protein can be transformed to the harmful disease-causing $PrP^{Sc}$ prion by contact with the harmful prion.** This contact causes the **$PrP^{C}$** protein to unfold and then re-fold abnormally so that its secondary structure is converted to that of the harmful **$PrP^{Sc}$** prion.

As each new harmful prion is formed, it too can convert other normal protein into harmful prions, and so on. This sets up a chain reaction that rapidly multiplies the numbers of harmful prions. This process is not a true biological reproduction, but this multiplication process gives the same end result. Figure 6.31 shows a representation of the chain reaction that multiplies the numbers of harmful prions.

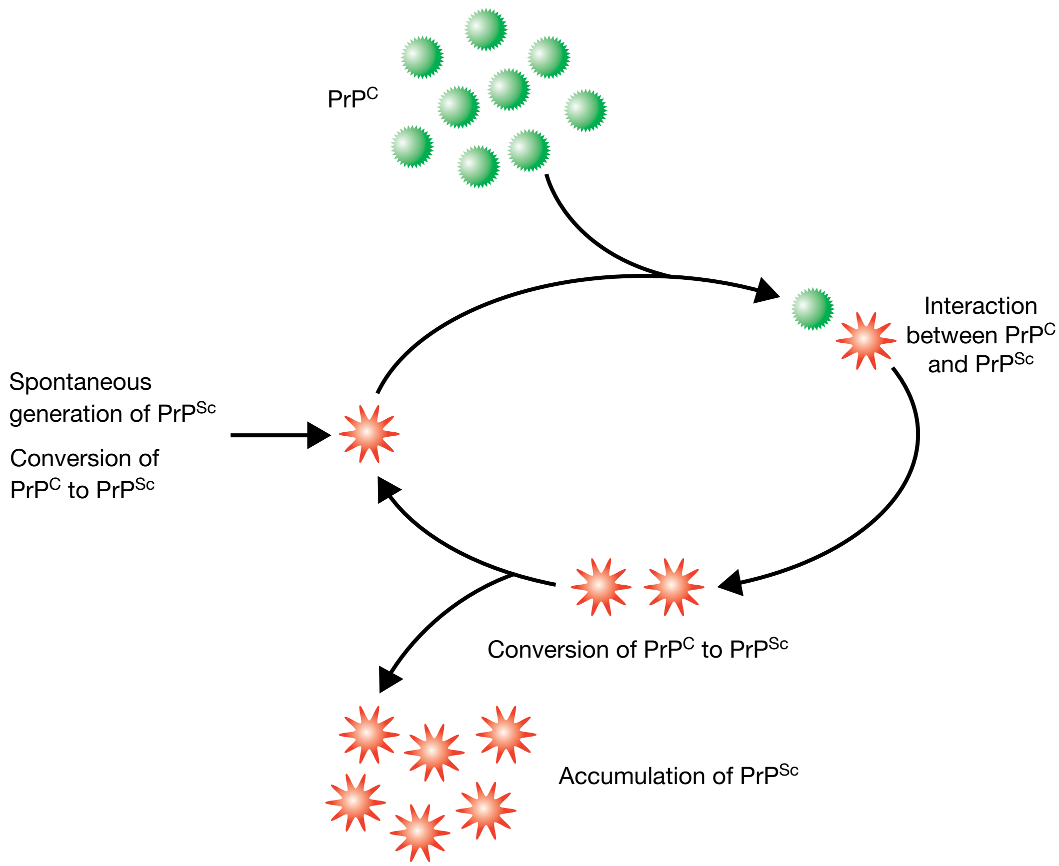



**FIGURE 6.31** Harmful **$PrP^{Sc}$** prions do not remain as monomers (single molecules) but aggregate into rods that form plaques in nerve cells. The accumulation of **$PrP^{Sc}$** prions in the brain causes progressive nerve cell death that is most visible as 'holes' or lesions in the brain, giving a spongiform (sponge-like) appearance to the brain tissue (see figure 6.32).

**FIGURE 6.32** Light microscope image of a section through the brain of a sheep diagnosed with scrapie. The hole in the brain (appearing here shaped like an arrow) results from the deaths of nerve cells. Such holes create a spongiform or sponge-like appearance that is typical of degenerative diseases of the brain, known as encephalopathies.

# Other prion diseases

We have just met scrapie, a prion disease of sheep. In the following section we will explore other prion diseases: 'mad cow disease' and some human prion diseases.

## 'Mad cow disease'

In April 1985, a British farmer reported that one of his cows was behaving unusually — she was nervous, was aggressive, and her movements were very uncoordinated. This cow was the first known to show the symptoms of 'mad cow disease', a disease more scientifically called **bovine spongiform encephalopathy (BSE)**. BSE is a prion disease. Just as occurs in scrapie, harmful prions aggregate in the nerve cells of cattle, damaging brain tissue and producing the spongiform appearance seen when the brain is examined microscopically.

bovine = relating to cattle; spongiform = sponge-like; encephalopathy = encephalo- = brain; pathos = disease.

**FIGURE 6.33** A scientist examines a sample containing brain tissue from a cow during prion-testing for BSE.

This disease, seen in a single cow in 1985, soon began to be seen in cattle on farms across England and Wales. By 1988, more than 2000 cases had been confirmed. The peak occurred during 1992 with an annual total of 36 682 confirmed cases. In 1988, the Southwood committee set up by the British government concluded, based on available information, that '*the risk of transmission to humans appears remote*'. However, this committee recommended that action be taken to prevent any part of an animal suspected of having the BSE disease from entering the food chain.

How had this disease spread to cattle on farms in isolated locations? It is highly probable that the source of the infection was a protein-rich supplement, known as meat and bone meal (MBM) that was fed to cattle. MBM was made from the cooked and ground offal (fat, blood and bones, feathers, organs) of sheep, cattle, pigs and poultry not fit for human consumption, plus the carcasses of dead animals. The procedure for processing MBM was changed in the late 1970s. This change involved the removal of a steam-heat treatment step. It is possible that this change meant that any infectious prions present in the MBM were not inactivated and so could infect cattle that ate the protein supplement.

Prions are resistant to destruction. In a furnace, prions are destroyed by incineration at a temperature of about 480 °C for four hours. In an autoclave, prions are destroyed in about 90 minutes at a temperature of about 130 °C and at a pressure 145 kPA. One effective process to deactivate harmful prions on surgical instruments is a combination of chemical treatment (soaking in 1M sodium hydroxide) and steam autoclaving (30 minutes at 121 °C). Figure 6.33 shows a scientist testing for prions.

The outbreaks of BSE in Britain and in some European countries led to the introduction of bans against the import of beef or beef products from these countries. The practice of feeding MBM is banned in all Australian states and territories.

**ODD FACT**

In 1994, the European Union banned the use of meat and bone meal (MBM) in feeds for cattle and other ruminants. In 2000, this ban was extended to all animal feeds. In 2013, this ban was eased so as to allow MBM from processed poultry and pigs to be fed to farmed fish.

## Human prion diseases

Table 6.10 summarises some human prion diseases and the avenues through which harmful prions may be transmitted or acquired.

**TABLE 6.10** Some human prion diseases.

| Human prion disease | How prion is transmitted | First described case |
|---|---|---|
| **classic Creutzfeldt-Jakob disease** | | 1920 |
| **sporadic CJD (sCJD)** | somatic mutation or spontaneous conversion of $PrP^{C}$ into $PrP^{Sc}$ (?) | |
| **familial CJD (fCJD)** | inherited mutations of **PNPR** gene on human chromosome 20 | |
| **Iatrogenic CJD (iCJD)** | acquired in medical procedures from prion-infected materials | |
| **variant Creutzfeldt-Jakob disease variant CJD (vCJD)** | acquired by ingestion of BSE contaminated beef | 1996 |
| **kuru** | ritual cannibalism at funeral | early 1900s# |

#Date based on oral history of the Fore people.

All of these prion diseases show the following features, but each is distinct:

- a long incubation period, sometimes measured in years
- a progressive deterioration of brain function with an inevitable fatal outcome
- distinctive changes to the brain including loss of neurons and development of lesions ('holes') in specific regions of the brain that produce a spongiform appearance that is visible on microscopic examination.

No treatment is available for people with prion diseases.

### Classic CJD: a human prion disease

The story of one human prion disease began in Breslau, Germany, in June 1913, when a 23-year-old woman, Bertha Elschker, was examined by a junior doctor, Hans Creutzfeldt (1885–1964). Bertha walked unsteadily, her arms jerked uncontrollably, her eyes twitched, and she had a dazed expression — all signs of brain damage. Bertha was admitted to hospital but both her mental and physical condition continued to deteriorate and she experienced many epileptic seizures that continued until her death in August.

Creutzfeldt carried out an autopsy and found that Bertha's brain showed extensive damage and loss of nerve cells. He recognised her condition as a new disease, but, because of the outbreak of World War I, his findings were not published until 1920. Alfons Jakob (1884–1931), a doctor in Hamburg, read Creutzfeldt's paper and recognised that he had patients with similar symptoms and brain damage. Jakob published a description of one of his patients in 1921.

This degenerative disease of the brain is now called **Creutzfeldt-Jakob disease** (CJD) and the young woman, Bertha Elschker, was the first case of classic CJD to be described. This disease is characterised by the loss of nerve cells and the presence of 'holes' in the brain, creating a spongiform (sponge-like) appearance that is seen particularly in the cerebral cortex. Figure 6.34 shows the characteristic holes in a thin section of brain taken from a CJD patient at autopsy.

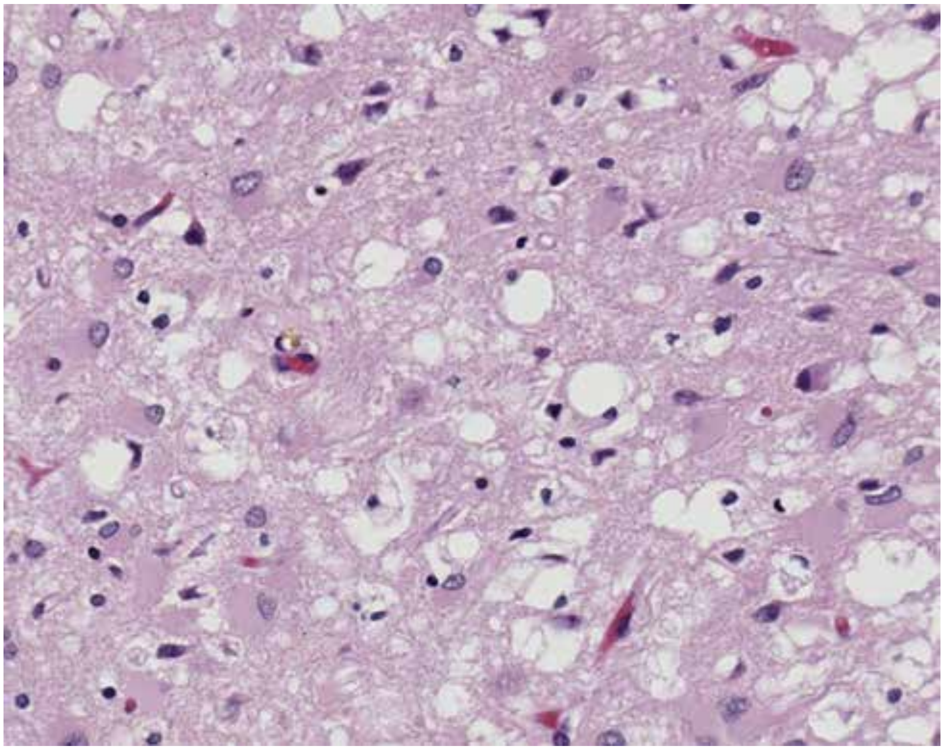


**FIGURE 6.34** Light microscope image of a section of brain from a person who died of classic CJD. Note the 'holes' or lesions — visible as clear areas — that give a sponge-like (spongiform) appearance to the brain, a distinctive feature of this CJD.

Three types of **classic CJD** are identified that differ in their mode of transmission:

- **sporadic**, with no apparent cause (about 85%)
- **familial**, due to inheritance of a mutant allele of the **PRPN gene** on chromosome 20 that encodes the prion protein (about 15%)
- **iatrogenic**, acquired through medical procedures (about 1%).
- **Iatrogenic CJD** was acquired in various ways, including the following:
- From 1958 to 1980, human growth hormone (HGH) for use in medicine was extracted from the pituitary glands of cadavers. This HGH was used, for example, to treat short children who were deficient in the production of this hormone. Years later, some of these children developed CJD because of prion infection of the hormone extract. (From 1980, a genetically engineered HGH became available and removed this risk.)
- Similarly, some cases of CJD have been reported in women who were treated for infertility with the hormone gonadotrophin, extracted from the pituitary glands of cadavers.
- Prion-infected grafts of brain membrane (dura mater) and corneas.
- Use of prion-contaminated neurosurgical instruments and probes.
- Transfusion of prion-infected blood or blood products.

**ODD FACT**

In Australia from 1993 to 2013, a total of 493 cases of sporadic CJD were reported.

## A new prion disease emerges

In mid-1994 in Britain, nearly ten years after 'mad cow disease' was first identified in cattle, 18-year-old Stephen Churchill was studying for his A levels in England with the ambition of joining the Royal Air Force. His studies began to deteriorate, he became depressed and forgetful, and his walking became unsteady. In February 1995, Stephen was diagnosed with a strange neurological disease that was progressive and degenerative. Stephen became unable to communicate and required 24 hour-a-day care, and died on 21 May 1995.

Stephen's death was followed by reports of other cases of this new disease, predominantly in Britain. **This new disease was termed variant CJD (vCJD),**

**because it is different from classic CJD.** Stephen was the first person to die from vCJD. In Britain in the period from 1995 to 2014, 177 cases of variant CJD were reported, all of which ended in death. In the same period in France, 27 cases of vCJD were reported and, for the rest of the world, 25 cases. No cases were reported in Australia.

What was the cause of this new neurological disease? The following box describes the evidence that established a **causal link** between the new prion disease, vCJD, and the consumption of meat from BSE-infected cattle.

## VARIANT CJD: A NEW HUMAN PRION DISEASE

In May 1990, in an attempt to allay public concerns about the increasing numbers of cattle with 'mad cow disease', the British Minister for Agriculture invited reporters and camera crews to photograph him trying to feed a beef burger to his four-year-old daughter. The Minister took a bite from the beef burger stating that it was 'absolutely delicious'. In a formal statement in May 1990, the Chief Medical Officer reassured the British public that 'beef can be eaten safely by everyone, both adults and children, including patients in hospital'. Sadly, this was to prove far from the case, and the deaths of 177 people in Britain from variant CJD (vCJD) were the tragic outcome of eating beef infected with 'mad cow disease' or bovine spongiform encephalopathy (BSE)

Clues to the possible cause of vCJD came from the patterns of occurrence of cases of BSE and those of vCJD (see figure 6.35). Likewise, the occurrence of cases of vCJD in Britain and its absence in countries where BSE was absent also suggested a possible link between BSE and vCJD. However, these patterns do not prove a causal link between 'mad cow disease' and vCJD. More substantive evidence is required.

It was the result of careful experiments that demonstrated a causal link between the occurrence of vCJD in people and their consumption of beef from cattle with bovine spongiform encephalopathy (BSE) or 'mad cow disease'. These experiments are outlined below.

**Experiment 1:** Laboratory mice were injected with samples of infectious brain samples from BSE-infected cows, from patients with vCJD, and from patients with classic CJD.

After this, observations were made of the following:

(i) incubation time

(ii) the particular area of the brain that showed damage

(iii) the type of brain damage observed.

**The results show that the vCJD that developed in mice is identical to BSE** in terms of the range of symptoms, the course of the disease and the microscopic appearance of the brain. Further, these results also confirmed that vCJD was a different disease from classic CJD.

**Experiment 2:** Prion proteins were visualised using a technique known as Western blotting. The prions were those from samples of iatrogenic CJD, sporadic

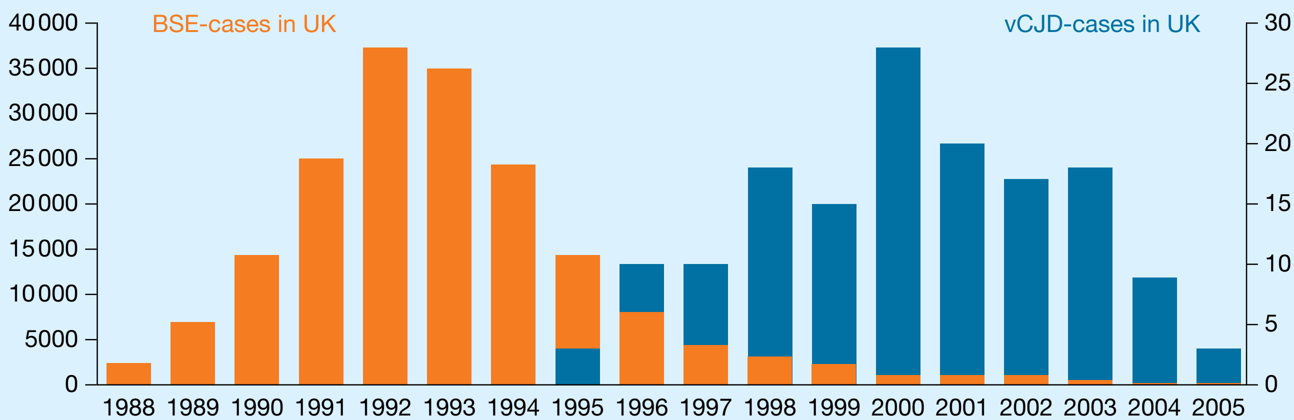



**FIGURE 6.35** Year of onset of BSE (bovine spongiform encephalopathy) in cattle, and the number of cases and year of onset of vCJD (variant Creutzfeldt-Jakob disease) and the number of cases in Britain. The timing of these cases suggested that vCJD was the human form of BSE.

*(continued)*

CJD, variant CJD, scrapie and BSE (see figure 6.36). Note that the pattern seen for human vCJD is identical to that for BSE. This experiment demonstrated that the prion protein in vCJD was identical to that in BSE. **These results support the conclusion that vCJD is a form of human BSE.**

These findings mean that, somehow, the BSE prion from infected cattle had been transmitted to people. It appears most likely that vCJD was a result of the consumption of beef products contaminated by prion-infected brain tissue of cattle. This was the first demonstrated example of the transmission of infectious prions from a non-human species (cattle) to people.

Variant CJD differs from classic CJD in the clinical signs and in other measures such as the median duration of the illness and the median age at death. For vCJD patients, the median duration of illness is 13 to 14 months and the median age at death is 28 years. For classic CJD patients, the median duration of the illness is 4 to 5 months and the median age at death is 68 years.

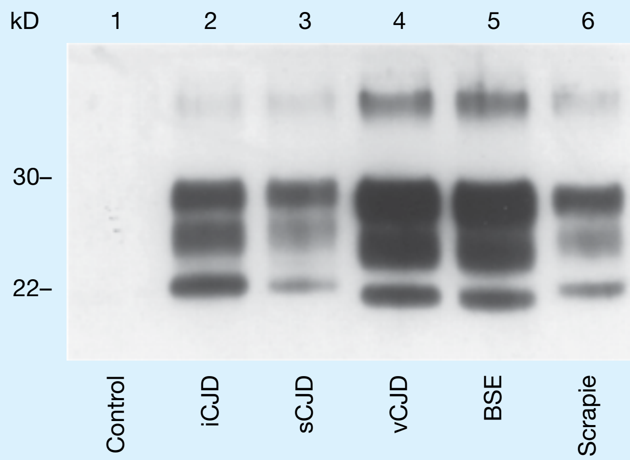



**FIGURE 6.36** Patterns of prion proteins from brains of mice infected by samples from people with iCJD (lane 2), with sporadic CJD (lane 3), with variant CJD (lane 4), from BSE-infected cattle (lane 5), and from scrapie-infected sheep (lane 6). The control sample in lane 1 comes from the brain of a healthy uninfected mouse. (Credit: National Academy of Sciences, USA.)

## KEY IDEAS

- Prions are infectious particles made of protein and lacking nucleic acids that are found mainly in nerve cells.
- The normal $PrP^{C}$ protein can be transformed to the harmful disease-causing $PrP^{Sc}$ prion by contact with the harmful prion.
- Prion diseases of farm animals include scrapie in sheep and bovine spongiform encephalopathy (BSE) in cattle.
- Creutzfeldt-Jakob disease is one example of a human prion disease.
- Variant CJD (vCJD) in people was caused by consumption of BSE-infected beef.

## QUICK CHECK

**17** Brain tissue with a prion infection shows a spongiform appearance. What is meant by the term ‘spongiform’?

**18** Identify the following as true or false:

**a** Prions reproduce in a similar manner to viruses.
**b** A chemical test of prions would show that they are composed of protein.
**c** A normal prion and a harmful prion differ in their tertiary structure.
**d** All cases of classic CJD have the same cause.
**e** Variant CJD (vCJD) resulted from the consumption of BSE-infected beef.

# Self or non-self?

You have now met various pathogens. For them, the human body is a warm and nutrient-rich environment in which the cellular pathogens can thrive and reproduce. The human body also provides an ideal environment for viruses, with access to cell organelles for viruses to make multiple copies of more viruses. Because of this, **the human body is a constant target of infectious agents and we are constantly at risk from a variety of infectious agents.**

So, infections are common events — people around you sneeze and you may breathe in airborne pathogens, food you eat and water you drink may carry pathogens into your gut (see figure 6.37), objects you handle may be contaminated and introduce pathogens onto your skin that you might then transfer into your mouth or nose, and a cut on your finger may allow pathogenic bacteria or their spores to reach sub-cutaneous tissue.

**FIGURE 6.37** Photo of an agar plate that shows growth of *Salmonella* bacteria, which cause severe food poisoning. These bacteria were isolated from contaminated food samples and cultured on an agar plate for 24 hours. Foods most likely to carry *Salmonella* bacteria include raw or undercooked eggs and meat, and faecal-contaminated water or food items washed in such water. *Salmonella* outbreaks occurred in Melbourne in August 2015 (with 130 cases) and in February 2016 (with 200 cases).

**studyon**

Unit 3 | AOS 2 | Topic 2 | Concept 4

**Defence against infection: humans**
Summary screen and practice questions

While infections are common, the transition from infection to disease is far less common. In reality, most people are usually healthy and free of disease most of the time. **This indicates that, in general, the human body has defence mechanisms that are highly effective against infectious agents**. These defence mechanisms are functions of the remarkable **immune system** of the human body.

Defence against infection depends on the ability of the body's immune system to:

1. **identify cells or molecules that are foreign or 'non-self' and react to and eliminate them.**
   These non-self cells and molecules include bacteria, bacterial toxins, viruses, and even one's own altered body cells, such as cancer cells. Non-self markers on cells and foreign molecules are examples of foreign **antigens**, which are like labels signalling to the body's immune system: *Eliminate me!*
   If process 1 fails, infectious pathogens can become established on or in body cells, reproducing and causing damage so that the infection develops into a specific **disease**.
2. **recognise the body's own cells and the compounds they produce as 'self' and *not* react against them.**
   If process 2 fails and the immune system attacks the body's own cells, the damage produces a so-called autoimmune disease.

Prions are 'self' proteins. Even though infectious prions are an abnormally folded form of one of the body's own proteins, this change does not make them foreign antigens. As a result, the immune system of an affected person does *not* recognise infectious **PrP$^{Sc}$** prions as non-self, and so does not react against them.

**studyon**

Unit 3 | AOS 2 | Topic 2 | Concept 1

**Antigens**
Summary screen and practice questions

## Antigens produce an immune response

Normally, an immune response is produced only when the immune system recognises the presence of non-self markers on cell surfaces and non-self molecules in the blood stream. Substances that can stimulate an immune response

study on

Unit 3
AOS 2
Topic 2
Concept 1

Do more
Antigens

Any molecule that can initiate an immune response is strictly termed an immunogen, while an antigen is any molecule (or part of a molecule) that can stimulate a specific immune response, namely the production of antibodies (Note: antigen = ***anti***body ***gen***erator.) All antigens are immunogens, but not all immunogens are antigens. However, in this book we are using antigens in the sense of any molecule that initiates an immune response.

Rarely, people's immune systems react abnormally by producing **autoantibodies** against their own molecules. This immune response results in the so-called autoimmune diseases (see chapter 8, page 351).

**self-antigen**: a molecule in person A that can produce an immune response in another person, but which is normally tolerated by the immune system of person A.

**auto-antigen**: an antigen that is part of the normal chemical makeup of a person but which produces an abnormal immune response in that person, typically by the production of antibodies, known as autoantibodies.

by the body are commonly termed **antigens**. In the strictest sense, these substances are immunogens.

Most antigens are proteins, but some are lipoproteins, polysaccharides and lipopolysaccharides, lipids, nucleic acids, and even some metals, such as mercury. An antigen may be a molecule or part of a molecule. Some antigens originate from the external environment, such as proteins in or on bacteria and viruses or even a short segment of these proteins, chemicals in snake venom, drugs, dust, proteins in food or protein markers on cells from an unrelated person. If these non-self antigens gain entry to a person's internal environment, the person's immune system recognises them as foreign and an immune response is activated to eliminate them (see figure 6.38). Other antigens are produced within infected cells of the body.

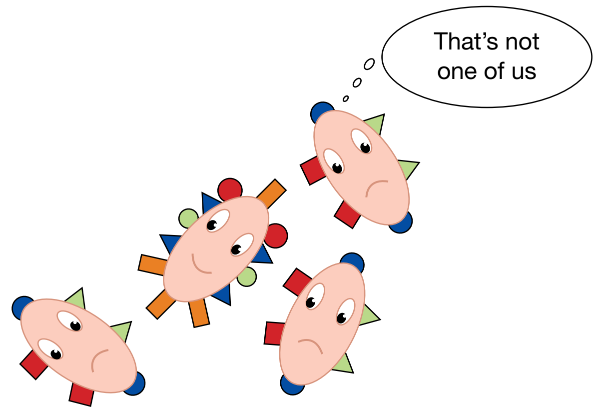



**FIGURE 6.38** The 'self cells' of a person carry protein markers on their surfaces. Cells with different markers are identified as 'non-self' and come under attack by the immune system.

Antigens produce immune responses in your body. However, the many functional and structural molecules in and on your own healthy cells and tissues do not normally cause an immune response. You do not normally react against your own cells and molecules. Instead, your immune system recognises these cells and molecules as **self antigens** and they are tolerated by your immune system. Of course, your self antigens would be recognised as non-self antigens by the immune system of another person, apart from an identical sibling.

Figure 6.39 shows a stylised diagram of an immune cell that checks for non-self invaders. Note that the plasma membrane of this immune cells carries:

- self antigens that identify this immune cell as 'self'
- receptors for self antigens so that this cell can identify and not attack other body cells
- receptors for foreign antigens so that the immune cell can identify foreign material and signal other immune cells to eliminate it.

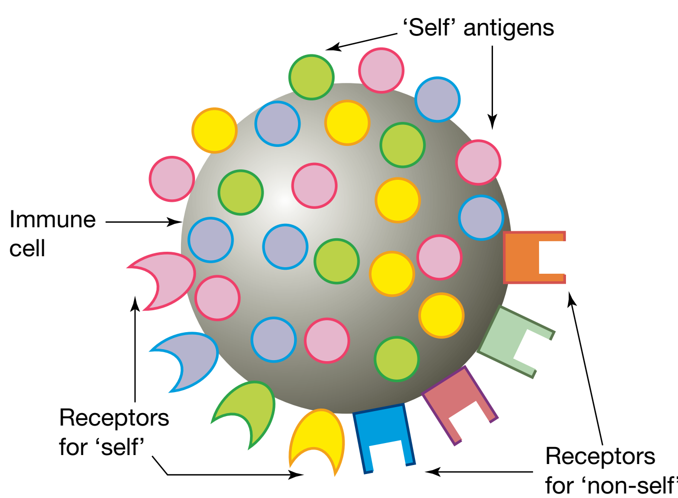



**FIGURE 6.39** Stylised and simplified diagram of an immune cell with 'self' antigens on its surface as well as many receptors. Some receptors identify self antigens, and other receptors recognise non-self antigens.

studyon

Unit 3
AOS 2
Topic 2
Concept 1

**See more**
Antigens

## Distinguishing self from non-self

Let's now look at how the immune system distinguishes 'self' from 'non-self'.

The immune system of a person identifies cells as either self or non-self, depending on particular protein markers (antigens) that they carry on their surfaces. The presence of self antigens on cells indicates that they belong to that particular person and they are tolerated. However, the presence of non-self antigens on a cell indicates that the cell is foreign and these antigens stimulate an immune response. Making the correct distinction between self and non-self antigens is critical and we will explore this issue in the context of organ and tissue transplants (see pages 265–270) and later, through autoimmune diseases (see chapter 8, page 351).

In the following section we will meet a major group of antigens that identify all your nucleated cells as 'self' and that are important in organ and tissue transplantation.

### HLA markers distinguish self from non-self

In 1958, the French researcher Jean Dausset (1916–2005) discovered that the immune system uses surface antigens on human cells to identify the cells that are self and the cells that are foreign or non-self. Dausset was the first to recognise that survival rates of transplanted organs were greatest when donor and recipient were identical twins, as compared to the situation when donor and recipient were unrelated. This observation pointed to the importance of matching the key antigens of the donor organ with the transplant recipient.

**ODD FACT**

The HLA genes are the human versions of a group of genes present in virtually all vertebrates. These genes are known as the major histocompatibility complex (MHC). In people, these genes form a very closely linked cluster on the number-6 chromosome.

The key antigens are called the **human leucocyte antigens (HLA)**. The various antigens fall into six major groups, namely A, B, C, DP, DQ, and DR. For each group, many different versions of the antigen exist, such as A1, A7, A9 and so on. **The various antigens encoded by a person's particular HLA genes identify the 'self' status of that person's cells.**

*Where are the different HLA markers found?*

- The various **A**, **B** and **C** antigens are present on the surfaces of all nucleated cells of your body. These proteins are called **Class I HLA markers** (also known as Class I MHC markers).
- The antigens encoded by genes **DP**, **DQ** and **DR** are found only on the cell surfaces of special white blood cells, known as T cells and B cells. These proteins are called **Class II HLA markers** (also known as Class II MHC markers).

### Transplantation and tissue typing

Paul is suffering from end-stage kidney disease and he is on the waiting list for a kidney transplant. His details are included in a centralised computer list of people waiting for a kidney from a deceased donor. Information about Paul on this list includes his tissue type or HLA type, as well as his ABO blood type.

To obtain his tissue type, a sample of Paul's white blood cells went through a procedure called **tissue typing** or **HLA typing.** Tissue typing reveals the major HLA antigens present on the surfaces of his body cells and tissues. Paul's tissue type for four of his HLA markers was identified as:

A1 A28, B27 B40, C3 C8, DQ5 DQ6, and DR3 DR15

Can you suggest why Paul's white blood cells were used to identify his tissue type, rather than his red blood cells or cells from the lining of his oral cavity (mouth)?

### Why is tissue typing done?

Tissue typing or HLA typing is done to check if the antigens present on the cells of a potential recipient are compatible with and match those present on cells of the donor organ. In general, the closer the HLA match between donor and recipient, the greater the chance of a successful transplant.

If the tissue typing reveals a poor match between the HLA markers on Paul's cells and on those of a deceased donor kidney, that kidney would be

**ODD FACT**

Matching of both the DNA sequences of the alleles of the HLA gene of the donor and recipient and the protein markers present on their cells ensures the best possible match and chance for a successful transplant.

allocated to a compatible, that is, a well-matched recipient on the waiting list. An immune attack by the recipient's immune system on a transplanted kidney because of a mismatch between HLA markers is one of several reasons that a kidney transplant may not be successful (see figure 6.40).

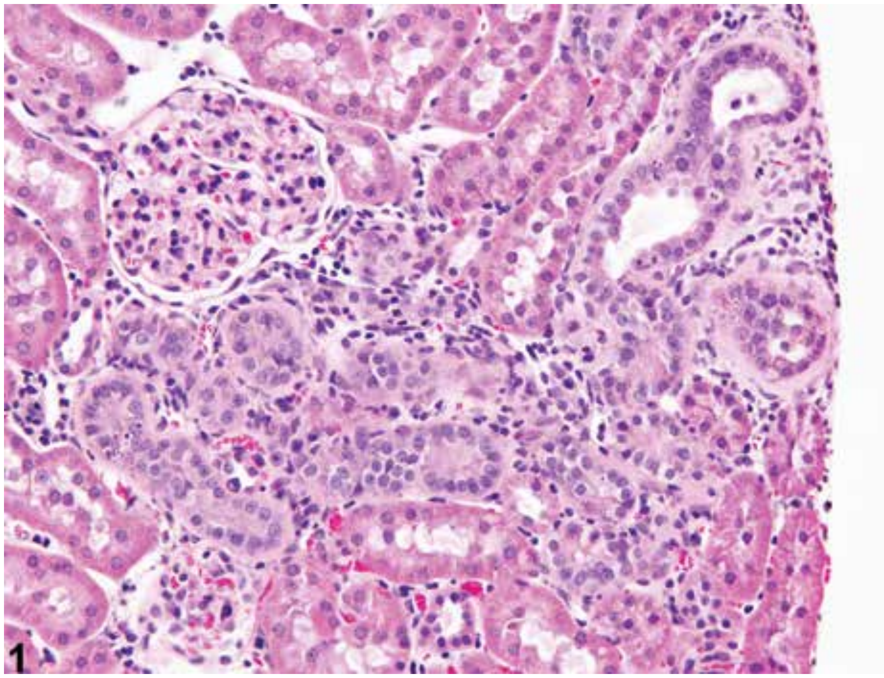



**FIGURE 6.40** Light microscope image of a section through a transplanted kidney that shows signs of acute rejection. The numerous purple dots are the nuclei of cells of the immune system. These cells have migrated to the transplanted kidney and are mounting an immune attack on the kidney cells. (Image courtesy of the National Toxicology Program, US Department of Health and Human Services.)

HLA tissue typing of the white blood cells is routinely performed for potential recipients of transplants of donor organs such as kidney, liver, pancreas, lungs and heart, and other tissues, such as bone marrow.

**ODD FACT**

In Australia, about 1700 corneal transplants are performed each year. Cornea transplants, however, do not require tissue typing because the cornea is composed of layers of the protein collagen, and significantly, *the cornea has no blood supply*, so that immune cells cannot travel to the transplant and attack it.

## How is tissue typing done?

Tissue typing of transplant patients is performed using a sample of a potential recipient's white blood cells, and the particular HLA markers present are identified. DNA-based tests are also used to identify the particular alleles of the major HLA genes that are present.

In some cases, tissue typing identifies a minimum of six major HLA markers: two A markers, two B markers and two DR markers. In other cases, tissue typing identifies an additional four HLA markers: two C and two DQ markers.

Each person can have a maximum of two different antigens for each HLA marker. Table 6.11 shows the HLA tissue types of three recipients on the waiting list for a deceased donor kidney, including Paul. Which of these three possible recipients has the best match between their tissue type and that of a deceased donor kidney that has become available? Will Paul finally receive a deceased donor kidney transplant?

**TABLE 6.11** Tissue type of a deceased donor kidney and those of three potential recipients, including Paul. Tissue types are shown for five pairs of HLA markers. Green shading denotes a match between donor and recipient for a particular antigen.

| Major HLA gene | Donor kidney Antigens present | Potential recipient 1 Antigens present | Potential recipient 2 Antigens present | Paul Antigens present |
|---|---|---|---|---|
| **A** | A1, A28 | A3, A28 | A10, A28 | A1, A28 |
| **B** | B3, B40 | B7, B11 | B7, B15 | B27, B40 |
| **C** | C3, C18 | C6, C8 | C8, C18 | C3, C8 |
| **DQ** | DQ5, DQ6 | DQ2, DQ4 | DQ2, DQ5 | DQ5, DQ6 |
| **DR** | DR3, DR15 | DR4, DR15 | DR8, DR15 | DR3, DR15 |

Examine table 6.11 above. Note that Paul's tissue type is the best match with that of the available deceased donor kidney. Eight of his protein markers (out of a possible 10) match the donor kidney. So, Paul has an 8/10 match. In contrast, potential recipient 1 has a poor match, with just a 2/10 match. In general, the lower the match between a recipient and a deceased donor kidney, the greater the risk of rejection of the transplant by the recipient's immune system.

Non-self antigens on a donor organ can stimulate a recipient's immune system to reject a transplant. However, transplantation of organs can proceed even if a perfect tissue type match is not available. In order to prevent rejection, transplant recipients are treated with **immunosuppressive drugs**, also known as anti-rejection drugs, to reduce their body's ability to reject a transplanted organ. However, these drugs can also make it more difficult for the patients to fight infection. Can you suggest why?

**ODD FACT**

Cyclosporine is one of many immunosuppressive drugs. Cyclosporin inhibits the immune cells that destroy 'non-self' cells.

**KEY IDEAS**

- The human immune system can identify cells and cell products as self or non-self.
- Antigens are any molecules or parts of a molecule that initiate an immune response.
- HLA markers on the surfaces of all nucleated cells are self antigens that are normally tolerated by a person's immune system.
- Cells with HLA markers that differ from a person's own 'self' markers are identified by that person's immune system as non-self and come under immune attack.
- For organ and tissue transplantation, the HLA markers of a donor organ are matched to those of a potential recipient.

**QUICK CHECK**

**19** Identify the following statements as true or false:
- **a** Unless a perfect match is found between donor organ and recipient, a transplant of that organ to that recipient cannot proceed.
- **b** Tissue typing is the same as HLA typing.
- **c** HLA typing identifies a person's ABO blood group.
- **d** 'Self' is determined by the HLA markers on a person's nucleated cells.
- **e** An example of a non-self cell is a bacterial cell that enters the body through a cut in the skin.

**20** A deceased donor kidney has the following HLA markers: A7 A8, B27 B30, DR4 DR5.
- **a** What protein markers would be expected on a recipient with a perfect match?
- **b** Write a possible HLA tissue type for a recipient with a 3/6 match for the HLA markers A, B and DR.

# Transplantation in Australia

**1. Organ transplants**

In Australia in 2014, a total of 1193 organs were transplanted from a total of 378 donors (see table 6.12). Except for kidneys, all organs for transplantation come from deceased donors. Kidneys may come from either deceased or living donors. In every case, HLA tissue typing was carried out to identify the best possible matches between donor organs and potential recipients.

**TABLE 6.12** Transplanted organs in Australia in the period Jan–Dec 2014. Data source: Australian and New Zealand Organ Donor Registry (ANZOD).

| Organ | Number |
|---|---|
| kidney | 904 |
| liver | 236 |
| lungs | 162 |
| heart | 79 |
| pancreas/islet cells | 54 |
| heart & lung | 4 |
| intestines | 1 |

Table 6.13 below shows the Australian data on the numbers of kidney transplants received by recipients in different age groups. In Australia, kidneys are the only *organs* that are commonly transplanted from a live donor. Which is more common — deceased donor kidney transplants or living donor kidney transplants?

**TABLE 6.13** Numbers of recipients of kidney transplants by age group. Data source: 2014 Annual Report of the Australian and New Zealand Dialysis and Transplant Registry (ANZDATA).

| Country | Donor Type | Graft | 0–4 | 5–14 | 15–24 | 25–34 | 35–44 | 45–54 | 55–64 | 65–74 | 75–84 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Australia | Deceased | 1 | 4 | 7 | 23 | 50 | 96 | 133 | 168 | 82 | 2 |
| | | 2 | 0 | 2 | 4 | 5 | 11 | 12 | 17 | 8 | 0 |
| | | 3 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 0 | 0 |
| | | 4 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| | Living | 1 | 6 | 11 | 26 | 27 | 32 | 51 | 52 | 18 | 0 |
| | | 2 | 0 | 0 | 3 | 2 | 6 | 7 | 6 | 2 | 0 |
| | | 3 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 0 |

The success rate one year after transplantation is 97 per cent for live donor kidney transplants and 91 per cent for deceased donor kidney transplants. (Source: The Australia and New Zealand Dialysis and Transplant Registry (ANZDATA))

**2. Tissue transplants**

In addition to organs, certain tissues are also transplanted. The most common of these tissues is bone marrow. In Australia each year, about 100 children and adults have bone marrow transplants that are effectively stem cell transplants.

Stem cell transplants typically come from matched living donors. In the past, bone marrow, which is a rich source of stem cells, was taken from the pelvic (hip) bone of a living donor under a general anaesthetic. However, today, peripheral blood stem cell transplants (PBSCTs) have replaced bone marrow transplants. The research that led to this advance was carried out at Walter and Eliza Hall Institute in Melbourne (see box at end of chapter 8).

Peripheral blood stem cell transplants are transplants of stem cells that have been stimulated to mobilise from the bone marrow into the blood. Blood is then collected from the donor who has received the stimulating factor, and the stem cells are separated. The donated stem cells are infused directly into the recipient's vein and, from there, the stem cells migrate to the bone marrow space (see figure 6.42; also refer to chapter 7, page 281).

When she was only two and a half years old, Ivy Steel from Hamilton, Victoria, was diagnosed with a rare form of blood cancer: acute lymphoblastic leukaemia. After two years of treatment and long periods in hospital, Ivy was in remission, but just ten days later, her cancer returned. At this stage, the only possible treatment for Ivy was a transplant of immune stem cells to replace her defective bone marrow with healthy tissue. Ivy's younger brother, Van, was found to be a perfect tissue type match and he was thrilled to be able to help

**ODD FACT**

Other tissues that may be transplanted include corneas, heart valves, skin, blood vessels and bone.

his big sister (see figure 6.41). In September 2015, the transplant was successfully performed. Ivy still requires regular checkups and daily medications but the healthy stem cells from Van have given her a new chance of a normal life.

**FIGURE 6.41** Five-year old Ivy Steel and her three-year-old brother, Van, share the same tissue type markers. Van was the donor of stem cells to his sister as part of the only option available to treat Ivy's rare blood cancer. (Image courtesy of David Caird, Newspix.)

**FIGURE 6.42** Stem cells are rarely seen in the bloodstream. However, stem cells can be recruited into the bloodstream by treatment of a donor with stimulating factors called colony stimulating factors (CSFs). The blood is filtered through a machine and the stem cells are separated off.

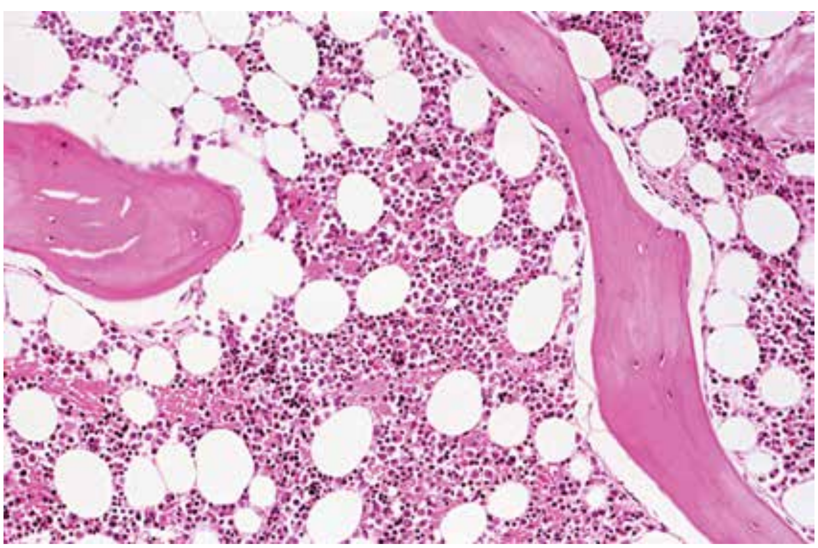


**FIGURE 6.43** Light microscope image of bone marrow tissue showing its cell population consisting of many small stem cells with darkly stained nuclei. The open areas are fat cells that have not taken up the stain.

## Finding living donors

Bone marrow is a rich source of the stem cells that can differentiate into the various specialised blood-forming cells, including red blood cells, and the various white blood cells that are the body's immune cells (see figure 6.43).

People with blood disorders, such as various forms of leukaemia, and immune disorders, such as severe combined immunodeficiency disease (SCID), may undergo treatment that involves high-dose chemotherapy or radiotherapy to kill their faulty bone marrow cells. This is followed by a transplant of bone marrow tissue taken from a matched healthy living donor. The transplanted cells populate the patient's bone marrow with new healthy stem cells. Such a bone marrow transplant effectively gives the recipient a new immune system.

How easy is it to find a tissue type matched for the major HLA markers between a live donor and a recipient? The answer to this question depends on whether the two people are biologically related or not, and, if related, the degree of relationship.

**1. Biologically related people**

For two biologically related people, the chance of the two having identical matching HLA markers is as follows:

- for monozygous (identical) twins or triplets: 100 per cent
- for siblings (brothers and sisters): 25 per cent or 1 in 4
- for first cousins: about 6 per cent, or 1 in 16.

Typically, there is no chance that a parent will have identical HLA markers to those of a child. Parents share just half of their HLA markers with each of their children. The following box shows the inheritance of the HLA protein markers.

As the degree of relationship decreases, the chance of a HLA tissue type match also decreases. As expected, because the chance of a HLA match between siblings is high, live donors of bone marrow are often the siblings of recipients. Otherwise, another biological relative of the recipient, such as a parent, may become the living donor. In the case of Ivy Steel, the donor for her new stem cells was her brother, Van.

## MENDELIAN INHERITANCE OF HLA GENES

Consider the major HLA genes on the number-6 human chromosome. Because these genes are closely linked, crossing over between the genes during egg and sperm formation by meiosis is very low (typically less than 1%). This means that the alleles of the major HLA genes tend to be transmitted from each parent to a child as an unchanged block of parental alleles (see figure 6.44). A set of closely linked alleles is called a **haplotype**.

This is in contrast to the inheritance of the alleles of unlinked genes that can re-assort as a result of independent assortment of chromosomes during meiosis.

Examine the following figure, which shows a cross of two parents for three of the HLA genes.

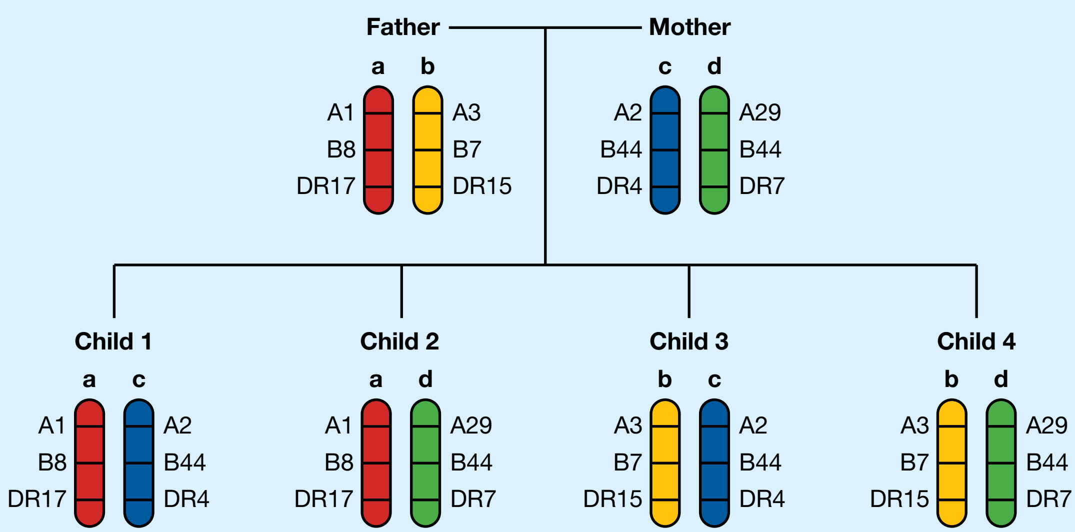



**FIGURE 6.44** Children inherit their HLA genes on one number-6 chromosome from the mother and one number-6 chromosome from the father. Because of the close linkage of these genes, they are inherited almost always without being recombined by crossing over.

Two siblings in one family have a chance of a tissue type match, either in full or in part. There is a 1 in 4 chance that the HLA genotypes of two siblings will be identical, a 1 in 2 chance that they will share half their HLA genes, and a 1 in 4 chance that their HLA genotypes will be different.

**ODD FACT**

The first unrelated bone marrow transplant took place in 1973 in the Memorial Sloan-Kettering Cancer Center in New York City. The recipient was a five-year-old patient with severe combined immunodeficiency syndrome. The matched donor was found in Denmark through the Blood Bank at Rigshospitalet in Copenhagen.

**eBook*plus***

**Weblink**
ABMDR

**2. Non-biologically related people**

The combinations of HLA markers in human populations are very diverse because the HLA genes have many different alleles. For example, for two biologically unrelated people, the chance of having identical HLA markers is low, even more so if these two people have different ethnic backgrounds. For example, the chance of their having matching HLA markers for three major genes is about one in 100 000.

The very low chance of a match between unrelated people reflects the fact that each major HLA group has a large number of different variants. There are nearly 70 common A variants, about 125 common B variants and nearly 50 common C variants. As a result, in human populations, a great diversity of HLA markers exists. In reality, the testing of populations worldwide has revealed many more HLA variants, some of them very rare.

About three quarters of people in need of a bone marrow transplant do not have access to a matched living donor. The hope for such people is that an unrelated matched donor can be identified through a database, either nationally or internationally.

**Bone Marrow Donors Worldwide (BMDW) is a global electronic file that lists information about the HLA tissue markers of more than 26 million potential bone marrow donors.** BMDW was set up in 1989 and is maintained in Leiden, Holland. BMDW collates HLA tissue types from anonymous bone marrow donor registers in 53 countries, including Australia. In Australia, the Australian Bone Marrow Donor Registry (ABMDR) has more than 160 000 registered bone marrow donors (see figure 6.45). Transplant centres worldwide can access the BMDW file to see if any potential matches are available for their patients needing a bone marrow transplant who do not have a suitable sibling donor. Through the international cooperation that supports BMDW, a patient in Australia might become a recipient of bone marrow from a compatible donor in Germany.

ABMDR
Australian Bone Marrow Donor Registry

**FIGURE 6.45** The Australian Bone Marrow Donor Registry (ABMDR) holds the records of more than 160 000 people registered as volunteer bone marrow donors. (Image courtesy of ABMDR.)

## COLONY STIMULATING FACTORS — CSFs

Cells in bone marrow must reproduce continuously because the immune cells in the blood that they produce have a relatively short life. Scientists showed that the continued division of bone marrow cells does not occur spontaneously but is regulated by some factor that was given the name **colony stimulating factor (CSF)**. In 1977, Professor Donald Metcalf (see figure 6.46), an Australian scientist at the Walter and Eliza Hall Institute of Medical Research in Melbourne, discovered CSFs. (CSFs are glycoproteins belonging to the group of proteins called cytokines.)

It was shown that injections of CSFs into mice and primates stimulated production of neutrophils and macrophages without significant toxic effects. Clinical trials commenced in 1987, and today two CSFs are used to treat people who have very low levels of white blood cells in their blood.

What application has this discovery had?

In some diseases, and as a result of certain forms of treatment, such as chemotherapy for cancer, the number of neutrophils and macrophages produced by the body declines significantly. This reduction in immune cells causes reduced resistance to infection.

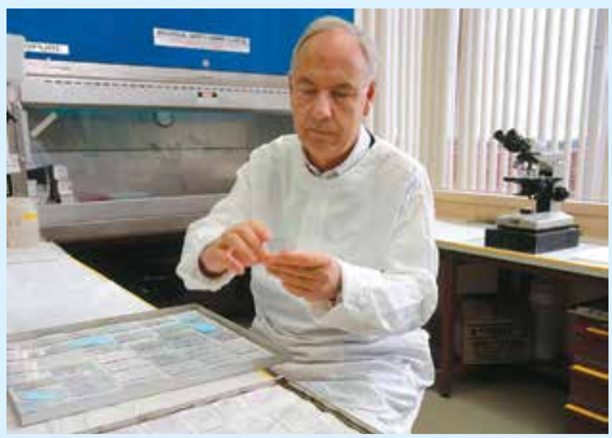


**FIGURE 6.46** Professor Donald Metcalf, scientist, Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia. Professor Metcalf discovered CSFs, which are used in the treatment of patients with various diseases, including blood cancers. He is a leading scientist whose contribution has made a difference not only in extending basic knowledge, but also in the treatment of diseases. (Image courtesy of the Walter and Eliza Hall Institute.)

*(continued)*

CSFs are now used to prevent the fall in neutrophils, thus preventing infections that could otherwise occur following treatment with chemotherapy.

Leukaemia is a cancer, or malignancy, of the blood. **Acute myeloid leukaemia** is a cancer in which a person produces excessive numbers of faulty white blood cells. Something has gone wrong with the blood-forming tissue. There is no point in using CSFs as a first treatment for patients with leukaemia. The faulty bone marrow could respond by producing even more faulty white blood cells. However, after a patient with leukaemia has received other treatments, CSFs may be given to assist their recovery.

Most patients with acute myeloid leukaemia initially receive chemotherapy to kill the faulty white blood cells. For some patients, a few faulty cells remain and can reproduce, causing leukaemia again. These patients can be treated with irradiation to kill all the faulty cells. However, this also kills all the normal bone marrow cells (including blood stem cells), meaning the patient cannot make their own blood. Therefore, in a second step of treatment, the patient is given a transplant of compatible bone marrow (including blood stem cells). This transplant often comes from a relative.

Bone marrow transplants can be very successful. The transplanted bone marrow behaves as the patient's own and produces healthy blood cells. Provided sufficient white blood cells grow, the patient regains normal resistance to infection. Sometimes the bone marrow cells do not divide at a sufficiently high rate to prevent infection. A third step of treatment can help to solve this problem. Injection of CSFs significantly increases the number of neutrophils and macrophages produced in the bone marrow recipient (see figure 6.47). This, in turn, significantly increases the chance that a patient is restored to full health. This was the case with Jose Carreras, a world-famous opera singer, who was one of the first people to be treated with CSFs.

Today, many millions of cancer patients worldwide have received CSF to increase their dangerously low white blood cell levels following their chemotherapy treatment. Australians can be proud of this important medical research.

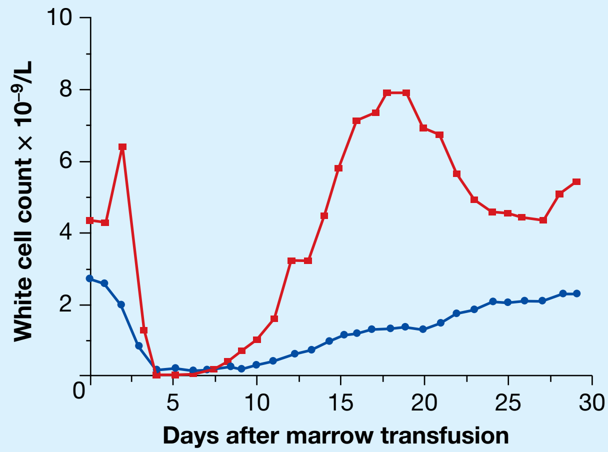



**FIGURE 6.47** Patients (15) who received a bone marrow transplant after irradiation of their own bone marrow and who also received CSFs. The red curve shows the average number of white blood cells in these patients. The blue curve shows the average white cell number in a sample of patients (18) before CSFs were available. Most of the increase was due to an increase in neutrophil production.

In another world first, scientists and physicians in Melbourne discovered that CSFs cause blood stem cells that normally reside in the bone marrow to move (or 'mobilise') into the blood. This was completely unexpected. They also discovered that these blood stem cells could be used instead of bone marrow cells in bone marrow transplantation. This technique, called **peripheral blood stem cell transplantation (PBSCT)** is much safer than traditional bone marrow transplantation because blood can be collected instead of bone marrow, removing the requirement for an anaesthetic. In addition, these mobilised blood stem cells work much better than bone marrow cells, and recovery is much faster after transplantation. For these reasons, mobilised blood stem cell transplantation has overtaken bone marrow transplantation around the world.

## KEY IDEAS

- For siblings, the chance of having matching HLA markers is 1 in 4.
- Many organ transplantations occur in Australia, the most common being kidneys from deceased donors.
- Various tissue transplants occur in Australia, the most common being that of bone marrow from living donors.
- Cooperation between bone marrow registries worldwide enables matching of unrelated donors to recipients.
- For two unrelated people, the chance of matching HLA markers is extremely low.

**eBook*plus***

**Weblinks**
CSF
FasL apoptosis
WHITE NIGHTS

## QUICK CHECK

**21** Identify the following as true or false:
   **a** A parent would be an ideal donor of bone marrow for one of their children.
   **b** A person with SCID would be an ideal candidate for a bone marrow transplantation.
   **c** Identical twins have a 100 per cent chance of a HLA match.
   **d** A mother shares 50 per cent of her HLA markers with her child.
   **e** A bone marrow transplant is effectively a stem cell transplant.

**22** Briefly identify the importance of the Bone Marrow Donors Worldwide database.

**23** Identify three tissues that might be involved in transplantation.

# BIOCHALLENGE

End-stage kidney disease (ESKD) (also termed stage 5 chronic kidney disease (CKD)) is a chronic disease in which kidney function deteriorates to a point where it cannot sustain life. ESKD can only be treated by dialysis or by a kidney transplant. (Refer back to *Nature of Biology Book 1 5th edition*, pages 134–7 for details of dialysis.) ESKD results from kidney damage, and the risk factors include poorly managed diabetes, high blood pressure, obesity and smoking.

Figure 6.48 shows the percentage of people in various age groups receiving treatment for ESKD by dialysis and by kidney transplants. (Data from Australian Institute of Health and Welfare.)

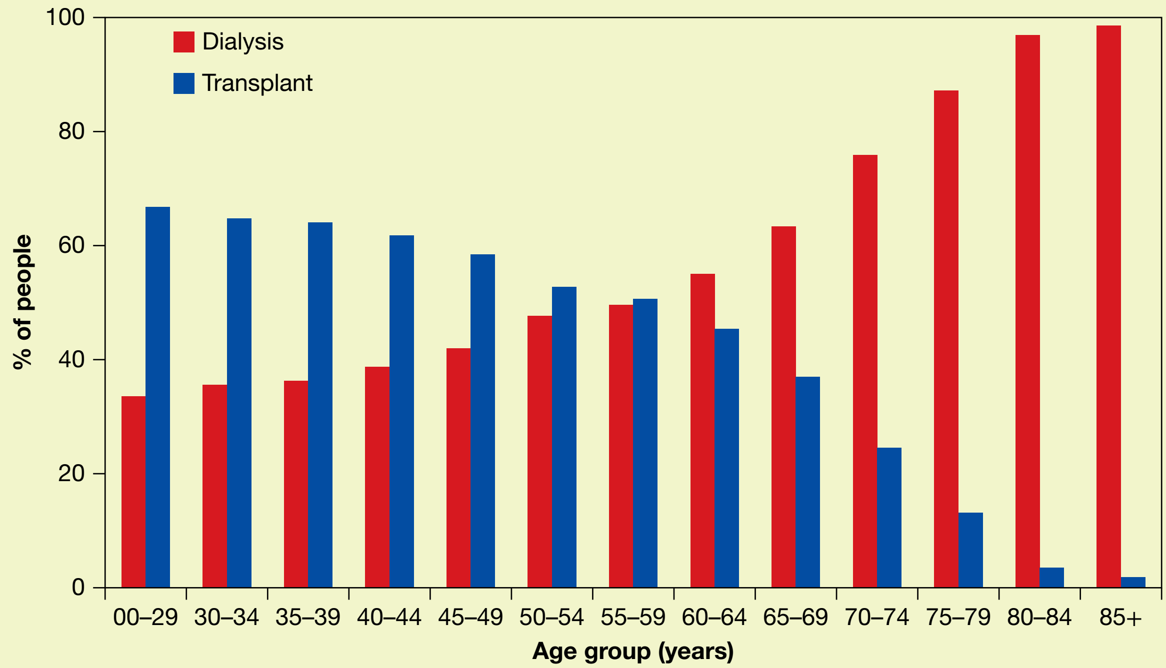



**FIGURE 6.48**

1 Examine these data and answer the following questions:
   a What trend is apparent from these data?
   b What is the more common treatment for ESKD for people aged 0 to 54 years?
   c What limits the greater use of kidney transplants for these age groups?
   d Suggest why treatment by dialysis, rather than kidney transplants, is far more common in the older age groups, from 70 to 85+ years?
   e What risk factors are associated with ESKD?

2 The number of kidneys available for transplantation is far less than the number of people who could benefit from a transplant. Because of this, inclusion criteria for a kidney transplant have been developed as follows:
   1 person has end-stage kidney failure requiring dialysis
   2 anticipated perioperative mortality of person is low
   3 person has a reasonable post-operative life expectancy, defined as having an 80 per cent likelihood of surviving for at least five years after the transplant.
   a What is meant by ‘inclusion criteria’?
   b Refer back to your suggestions in response to question 1d. Did you identify any of these inclusion criteria?

3 a What is the role of tissue typing in kidney transplantation?
   b Can a kidney transplant only occur when a perfect tissue type match exists between the donor kidney and the intended recipient? Explain.

Table 6.14 shows the data for kidney transplants in Australia for 2005–2014 (Data from ANZOD). The source of kidneys from either deceased donors or from live donors is also shown.

**TABLE 6.14** Total kidney transplants in Australia in the period 2005–2014 and the source of donor kidneys.

| | 2014 | 2013 | 2012 | 2011 | 2010 | 2009 | 2008 | 2007 | 2006 | 2005 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Kidney transplants:** | 904 | 897 | 844 | 825 | 846 | 773 | 813 | 613 | 643 | 623 |
| **from deceased donor** | 637 | 630 | 607 | 570 | 550 | 446 | 459 | 342 | 370 | 377 |
| **from live donor** | 267 | 249 | 237 | 255 | 296 | 327 | 354 | 271 | 273 | 246 |

4 Examine these data and answer the following questions:
   a What trend is apparent for deceased donor kidneys in the period from 2005 to 2014?
   b Suggest a possible explanation for this trend.
   c Calculate the percentage of transplants from live donors and from deceased donors for 2014 and for 2005 only, and show these data in the form of a bar graph.

5 A deceased donor kidney has the tissue type: **A2 A26, B8 B27, DR4 DR12**.
   a Which of the following people, both meeting the inclusion criteria for a kidney transplant, has the better tissue type match in terms of these three HLA markers?
   Person M: **A2 A11, B2 B18, DR3 DR10**
   Person N: **A2 A6, B8 B53, DR4 DR12**
   b Give a reason for your decision.
   c Is a cell-surface marker, such as A2, an example of an antigen or an antibody? Briefly explain.
   d Could one of the parents of person M have the tissue type **A2 A17, B3 B18, DR10 DR12**? Explain your decision.

# Chapter review

studyon

Unit 3 | AOS 2 | Topic 2 — Responding to antigens

Practice questions

## Key words

antigen
asymptomatic carrier
autoantibodies
auto-antigen
bacteria
bacteriophage
biomass
bovine spongiform encephalopathy (BSE)
buboes
bubonic plague
budding
capsid
capsule
causal link
communicable disease
Creutzfeldt-Jakob disease (CJD)
direct transmission
disease
enveloped virus
environmental disease
exotoxins
facultative anaerobes
germ theory of disease
Gram-negative bacteria
Gram-positive bacteria
haplotype
HLA typing
human leucocyte antigens (HLA)
immune system
immunosuppressive drugs
incubation period
indirect transmission
infectious disease
inherited disease
lysis
miasma
miasma theory of disease
microbe
microbiota
naked
non-communicable disease
non-infectious disease
nucleocapsid
nutritional deficiency disease
obligate anaerobes
pasteurisation
pathogenic
peptidoglycan
peripheral blood stem cell transplantation (PBSCT)
PRPN gene
RNA polymerase
self antigens
septicemia
spores
tissue typing
transcription
transmissible diseases
vector
virion
virulence
zoonose
zoonotic disease

## Questions

1 **Making connections** → Use at least eight chapter key words to construct a concept map. You may choose to add additional words.

2 **Recognising differences and similarities** → Identify one key difference between the members of the following pairs:
- a infection and disease
- b exotoxin and endotoxin
- c Gram-negative and Gram-positive bacteria
- d cellular and non-cellular pathogens
- e a harmless prion protein $PrP^{C}$ and a harmful infectious prion protein $PrP^{Sc}$
- f an antigen and an antibody
- g a self antigen and an auto-antigen.

3 **Using knowledge to make predictions** → Various kinds of bacteria can be grown in a special medium known as thioglycolate broth, which contains a chemical that absorbs oxygen. However, oxygen diffuses into the broth from the air so that the highest concentration of oxygen is present in the broth at the top of the tube but the concentration decreases with depth, with a zero concentration of oxygen at the bottom of the tube.

This broth is used to identify different kinds of bacteria in terms of their oxygen requirements.

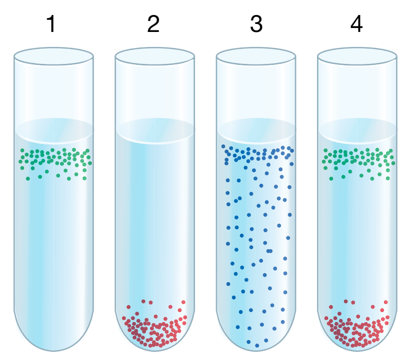



**FIGURE 6.49** Distribution of different bacteria growing in tubes of thioglycolate broth.

Consider the following kinds of bacteria in terms of their oxygen requirements

- A. obligate anaerobic bacteria
- B. obligate aerobic bacteria
- C. a mixture of obligate anaerobes and obligate aerobes
- D. facultative aerobic bacteria.

a Match each group of bacteria (A to D) to its predicted pattern of distribution in one of the tubes of thioglycolate broth (1 to 4) in figure 6.49 above.

b Other bacteria are termed aerotolerant — these bacteria do not require oxygen for their energy production, but they are *not* poisoned by oxygen as is the case for obligate anaerobes. Draw a tube showing a likely distribution of aerotolerant bacteria in a tube of thioglycolate broth.

- **c** A sample of bacteria from a suspected *Clostridium* infection was placed in thiogylcolate broth.

  What pattern would be expected? Explain your choice.

**4 Evaluating data and communicating information →** The proposal that a misfolded protein or prion could cause disease was controversial when it was first proposed in 1980 by Stanley Prusiner. He wrote: '*While it is quite reasonable for scientists to be skeptical of new ideas that do not fit within the accepted realm of scientific knowledge, the best science often emerges from situations where results carefully obtained do not fit within the accepted paradigms*.'

- **a** Why was the concept of disease-causing prions controversial when it was first announced?
- **b** Explore the web and try to identify another proposal that was controversial when it was first proposed but is now mainstream science. (Many examples exist, but if you want a pointer to another proposal, check out the Nobel Prize in Medicine or Physiology awarded in 2005.)
- **c** One clue that assisted Prusiner to conclude that prions were infectious particles made of protein and lacked nucleic acid was the observation that the causative agent of scrapie was extremely resistant to inactivation by radiation. The same resistance was later shown to be present in the causative agent of classic CJD.

  Suggest why this observation was important in identifying the chemical nature of prions.
- **d** What other line of evidence gave support to the hypothesis that prions were made of protein and not nucleic acids?

**5 Formulating and communicating explanations →** A young child was rushed to hospital and was found to be suffering from meningococcal meningitis resulting from a bacterial infection by *Neisseria meningitidis.* An antibiotic known to be effective against these bacteria was administered. The doctors informed the child's parents that the antibiotics would cure their child. However, they also explained that, after antibiotic treatment, the child would become more ill before showing any signs of improvement.

- **a** Identify the cause for this 'turn for the worse' in the child after administration of the antibiotic?
- **b** How might you explain this cause to the parents in terms that they might understand?

**6 Demonstrating knowledge and understanding →** Identify a pathogen by its kind (bacteria, virus, etc.) and its scientific name that matches the following descriptions. (In some cases, only one pathogen will fit the description, but in other cases, several different answers are possible.)

- **a** can cause the major organ failure of streptococcal toxic shock syndrome
- **b** produces an exotoxin that causes a disease with one symptom being copious and watery diarrhoea
- **c** is non-cellular, naked, and its genetic material is single-stranded RNA
- **d** is the source of an exotoxin that is the most lethal toxin known
- **e** has a cell wall that includes a thick layer of peptidoglycan
- **f** is non-cellular, enveloped and its genetic material is double-stranded DNA.

**7 Analysing data and drawing conclusions →** *Listeria monocytogenes* causes the bacterial disease listeriosis, which is usually not serious in most healthy people, but can cause miscarriages and may be fatal in elderly people and in people whose immune systems are depressed. *L. monocytogenes* infection most commonly occurs via consumption of contaminated food. Unusually, *L. monocytogenes* can grow at temperatures as low as 0 °C, although relatively slowly. The generation time (doubling time) of *L. monocytogenes* growing in dairy products was found to be about 30 hours at 4 °C, about 2 hours at 21 °C and about 0.8 hours at 35 °C.

- **a** Starting with 50 *Listeria* cells in contaminated milk, about how many bacterial cells would be expected to be present after a period of about 16 hours:
  - i when the milk is stored at 21 °C
  - ii when the milk is left in the summer sun at 35 °C
  - iii when the milk is stored in the refrigerator?
- **b** Which of these situations is most similar to what can happen if *Listeria* bacteria gained entry to the human body?
- **c** Why is correct food storage an important public health consideration?

**8 Interpreting experimental results →** Figure 6.50 shows the results of Louis Pasteur's simple but elegant experiments with boiled broth. The S-bend in the tube prevents any particulate matter from outside the flask reaching the broth.

- **a** Why was the nutrient broth boiled in both experiments?
- **b** Briefly describe the difference between the results observed in experiments 1 and 2.
- **c** Explain why different results were obtained for the two experiments?

An important feature in these experiments is the presence of the S-bend in the two flasks.

- **d** What role did this S-bend serve?

e Nutrient broth was put in both flasks. Why not use microbe-free filtered water?
f What important conclusion can be drawn from these experimental results?

Consider an alternative to experiment A in which an open-necked flask containing nutrient broth is boiled and immediately sealed.

g Would microbial growth have been observed after waiting for the same period?
h Could the same conclusion as in **f**, above, be drawn from this result? Explain.

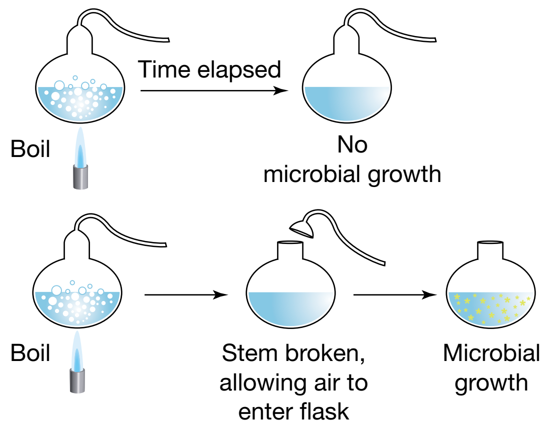



**FIGURE 6.50**

**9 Demonstrating knowledge and understanding →** Where would you expect to find the following pathogens in the human body:

a a misfolded prion protein
b a replicating Hepatitis B virus
c *Helicobacter pylori* bacteria
d an obligate anaerobic bacterial species
e a self antigen?

**10 Discussion question →** Joseph Lister (1827–1912), a Scottish surgeon, introduced antiseptic methods of killing microbes during surgery, including the use of a spray of carbolic acid around the area of surgical incision (see figure 6.51). Before introducing his antiseptic method, the death rate for Lister's surgical patients was 45 per cent. After introducing his antiseptic method, the death rate for Lister's surgical patients in the period 1867–1870 fell to 15 per cent.

In 1865, Lister read Louis Pasteur's paper, *Recherches sur la putrefaction*, that described how wine went bad because of microorganisms in the air. In a letter he sent to Pasteur, Lister wrote, *'I . . . tender you my most cordial thanks for having, by your brilliant researches demonstrated to me the truth of the germ theory of putrefaction, and thus furnished me with the principle upon which alone the antiseptic system can be carried out.'*

**FIGURE 6.51** Joseph Lister in surgery in the late 1860s. Note (at right) the person ready to use the spray of carbolic acid.

Discuss the following with your classmates:

a What 'eureka' moment did Lister have when he read of Pasteur's research on spoiled wine?
b The germ theory of putrefaction (infection) replaced an alternative explanation. What had been previously assumed to be the cause of putrefaction?
c The use of a carbolic acid spray was harmful to breathe and it can damage tissues. Lister stopped using the spray but made use of carbolic acid in other ways with his surgical patients. Suggest other possible ways in which Lister might have used carbolic acid in his operating theatre.
d Lister's antiseptic surgical procedures were initially regarded by some of his colleagues with skepticism and opposition. Discuss possible reasons for this opposition. Is this unusual in science?
e In surgical procedures today, the emphasis is not on antiseptic but on aseptic and sterile techniques. What is the key difference between these two techniques: antiseptic versus aseptic?

CHAPTER

# 7 Immune defences against pathogens

**FIGURE 7.1** A false-coloured scanning electron micrograph showing a white blood cell, known as a dendritic cell (grey) starting to engulf a yeast spore (yellow). Note the extensions (pseudopodia) of the dendritic cell that are spreading around the yeast cell at the start of the process of phagocytosis. The actions of phagocytic cells are just one of the many immune mechanisms that defend against infection. In this chapter, we will explore the two arms of the immune system, namely innate immunity and adaptive immunity.

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- understand the significance of immunity as a multipronged defence against disease
- recognise innate immunity and adaptive immunity as major subdivisions of the immune system
- understand the operation of innate immunity
- understand the operation of adaptive immunity
- become familiar with the cells of the immune system and their modes of operation.

# Inflammation of the gut

Julie, a normally busy young woman, is aware that she is suffering fatigue and is losing weight. Julie tells her friends that her gut is increasingly '*playing up from time to time*' and she is having stomach cramps. However, Julie thinks that her problem is due to something in her diet that is upsetting her gut. She knows that her fatigue is due to the fact that her sleep is frequently interrupted by her need to go to the toilet when her gut is 'playing up' and she has severe diarrhoea.

On one of her frequent visits to the toilet, Julie realises from stains on the toilet paper that blood is present in her faeces. Alarmed that this as a possible sign of bowel cancer, Julie goes to her local doctor, who takes her medical history and arranges for her to have a blood test to check for anaemia, and to have a test for the presence of blood in her faecal samples.

Based on the positive test results, Julie is referred to a specialist. She undergoes a procedure called **colonoscopy**. In this procedure, a colonoscope, a long thin flexible tube fitted with a small camera and light, is inserted into the gut via the anus for the purpose of examining the internal lining of the colon (see figure 7.2). The colon is part of the large intestine between the caecum and the rectum. This instrument can detect polyps (small growths), colon cancer, ulcers, and areas of inflammation and bleeding.

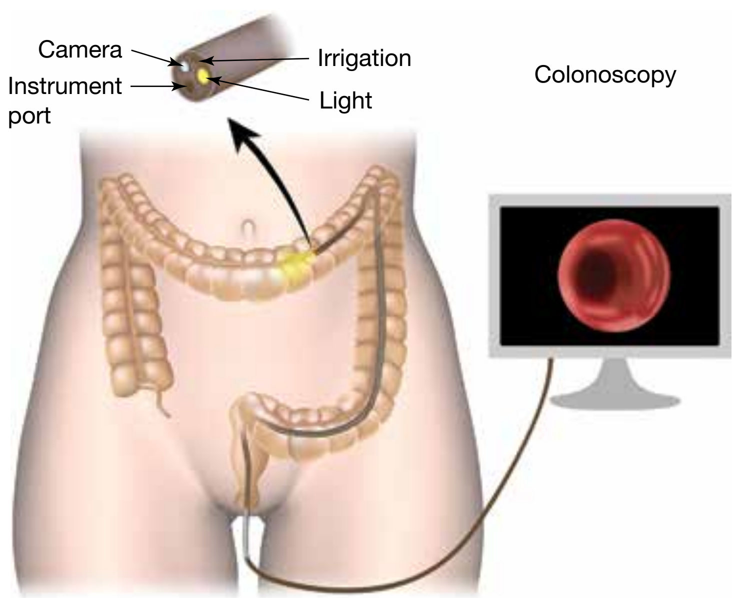



**FIGURE 7.2** In the procedure for a colonoscopy, a colonoscope is carefully introduced into the gut, via the anus, and is used to visualise the lining of the colon.

The results of the colonoscopy show that Julie is suffering from a condition called **ulcerative colitis** (UC) in which the lining of her colon is inflamed and has many small ulcers (see figure 7.3). Fortunately, Julie's condition is now successfully controlled by medication, including an anti-diarrhoea drug and prescription steroids to control the inflammation of the lining of her colon. For some sufferers of UC, surgical intervention may be required.

UC is a chronic inflammatory disease of the colon. In UC, the lining of the colon becomes inflamed and develops ulcers that produce pus and mucus. The combination of inflammation and ulceration causes abdominal discomfort and diarrhoea. Periods of active UC flare up from time to time and tend to be interspersed with periods of remission.

**ODD FACT**

A small percentage of people with ulcerative colitis may eventually need surgery to remove the colon and rectum. In one procedure, the lowest part of the small intestine is then attached to a hole (stoma) made in the abdominal wall that enables wastes to empty into an external bag.

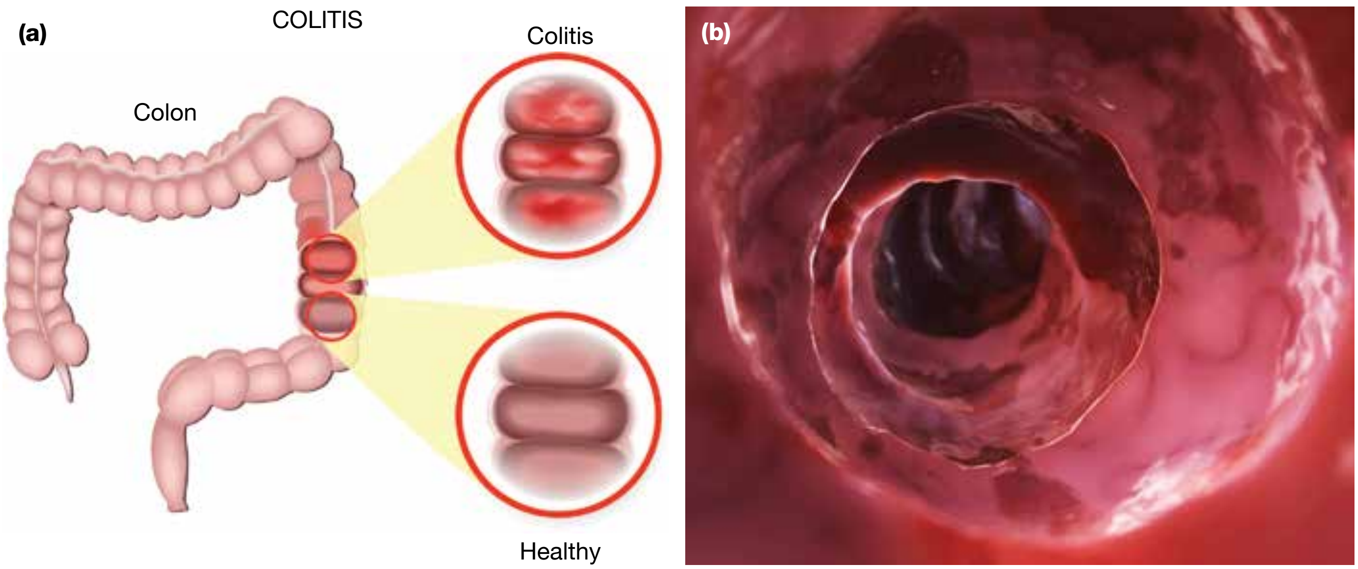



**FIGURE 7.3** **(a)** Diagram showing normal colon lining and one showing ulcerative colitis. **(b)** Colonoscopic image of an unhealthy colon. The dark regions are areas of inflammation and/or ulceration.

**ODD FACT**

Estimates of the number of people in Australia affected by ulcerative colitis at any given time range between 16 000 and 33 000 people.

'Fulminant' means occurring with rapidity and severity, and comes from the Latin *fulminare* = 'to strike with lightning'.

The mildest form of UC is a condition in which a person passes faeces, with or without blood, less than four times daily. The most severe form is fulminant UC, a very serious condition in which a person passes faeces more than ten times daily, has continuous rectal bleeding that requires blood transfusions, has abdominal swelling and tenderness, and shows system-wide toxic effects.

What is the cause of UC? The normal response of the innate immune system to infection involves acute (short-term) **inflammation** that isolates and eliminates invading pathogens. However, after the pathogens have been eliminated, this acute inflammatory response must be turned off. If this does not happen, healthy tissue becomes damaged by the active molecules released by immune cells.

UC arises when the body's immune system malfunctions by failing to 'turn off' an acute inflammatory response. In UC, the immune system sets up an inflammatory response, perhaps mistakenly in response to harmless material in the colon, and it is not switched off, so that the inflammation persists and becomes chronic (of long duration). (This of course begs the question: What causes the failure to switch off the acute inflammatory process?)

Let's now look at immunity and then explore the two major subdivisions of the immune system, namely innate immunity and adaptive immunity, which defend us against infectious diseases.

**ODD FACT**

The remarkable efficiency of the immune system in distinguishing between 'self' and 'non-self' markers is unfortunately demonstrated when the immune system of the recipient of a transplanted organ recognises the transplant as carrying non-self markers and attacks it, causing its rejection (refer back to chapter 6, pages 263–270).

## Immunity: defence against infection

**Immunity** is resistance to infectious disease. The term 'immunity' is derived from the Greek *immunitas* = exemption from duty. However, in a biological sense, immunity means 'exemption from disease'.

The cells and tissues involved in resistance to infection are part of the body's immune system. The human immune system is the collection of organs, tissues, cells and molecules that protects us from various damaging agents around us, including pathogens, toxins and other foreign molecules. The essence of immune defence is the ability of the immune system to distinguish between the body's own cells and molecules, and the foreign cells and molecules that carry distinctive 'non-self' antigens.

## Cells of the immune system

All the immune cells that provide us with defence against infection are various kinds of white blood cells that are found in lymph, blood, lymphoid organs and tissues. White blood cells are greatly outnumbered in the blood by the oxygen-carrying red blood cells. White blood cells are present in normal blood at concentrations in the range of 4000 to 10 000 cells per microlitre. Contrast this with red blood cells, which are present at concentrations ranging from 4.2 to 5.4 million cells per microlitre for adult females and 4.7 to 6.1 million cells per microlitre for adult males.

The different kinds of white blood cells can be distinguished by features such as the shapes and sizes of their nuclei and the presence (or absence) of prominent granules in their cytoplasm. Figure 7.4 shows some examples of different kinds of white blood cells. See if you can identify one or more kinds of white blood cells based on the descriptions in the figure caption. Later in this chapter, you will meet these various cells and learn about their particular roles in innate immunity and in adaptive immunity.

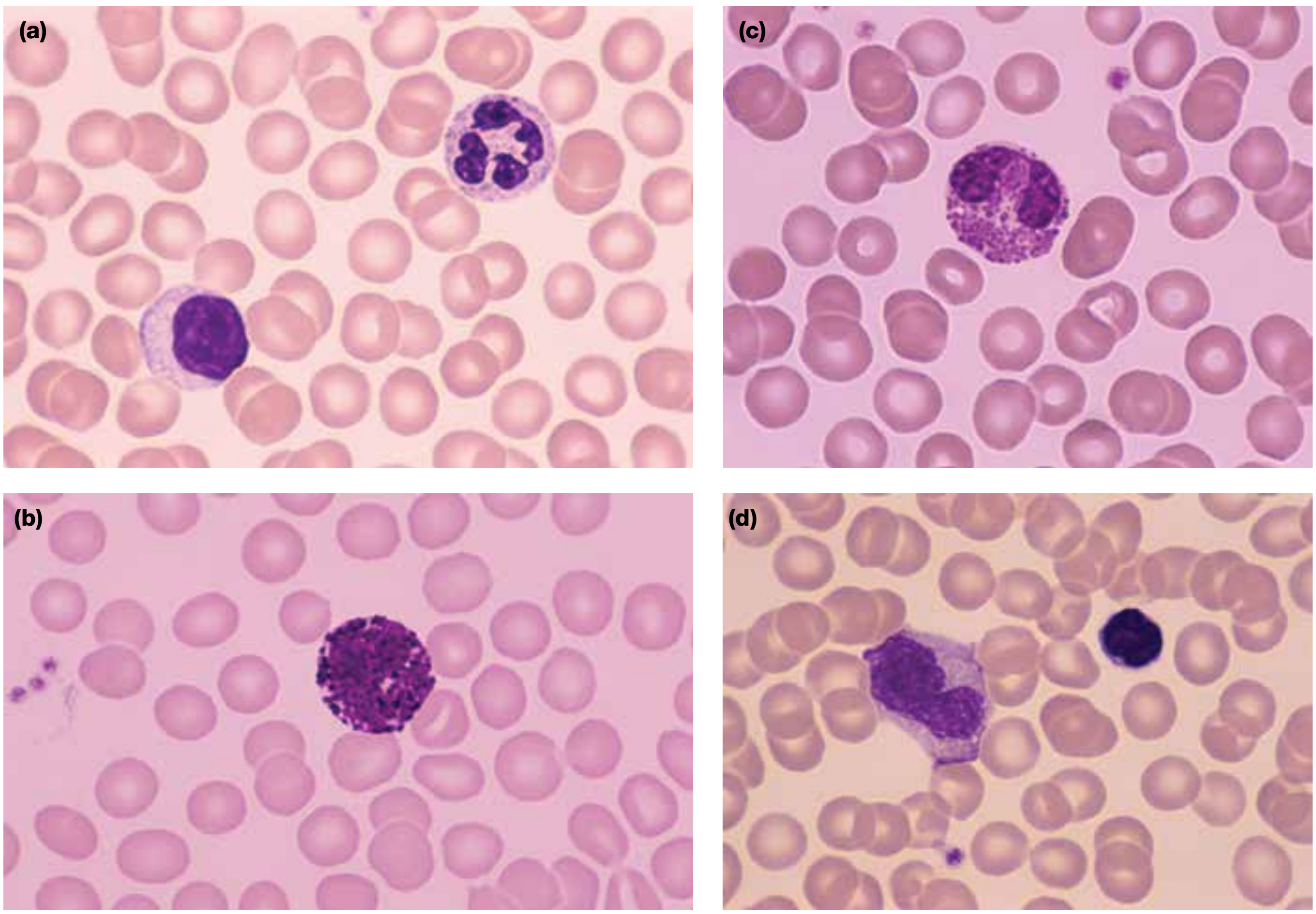



**FIGURE 7.4** Photomicrographs of human blood smears (400×) stained with Wright's stain and showing one or more white blood cells against a background of smaller pale red blood cells. Can you identify the neutrophil that has a distinctive multi-lobed nucleus? What about a basophil with its large number of densely stained cytoplasmic granules? What about a lymphocyte — either a B cell or a T cell — that is a smaller white blood cell with a relatively large round nucleus typically surrounded by a thin ring of cytoplasm? What about an eosinophil with its bilobed nucleus and many granules? Can you identify a monocyte — a larger cell with a kidney-shaped or notched nucleus? (Check bottom of page 282 for answers.)

All the cells of the immune system are white blood cells derived from multipotent stem cells in the **bone marrow** (see figure 7.5). One subgroup of immune cells are the lymphocytes, which include B cells and T cells found in large numbers in lymph nodes. Other immune cells, such as neutrophils, circulate in the bloodstream; others, for example macrophages and dendrites, are located close to points where pathogens can gain entry to the body, such as in the mucous membranes of the throat, airways and gut.

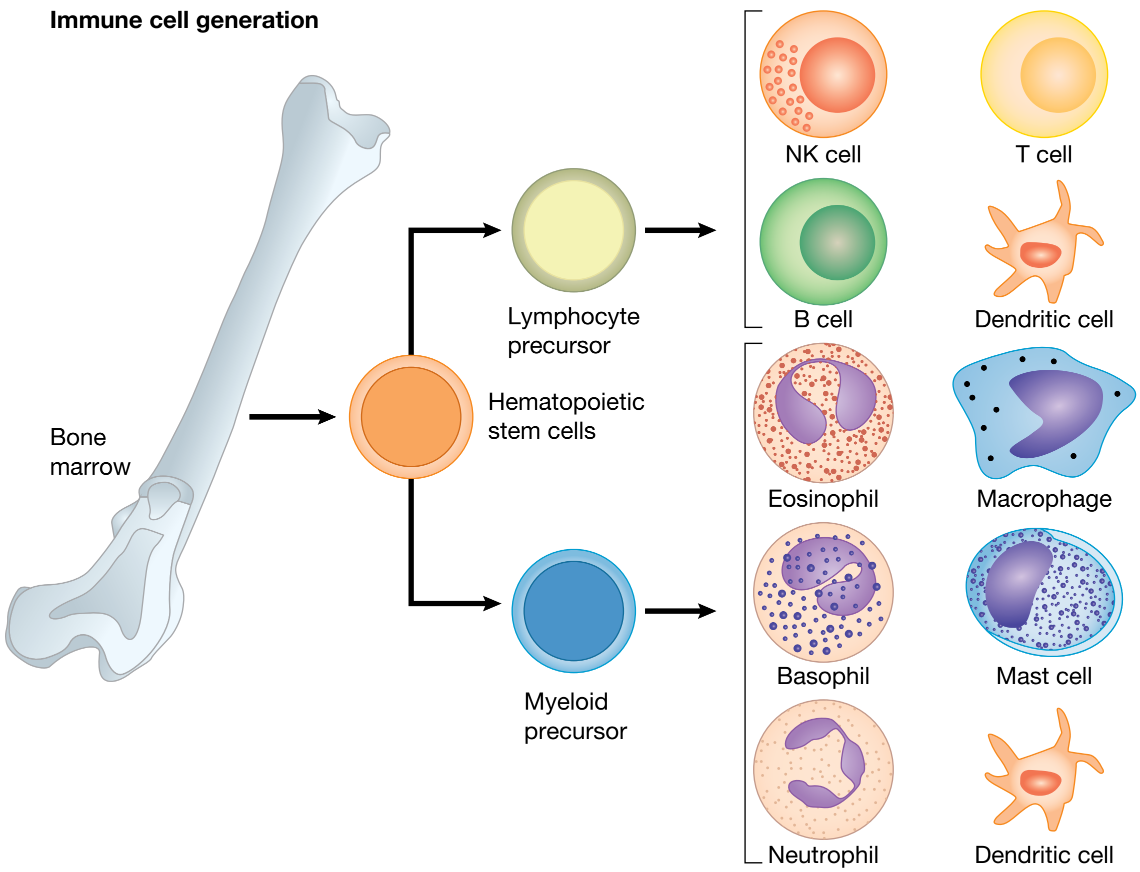



**FIGURE 7.5** Diagram showing the production of immune cells from multipotent stem cells in the bone marrow. All these cells, except for B cells and T cells, contribute to innate immunity. NK cell = natural killer cell. The B cells and T cells are the key players in adaptive immunity.

**ODD FACT**

About three litres of lymph are collected by lymphatic vessels each day and drain into the blood circulatory system.

## Role of lymphatic system in immunity

The lymphatic system consists of a network of thin-walled lymphatic vessels containing lymph that reaches all tissues of the body and interconnects the lymphoid organs. These organs include bone marrow, thymus, spleen, and lymph nodes and they are composed of lymphoid tissue — so called because they contain large numbers of lymphocytes. The distribution of lymphoid tissue throughout the body, in particular in the hundreds of lymph nodes, enables foreign antigens to be rapidly detected (see figure 7.6).

The lymphatic system plays important roles in immunity:

- The bone marrow and the thymus are called **primary lymphoid organs** because they are the sites where mature lymphocytes (B cells and T cells) develop from precursor cells.
- The lymph nodes and spleen are composed of **secondary lymphoid organs.** These organs are the sites where mature B cells and T cells are activated by meeting their complementary antigens and developing into effector cells. **Lymph nodes are where adaptive immune responses occur.**

studyon

Unit 3 | AOS 2 | Topic 2 | Concept 8

**The lymphatic system**
Summary screen and practice questions

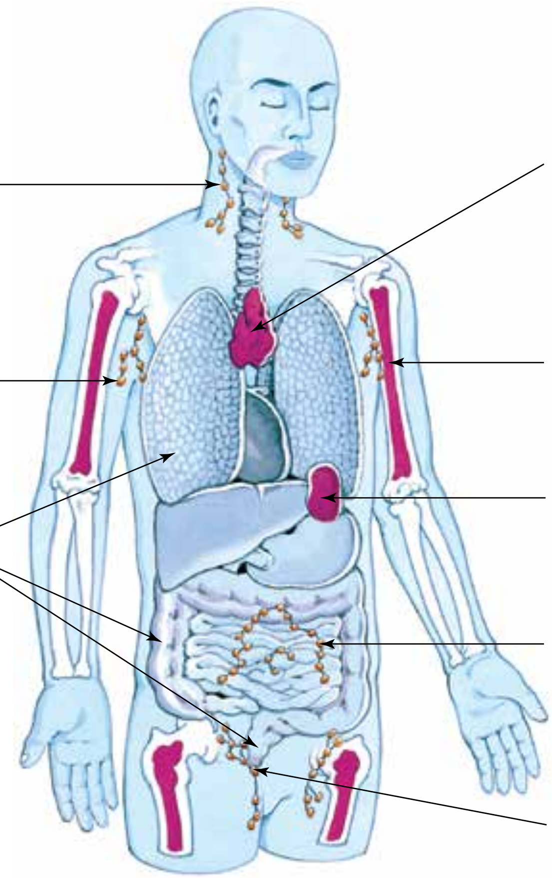



**FIGURE 7.6** Distribution of organs and tissues of the lymphatic system in the human body. These organs contain lymphoid tissue with large numbers of lymphocytes involved in various aspects of immunity.

The various organs of the lymphatic system are listed here.

**Bone marrow:**
- forms the soft tissue in the hollow centre of long bones
- is the source of pluripotent stem cells, from which all the cells of the immune system (and other blood cells, such as red blood cells and platelets) originate
- is the site of development of B cells.

**Thymus:**
- is located just behind the sternum (breastbone) of the rib cage
- shrinks with age, from about 70 grams in infants to about 3 grams in elderly people
- is the site where T cells develop after being released from the bone marrow.

**Spleen:**
- is a flattened organ lying in the upper-left sector of the abdomen
- filters the blood passing through it, clearing the blood of bacteria and viruses as well as worn-out red blood cells

Cells in figure 7.4 on page 280: (a) Upper right: neutrophil; lower left: lymphocyte, probably a T cell or a B cell. (b) basophil (c) eosinophil (d) monocyte (at left) and lymphocyte, probably a T cell or a B cell (at right).

- contains T cells and B cells that detect and respond to infectious agents in the blood
- contains other immune cells, including macrophages and dendritic cells.

**Lymph nodes:**

- are small bean-shaped structures located along lymphatic vessels
- are present in high numbers in strategic locations within the body — in the armpits, groin, neck and abdomen
- are located along blood vessels and lymphatic vessels, which enable B cells and T cells to enter and exit the lymph nodes. Most enter the lymph nodes via incoming arteries, but about 10 per cent enter via the incoming lymphatic vessel
- **are the sites where any 'new' foreign antigen meets and activates B cells and T cells and where immune responses occur**
- swell when infections occur, because the numbers of B and T cells in the lymph nodes increase — this produces the so-called 'swollen glands'
- consist of an internal structure comprising an outer cortex, an inner cortex and a central medulla (see figure 7.7):
  - The outer cortex contains follicles with large numbers of B cells that divide and diversify; antigen-presenting cells, such as dendritic cells, are also found here.
  - The inner cortex mainly contains T cells and other immune cells such as dendritic cells.
  - The medulla contains B cells, including special antibody-producing B cells called plasma cells.
- trap cancer cells or bacteria travelling in lymph vessels — for this reason, people who have a cancer surgically removed, such as a breast cancer, typically will also have the lymph nodes draining the affected organ removed (see figure 7.8).

**ODD FACT**

It is estimated that up to 99 per cent of bacteria that enter a lymph node are removed at this point.

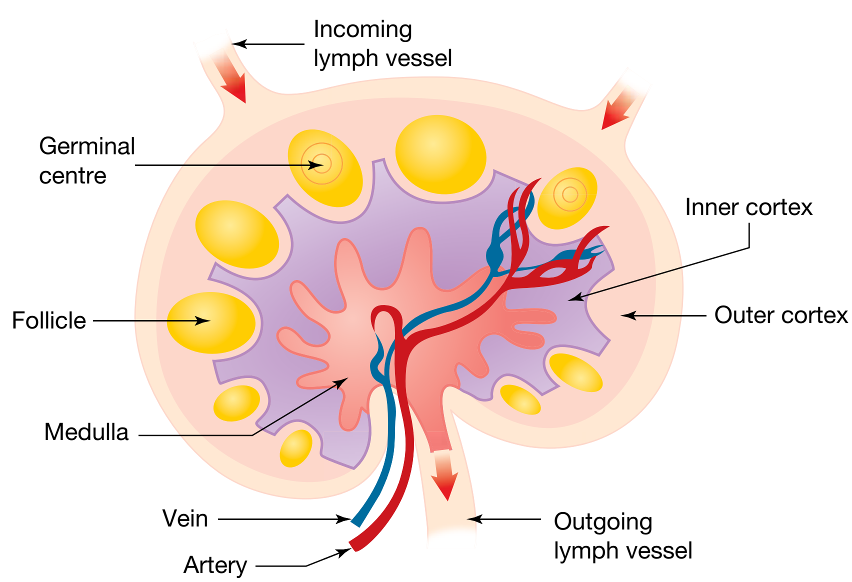



**FIGURE 7.7** Diagram showing the LS through a lymph node. Immune cells and foreign particles enter the lymph node either from the incoming lymphatic vessel or the incoming artery. Immune cells leave the lymph node in the outgoing lymphatic vessel. Lymph nodes contain clusters of B cells in the follicles of the outer cortex and clusters of T cells in the inner cortex.

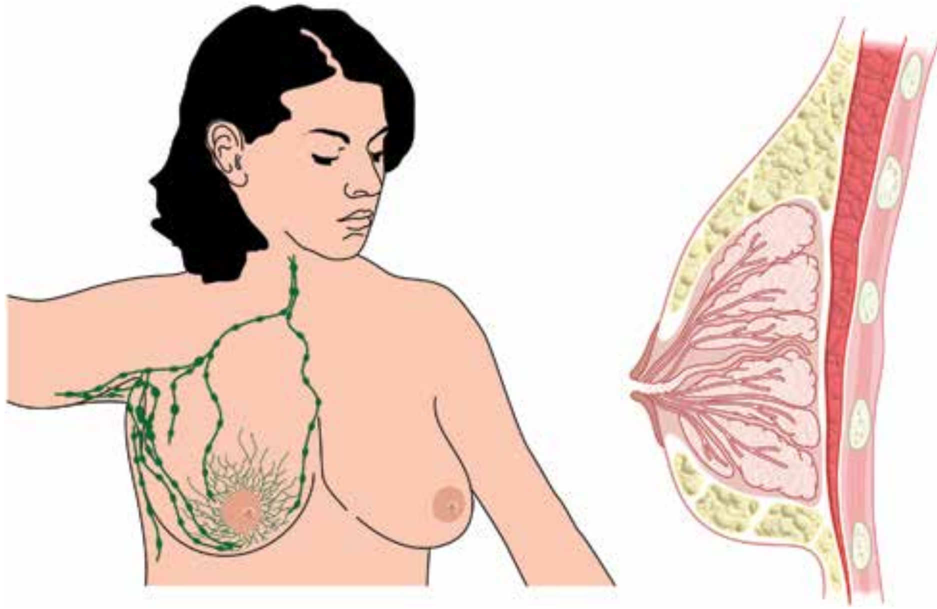


FIGURE 7.8 Diagram showing lymphatic drainage of the female breast. Note the numerous lymph nodes that are present along the lymphatic vessels.

## Subdivisions of the immune system

Immunity has two major subdivisions (see figure 7.9) whose combined operations protect the body from infectious diseases:

1. **innate immunity**, also known as non-specific or natural immunity
2. **adaptive immunity**, also known as specific or acquired immunity.

Different immune cells and active molecules, in particular soluble proteins, are involved in these two kinds of immunity.

The innate immune system is the body's pre-existing defence against invasions by pathogens and it comes into operation within minutes of an infection. In contrast, the adaptive immune system comes into operation only after the innate immune defences are evaded or overwhelmed, and this takes days to develop fully.

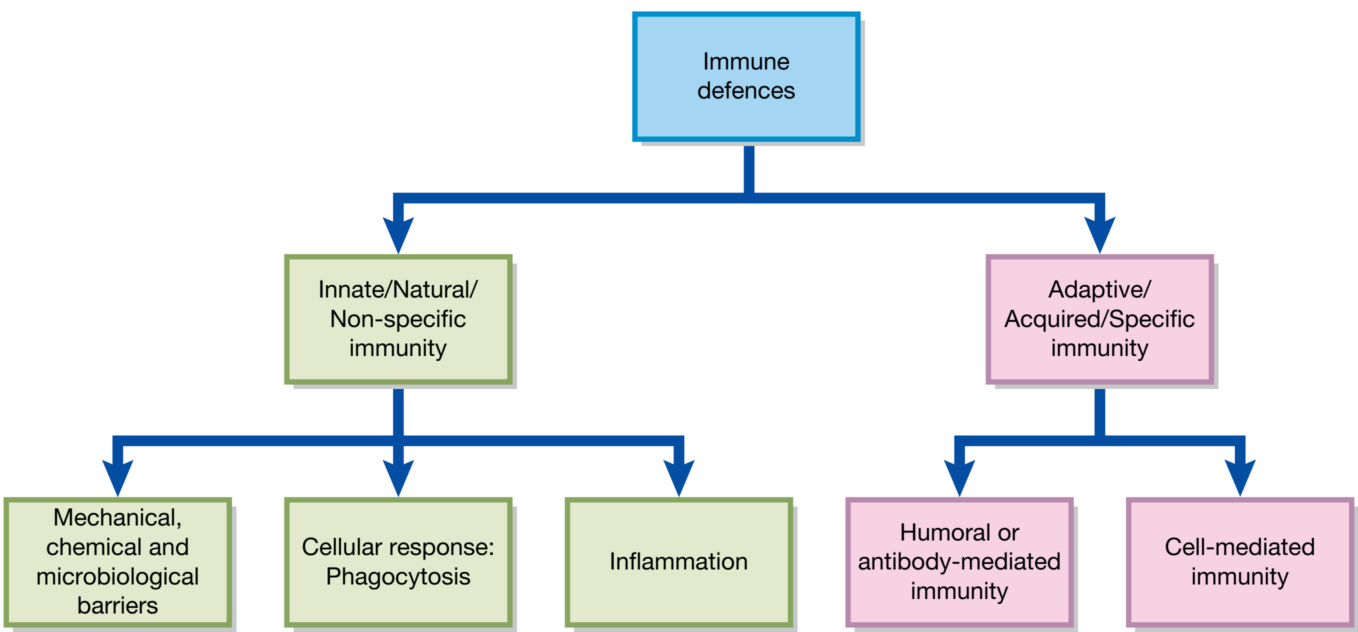



FIGURE 7.9 The immune system provides the body's defence against various pathogens. The immune defence system has two major subdivisions: (i) innate or non-specific immunity and (ii) adaptive or specific immunity.

**ODD FACT**

The term 'humoral' comes from the archaic term 'humor', which refers to the four body fluids or humors — black bile, yellow bile, blood and phlegm. In medieval times, the balance between these humors was believed to determine people's temperaments and drive their behaviours.

Both subdivisions of immunity protect against infection within the body through the actions of:

1. immune cells that directly attack invading pathogens. Direct attack by immune cells against pathogens form the *cellular* (or *cell-mediated*) immune responses. (Later in this chapter we will meet some of these cells — the natural killer (NK) cells of innate immunity and cytotoxic T cells of adaptive immunity.)
2. soluble molecules in the blood, lymph and interstitial fluid that disable pathogens. The actions of these molecules form the *humoral* immune responses. (Later in this chapter we will meet some of these soluble molecules, such as the complement proteins of innate immunity and the antibodies of adaptive immunity.)

Table 7.1 summarises differences between innate immunity and adaptive immunity. Check out this table.

**TABLE 7.1** Summary of differences between innate (non-specific) immunity and adaptive (specific) immunity.

| Innate immunity | Adaptive immunity |
|---|---|
| also known as non-specific and natural immunity | also known as specific immunity and acquired immunity |
| present in animals, plants, fungi and invertebrates | present only in jawed vertebrates* |
| response is not antigen specific | response is antigen specific |
| produces non-specific (generic) responses against classes of pathogens, not against specific pathogens | produces specific responses, with tailor-made antibodies against each particular microbe |
| immune responses occur mainly at sites of infection | immune responses occur mainly in secondary lymphoid organs, such as lymph nodes |
| only reacts against microbes | reacts against both microbes and foreign molecules such as toxins |
| present from birth | develops only after infection or after immunisation |
| activity is always present so that maximum response is rapid and immediate | normally inactive or silent so that maximum response is slower — within days or weeks |
| defence provided is not long-lasting | defence provided after an infection is long-lasting, and in some cases, even a lifetime |
| no 'memory' of prior infections, so that an identical response occurs with every infection | 'memory' of prior infections so that response is faster and stronger if the same microbe re-infects |
| Major responses are: | Major responses are: |
| 1. Cellular attack on bacteria and virus-infected cells | 1. Cellular responses attack infected cells (that is, intracellular pathogens) |
| 2. Attack by soluble proteins<br>These two responses (1 & 2) combine to produce inflammation | 2. Antibody responses target extracellular pathogens and non-self antigens |

*Among the vertebrates, only lampreys and hagfish do not have adaptive immunity; these are jawless vertebrates.

Note that innate immunity provides a number of fixed or *non-specific* defences against groups of pathogens. 'Non-specific' means that innate immunity identifies and responds to a limited number of features of pathogens. In contrast, adaptive immunity provides defences that are specifically tailored for attack against each particular pathogen. As its name suggests, the adaptive immunity recognises an unlimited number of antigens.

Note also that innate immunity retains no memory of pathogens that infect the body and, as a result, if the same kind of pathogen invades again, the innate immune responses are the same as for the first infection. In contrast, the adaptive immunity retains a 'memory' of past infections and, as a result, if the same pathogen invades again, the adaptive immune system produces a quicker and stronger response.

Immunity operates through three lines of defence against pathogens:

1. The **first line of defence** consists of physical and chemical barriers to prevent pathogens from gaining entry to the body.
2. The **second line of defence** consists of the actions of immune cells and soluble proteins mounting a rapid but non-specific attack against pathogens that gain entry to the body.

**The first and second lines of defence involve operations of the innate immune system.**

3. The **third line of defence** consists of the actions of immune cells and antibodies tailored specifically to attack each invading pathogen.

**The third line of defence comes into operation only if the second line of defence fails. This line of defence involves the operations of the adaptive immune system.**

The two subdivisions of the immune system do *not* operate in isolation, although the timing of their responses differs. The first to come into operation are the innate immune responses and, only after they have failed, do adaptive immune responses begin. Communication between cells of the two immune subdivisions ensures that immune responses are highly coordinated. Cells of the adaptive immune system are alerted, by innate immune cells, to the antigens present on particular pathogens that have overwhelmed the innate immune system.

In the following section, we will explore the innate immune system. Then we will examine the adaptive immune system (see page 308).

## KEY IDEAS

- The immune system comprises various organs, cells, and molecules that act to protect the body from infectious microbes.
- Immunity depends on the ability to distinguish between 'self' and 'non-self' markers (antigens) on cells.
- All the cells of the immune system originate as stem cells in the bone marrow.
- Primary lymphoid organs are sites where immune cells are produced and mature.
- Secondary lymphoid organs are the sites where immune cells are activated by meeting antigens and where immune responses occur.
- The immune system has two major subdivisions: innate (non-specific) immunity and adaptive (specific) immunity.
- Immunity provides three lines of defence against infectious disease.
- The first and second lines of defence involve operations of the innate immune system.
- The third line of defence involves operations of the adaptive immune system.
- Innate and adaptive immunity both operate through cellular responses and the actions of soluble proteins (humoral responses).

## QUICK CHECK

1 Give an example of the following:
   a a primary lymphoid organ
   b a secondary lymphoid organ
   c a lymphocyte.
2 Identify the subdivision of the immune system that:
   a is the first to respond to a pathogen
   b has a memory of previous infections
   c responds to each pathogen in a specific manner
   d includes physical barriers to the entry of pathogens to the body
   e provides long-lasting defence against pathogens that have caused disease.
3 Identify the component of the human immune system that:
   a is the source of all immune cells
   b is the site of maturation of T cells
   c filters the blood that passes through it
   d traps cancer cells that escape into the lymphatic vessels
   e is the site where 'new' foreign antigens are shown to T cells.

# The innate immune system

Your **innate immunity,** also known as **non-specific immunity** and **natural immunity**, provides immediate protection against pathogens through two lines of defence:

- The first and external line of defence is achieved by physical and chemical barriers that prevent entry of pathogens to the body. This first line of defence involves the immediate strategy: *'Keep pathogens out'*.
- The second line of defence is initiated if pathogens overcome the first line of defence and gain entry to the body. The second line of defence involves the strategy: '*Identify and eliminate pathogens.*'

## First line of defence: keep them out!

The best protection against infection is preventing pathogens from crossing the body surfaces. This is achieved by the **epithelial tissues** of the body and their secretions, which are the major components of the first line of defence.

Epithelial tissues include (i) the external layers of the skin, and (ii) the internal linings or **mucous membranes** of the cavities of the airways, the gut and the urogenital tract. These epithelial tissues form a *physical barrier* to the entry of microbes because their cells are densely packed, leaving no intercellular space between them. In addition, secretions of various epithelial cells contain anti-microbial agents that produce a *chemical barrier* against pathogens (see figure 7.10).

The major barriers that prevent entry of pathogens are the intact skin, mucous membranes and their secretions:

- ***Intact skin:*** The intact skin constitutes an important physical and chemical barrier to microbial infection. The epidermis of the skin is composed of many layers of cells (keratinocytes) with the outermost layer consisting of flattened dead cells. The constant shedding of these dead surface cells is an effective barrier against entry of pathogens (see figure 7.11). Sebaceous glands in the skin produce a secretion called **sebum**, which provides a protective and anti-microbial film on the skin. Sweat secreted onto the skin contains an anti-microbial protein called dermcidin, which acts against a wide range of pathogens.

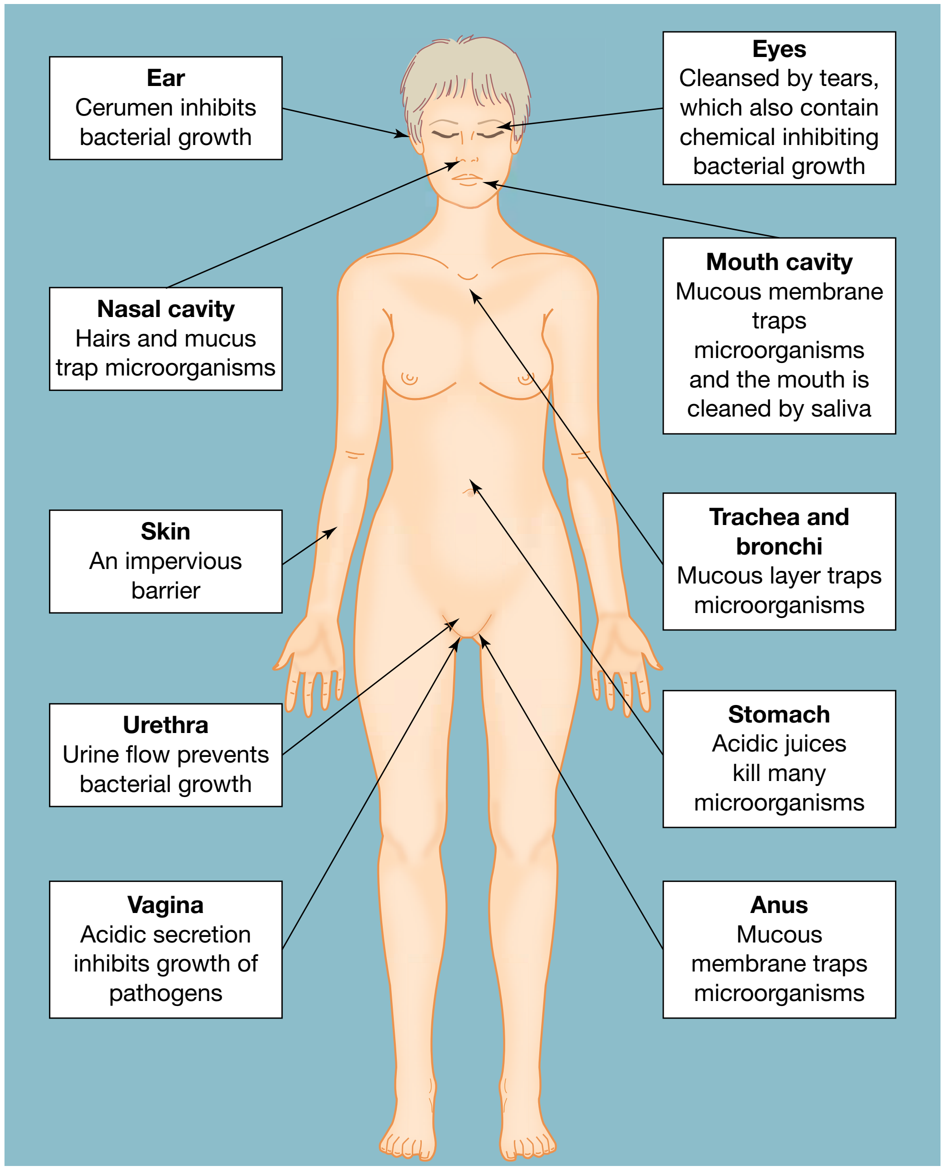



**FIGURE 7.10** Diagram showing the components of the first line of defence against pathogens, comprising the intact skin, mucous membranes and many secretions. The first line of defence prevents the entry of pathogens to the body through these various physical and chemical barriers, which form the external part of innate immunity.

- ***Mucous membranes:*** The inner spaces of the airways, the gut, and the urogenital tract are lined by mucous membranes consisting of epithelial cells. Adjacent cells have tight junctions between them that prevent the entry of microbes.

  Special cells in the mucous membranes of the gut, the genital tract and the airways secrete a thick gelatinous fluid called **mucus** that can trap particulate matter, including pathogens. In the airways, the mucous membranes include many cells with **cilia** on their outer surfaces. Cilia are fine hair-like outfoldings of the plasma membrane (see figure 7.12a). Regular beating of the cilia moves mucus from deep in the airways to the back of the throat (see figure 7.12b). Some mucus is swallowed, and the acid secretions of the stomach destroy any bacteria. Other mucus is expelled from the back of the throat when people cough or blow their noses.

  Some mucous membranes are washed constantly or regularly by secretions. Examples include the surfaces of the mouth and throat, which are washed by the constant flow of saliva, and the surfaces of the bladder and the urethra, which are regularly washed by urine that is sterile and slightly acidic. This surface-washing prevents pathogens from becoming established.

FIGURE 7.11 Photomicrograph of a stained longitudinal section through the skin. The epidermis consist of (i) the upper red cells, which are flattened dead cells that are constantly being shed and (ii) the underlying dark pink layers composed of densely packed living cells. The pale pink cells below this are part of the dermis. At the boundary between the epidermis and the dermis is the basal layer that contains stem cells.

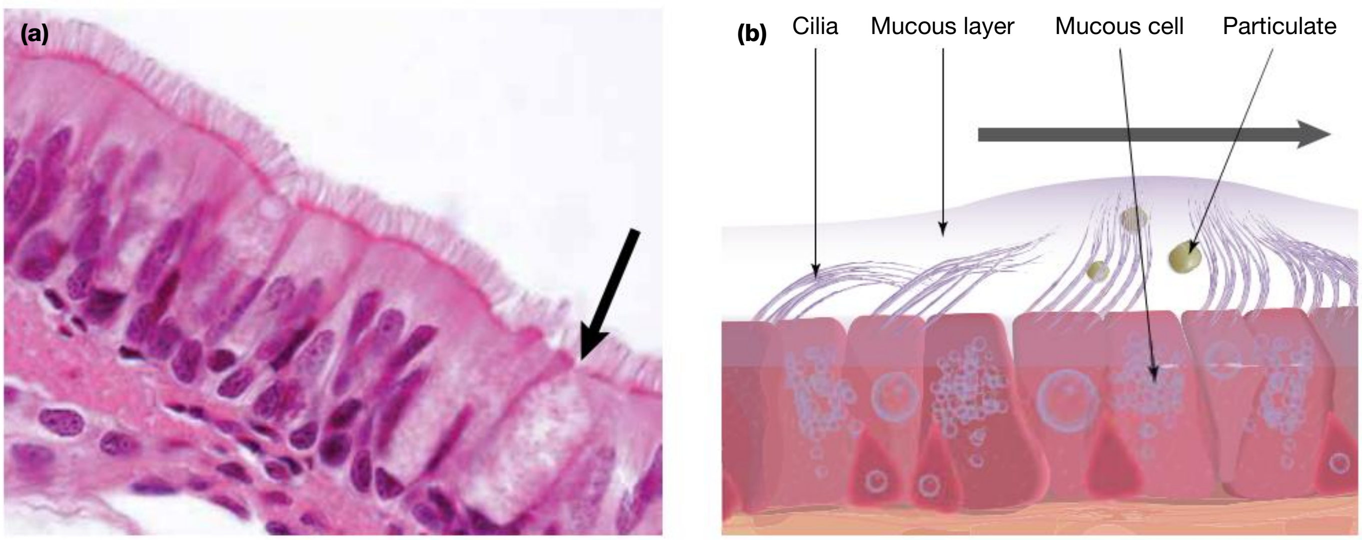



**FIGURE 7.12** **(a)** Light microscope image of ciliated cells in a section of the mucous membrane lining the airways. The cell near the lower right (arrowed) is a mucus-secreting cell or mucous cell. **(b)** Diagram showing the ciliated epithelium of the airways. Cilia beat synchronously at the rate of about 12 beats per second, and their movements shift mucus, including any trapped pathogens, upwards to the throat from where the mucus can be expelled.

Another barrier to the entry of pathogens is microbiological:

- ***Presence of normal flora:*** The term 'normal flora' refers to the non-pathogenic bacteria that are the normal residents in particular regions of the body, including the gut, the mouth and throat, and the genital tract. The presence of these harmless bacteria inhibits the growth of pathogenic microbes. For example, the normal flora of the vagina is *Lactobacillus acidophilus*. These bacteria produce lactic acid, which prevents the establishment of pathogenic microbes.

## KEY IDEAS

- The first line of defence is part of the innate immune system.
- The first line of defence consists of physical and chemical barriers that prevent entry of pathogens to the body.
- Epithelial tissues and their secretions are the major physical and chemical barriers of the first line of defence.
- Secretions that contribute to preventing entry include mucus, tears, saliva and urine.
- The so-called 'normal flora' that are the non-pathogenic bacterial residents of regions of the body act as a microbiological barrier to pathogenic bacteria.

## QUICK CHECK

**4** Give an example of:
  **a** an epithelial tissue
  **b** a mucous membrane.
**5** Where would you expect to find a ciliated mucous membrane?
**6** Briefly outline how mucus defends against entry of pathogens to the body.
**7** How does the 'normal flora' of the body contribute to the first line of defence?

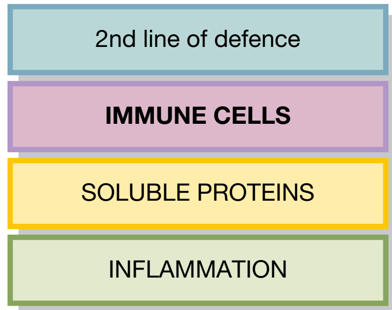
## Second line of defence: identify and eliminate!

Jason falls and grazes his knees so that the epidermis of his skin is eroded and the dermis is exposed (see figure 7.13). He brushes the dirt away from the injured area and sees that it is bleeding. His first line of defence against pathogens has been breached. Pathogenic bacteria in the soil have gained access to the internal environment of his body and this infection places Jason at risk of disease. However, when the first line of defence has been crossed, the second line of defence of the innate immune system comes into operation.

**FIGURE 7.13** The breaks in the skin created by the grazes enable bacteria to enter the tissues. The second line of defence of the innate immune system comes into action immediately and involves the operation of immune cells, including macrophages and neutrophils, and humoral factors, including soluble proteins.

The second line of defence involves the actions of immune cells and soluble proteins that produce inflammation, the major defensive response of innate immunity to infection (see figure 7.14).

**FIGURE 7.14** Overview of the key components of the second line of defence of innate immunity. The components include particular immune cells and soluble molecules present in the blood and tissue fluid or secreted by immune cells. The third component is the inflammatory process that results from the action of the immune cells and the soluble proteins.




The immune cells involved in innate immunity include:

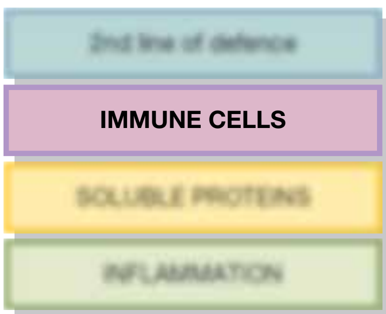
- phagocytic cells, which identify and destroy extracellular pathogens
- natural killer cells (NK cells), which identify and destroy cells infected with pathogens.

The various soluble proteins involved in innate immunity include:

- complement proteins in the bloodstream
- interferons secreted by many immune cells.

The inflammatory response is a localised protective response to infection that results from the action of immune cells and soluble proteins.

First, let's meet the cells of the innate immune system.




## Cells of the innate immune system

Different kinds of cells take part in the innate immune responses of the body to pathogens. Check out table 7.2, which summarises the functions of various cells involved in innate immunity. The cells of most interest to us are two **phagocytic** cells (neutrophils and macrophages), natural killer cells (NK cells) and mast cells. Phagocytic refers to various kinds of cells that can use the process of phagocytosis.

In the following section, we will explore how cells of the innate immune system eliminate pathogens.

The phagocytes and the natural killer (NK) cells of our innate immune systems can eliminate pathogens. The cell type and the mechanism used depend on whether the pathogen is extracellular or intracellular.

- Extracellular pathogens, such as bacteria in tissue fluid or blood, can be directly attacked by phagocytic cells, and their attack strategy is termed **phagocytosis**.
- Once pathogens have gained entry to body cells, they cannot be *directly* attacked by innate immune cells. Instead, infected body cells (along with their intracellular pathogens) are eliminated by immune cells called natural killer (NK) cells, using a strategy called **degranulation**.

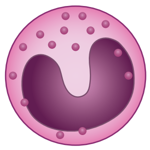
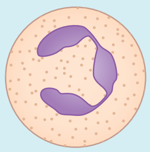
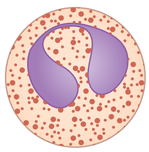
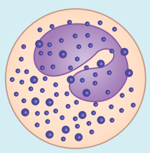
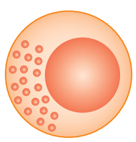
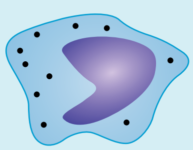
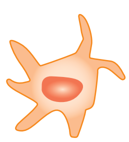
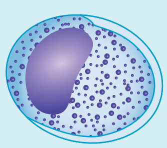
Each of these processes is described below.

### Phagocytosis: attacking extracellular pathogens

Phagocytes were first introduced in chapter 5, page 208.

Among the 'big guns' of the innate immune system is the group of innate immune cells known as **phagocytes** (*phago-* = eating; *kytos-* = cell). These cell types eliminate pathogens and cell debris by engulfing them in the process of phagocytosis. Phagocytic cells, such as macrophages and neutrophils, are like the 'vacuum cleaners' of the immune system.

**TABLE 7.2** Various cell types involved in the innate immune responses to pathogens. Of these cell types, the most important include (1) neutrophils and macrophages, which are the major cells involved in the phagocytic attack against pathogens, and (2) natural killer (NK) cells, which eliminate virus-infected cells and cancer cells. (Data on frequencies taken from STEMCELL™ technologies, document #23629)

| Cell type (a) in blood | Stylised image | Description & characteristics | Major role in innate immunity |
|---|---|---|---|
| • **monocytes 2–12% of circulating white blood cells** |  | circulating white blood cells (2–12%); can leave bloodstream and move into tissues where they differentiate into macrophages | precursors of macrophages |
| • **neutrophils *'first into the fray'* 30–80% of circulating white blood cells** |  | most abundant circulating white blood cells (30–80%) with distinctive multi-lobed nuclei; short-lived; one of the major cells involved in phagocytosis, which kills engulfed microbes by toxic chemicals; first cells to arrive at infection site in response to chemotactic signals from other cells | identify and mount phagocytic attack on microbes |
| • **eosinophils up to 7% of circulating white blood cells** |  | circulating white blood cells; bilobed nucleus; phagocytic; release toxic chemicals from granules; major defence against parasites that are too large to be attacked by phagocytosis, e.g. parasitic worms | defence against larger parasites; attack uses toxic chemicals released from cytoplasmic granules |
| • **basophils up to 2% of circulating white blood cells** |  | rare circulating white blood cells; phagocytic; contain cytoplasmic granules that are rich in histamine and heparin | release of histamine and other molecules as part of inflammatory response; play a role in allergic reactions |
| • **natural killer (NK) cells '*the assassins*' 1–6% of circulating white blood cells** |  | NK cells contain granules filled with potent chemicals; recognise and attack cells lacking 'self' markers; degranulation releases proteases and also perforin proteins, which insert holes in plasma membrane of foreign cells; induce programmed cell death (apoptosis) | elimination of virus-infected cells and cancer cells by degranulation |
| **(b) in tissues** | | | |
| • **macrophages *'the vacuum cleaners'*** |  | develop from monocytes that migrate into tissues; phagocytic; present in almost every tissue of body; eliminate microbes and cell debris; initiate acute inflammatory responses by secretion of various cytokines (also play a role as antigen-presenting cells) | identify and eliminate pathogens by phagocytosis; also remove dead cells and cell debris |
| • **dendritic cells *'sentinels on patrol'* in tissues** |  | phagocytic cells, resident in tissues; characteristic star-shaped cell; present in skin epidermis, and on surface of linings of airways and gut; migrate via lymphatic vessels to lymph glands where they act as antigen-presenting cells | mobile cells that identify pathogens and secrete antiviral cytokines |
| • **mast cells *'border guards'*** |  | contain granules rich in histamine and heparin; found in tissues close to the external environment; involved in early recognition of pathogens; release chemical signals that attract other immune cells to infection site; involved in acute inflammatory response | release of histamine and other active molecules during acute inflammation; play role in allergies |

studyon

Unit 3
AOS 2
Topic 2
Concept 5

**Phagocytosis**
Summary screen and practice questions

studyon

Unit 3
AOS 2
Topic 2
Concept 5

**See more**
Phagocytosis

The major phagocytic cells of the innate immune system are:

- **neutrophils**, which circulate in the bloodstream but can migrate into sites of infection in tissues
- **macrophages** (see figure 7.15), which are present in most tissues but are most concentrated in tissues close to epithelial layers. This is an ideal location for macrophages because this is close to potential entry points for pathogens to the body.

**FIGURE 7.15** False-coloured scanning electron microscope (SEM) image of a macrophage, a major phagocytic cell type present in many body tissues. What is the target of macrophages? How do they destroy their target?

Figure 7.16 illustrates the stages in phagocytosis:

1. The pathogen is identified by a pattern recognition receptor on the surface of the macrophage and is engulfed by outfoldings of the plasma membrane of the macrophage (also refer back to figure 7.1).
2. The pathogen is completely enclosed in a vesicle called a **phagosome**.
3. Lysosomes fuse with the phagosome (forming a phagolysosome) and release toxic chemicals that attack the pathogen. Toxic chemicals include oxygen radicals and nitric oxides that damage macromolecules, lysozymes that attack the cell walls of Gram-positive bacteria, and proteases and other digestive enzymes.
4. The pathogen undergoes digestion from this combined chemical attack. (See marginal note.)
5. Indigestible material is discharged from the phagocytic cell by a process of exocytosis.

Some partly digested proteins (peptides) of ingested pathogens are retained by macrophages and dendrites for the purpose of antigen presentation (see page 325).

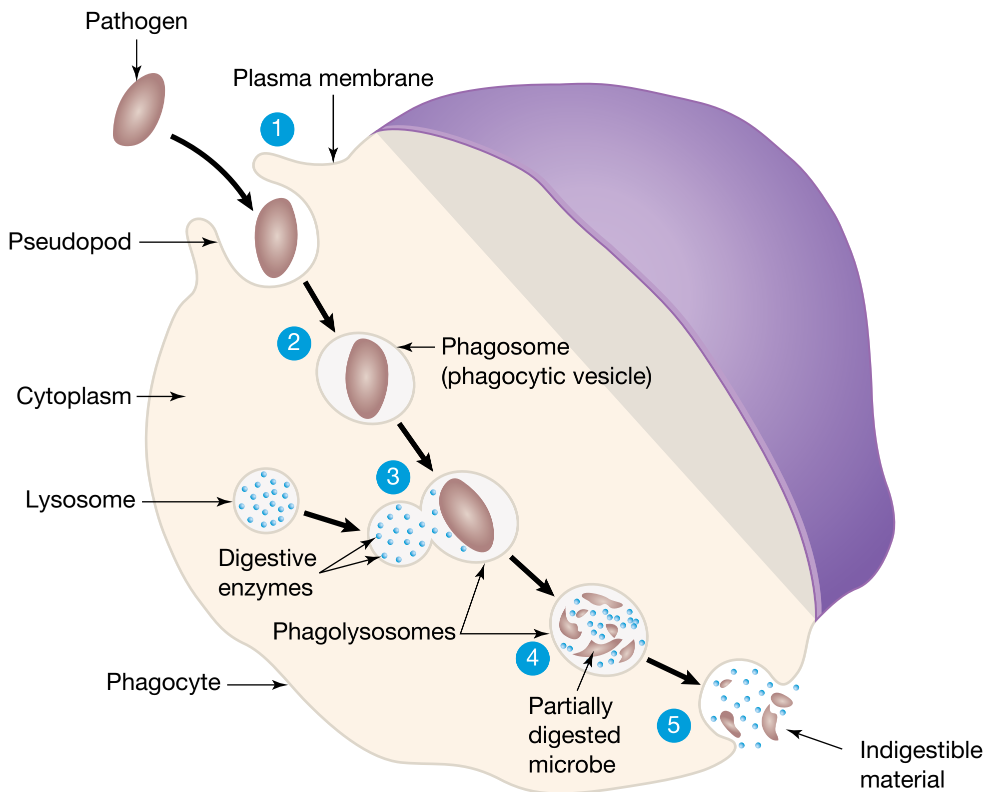



**FIGURE 7.16** Diagram showing the elimination of a pathogen by phagocytosis. The innate immune cells involved are mainly macrophages and neutrophils. The process of elimination by phagocytosis is effective against pathogens, such as bacteria, that are in body fluids and *outside* cells.

**ODD FACT**

Phagocytosis of a bacterial cell by a human neutrophil takes about nine minutes.

Some protein fragments from the digested pathogen are retained by macrophages and are used to signal cells of the adaptive immune system to initiate their defence against the specific pathogen. There is no replacement text.

## How do phagocytic cells identify extracellular pathogens?

In order to defend against pathogens, cells of the innate immune system must first identify pathogens as 'foreign' or non-self before they can eliminate them.

Phagocytic cells of the innate immune system, such as macrophages and neutrophils, have many receptors called **pattern recognition receptors (PRRs)**. Some PRRs are located on the surfaces of these immune cells, while other are present in their cytoplasm. PRRs enable cells of the innate immune system to identify pathogens by recognising and binding to non-specific **molecular patterns** found on groups of pathogens (see figure 7.17). This level of identification is a bit like being able to identify dogs and cats at the generic level, as, for example: 'That's a cat!' or 'No, it's a dog!'

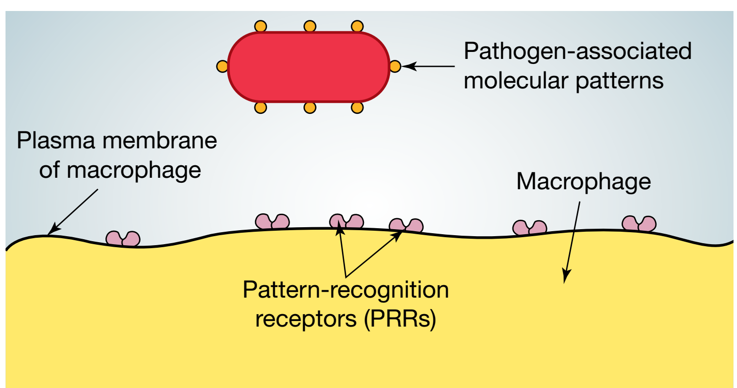



**FIGURE 7.17** Diagram showing the pattern-recognition receptors (PRRs) on the surface of a macrophage. The PRRs can recognise various molecular patterns present on pathogens that identify these cells as invaders to be eliminated by phagocytosis.

Contrast this with the level of pathogen identification that occurs with the adaptive immune system. We will see later in this chapter (see page 315) that the cells of the adaptive immune system can identify every different pathogen and antigen. This is a bit like being able to identify every dog and every cat by breed, colour, age, and every other distinguishing features, as, for example: 'That's a three-legged eight-year-old black Labrador named Lulu.'

Examples of molecular patterns that are recognised by PRRs and their level of identification are as follows:

- lipopolysaccharide (LPS) of cell walls of Gram-negative bacteria: 'That's a bacterium.'
- lipotechtoic acids on cell walls of Gram-positive bacteria: 'That's another bacterium.'
- flagellin protein present in bacteria with flagella: 'Yet another bacterium.'
- the glycoproteins of enveloped viruses: 'That's a virus.'

## Degranulation: eliminating intracellular pathogens

**Degranulation** is the release of anti-microbial and toxic molecules from membrane-bound granules stored in the cytoplasm of some innate immune cells. **The release of granules into target cells is the main weapon used by natural killer (NK) cells to eliminate infected body cells.** NK cells are like the assassins of the innate immune system (see figure 7.18). In particular, NK cells recognise missing or abnormal HLA markers on virus-infected cells and cancer cells and eliminate these cells by degranulation.

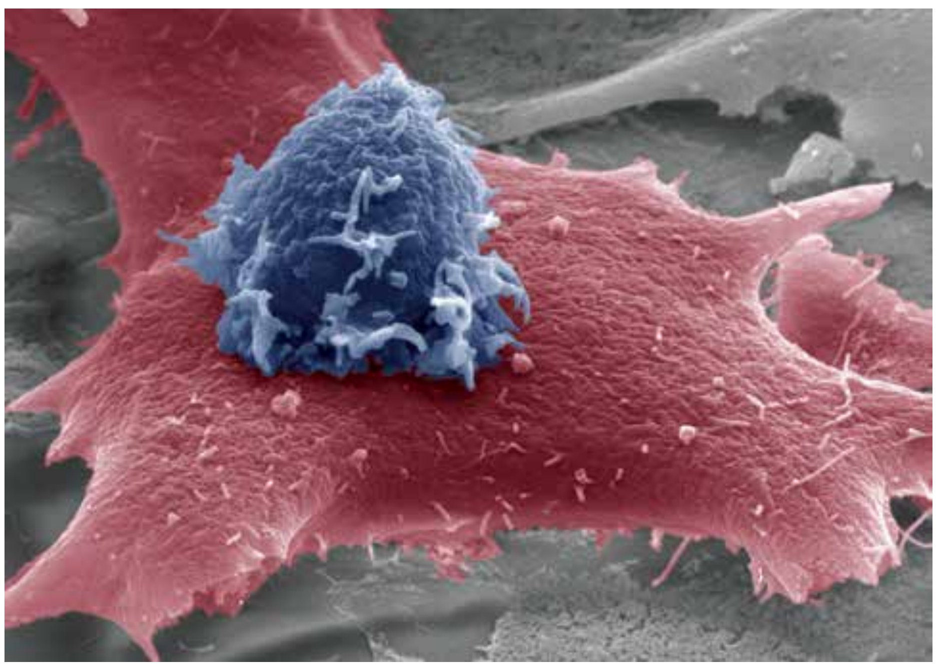


**FIGURE 7.18** False-coloured scanning electron microscope image of a natural killer (NK) cell (blue) attacking a cancer cell (pink). (Image courtesy of Dr Donna Stolz and Mr Marc Rubin of the Center for Biologic Imaging at the University of Pittsburgh.)

**Elimination of targets by NK cells:** Figure 7.19a shows a transmission EM image of an NK cell that contains several large granules (see lower left).

The granules in NK cells contain active protease enzymes, known as **granzymes**, and a pore-forming protein called **perforin** (see figure 7.19b). The combined action of these proteins destroys the target cell. Perforin molecules form a ring structure that punches a hole in the plasma membrane of the target cells, enabling entry of the proteases into the target cell. Once inside the cytoplasm of an infected target cell, the proteases induce apoptosis (refer back to chapter 5, page 209). Elimination of an infected cell or a cancer cell by degranulation of an NK cell is completed within hours.

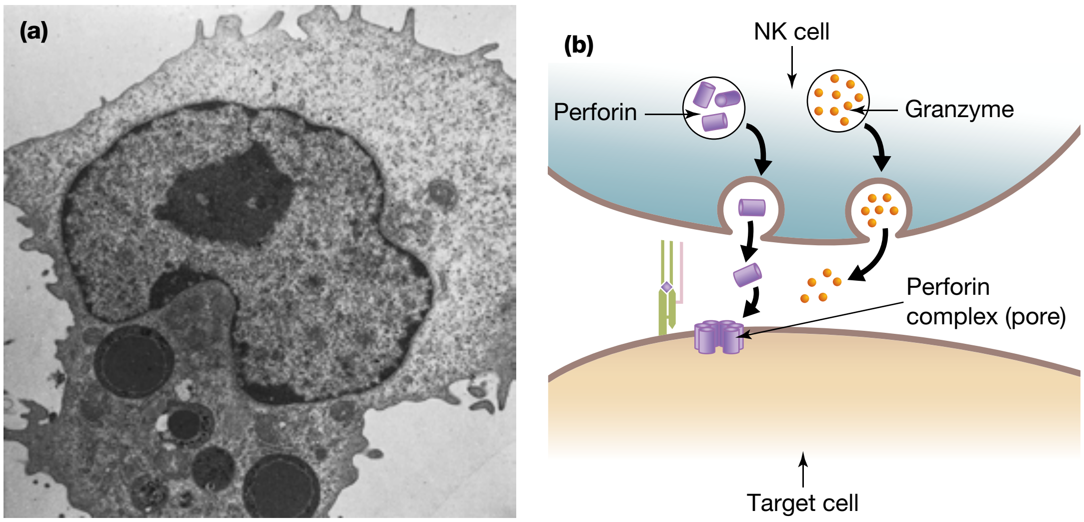



**FIGURE 7.19 (a)** Transmission electron micrograph of a natural killer (NK) cell. Note the large nucleus and, below it, several prominent membrane-bound granules that contain active chemicals, including proteases, known as granzymes, and the pore-forming protein, perforin. (Image courtesy of Professor Colin Brooks and Dr George Sale.) **(b)** Diagram showing the release of granzymes and perforin from membrane-bound granules in an NK cell into a target cell, such as a virus-infected cell or a cancer cell. When granzymes gain entry to the target cell via the pore created by perforin, they initiate a signal pathway for the cell to undergo apoptosis.

**ODD FACT**

As well as NK cells, other innate immune cells also release granules by exocytosis to kill target cells. These immune cells include neutrophils with granules containing cell-killing reactive oxygen species, and eosinophils with granules containing a protein (PRG2) that kills parasitic helminth worms.

Killing virus-infected cells by apoptosis, rather than by **lysis**, is important. Why? Lysis of a virus-infected cell simply explodes the cell, releasing the virus particles into the extracellular fluid so that the virus particles can infect other cells. In contrast, apoptosis destroys both the cell and any viruses it contains in a systematic manner, preventing the further spread of the virus.

## How do NK cells identify abnormal body cells?

Body cells infected with viruses and cancer cells are part of a person's body. How do natural killer (NK) cells recognise these infected and abnormal cells as targets for elimination?

NK cells identify targets for elimination even though the targets are body cells. The decision: '*kill or don't kill*' by an NK cell is believed to be determined by receptors on the plasma membrane of the NK cell: one receptor is for 'kill' and another is an inhibitory receptor that blocks the kill signal.

All body cells carry ligands that can bind to 'kill' receptors on NK cells. This binding activates the 'kill' receptors that signal an NK cell to eliminate any cell bound to it. However, this kill signal is inhibited by normal HLA markers on healthy body cells. The activation of the inhibitory receptors sends a signal that

HLA markers can also be written as MHC markers. There is no replacement text.

blocks the kill signal. As a result, an NK cell does not identify healthy body cells as targets for elimination (see figure 7.20a).

Virus infections of body cells suppress the expression of class I HLA markers. Likewise, cancer cells produce abnormal HLA markers. These missing or abnormal HLA markers make these cells susceptible to being killed by NK cells because they cannot inhibit the 'kill' signal. An abnormal cell with missing HLA markers is like a flashing red light that signals to an NK cell: 'Stop, degranulate and eliminate me.'

When contact is made with an infected cell that is missing or has abnormal HLA markers, the NK receptor for the inhibition of the 'kill' signal is not activated (see figure 7.20b). As a result, the NK cell responds to the 'active kill' signal and degranulates, spilling the lethal contents of its granules onto the target cell.

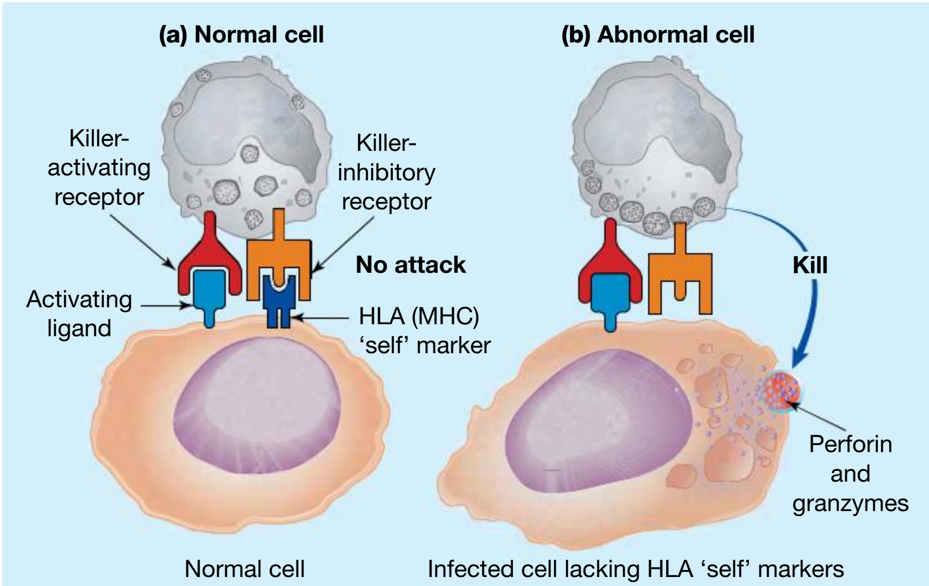



**FIGURE 7.20** Natural killer (NK) cells identify and destroy abnormal body cells such as virus-infected cells and cancer cells, but not normal body cells. **(a)** The NK cell does not attack a normal body cell because its inhibitory receptor is activated when it binds to a HLA self marker on this cell. **(b)** The NK cell attacks an abnormal body cell because the inhibitory receptor is not activated.

## KEY IDEAS

- The second line of defence of the innate immune system comes into immediate operation if pathogens overcome the barriers of the first line of defence.
- The second line of defence involves the operation of innate immune cells and humoral factors.
- The main cells involved in the cellular responses of innate immunity are macrophages, neutrophils and natural killer (NK) cells.
- Macrophages and neutrophils directly attack and eliminate extracellular pathogens through the process of phagocytosis.
- Pattern recognition receptors on phagocytic cells identify pathogens for elimination.
- Natural killer (NK) cells attack pathogens indirectly by identifying and eliminating pathogen-infected body cells and cancer cells through the process of degranulation.

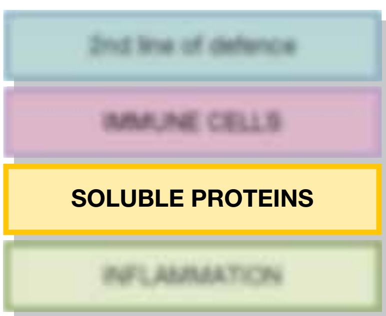
**QUICK CHECK**

**8** Give an example of the following that are part of innate immunity:
- **a** the most common circulating white blood cell
- **b** the ‘assassins’
- **c** a phagocytic cell
- **d** a cell that eliminates intracellular pathogens
- **e** a non-specific feature of a bacterial cell that could be recognised by a PRR.

**9** Identify two methods by which innate immune cells eliminate pathogens.

**10** What different outcome results from the lysis of a virus-infected cell as compared to the apoptosis of a virus-infected cell?

**11** Identify the following as true or false:
- **a** The level of identification of pathogens by the cells of innate immunity is more specific than that of cells of adaptive immunity.
- **b** Neutrophils are the first cells to arrive at a site of infection.
- **c** Natural killer (NK) cells eliminate pathogens by degranulation.
- **d** Macrophages are tissue-based cells of the innate immune system that eliminate pathogens by phagocytosis.
- **e** Macrophages eliminate infected body cells by phagocytosis.
- **f** Pattern recognition receptors enable innate immune cells to identify a toxin in the bloodstream.




# Humoral innate immunity

In the section above, we explored various cell-mediated responses to infection, namely phagocytosis and degranulation. The other important part of the innate immune system is the humoral component, which involves the action of soluble proteins and their derivatives.

The second line of defence against pathogens involves humoral responses to pathogens. These responses involve the actions of soluble active molecules. Key soluble molecules involved in innate immunity include proteins known as complement and interferons.

## Complement proteins

Dissolved in the plasma of your blood are proteins that form what is called **complement**. In total, the complement system comprises a large group of protease proteins (more than 25). The 11 major complement proteins are denoted C1 through C9, D and B, with the most abundant being C3. These plasma proteins are called ‘complement’ because they complement or add to the function of immune cells.

Complement proteins can also be activated by contact with antibodies bound to antigens (see page 311). The reaction cascade involved has a different starting point, but the end results are the same as outlined below.

The complement proteins circulating in the bloodstream are inactive enzymes. **Complement proteins are activated when they make direct contact with molecules on the surface of a pathogen**.

The activation of the first complement protein starts a sequence of reactions that take place on the surface of a pathogen (but *not* on a healthy human body cell). The first protein in the series enzymatically alters the next protein in the series. The product of the first reaction then activates the next enzyme in the series, which, in turn, activates the next protein, and so on. The activation of a complement protein occurs when the protein is cut (cleaved) into two fragments — a larger activated protein and a smaller peptide fragment. This sequence of reactions starts a cascade that can neither be stopped nor reversed.

Figure 7.21 shows **the complement cascade that begins when molecules of protein C3b bind to the surface of a pathogen**. Review this diagram and note, for example, that the C3 protein is cleaved into two products, C3a and C3b, and that the C3b protein then catalyses the next step in the sequence.

The complement cascade produces some larger active protein fragments (such as C3b and C5b) that remain on the surface of the pathogen, and also smaller active peptide fragments (such as C3a and C5a) that diffuse away from the surface of the pathogen.

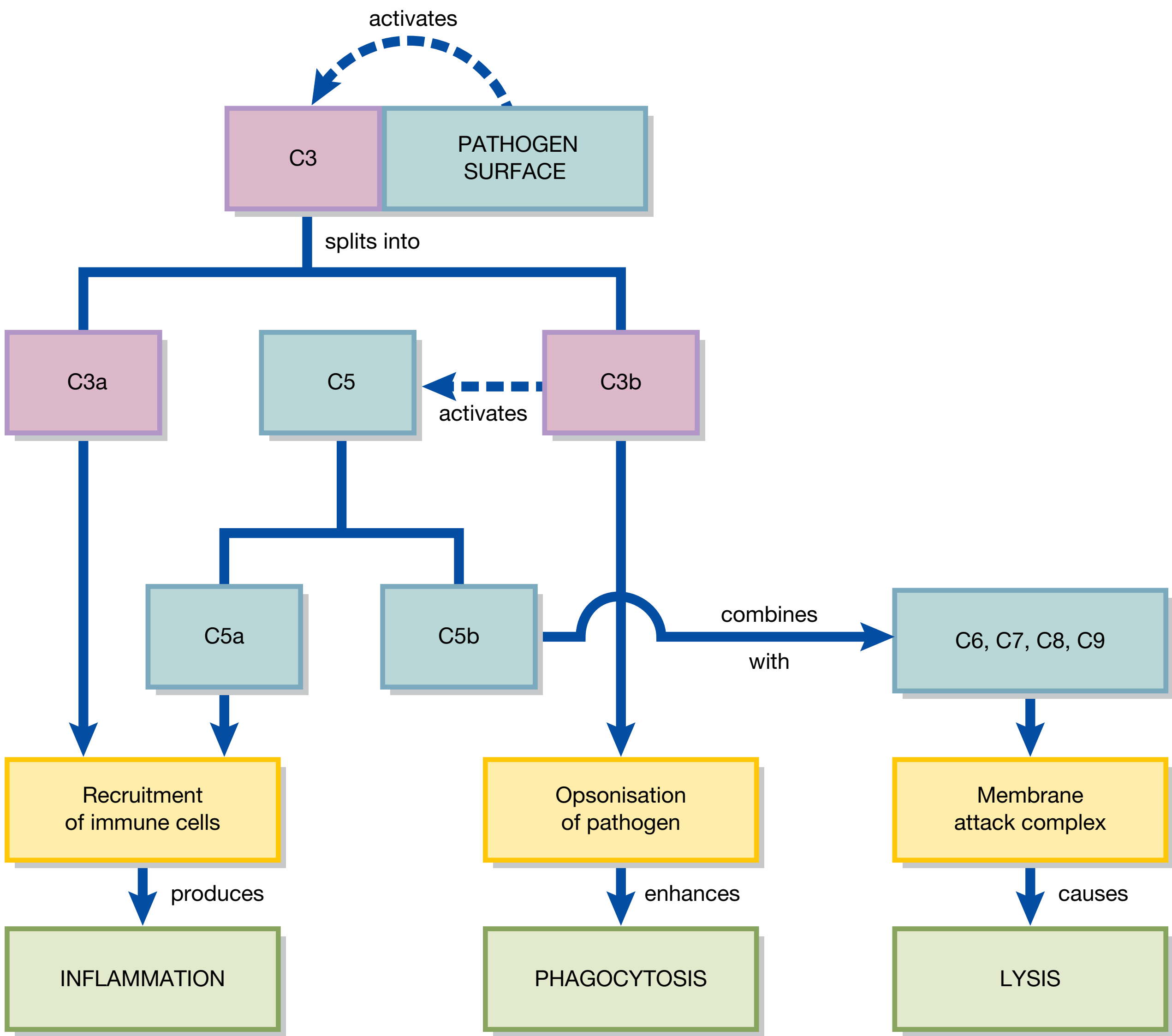



**FIGURE 7.21** Simplified diagram showing the cascade of reactions of complement proteins that starts when the protein C3 binds to the surface of a pathogen and is cleaved (broken) into two fragments, C3a and C3b. Protein C3b then activates C5 by splitting it into C5a and C5b. Note the three different actions of the various complement proteins (see green blocks).

Various activated complement proteins formed by the cascade contribute to innate immune defence in the following ways:

1. by opsonising pathogens, making them more susceptible to phagocytosis
2. by recruiting immune cells involved in an inflammatory response
3. by destroying bacterial pathogens by lysis.

*Opsono-* (ancient Greek) = to cater or to buy food.

Each of these actions of activated complement proteins is outlined below.

1. **Opsonisation:** Some complement proteins, particularly protein C3b, smother the surface of a pathogen and form a complex with the surface antigens on the pathogen. This process is termed **opsonisation**. Opsonisation makes pathogens more susceptible to elimination by phagocytosis because phagocytes, such as macrophages, have receptors for complement proteins on their plasma membranes and these bind to the opsonised microbes (see figure 7.22).

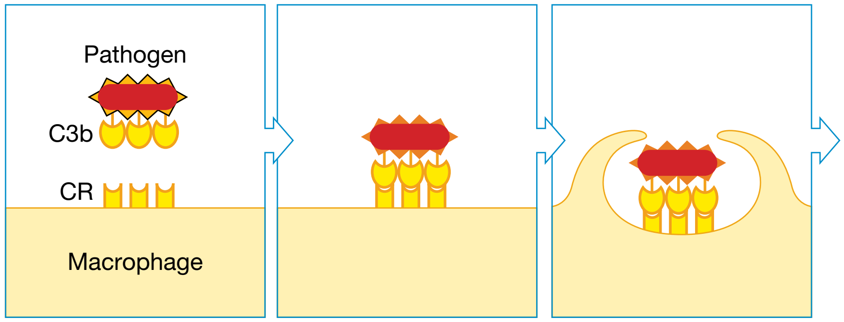



**FIGURE 7.22** Complement proteins (C3b) opsonise a pathogen by coating its surface. The presence of complement receptors (CR) on the macrophage enables it to bind to the pathogen and begin to eliminate the pathogen by phagocytosis. Could you draw the next image in this series?

2. **Chemotaxis**: Chemotaxis refers to the movement of cells in response to a chemical stimulus. Small complement peptides, such as C3a and C5a, that diffuse from the pathogen surface act as chemical signals attracting immune cells involved in the inflammatory response to the site of the infection.
3. **Lysis of pathogens**: Some complement proteins are involved in the direct 'explosive' killing of extracellular pathogens. This occurs when a **membrane-attack complex (MAC)** forms on the plasma membrane of the pathogen. The formation of a MAC starts with the C5b protein on the surface of the pathogen that binds to one molecule each of proteins C6, C7 and C8, and then binds to several C9 molecules (see figure 7.23). The MAC inserts into the plasma membrane of the pathogen and produces a pore that allows fluid to enter, causing the pathogen cell to swell and burst — explosive death by osmotic shock!

## ODD FACT

A single membrane-attack complex (MAC) is sufficient to kill one pathogen. However, a MAC attack is not effective against Gram-positive bacteria. Can you suggest why this might be the case? (Clue: Refer back to figure 6.18 on page 239.)

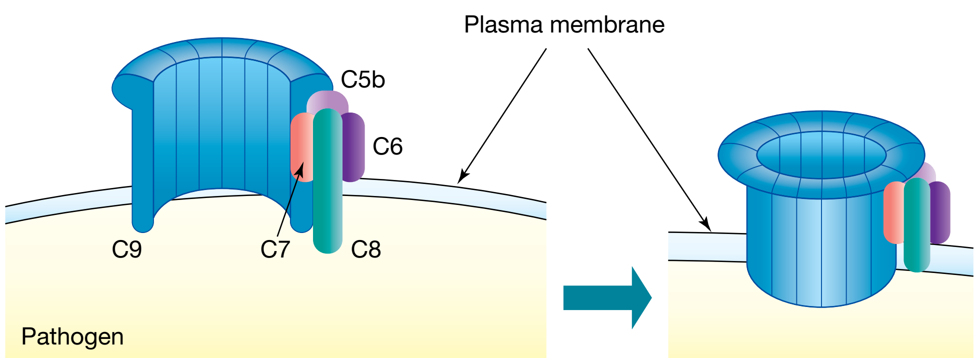



**FIGURE 7.23** A membrane-attack complex (MAC) is formed by a combination of several complement proteins. When the MAC is completed (see RH image), a hole is created in the plasma membrane of the pathogen. This hole allows for an inflow of fluid into the pathogen, creating an osmotic shock and causing the pathogen cell to burst (lyse).

Interferons are a subgroup within the larger group of signalling molecules called cytokines. Refer back to chapter 5, page 184.

**studyon**

Unit 3 | AOS 2 | Topic 2 | Concept 7

**Complement proteins and interferons**
Summary screen and practice questions

**studyon**

Unit 3 | AOS 2 | Topic 2 | Concept 7

**See more**
The interferon mechanism against viruses

## Interferons: antiviral agents

**Interferons (INFs)** are important in the innate immunity response to virus infections. The name of these proteins reflects the fact that they 'interfere' with viral replication. Because viral infections can occur quickly, interferons are critical in the body's initial responses.

Once a body cell is infected with viruses, the cell secretes interferons. That cell is doomed, but the interferons that it secretes into its surroundings act as warning signals to nearby cells so that they can prepare in advance for a possible virus infection.

Interferons bind to receptors on neighbouring cells, producing signals that:

1. induce transcription of a number of specific genes that encode production of *inactive* forms of antiviral enzymes inhibiting protein synthesis and destroying RNA. These enzymes are activated only if a virus succeeds in infecting the cell and, if activated, they block the synthesis of viral proteins
2. make the plasma membrane less fluid, making its fusion with viral particles more difficult and so reducing the chance of viral infection in these cells
3. cause virus-infected cells to undergo apoptosis
4. activate immune cells, such as natural killer (NK) cells, that eliminate virus-infected cells by apoptosis.

Figure 7.24 shows the various responses produced by interferon signals.

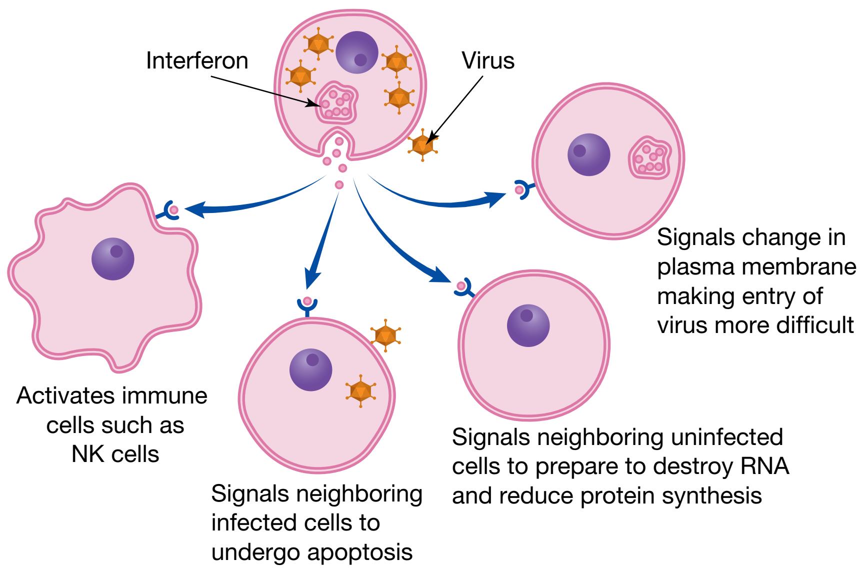



**FIGURE 7.24** Diagram showing the various actions on nearby cells of interferons released from a virus-infected cell.

Different viruses infect different tissues. The further a virus must travel to reach its target cell, the more likely it will be destroyed before it gets there. Interferons are particularly important if a virus gains entry to the body close to its target cell. This is the case with cold and influenza viruses that infect cells in the nose and throat. Because infection occurs quickly, the body sometimes does not have the time to develop antibodies against these viruses and relies on interferons for its defence. If a person develops a cold or flu, interferons have failed to prevent the infection and the infection has developed into a disease that must be eliminated by the adaptive immune system.

## KEY IDEAS

- Complement consists of a large group of inactive protease proteins in the bloodstream.
- Activation of complement proteins involves a cascade of reactions that occurs on the surface of pathogens.
- Activated complement proteins contribute to the elimination of bacterial pathogens by opsonisation and by formation of membrane-attack complexes (MACs).
- Interferons secreted by a virus-infected cell signal nearby cells to make various advance preparations to reduce the chance of a virus infection.
- Interferons signal virus-infected cells to undergo apoptosis.

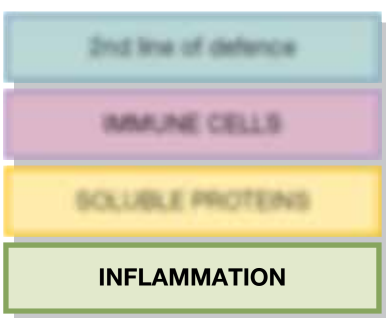
## QUICK CHECK

**12** Identify the following as true or false:
  **a** Soluble proteins form the humoral component of innate immunity.
  **b** Complement proteins are involved in the elimination of intracellular pathogens.
  **c** Activation of complement C3 involves its cleavage into two fragments, C3a and C3b.
  **d** Interferon released by a virus-infected cell can signal nearby cells to produce antiviral proteins.

**13** A bacterial cell is 'opsonised'. Briefly explain what has happened to this bacterial cell.

**14** How many different complement proteins are involved in the formation of a membrane-attack complex (MAC)?

**15** Identify the following statements as true or false:
  **a** Interferons cure virus-infected cells.
  **b** A signal from interferons that is received by a virus-infected cell results in the death of that cell by apoptosis.
  **c** Opsonisation of a bacterial cell prevents it from being eliminated by phagocytosis.
  **d** Complement refers to a collection of various immune cells that circulate in the bloodstream.




# The inflammatory response

We have now met the cells and the soluble proteins that are involved in responses to infection by the innate immune system. Now let's examine **inflammation**, the major and rapid response of the innate immune system to an invading pathogen.

Inflammation is an early response of the human body to infection, such as the entry of bacteria via an open cut into a region of a finger (see figure 7.25). Inflammation is normally a short-term (acute) immune response that is localised around the site of entry of a pathogen so that the local area involved becomes red, swollen, hot and often painful. An acute inflammatory response is rapid in onset, appearing within minutes or hours of the onset of an infection.

Inflammation is an important protective response of the innate immune system to eliminate an infectious pathogen. (Note that inflammation may also occur when cells are damaged by agents other than pathogens, as, for example, thermal burns to the skin, corrosive chemical spills, frostbite and sunburn.)

**FIGURE 7.25** When the skin is cut with a piece of glass or a knife, the entry of pathogenic bacteria into surrounding tissues can result. Acute inflammation is a rapid response of the body's innate immune system to this invasion.

When bacteria enter the body via a cut in the skin, the inflammatory response:

- localises and prevents the spread of the pathogen
- eliminates the pathogen and also removes any tissue debris
- repairs the damaged tissue (healing) and restores the normal (non-inflammatory) state.

Let's look at what happens when a person cuts her finger on a piece of glass, and bacteria on the glass gains entry to body tissues. The person's immune system rapidly initiates an inflammatory response. The first stage of the acute inflammatory response is the vascular stage, which involves changes to the blood vessels; the second stage is the cellular stage, which involves actions of various immune cells; the third stage is the resolution stage, which switches off the inflammatory response.

### 1. The vascular stage of inflammation

1.1 *Dilation of blood vessels:* The trigger for the dilation of blood vessels is the release of chemicals, including histamine and prostaglandins, from cells damaged by the cut. The arteriole and venule of the local capillary bed dilate so that the blood flow to the damaged area increases (see figure 7.26). This vasodilation produces the *redness* typical of acute inflammation, and the increased blood flow produces *heat.*

1.2 *Increased permeability of local capillaries:* The capillaries in the area around the cut become more 'leaky' so that protein-rich fluid (exudate) escapes from the capillaries into the infected region. The exudate causes the *swelling* (*edema*) that is typical of inflammation. The swelling causes pressure on the surrounding tissue, which stimulates pain receptors and contributes to the *pain* that is characteristic of inflammation. Clotting agents in the exudate assist in clot formation, which isolates the infection.

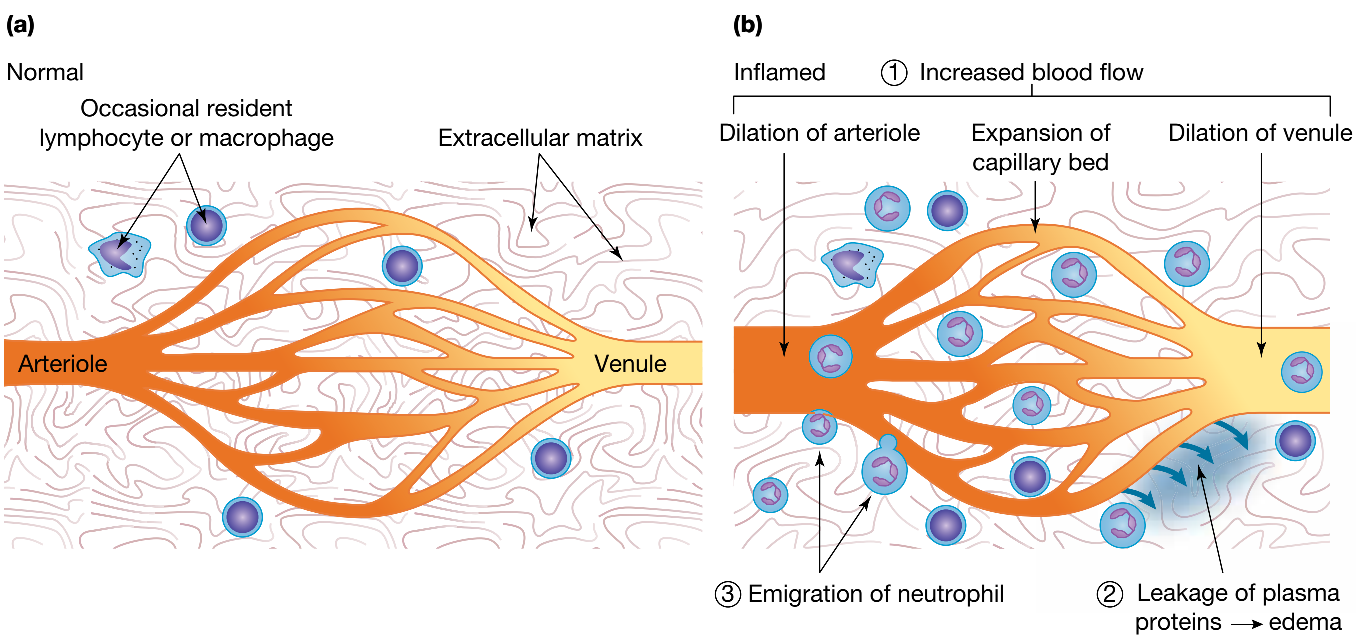



**FIGURE 7.26** Diagram showing **(a)** normal situation and **(b)** the dilation of the blood vessels in a region of localised acute inflammation, which increases the blood flow in the region. What is the trigger for this increase in diameter of the blood vessels? Note also the net out flow of fluid (exudate) from the blood into the surrounding extracellular matrix, which causes localised swelling.

studyon

Unit 3 | AOS 2 | Topic 2 | Concept 6

**Inflammation**
Summary screen and practice questions

studyon

Unit 3 | AOS 2 | Topic 2 | Concept 6

**See more**
Defence adaptations

### 2. The cellular stage of inflammation

2.1 *Escape of immune cells from capillaries:* The expansion of the capillary bed enables neutrophils to squeeze between the endothelial cells that line the capillaries (see figure 7.27).

2.2 *Migration of neutrophils to infection site:* Neutrophils are the first immune cells to arrive at the infection site, attracted by cytokines released by damaged cells. Other immune cells, including macrophages from nearby tissues, follow. These cells release signals, such as histamine and more cytokines, that attract more phagocytic cells to the infection site.

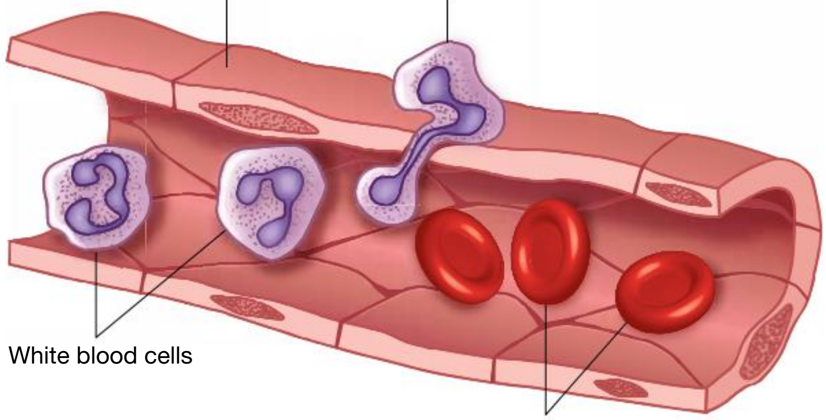



**FIGURE 7.27** Diagram showing a neutrophil squeezing between the endothelial cells that line capillaries and the post-capillary venules. The neutrophils escape into the interstitial fluid and migrate to the infection site.

2.3 *Phagocyte attack on bacteria:* Both neutrophils and macrophages attack bacterial pathogens by engulfing them in a process called phagocytosis (refer back to page 291). Phagocytic cells also remove cell debris from the infection site by a similar process. Pus consists mainly of dead phagocytic cells and other immune cells, and also contains living cells and cell debris.

Figure 7.28a summarises the events in acute inflammation, and figure 7.28b shows the effects in real life.

**ODD FACT**

An abscess is a collection of pus in a confined space.

## The resolution stage: returning to normal

While the infection is present, pro-inflammatory cytokines are released to maintain the inflammatory response. This response continues until the pathogen has been eliminated. However, once the infection is under control and tissue repair underway, it is important that the normal state is restored. This is termed the resolution stage.

Resolution is a complex process that includes a reversal of all the processes that produced the acute inflammation, as, for example, the reversal of capillary dilation and cessation of release of pro-inflammatory cytokines. Resolution involves the release of many active molecules or mediators, including anti-inflammatory cytokines, the appropriately named lipids resolvins and protectins.

If normality is not established, a situation of **chronic inflammation** is created, which is damaging to the body and is the cause of several diseases. This is the 'dark side' of inflammation, occurring when the immune system does not shut off in the absence of infection but continues to produce active immune cells and pro-inflammatory molecules.

Long-term chronic inflammation is seen in several diseases, as, for example, ulcerative colitis (UC) (refer back to page 278), and rheumatoid arthritis (RA), which is an autoimmune disease where chronic inflammation of the tissues around the joints occurs, causing swelling, pain and stiffness.

**ODD FACT**

After engaging in phagocytosis, the neutrophils that are active in the inflammatory response undergo apoptotic cell death and their remains are removed by phagocytic cells.

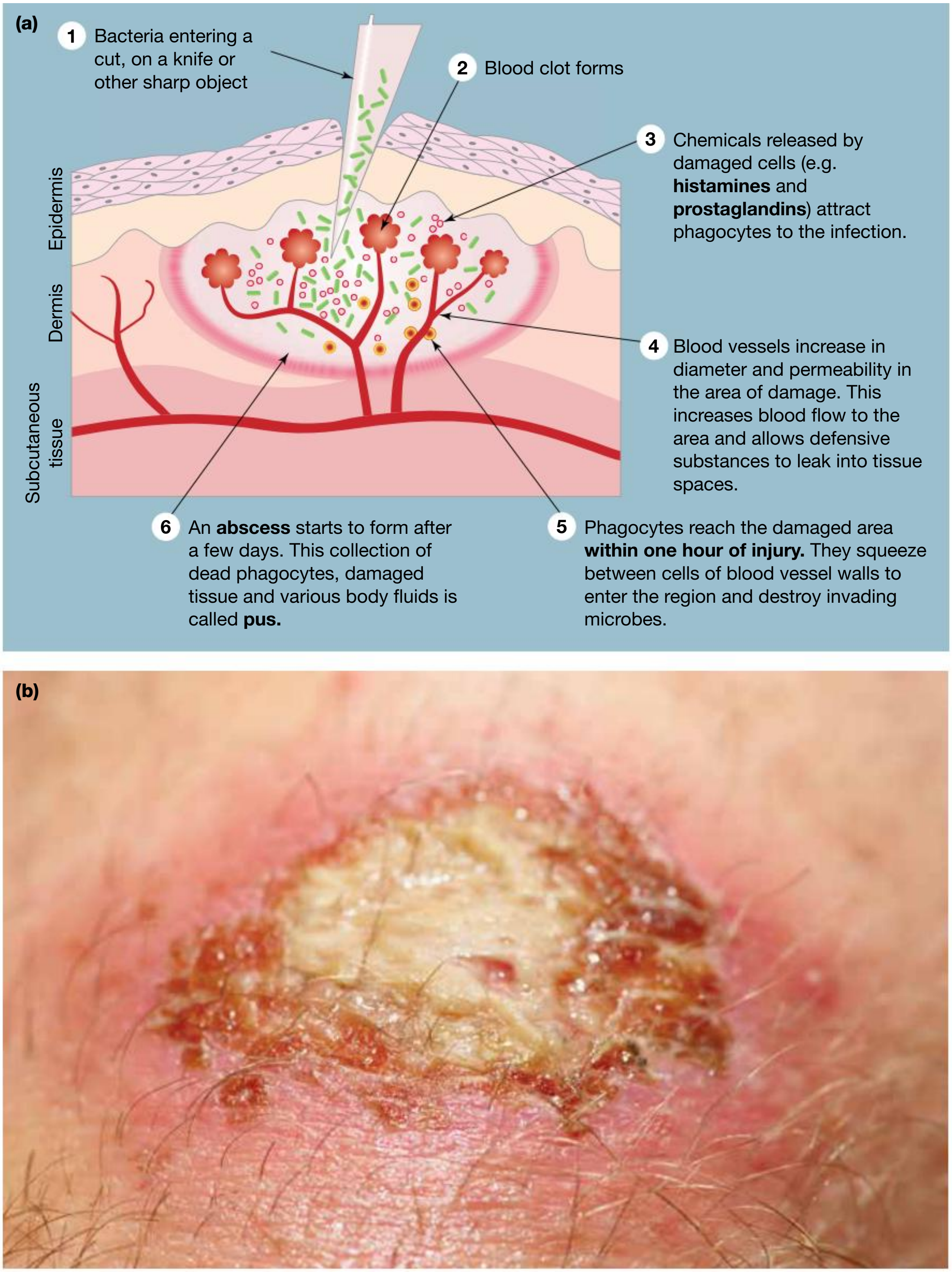



**FIGURE 7.28** Acute inflammation. **(a)** Diagram summarising the events of acute inflammation. **(b)** Acute inflammation in real life. This may look nasty, but this image shows the innate immune system producing an acute inflammation that has contained the infection. Note the classic signs of inflammation — redness, swelling, and you can almost feel the pain. What is the major component of the pus?

## KEY IDEAS

- The three stages of an inflammatory response are the vascular stage, the cellular stage and the resolution stage.
- The signs of inflammation are redness, swelling, heat and pain.
- The vascular stage of inflammation involves changes in the capillaries supplying the site of infection.
- Actions by cells in the cellular stage of inflammation include various cell movements, as well as phagocytic activity at the infection site.
- If the inflammatory response is not turned 'off' when an infection is contained, a damaging situation of chronic inflammation can arise.

## QUICK CHECK

**16** What is pus?
**17** Briefly explain why the area around an infection becomes swollen.
**18** Identify two key changes in capillary vessels that occur during the vascular stage of inflammation.
**19** Identify the following statements as true or false:
  **a** The main response of the innate immune system to invading pathogens is acute inflammation.
  **b** The stages in an inflammatory response in order are: cellular stage, resolution stage and vascular stage.
  **c** Resolution of an inflammatory response requires the production of more pro-inflammatory molecules.
  **d** Cytokines released by cells at the infection site attract more immune cells to the site.
  **e** The redness around an area of inflammation is the result of dilation of the associated capillaries.
**20** Give an example of a human disease that is believed to be the result of chronic inflammation.

# The common cold

Kylie is standing in a crowded train and a man near her sneezes but fails to cover his mouth. Invisible airborne particles carrying pathogens rush past her face. As she breathes in, several pathogens enter her nasal passages and move down the airways towards her lungs, while other pathogens adhere to the skin of her face. Among these unwelcome entries to her nasal passages are some rhinoviruses.

Her innate immune system immediately comes into action to eliminate the pathogens in her throat and lungs. At first, she shows no signs of infection. Some days later, Kylie is feeling unwell (see figure 7.29a) and is showing the first symptoms of the common cold — a sore throat and a runny nose (see figure 7.29b). Her innate immune system has been overcome by the viral infection in the epithelial cells of her upper airways. Now her adaptive immunity must come into action to deal with the virus. Kylie's adaptive immune responses require several days to become fully effective and, during this period, the virus continues to replicate; she continues to show symptoms and feels ill. After this period, her adaptive immune responses progressively eliminate the rhinovirus, a process taking seven to ten days.

### ODD FACT

When you have a cold, you initially produce clear watery mucus to wash the viruses from your nose and sinuses (the 'runny' nose). Later, the mucus is thicker and may be white or yellow (the 'stuffy' nose).

### ODD FACT

Influenza is caused by one virus, *Haemophilus influenzae*, but the common cold is caused by any one of hundreds of different kinds of viruses that can infect the cells of the upper airways.

### Time course of an infectious disease

Let's look at what might happen when a person is infected by viruses or bacteria. Figure 7.30 shows the time course of an infection that overcomes the innate immune defences, develops into a disease and is then resolved.

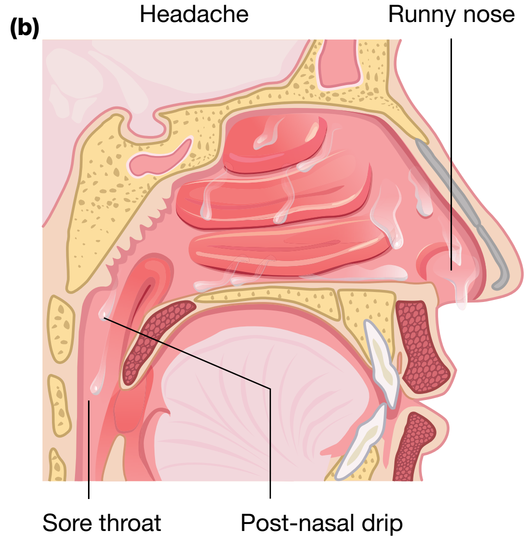




**FIGURE 7.29** **(a)** On average, adults have two to three colds each year, while children have even more. Greater than 200 different viral infections can develop into the common cold disease, the most common being rhinovirus, which is transmitted as an aerosol (airborne) or by direct contact. **(b)** Early symptoms of the common cold.

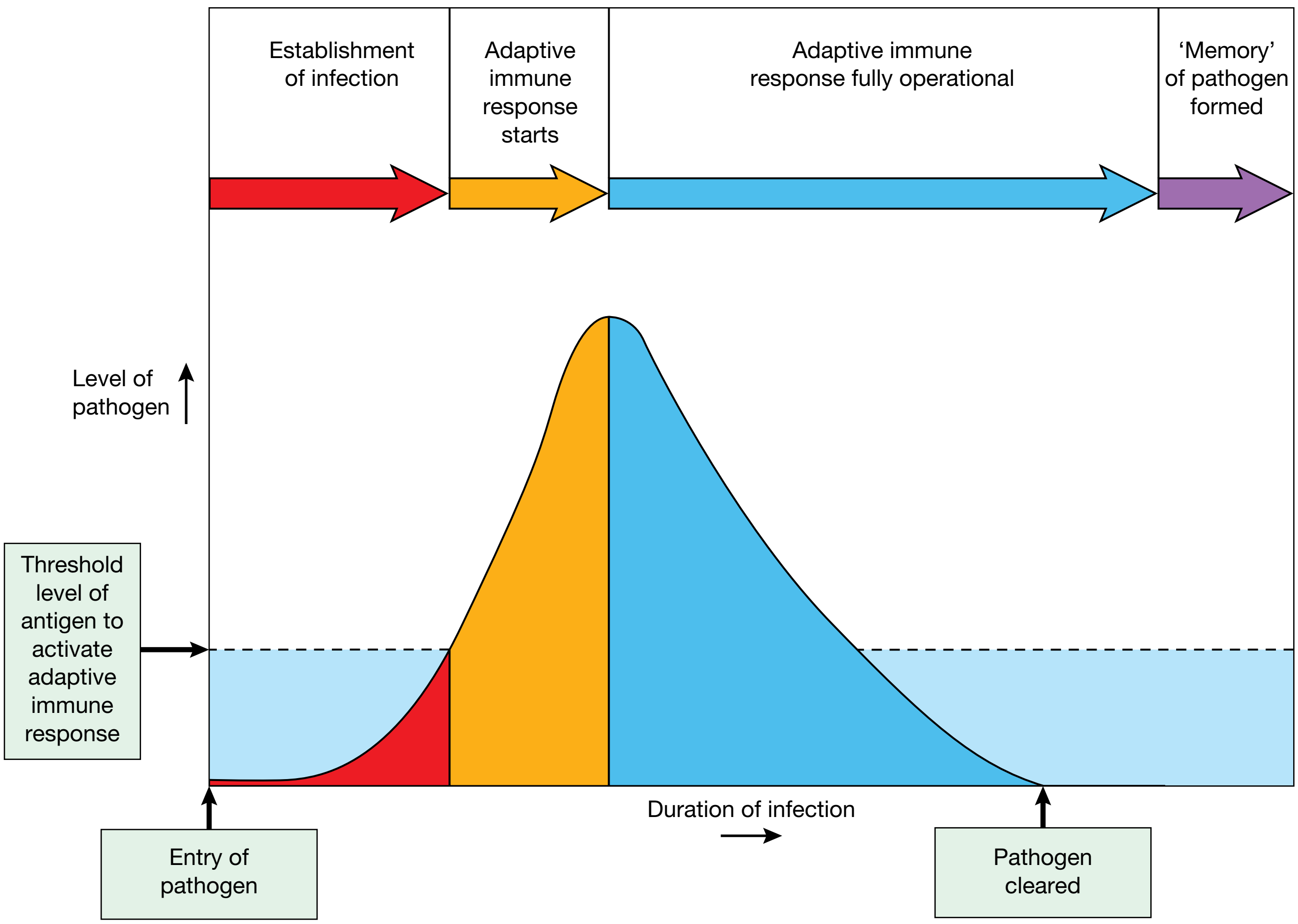



**FIGURE 7.30** Diagram showing the changing levels of an invading pathogen. If the innate immune responses fail to contain the pathogen, the pathogen's numbers increase and reach a level when the adaptive immune responses are called into action. Several days are needed for these adaptive defences to become fully effective and the pathogen continues to replicate. After this time, adaptive immune responses eliminate the pathogen.

studyon

Unit 3
AOS 2
Topic 2
Concept 6

**Do more**
Overview of defence mechanisms

Initially, the innate immune defences respond to the infection (period shown in red). However, if the viruses reach a certain level, the innate immune defences are overcome. Then the adaptive immune defences are called into action. The responses of adaptive immunity typically take days to develop (period shown in yellow) and the pathogens continue to increase in number. (This is the period during which Kylie felt ill and miserable.) However, after several more days, the adaptive immune responses are fully developed and they progressively eliminate the pathogen (period shown in blue).

The situation described above, in which the pathogen is progressively eliminated by adaptive immune defences, is the most common outcome. If, however, the adaptive immune defences are overcome, the infection may become long-lasting or chronic. In other cases, the infection can develop into a life-threatening situation requiring emergency intervention, such as can occur with *Staphylococcus pyogenes* toxic shock syndrome (refer back to Matthew's story on pages 240–241, in chapter 6).

Let's now look in more detail at adaptive immunity, the third line of defence that protects us against both pathogens and toxins.

# Adaptive immunity

Refer back to table 7.1 to compare the features of the innate immune system and the adaptive immune system.

**Adaptive immunity** is also called the **acquired immunity** because we acquire or develop this form of immunity through contact with the various pathogens we meet during our lifetimes. This is in contrast to our innate immune system — we are born with that. The adaptive immune responses are initiated after innate immunity has failed to check an infection, but adaptive immune responses are slower to come into full operation.

Three distinguishing features of adaptive immunity are:

- specificity
- self-tolerance
- memory.

One of these features is shared with innate immunity. Which is it?

***Specificity***: Adaptive immunity is specific because its responses are tailor-made to recognise each antigen involved with an infection. When an infection occurs, only the specific adaptive immune cells that can recognise the pathogen are selected for action and proliferate. This process corresponds to the yellow sector in figure 7.30 above and involves **clonal selection** and **clonal expansion** (see page 319).

***Tolerance***: Adaptive immunity is self-tolerant and immune cells will not normally respond to the antigens on healthy body cells (self antigens). If self-tolerance breaks down, one of the so-called autoimmune diseases can result (see chapter 8, page 351).

***Memory***: As well as defending against an established infection, adaptive immunity has the ability to remember previous antigens to which it has responded. This is termed **immunological memory** and it enables more rapid responses in the case of future infections by the same pathogen. Most of the cells involved in the adaptive immune response to infection are removed by apoptosis after a particular pathogen has been eliminated. However, a small number of so-called **memory cells** remain (refer back to figure 7.30 and see the purple arrow).

Memory cells are already primed to produce a more rapid and stronger adaptive immune response when a pathogen re-infects. The pathogen is often eliminated before any symptoms appear. Refer to figure 7.31, which shows the antibody levels produced during the body's first and second responses to surface antigens on an infectious pathogen. The action of memory B cells can be seen in the faster and greater production of antibodies on the second exposure to an antigen as compared to the initial exposure.

studyon

Unit 3
AOS 2
Topic 2
Concept 9

**Adaptive (specific) immune response**
Summary screen and practice questions

studyon

Unit 3
AOS 2
Topic 2
Concept 9

**Do more**
Third line of defence: the immune response

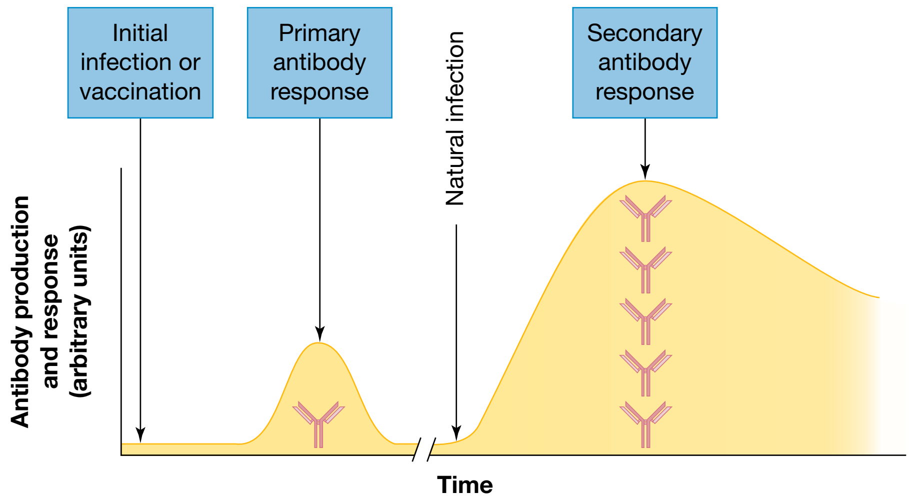



**FIGURE 7.31** An initial infection (or vaccination) produces a primary response about 10 days after infection. This response is shown by the appearance of antibodies specific to antigens of the pathogen concerned. In the secondary response to the same pathogen, memory immune cells are activated more rapidly, and antibodies are produced more quickly and in greater quantities than in the primary response.

## Components of adaptive immunity

The key components of adaptive immunity include:

- special white blood cells (lymphocytes) — T cells and B cells
- special antigen-binding proteins — antibodies, also known as immunoglobulins
- lymph nodes — secondary lymphoid organs where B cells and T cells meet foreign antigens and are activated, and where adaptive immune responses occur.

Adaptive immunity operates in two different ways (see figure 7.32):

1. **Humoral immunity** involves the actions of antibodies that identify and bind to **extracellular** pathogens, to toxins and to other extracellular foreign antigens. Antibodies are products of special B cells termed plasma cells.
2. **Cell-mediated immunity** involves various actions of T cells. Cytotoxic T cells eliminate body cells that are infected by pathogens or have abnormal or missing self markers. Note that cell-mediated responses do not directly attack pathogens. Instead, they remove infected cells and, by this means, eliminate **intracellular** pathogens.

Let's look first at the antibodies that are the antigen-binding proteins of the humoral component of adaptive immunity. Then we will examine the cells of adaptive immunity.

## Antibodies: weapons against extracellular antigens

**Antibodies,** also known as **immunoglobulins**, directly identify and bind to extracellular foreign antigens, either neutralising them or tagging them for destruction. Many antibodies are free molecules in solution in the blood or lymph, or in secondary lymphoid organs. Other antibodies are present as surface receptors on B cells (see figure 7.33) **Each type of antibody has a unique antigen-binding site so that it can recognise and bind to just one specific antigen.** The major antibodies involved in these immune responses are immunoglobulin G (IgG) and immunoglobulin M (IgM).

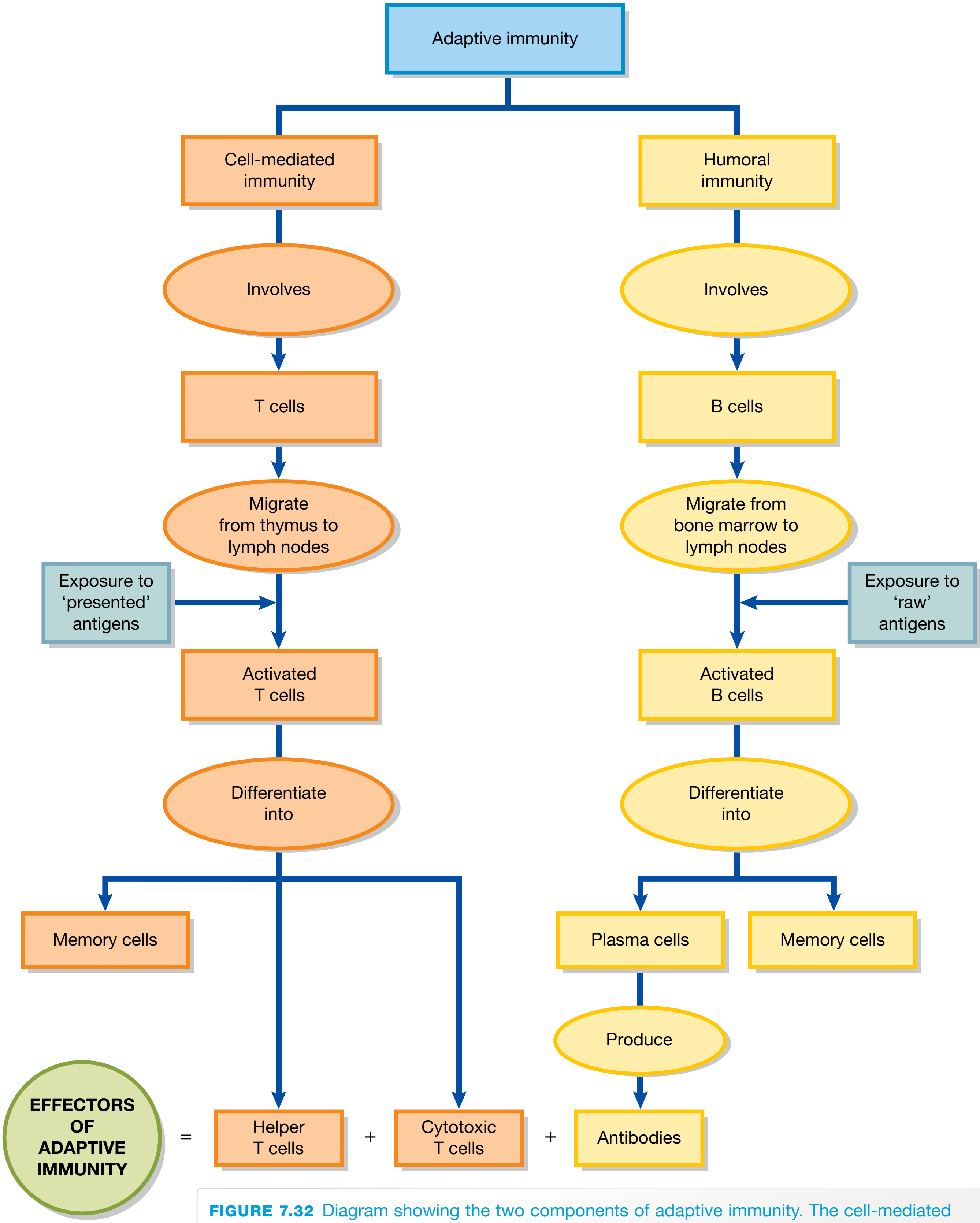



**FIGURE 7.32** Diagram showing the two components of adaptive immunity. The cell-mediated component (shown at left) is delivered by helper T cells and cytotoxic T cells, and the humoral component (shown at right) is delivered by antibodies that are the product of special B cells, called plasma cells. 'Raw' refers to antigens as they exist without any processing. 'Presented' refers to antigens displayed on the surface of a cell that presents them to T cells.

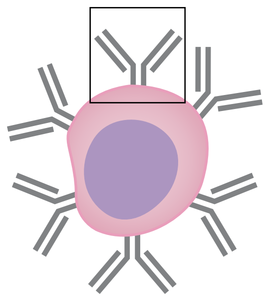



**FIGURE 7.33** Diagram showing a B cell with its surface receptors, each receptor composed of a Y-shaped antibody (immunoglobulin) molecule. Each receptor has two antigen-binding sites, one at the end of each arm of the Y. Activated B cells also secrete the same antibodies into extracellular fluids where they can bind directly to soluble antigens or they can bind to surface antigens on an extracellular pathogen.


Refer back to page 298 to check on how activated complement proteins assist in the elimination of pathogens.


**studyon**

| Unit 3 | **Production and action of antibodies** Summary screen and practice questions |
|---|---|
| AOS 2 | |
| Topic 2 | |
| Concept 10 | |

**studyon**

| Unit 3 | **See more** Antibody activity |
|---|---|
| AOS 2 | |
| Topic 2 | |
| Concept 10 | |

The antigens that are targeted by antibodies may be on the surface of extracellular pathogens, or they may be free-floating foreign molecules, such as toxins in the blood, lymph or extracellular fluid. Antigens can even be part of a molecule, such as a small peptide fragment from a protein molecule.

Antibodies do not directly destroy pathogens but defend against infectious pathogens and toxins as follows:

1. Antibodies bind to surface antigens on pathogens and form a coating that neutralises pathogens by blocking their receptors so that the pathogens cannot attach to healthy body cells and infect them.

   Likewise, antibodies bind to bacterial toxins, animal toxins and venoms. The antibodies bind to and neutralise the harmful effects of the toxin or venom. For example, a person bitten by a venomous snake is given antivenom by intravenous injection. Antivenom is a solution of specific antibodies that combine with the venom molecules, preventing them from binding to cell-surface receptors — this renders the venom harmless.
2. By binding to surface antigens on pathogens, antibodies tag pathogens for destruction. The elimination is carried out either by complement proteins or by phagocytic cells such as macrophages.

   For example:
   - Complement proteins activated by antibodies on the surface of pathogens destroy the pathogens by lysis as a result of forming MACs.
   - Antibodies coating the surface of pathogens opsonise them, making them more susceptible to elimination by phagocytic cells (see figure 7.34).

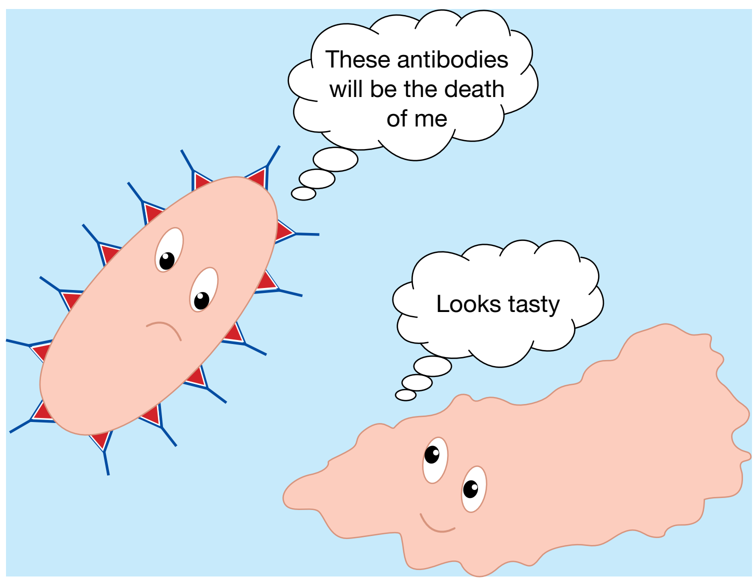



**FIGURE 7.34** A coating of antibodies bound to the surface antigens of an extracellular pathogen tags the pathogen for destruction by a phagocyte, such as a macrophage. Surface antigens of the pathogen are shown as red triangles, and the matching antibodies produced by B cells are shown as blue Y shapes.

## Structure of antibodies

**Antibodies, also known as immunoglobulins, are antigen-binding proteins produced by B cells.** All antibody molecules have the same basic structure consisting of two identical heavy polypeptide chains and two identical light polypeptide chains. This means that antibodies are proteins with a quaternary structure (refer back to chapter 2, page 53).

In terms of their amino acid sequences, both the heavy chain and the light chain have two distinct regions: (1) a constant region that does not vary, and (2) a variable region that differs between antibodies.

Examine figure 7.35. An antibody forms a Y-shaped structure. The variable regions are located at the end of each arm of the Y and these regions are the **antigen-binding sites**, where the antibody binds to its specific antigen. (This region is denoted by the symbol **Fab** = fragment antigen binding.) The shape of the antigen-binding site determines whether an antibody and antigen can bind. The remainder of the antibody forms its constant region (shown in purple for the H chain and in yellow for the L chain, and denoted **Fc**). Complement proteins can bind to the constant region of an antibody.

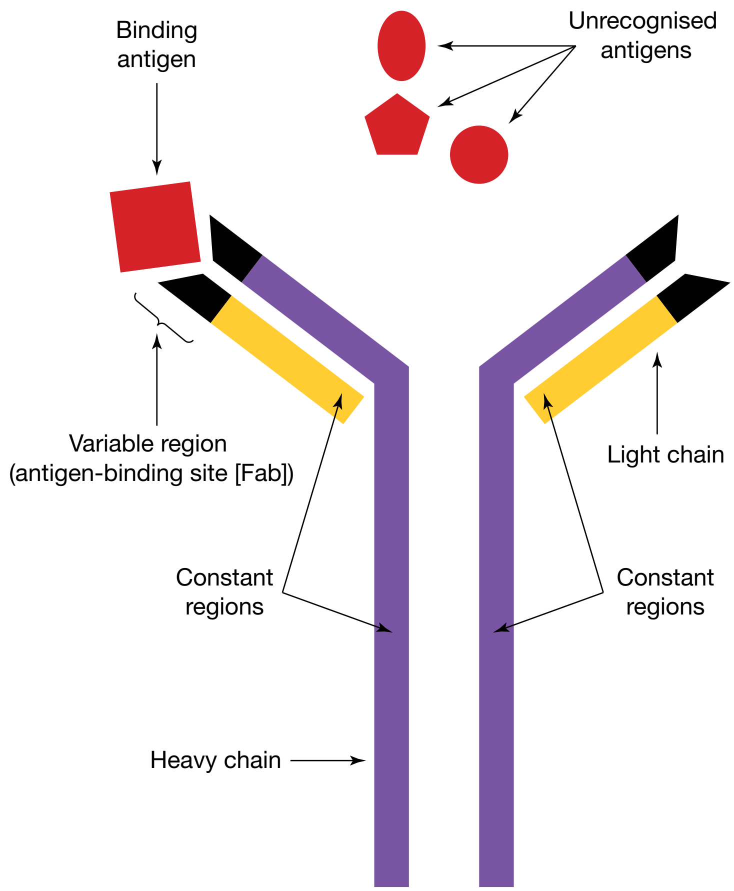



**FIGURE 7.35** Diagram showing the basic structure of an antibody molecule — two identical heavy (H) chains and two identical light (L) chains. Both the H chains and the L chains have a variable region (shown in black and about 110 amino acids long) where the antigen specific to that antibody can bind. The amino acid sequences in the variable regions of each chain differ between antibodies that can bind to different antigens.

Figure 7.36 shows a ribbon model of immunoglobulin G, the main antibody in the blood. A hinge at the region of the disulfide bond gives flexibility to the antibody molecule, enabling it to maintain the best fit with its antigen.

## Classes of antibodies

**The five classes of antibody (immunoglobulin) differ in the particular kind of heavy chain in their structure and differ in their specific functions.** Table 7.3 identifies these five classes of antibody and lists some of their characteristics.

Note that:

IgG is the main antibody in the blood and provides the key antibody-based defence against pathogens.

IgM is the first antibody to appear in an infection and provides defence until sufficient IgG is formed.

IgA attaches to the surfaces of the mucous membranes of airways, gut and urogenital tract.

IgE stimulates the release of histamine and other chemicals that cause allergies.

**eBook*plus***

**Weblink**
Dr Song Tan's image gallery

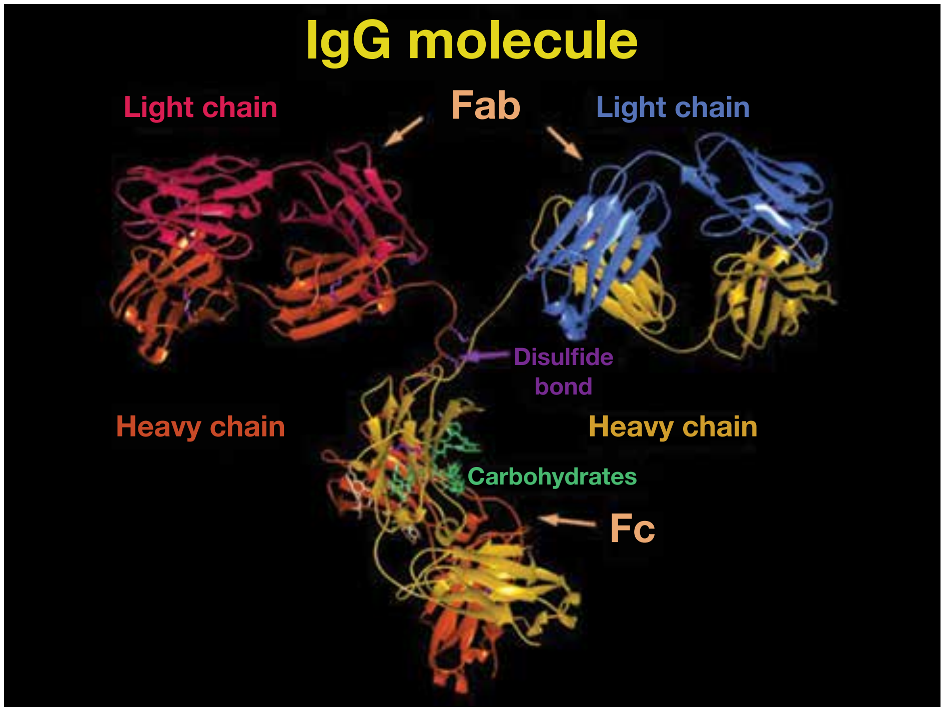



**FIGURE 7.36** Ribbon model of an antibody, immunoglobulin G (IgG). This 3D representation shows the protein backbone of this antibody. The two heavy chains of this antibody are colour-coded in orange and yellow, and the two light chains are in red and blue. (Image courtesy of Dr Song Tan, Penn State University.)

**TABLE 7.3** Different classes of antibodies and some of their features. The identity of the heavy chain determines the class to which an antibody belongs. Every antibody molecule has either two lambda chains or two kappa chains in its structure. Which class of antibody gives passive adaptive immunity to a newborn baby until its own immune system becomes operational?

| Feature | IgG | IgA | IgM | IgD | IgE |
|---|---|---|---|---|---|
| molecular weight | 150 000 | 320 000 | 900 000 | 180 000 | 200 000 |
| approx. percentage of total immunoglobulins | 75 | 15 | 1 | 0.2 | 0.002 |
| concentration in blood (mg/mL) | 12 | 2 | 1 | 0.04 | <0.001 |
| heavy chain type | gamma | alpha | mu | delta | epsilon |
| light chain type | kappa or lambda | kappa or lambda | kappa or lambda | kappa or lambda | kappa or lambda |
| number of units in functional antibody | 1 | 2 | 5 | 1 | 1 |
| able to cross placenta | yes | no | no | no | no |
| found in saliva, tears, sweat and on mucous membranes | no | yes | no | no | no |
| present in breast milk | yes | yes | no | no | no |
| present on surface of B cells | no | no | no | yes | no |
| active against viruses | yes | yes | some | no | no |
| active against some bacteria | yes | yes | yes | no | no |
| activate complement | yes | | yes | | |
| attach to surfaces of mast cells (involved in allergic reactions) | no | no | no | no | yes |

## ODD FACT

IgD acts as the surface receptor on B cells, and each B cell is covered with several hundred thousand of these receptors.

Figure 7.37 is a diagram showing the different classes of antibodies. These classes differ in their heavy chains and some differ in the number of units in the functional antibody: 1, 2 or 5.

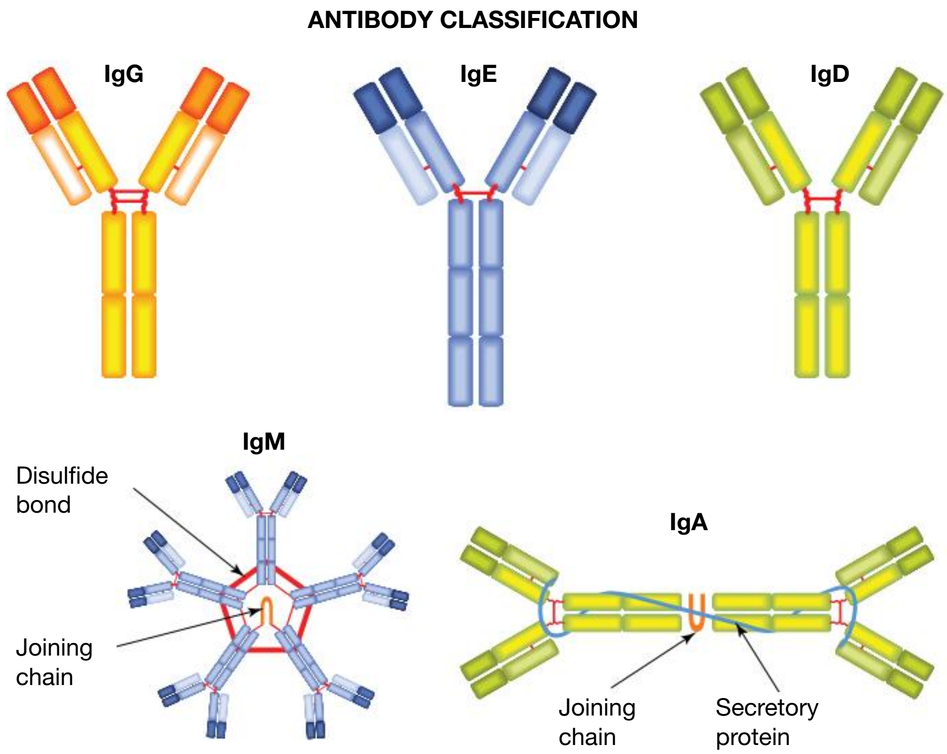



**FIGURE 7.37** The five classes of antibodies. Three classes, IgG, IgE and IgD, consist of single units, while IgA and IgM consist of 2 and 5 units respectively. Can you suggest why IgG is able to cross the placenta, but not IgA?

### Antigen-binding sites for different antigens are unique

Cells of adaptive immunity are able to produce a remarkable diversity of antibodies that can respond to billions of different antigens. The antigen-binding sites on one antibody molecule have a unique shape that recognises just one particular antigen (see figure 7.38a). Binding can only occur between an antibody and an antigen that includes a shape that is complementary to (fits into) the shape of the antigen-binding site of an antibody — a bit like a key and a lock. Different antibodies have differently shaped antigen-binding sites.

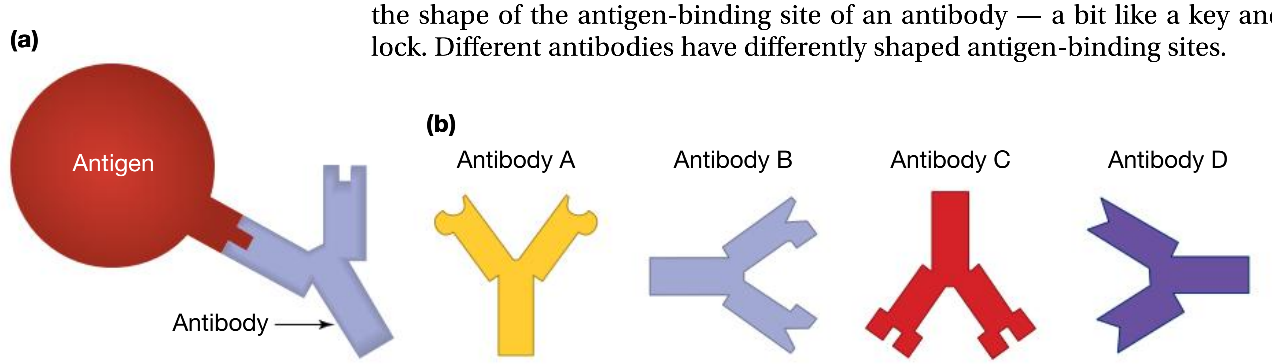



**FIGURE 7.38** **(a)** Diagram showing a stylised antigen bound to an antibody. Note that the antigen includes a shape that can fit the antigen-binding site of the antibody. **(b)** Diagram showing stylised antibodies with differently shaped antigen-binding sites. The shape of the site determines the antigen that can be bound to the antibody.

Given that we are capable of producing an enormous number of different antibodies in terms of their antigen-binding sites, how is this achieved? Check out the next section.

## Diversity of antibodies

We cannot possibly know all the antigens that we might meet during our lives, yet our adaptive immune system has the remarkable ability of responding to every possibility. This is achieved because our B cells can produce approximately ten billion different kinds of antibody (immunoglobulin), with each different antibody able to recognise and bind to one specific antigen.

Each of these billions of different antibodies has a different amino acid sequence in the variable regions of both its heavy and its light chains. These differences produce the extraordinarily large number of different shapes of antigen-binding sites on antibody molecules. The human genome comprises an estimated 21 000 different genes. Clearly, each different antibody cannot be the product of one different gene. How can the human genome encode billions of different amino acid sequences in these different antibodies?

The mystery was not solved until 1976 when the Japanese scientist Susumu Tonegawa (1939–) discovered how this incredible multitude of different antibodies was created. Tonegawa showed that the variable regions of each antibody are encoded by combinations of several genes randomly picked from a few larger pools of genes.

For example, the variable region of the heavy chain of a human antibody is encoded by three genes (V, J and D) that are randomly picked, one from each of a pool of 51 V genes, 6 J genes and 27 D genes. This is a bit like picking one ball at random from each of three barrels (see figure 7.39). This process generates a possible $51 \times 6 \times 27 = 8262$ different genetic combinations to encode variable regions of the heavy chains of an antibody.

So, just three genes picked at random from a total of 84 genes creates more than 8000 different genetic combinations that encode the variable regions of the heavy chains. This process of gene shuffling is termed **V(D)J recombination** and it takes place in B cells while they are in the bone marrow.

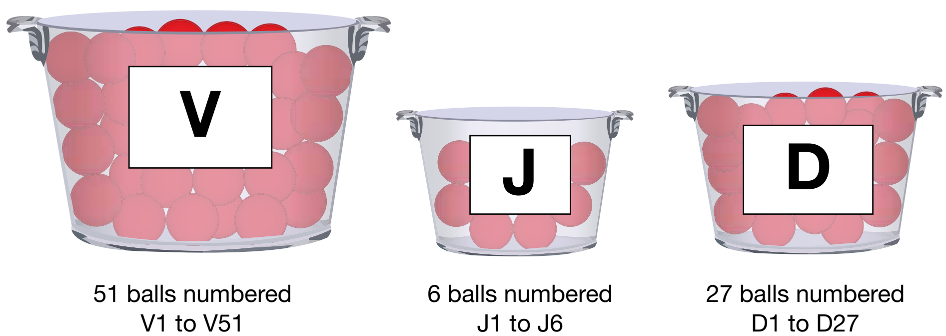



**FIGURE 7.39** The diversity of the genes that encode the heavy chain of antibody molecules produced by B cells is created by randomly picking three genes, $V_n$, $D_n$ and $J_n$, from a larger pool of each of these genes. This is like picking one numbered ball from each of three different barrels. More than 8000 different combinations of three balls can be picked at random.

A similar process involving two genes (V and J) occurs for variable regions of the light chains of antibody molecules. About 320 different gene combinations are possible for the light chain of an antibody molecule. Combining the possibilities for both the heavy and the light chains produces

T cell surface receptors consist of two protein chains (alpha and beta). Random gene rearrangement (similar to that occurring with B cells) generates the diversity seen in T cell receptors.

## ODD FACT

In 1987, Susumu Tonegawa was awarded the Nobel Prize in Medicine or Physiology for his discovery of 'the genetic principle for generation of antibody diversity'.

$8262 \times 320 = 2.6$ million different variations. More antibody diversity is produced because of **somatic mutations** in the genes concerned. In addition, variations in how the genes are joined create even more diversity.

Through these gene re-arrangements, the B cell population in the bone marrow has the capability of producing billions of antibodies with different antigen-binding sites. This antibody diversity can be expressed both in the soluble antibodies secreted by B cells into extracellular fluids, as well as in the antibodies that form the surface receptors on B cells. It would, of course, be wasteful to produce all these antibodies since most would never be needed and others would only be needed at particular times. As we will see (page 319), the adaptive immune system has a neat solution and, when a specific antibody is required, only the B cells that produce that particular antibody become active and multiply.

## KEY IDEAS

- Adaptive or acquired immunity forms the third line of defence against infection.
- Adaptive immunity has a humoral component and a cell-mediated component.
- Humoral adaptive immunity targets extracellular antigens and is achieved through the action of antibodies.
- Cell-mediated immunity attacks infected or abnormal cells and is brought about by T cells.
- Antibodies, also known as immunoglobulins, are a diverse group of antigen-binding proteins consisting of two identical heavy chains and two identical light chains.
- Antibodies are the products of B cells.
- Antibodies can only bind to those antigens that include a shape that is complementary to the shape of the antigen-binding sites of the antibodies.
- Antibody diversity is produced by processes including V(D)J recombination and somatic mutation of B cell genes.

## QUICK CHECK

**21** Identify the following statements as true or false:

- **a** At birth, a baby has no adaptive immune defences.
- **b** Antibodies are antigen-binding proteins produced by B cells.
- **c** Cell-mediated adaptive immunity is produced by T cells.
- **d** An antibody can bind only to an antigen that includes a shape that is complementary to the antigen-binding sites of the antibody.
- **e** Adaptive immunity, unlike innate immunity, has a 'memory'.
- **f** Innate immune responses occur after adaptive immune responses have been overcome.
- **g** Antibodies are released in advance of possible invasions by infectious pathogens.
- **h** V(D)J recombination of antibody genes occurs in B cells when they are in the bone marrow.

**22** Give an example of:

- **a** a class of antibody that can cross the placenta
- **b** a class of antibody that is found on the surface of mucous membranes
- **c** a type of immune cell that eliminates infected body cells
- **d** a randomly picked gene that contributes to the diversity of the heavy chains of an antibody.

## Cells of adaptive immunity

The main cells of the adaptive immune system are members of the class of white blood cells known as **lymphocytes**. **The lymphocytes that play the key roles in adaptive immunity are T cells and B cells.** Lymphocytes are small cells (about 7 to10 μm in diameter, similar in size to red blood cells). They are found in the blood and the lymph, and are highly concentrated in secondary lymphoid organs such as lymph nodes. Figure 7.40 shows a scanning electron microscope (SEM) of whole blood that includes two lymphocytes. (Refer back to figure 7.4 on page 280 to see light microscope images of lymphocytes in blood smears.) The contrast between these images shows the power of SEM in providing vivid 3D images.

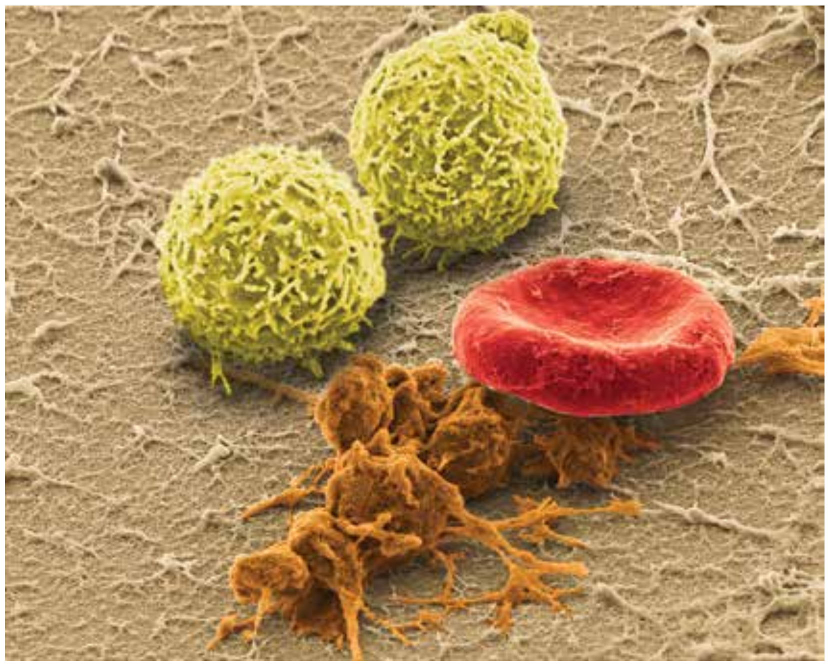


**FIGURE 7.40** False-coloured scanning electron microscope image of harvested whole blood showing two lymphocytes (yellow), a single red blood cell (red) and a cluster of activated platelets (orange). (Image courtesy of Jonathan M. Franks and Dr Donna Stolz, University of Pittsburgh Center for Biologic Imaging.)

**ODD FACT**

B cells and T cells also differ in other cell-surface markers. Mature B cells carry markers known as CD 19 and CD 20, while mature T cells have either CD4 or CD8 surface markers.

Using conventional staining, B cells cannot be visually differentiated from T cells. However, they can be distinguished by the different cell-surface markers they carry, including their specific receptors. Each B cell has a large number of identical B cell receptors, while each T cell has multiple identical copies of T cell receptors.

### B cells and T cells: different roles in adaptive immunity

The roles of these cells in adaptive immunity differ:

- T cells deliver the cell-mediated immune defences that include the direct elimination of pathogen-infected cells and other abnormal cells, such as cancer cells.
- B cells deliver the humoral immune defences by secreting **antibodies** that bind to surface antigens on pathogens and label them for elimination; antibodies also bind to soluble toxins.

Figure 7.41 shows a dense population of lymphocytes, mainly B cells, in the cortex of a lymph node.

**ODD FACT**

The combined mass of all the lymphocytes in the blood and lymphoid tissue of an adult is equivalent to the mass of the human brain.

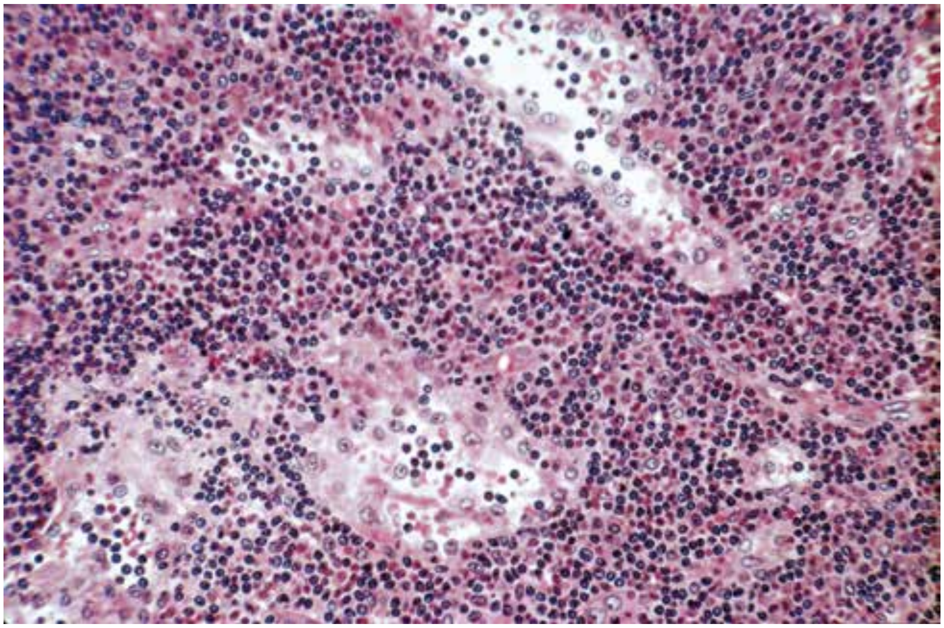


**FIGURE 7.41** Light microscope image of a section of the cortex of a lymph node showing the dense population of lymphocytes, mainly B cells. Each B cell has a heavily stained nucleus surrounded by a narrow ring of cytoplasm. Each B cell carries multiple copies of just one kind of receptor that can recognise and bind to just one specific antigen.

The diversity of the B cell receptors (BCRs) results from gene shuffling as outlined on pages 315–316. The diversity of T cell receptors (TCRs) is produced in a similar manner.

A remarkable diversity of surface receptors is present in the large populations of B cells and T cells in the lymph nodes. This diversity exists even though the majority of antigens that these cells can recognise may never gain entry to the body. The body is armed in advance against invasion by every possible harmful antigen, but the numbers of B and T cells in the populations that can respond to one specific antigen are very small. However, when an infection occurs, the numbers of the specific T cells and B cells that can respond to the infection are greatly increased.

Table 7.4 shows a summary of the different roles of T and B cell lines of adaptive immunity. Also refer back to figure 7.32 on page 310. Note that:

- B cells recognise *extracellular* pathogens and label them for destruction with antibodies, while T cells recognise and eliminate *intracellular* pathogens.
- B cells respond to 'raw' antigens, while T cells respond to antigens that are 'presented' to them by other cells.

**TABLE 7.4** Summary of the two cell lines of adaptive immunity. From each cell line, mature functional immune cells develop that, when activated by exposure to their specific antigens, deliver their immune responses.

| Cell line | Features | Function |
|---|---|---|
| • **T cells** ***'helpers and hunters'*** | mature in thymus, then migrate to lymph nodes; each T cell has many surface receptors that recognise one specific antigen; activated by exposure to antigens presented to them on the surface of other cells; differentiate into various types of T cells, including helper T cells and cytotoxic T cells | Activated cytotoxic T cells eliminate infected and abnormal cells. Activated helper T cells send signals that stimulate B cells to secrete specific antibodies. Memory cells retain a 'memory' of antigens met previously. |
| • **B cells** ***'antibody factories'*** | mature in bone marrow, then migrate to lymph nodes as naïve B cells; each B cell has many receptors that recognise just one kind of antigen; activated by direct exposure to raw antigens; develop into plasma cells that produce antibodies that defend against the specific antigen that activated the B cell | Activated B cells produce antibodies against extracellular antigens, either free-floating or on the surface of pathogens. Memory cells retain a 'memory' of antigens met previously. |

## B cells and T cells must pass a test

T cells and B cells develop in the primary lymphoid organs — T cells in the thymus and B cells in the bone marrow.

Each single B cell and T cell carries many copies of one surface receptor that can recognise and bind to just one specific antigen. When B cells and T cells are in the primary lymphoid organs, they are put through a test. What test? Cells are tested by being exposed to self antigens. Any cells that do not recognise self antigens fail the test and are eliminated by apoptosis. Cells that fail to respond to non-self antigens are also eliminated. Testing ensures that the surviving T cells and B cells are self-tolerant and will not attack the body's own healthy cells carrying self antigens, but will attack foreign antigens.

**naïve**: refers to an immune cell that has not yet met and been activated by its complementary antigen

Cells that pass the test graduate as **naïve** T cells and B cells. These cells migrate to secondary lymphoid organs, such as the lymph nodes. This is where the naïve B and T cells are activated by exposure to their complementary antigens and undergo cycles of cell division and differentiate into effector cells. Lymph nodes are the location where the various adaptive immune responses to pathogens take place.

In a sample of normal blood, lymphocytes (T cells, B cells of adaptive immunity, and the NK cells of innate immunity) constitute from 14 to 47 per cent of the total white blood cells. The majority (about 80%) of the circulating lymphocytes are T cells, with B cells making up a smaller proportion (5 to 15%).

Let's now look at how B cells operate.

# B cells: antibody factories

The critical role of B cells in adaptive immunity is to produce antibodies that are specific to foreign antigens and toxins. These antibodies defend only against *extracellular* antigens, such as antigens on the surface of a pathogen in the blood, or soluble antigens, such as toxins or venoms, in extracellular body fluids. After an infection starts, about one week is required before the specific antibodies needed to counter the infection are present in sufficient amounts. While this is happening, a person can feel quite ill. During this period, the B cells with receptors that match the antigen concerned go through cycles of cell division and differentiate into special antibody-secreting cells, called plasma cells.

Within the population of B cells in the lymph nodes, only a few cells have one particular receptor. This small number of B cells with receptors specific to one particular antigen is not adequate to counter a specific bacterial infection. So, if a new pathogen gains entry to the body and multiplies rapidly, how can a few B cells produce enough antibodies to defend against the new pathogen? The answer is: 'By clonal selection and clonal expansion.'

### Clonal selection and clonal expansion of B cells

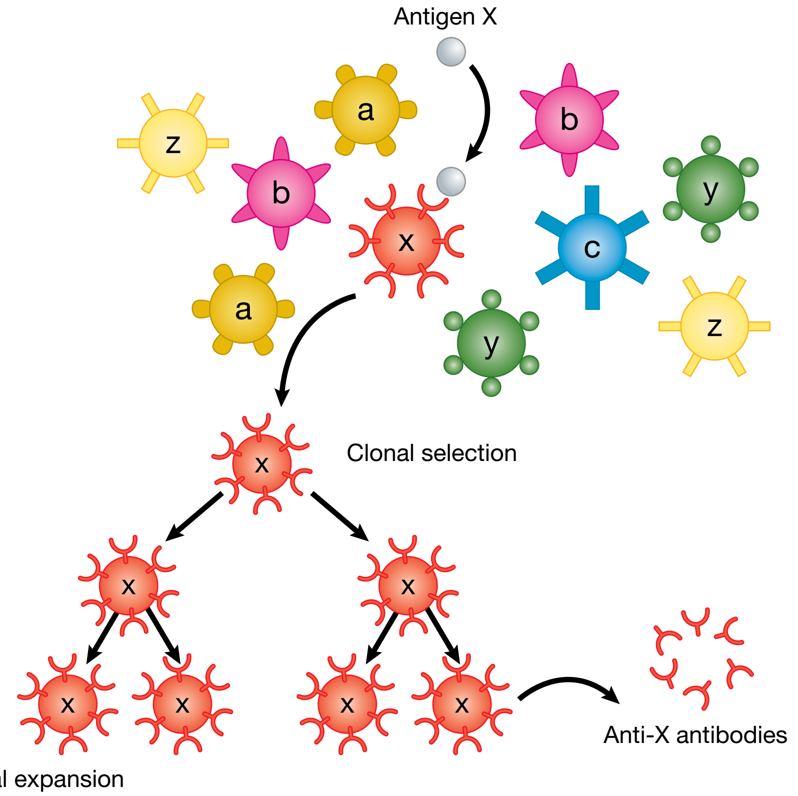
Consider a 'new' antigen that gains entry to the body, such as an antigen on the surface of a microbe. When this antigen reaches the lymph nodes, it comes into contact with many naïve B cells that do not recognise it. Eventually, however, the antigen meets a B cell that can recognise and bind to it. In effect, this antigen has 'selected' the specific B cell that can eliminate it (and the pathogen on which it is present). This is **clonal selection.**

**ODD FACT**

After being 'selected', the B cell repeatedly divides, doubling in number every six hours. Within a week, a clone of about 20 000 B cells has been built up.

The binding of the antigen to its 'selected' B cell activates the B cell and initiates cycles of cell division. The daughter cells from each division also replicate, and so do their daughter cells, and so on. The result is the production of a large **clone** of B cells, all having identical antigen-binding receptors — this is called **clonal expansion**. Most of these cells differentiate into short-lived plasma cells that secrete soluble antibodies against the specific antigen (see figure 7.42). B cells can produce antibody molecules at an estimated rate of up to 200 per second.

FIGURE 7.42 Diagram showing clonal selection. The B cells shown at the top are various B cells with different cell-surface receptors. Of these B cells, only the cell labelled X has cell-surface receptors (antibodies) that recognise antigen X. By binding to this B cell, antigen X 'selects' this cell and activates it, causing it to undergo several cycles of cell division and differentiate into antibody-secreting B cells or plasma cells. Many identical B cells are produced, all of which produce identical anti-X antibodies. The production of a large number of identical cells from a single cell is termed **clonal expansion**.



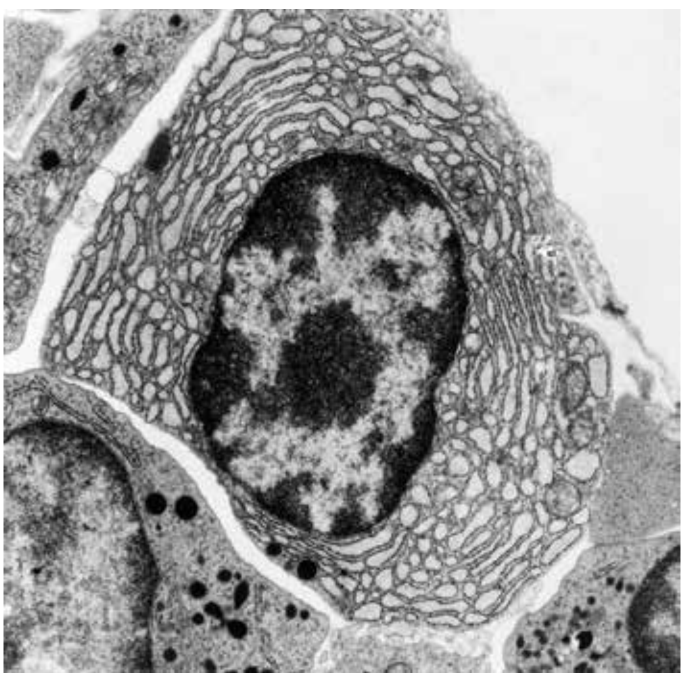



FIGURE 7.43 TEM of an activated B cell or plasma cell. Note the prominent channels of the rough endoplasmic reticulum to which are attached ribosomes, which are the sites of protein synthesis. Other channels form part of the structure of the Golgi complex.

This explanation as to how the 'right' antibodies are produced just when they are needed is termed the **clonal selection theory**. This theory provides an explanation for the specificity of adaptive immunity, that is, how the antibodies produced at a given time are those that specifically defend against the invading pathogen.

The clonal selection theory was first proposed in 1955 by an Australian, Sir Frank Macfarlane Burnet (see box at the end of this section). Evidence in support of Burnet's clonal selection theory came from later research findings of Sir Gustav Nossal (1931–) who demonstrated that each B cell carries multiple copies of just one kind of cell-surface receptor. Can you suggest why this finding gave support to Macfarlane Burnet's theory?

Figure 7.43 shows a transmission electron micrograph of an activated B cell or plasma cell. These cells are very active in protein synthesis — remember that antibodies are composed of polypeptide chains. Because of this, cell organelles such as the rough endoplasmic reticulum are very prominent features of activated B cells. Newly synthesised antibodies are not retained within plasma cells but are transported out of these cells as free-floating molecules in the bloodstream and lymph. For this reason, another organelle expected to be prominent in plasma cells is the Golgi complex.

In the next sections, we will explore the T cells of adaptive immunity.

## BIOLOGIST AT WORK

### An Australian Nobel Laureate: Sir Frank Macfarlane Burnet

**FIGURE 7.44** Sir Frank Macfarlane Burnet, Nobel Laureate, working in his laboratory at the Walter and Eliza Hall Institute of Medical Research in the 1950s. (Image courtesy of the Walter and Eliza Hall Institute.)

The highest awards for achievement in science are the Nobel Prizes, set up under the will of Alfred Nobel, a Swedish chemist and engineer who died in 1896. The distinguished Australian scientist Frank M. Burnet, later Sir Macfarlane Burnet (see figure 7.44), received a Nobel Prize. He was born in the country town of Traralgon in Victoria and, after graduating in medicine from the University of Melbourne, chose to carry out research on viruses, particularly animal viruses. After working in England, Dr Macfarlane Burnet returned to the Walter and Eliza Hall Institute of Medical Research (refer to chapter 8, The Walter and Eliza Hall Institute of Medical Research) where he later became Director from 1944 to 1965. He experimented on influenza viruses by growing them in fertilised hens' eggs.

Although their main research was on the influenza virus, Sir Macfarlane Burnet and his co-workers also did important work on other viruses including poliomyelitis, herpes, Murray Valley encephalitis, myxomatosis and smallpox-like viruses. They also worked on organisms half-way between viruses and bacteria, the rickettsias. The causative agent of Q-fever, *Coxiella burnetii*, is named after Burnet.

Sir Macfarlane Burnet became interested in how the body defences act against infection by micro-organisms and extended his research to include the immune system. He realised that the essential basis of the body's defence was its ability to identify foreign material as being different from its own constituents. He proposed a radical new theory, the 'clonal selection theory' (see page 319). Sir Macfarlane Burnet received many awards for his work. He received numerous honorary degrees, fellowships and prizes, leading to the ultimate achievement of the Nobel Prize for Medicine in 1960. The citation for his Nobel Prize is shown in figure 7.45. The Nobel Prize was shared with Sir Peter Medawar (1915–1987) of University College, London. Sir Macfarlane Burnet was one of the twentieth century's great scientists.


**FIGURE 7.45** Sir Macfarlane Burnet's Nobel Prize citation and medal. (Images courtesy of the Walter and Eliza Hall Institute.)

## KEY IDEAS

- Primary lymphoid organs are where B cells and T cells of the immune system are tested for their ability to tolerate self antigens but react to foreign antigens.
- Cells passing the test are selected as naïve cells and released to migrate to lymph nodes.

*(continued)*

- B cells and T cells that fail the test are removed by apoptosis.
- Each B cell is pre-programmed to recognise and bind to one particular antigen by virtue of its specific B cell receptor.
- A new antigen 'selects' the matching naïve B cells, binds to them and activates them.
- Activated B cells proliferate, producing several generations of daughter cells, resulting in a large clone of B cells.
- Activated B cells differentiate in antibody-producing plasma cells that secrete free-floating antibodies into blood and other body fluids.

**QUICK CHECK**

**23** Identify the two groups of lymphocytes involved in adaptive immunity.
**24** Where would you expect to locate the following:
  **a** B cells undergoing clonal expansion
  **b** B cells undergoing testing for self-tolerance
  **c** B cells differentiating into antibody-producing plasma cells?
**25** Consider the clonal selection theory.
  **a** Who proposed the clonal selection theory?
  **b** What is meant by the term 'clone'?
  **c** What is 'selected'?
**26** Identify the following statements as true or false:
  **a** All B cells are active in producing antibodies at any time.
  **b** One B cell carries several different receptors that can identify different antigens.
  **c** B cells that react to self antigens fail the self-tolerance test.
  **d** Plasma cells differentiate from activated B cells.

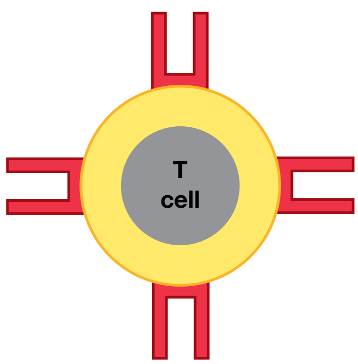



**FIGURE 7.46** Stylised diagram showing four T cell receptors (red) on a T cell. The cleft between the two arms forms the single antigen-binding site of a T cell receptor. In reality, a T cell has thousands of these surface receptors. Refer back to figure 7.33 to compare these receptors with B cell receptors.

The HLA self markers (also known as MHC markers) are outlined in chapter 6, page 263. HLA/MHC class I markers are on all nucleated cells, while HLA/MHC class II markers are restricted to a few immune cells only, including dendritic cells. In this section, the term MHC will be used.

# T cells: helpers and killers

T cells make up the majority (about 80% of circulating blood lymphocytes and are critical in the processes of cell-mediated adaptive immunity.)

Precursor T cells are produced in the bone marrow and migrate to the thymus where they mature into naïve T cells. Naïve T cells are activated to effector T cells in the lymph nodes by exposure to their matching antigens. T cell receptors consist of two polypeptide chains (alpha and beta), and each receptor has a single antigen-binding site (see figure 7.46).

Any one T cell carries thousands of T cell receptors, all with identical antigen-binding sites. However, the genes that encode T cell receptors are produced by a process of random gene shuffling (such as occurs for the polypeptide chains of antibodies (refer back to page 315)). As a result, the receptors present in a population of T cells are extremely diverse and can recognise a great variety of antigens. It is important to note that T cells can only recognise their matching antigens when they are linked to an MHC self marker on the surface of another cell.

The key role of T cells is in cell-mediated immunity. Once a pathogen gains entry to a cell, it is 'hidden' from any defence by antibodies. **Only T cells are able to take action against intracellular pathogens**.

Several kinds of T cells exist, including **cytotoxic T cells** and **helper T cells**. These two kinds of T cell look the same but can be distinguished because they have different cell-surface markers — cytotoxic T cells have CD8 markers on their surfaces and are sometimes called $CD8^+$ T cells. Helper T cells have CD4 markers and are sometimes called $CD4^+$ T cells.

CD4 glycoprotein markers on helper T cells are discussed again in relation to AIDS in chapter 8, page 360.

## Cytotoxic T cells: the killers

The major role of cytotoxic T cells (Tc cells) is to monitor body cells for the presence, on their surface, of foreign antigens that have been generated within the body cell by intracellular pathogens. Such antigens include viral proteins produced during the multiplication of a virus within a cell and bacterial compounds resulting from metabolic activity of the bacteria within a cell. Infected body cells advertise their infected state on their surfaces by displaying pathogen antigens that are linked to MHC class I molecules.

The monitoring of body cells by Tc cells requires that contact be established between the T cell and the cell to be monitored. If any cells are identified as being infected, cytotoxic T cells set about destroying them. By destroying infected cells, Tc cells stop infections from spreading to nearby body cells.

Cytotoxic T cells recognise the foreign antigens that are linked to MHC class I molecules on the surface of an infected body cell. Figure 7.47 shows a diagram of a cytotoxic T cell making contact with an infected body cell. Note the foreign antigen (shown in red) that is linked to an MHC class I marker (shown in green). The T cell receptor (TCR) (shown in grey) recognises this MHC–linked antigen as a signal that this body cell is infected.

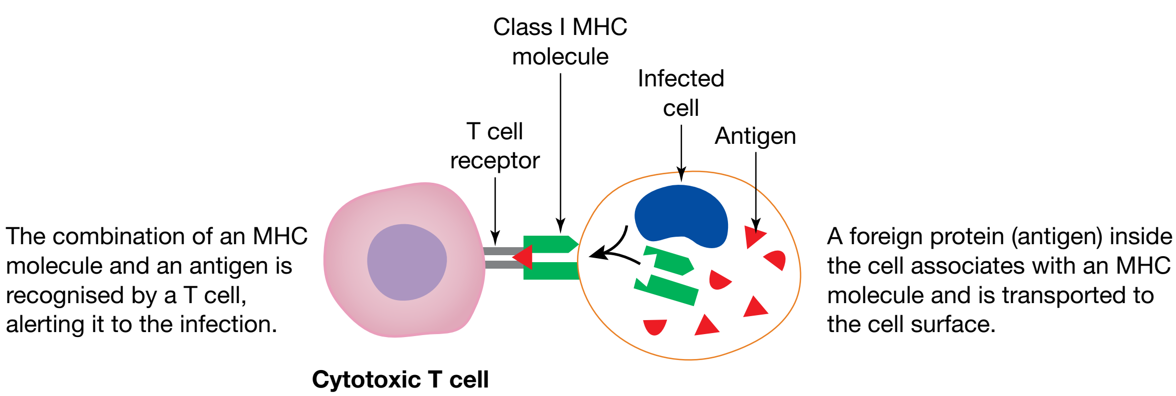



**FIGURE 7.47** A cytotoxic T cell (at left) recognises and binds to a foreign antigen (shown in red) on the surface of an infected cell. To be recognised, this antigen must be linked to an MHC class I molecule.

This binding of the antigen to the T cell receptor activates these cytotoxic T cells. These activated cells undergo cycles of cell division and form effector cytotoxic T cells. These cell divisions brings about a several thousand-fold increase in the number of effector cytotoxic T cells that can recognise the particular antigen. Although this clonal expansion may take a week to occur, the adaptive immune system is now well equipped to destroy infected cells.

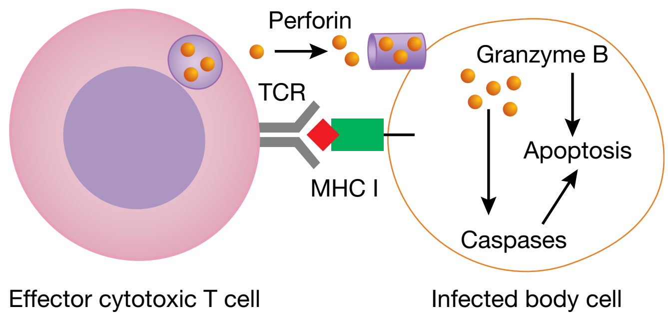
### How do cytotoxic T cells kill?

Effector cytotoxic T cells attack infected body cells by releasing perforin, a protein that inserts into the plasma membranes of target cells and creates pores. Cytotoxic T cells release destructive granules, called granzyme B, that enter the infected cell via the pore and initiate cell death by apoptosis (see figure 7.48). The cytotoxic T cells are then free to attack other infected cells that display the same foreign antigen.

Figure 7.49 shows a fluorescence image of a Tc cell that has recognised and will destroy a pathogen-infected cell.

As well as recognising viral and bacterial antigens, Tc cells can also recognise abnormal proteins such as those resulting from DNA mutations in cancer cells. Through this recognition, cytotoxic T cells can target and eliminate cancer cells.

**FIGURE 7.48** Diagram showing the mode of action of a cytotoxic T cell when it makes contact with an infected body cell carrying an antigen that it recognises. Note that the antigen must be presented linked to an MHC class I molecule. A pore forms in the plasma membrane of the infected cell. What creates the pore? Note the release of granzyme granules signalling death by apoptosis of the infected cell.



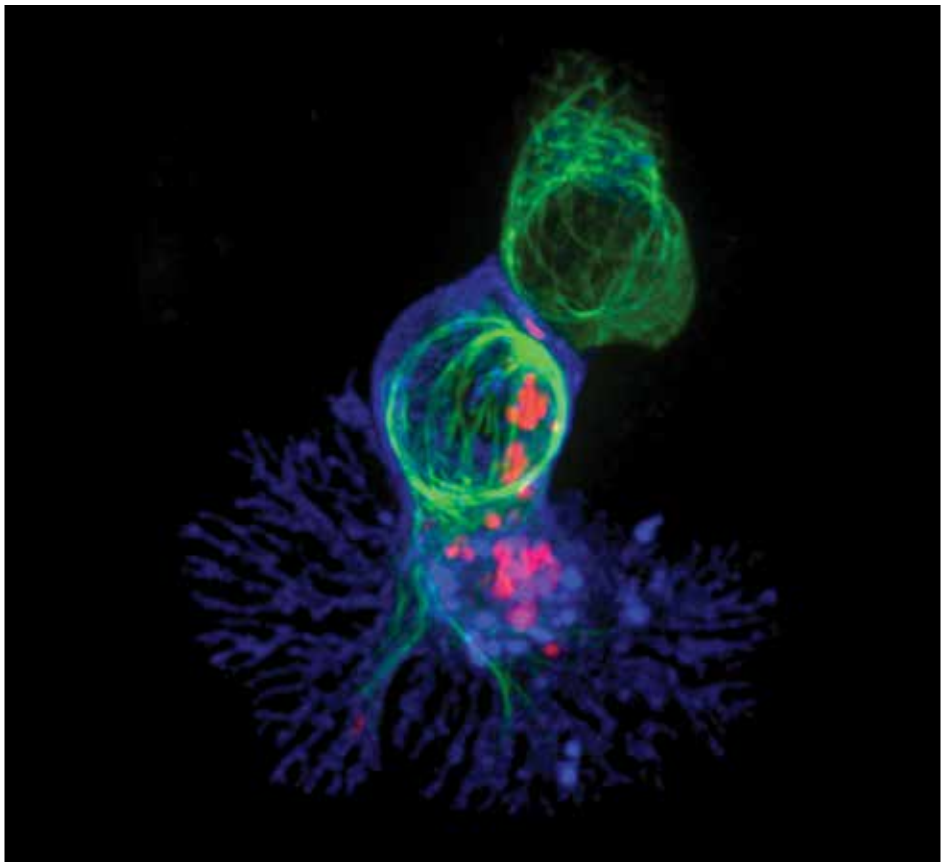



**FIGURE 7.49** A cytotoxic T cell with the CD8 markers on its plasma membrane in purple, microtubules in green, and its collection of lysosomes containing destructive granules in orange/pink. These lysosomes are being channelled towards the junction between the cytotoxic cell and its target cell (at top right), and their contents will be released to destroy the target. (Image courtesy of Alex Ritter and Gillian M. Griffiths, Cambridge Institute for Medical Research.)

The other major effector cells in adaptive immunity are the **helper T cells**. To appreciate the role of helper T cells in adaptive immunity, we first need to meet dendritic cells in their role as antigen-presenting cells (APCs).

## KEY IDEAS

- T cells deliver the cell-mediated components of adaptive immunity.
- Each T cell carries multiple copies of one specific T cell receptor that can recognise just one antigen.
- T cells include cytotoxic T cells and helper T cells.
- T cells can only recognise antigens when they are located on the surface of another cell and are linked to MHC molecules.
- Cytotoxic T cells attack pathogen-infected cells (and also cancer cells) using chemical compounds, resulting in the death of the target cell by apoptosis.

## QUICK CHECK

**27** Identify the following:
- **a** the event that activates a naïve T cell
- **b** the molecule that can create holes in the plasma membrane of an infected cell
- **c** the location where T cells are activated.

**28** Identify the following statements as true or false:
- **a** One T cell receptor has a single antigen-binding site.
- **b** Cytotoxic T cells produce antibodies to kill pathogens.
- **c** A population of naïve T cells in a lymph node would be expected to carry a great diversity of T cell receptors.
- **d** A single T cell has multiple copies of one kind of T cell receptor.
- **e** Cytotoxic T cells can destroy both extracellular and intracellular pathogens.

# Antigen-presenting cells

In 2011, the Nobel Prize in Medicine or Physiology was awarded to Ralph Steinman (1943–2011) '*for his discovery of the dendritic cell and its role in adaptive immunity*'. (Sadly, Steinman, who discovered dendritic cells in 1973, died before receiving news of this award.) Dendritic cells are cells of the innate immune system, but they also play an important role in activating naïve T cells of the adaptive immune system, both cytotoxic T cells that kill infected body cells, and helper T cells that direct the activities of other immune cells.

Cells that can move antigens to their surface and display these antigens to other immune cells are called **antigen-presenting cells (APCs)**. **The principal cells involved in antigen presentation to other immune cells are dendritic cells.** Figure 7.50 shows a dendritic cell interacting with a lymphocyte. Macrophages can also engage in antigen presentation.

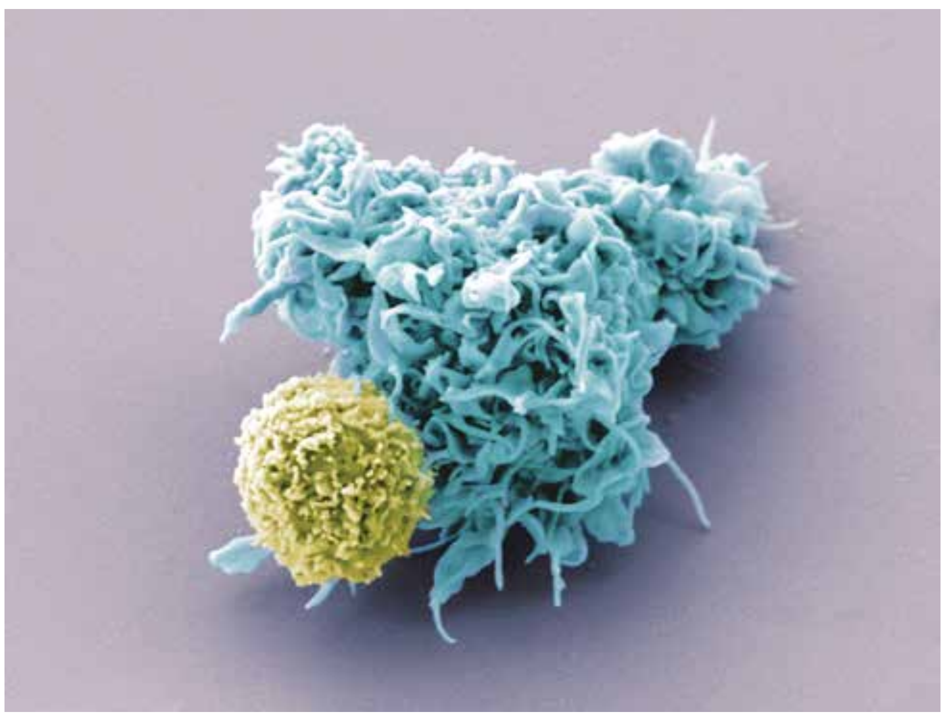


**FIGURE 7.50** Image of a dendritic cell (at right) interacting with a lymphocyte. Dendritic cells are the major cell type that presents antigens to cells of the adaptive immune system. (Image courtesy of Olivier Schwartz and Institut Pasteur.)

Let's look at the process of antigen presentation by APCs, such as dendritic cells.

## Dendritic cells: major antigen-presenting cells

- *From where do dendritic cells obtain the antigens they present?*
  Pathogens can enter the human body by ingestion, injection and by inhaling. Typically, these pathogens are met close to the entry point by phagocytic cells that engulf and eliminate them. In particular, dendritic cells are active in this role because they are located on and in epithelial tissues of the skin and the linings of the airways and the gut — all potential points of entry of pathogens to the body. Dendritic cells capture these pathogens from the external environment, and they are the source of the antigens that are presented by dendritic cells.
- *How do dendritic cells obtain these antigens?*
  Dendritic cells engulf pathogens by phagocytosis and degrade them. However, **dendritic cells retain, in their cytoplasm, some peptide fragments from degraded pathogen proteins**. These peptide fragments are the antigens that will be presented.
- *How are the antigens presented?*
  Dendritic cells link these foreign antigens to MHC class II molecules in their cytoplasm and then transfer them to their cell surfaces, putting them 'on display'.
- *Where does antigen presentation take place?*
  Dendritic cells transport these antigens to the nearest lymph nodes and present them to naïve helper T cells that carry receptors that recognise and then bind to the displayed antigens (see figure 7.51).

Figure 7.51 shows how a foreign antigen from a degraded microbe is processed by a dendritic cell for presentation to a naïve helper T cell. Note that the foreign antigen is linked to an MHC class II molecule on the surface of the dendritic cell. By presenting antigens to T cells, APCs such as dendritic cells initiate specific immune responses.

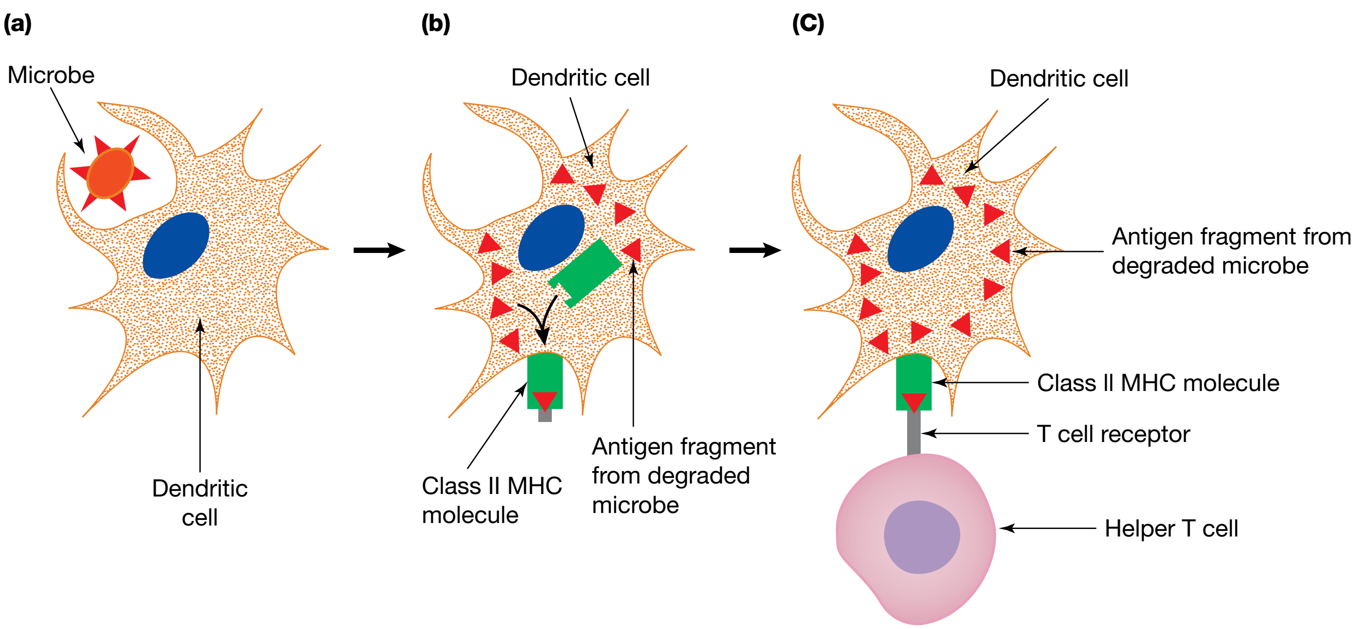



**FIGURE 7.51** **(a)** A dendritic cell engulfs a microbe by phagocytosis — this microbe will be degraded. **(b)** Some peptide fragments from degraded microbial proteins are retained by the dendritic cell. These foreign antigens are linked to MHC II molecules and shifted to the surface of the dendritic cell where they are 'on display'. **(c)** In the lymph nodes, dendritic cells display the foreign antigen on their surfaces to naïve helper T cells. Any helper T cell with a matching receptor can bind to the antigen. The antigen T cell receptor binding activates helper T cells to undergo cycles of cell division and differentiate into effector helper T cells.

# Helper T cells

Helper T cells are regarded as the most important cells in adaptive immunity because they are involved in all adaptive immune responses. The importance of helper T cells in defence against infectious microbes can be seen in people with untreated AIDS. The helper T cell population of such people is progressively lost, and they become susceptible to infections by microbes that are normally harmless (see chapter 8, page 360).

As shown in figure 7.51 immediately above, antigens are presented by dendritic cells to naïve helper T cells. Binding of the antigen to T cell receptors starts a cascade of intracellular signalling that activates the T cells.

Activated helper T cells proliferate, and the effector helper T cells stimulate immune responses in the secondary lymphoid organs where adaptive immune cells are concentrated.

Activated effector helper T cells secrete various cytokines, including interleukin 2, that:

- helps activate cytotoxic T cells to seek and destroy cells infected by pathogens carrying that antigen
- helps activate B cells into becoming antibody-producing plasma cells that will produce and release antibodies to defend against pathogens carrying that antigen
- helps activate macrophages to remove antibody-coated pathogens by phagocytosis.

It may be seen that effector helper T cells simulate the activities of both the cell-mediated arm of adaptive immunity and the humoral arm of adaptive immunity.

Figure 7.52 is a diagram summarising the central role of effector helper T cells in adaptive immunity. Note that the dendritic cell initiates helper T cell activity through its antigen presentation role.

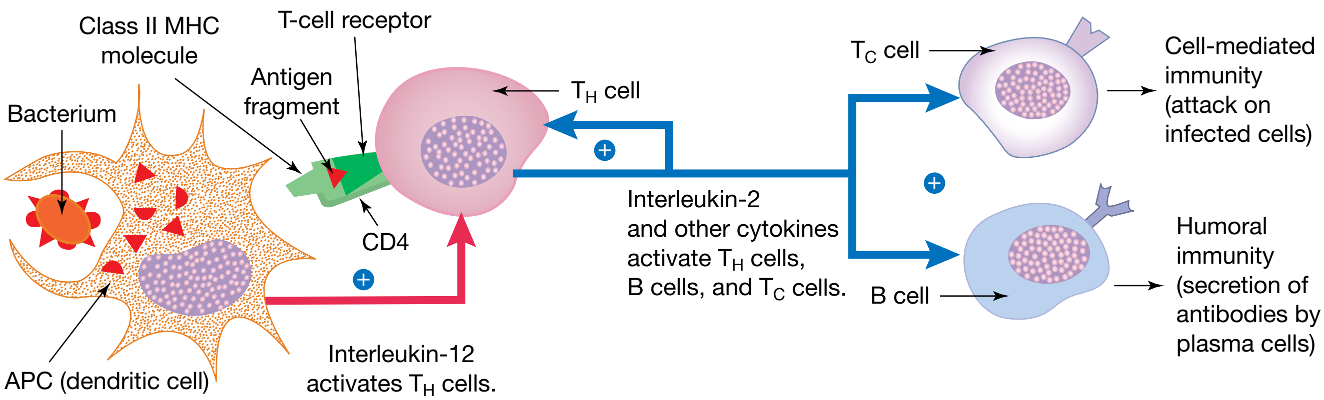



**FIGURE 7.52** Diagram showing the central role of helper T cells in adaptive immunity. Note that cytokines secreted by effector helper T cells ($T_H$ cells) activate both cytotoxic T cells ($T_C$ cells) that kill infected body cells and B cells that are involved in producing antibodies that target extracellular pathogens and toxins.

We have explored the remarkable immune defences of the human body, both the innate immune system and the adaptive immune system. The following box outlines some of the strategies that have evolved in pathogens in order to avoid these defence mechanisms.

## PATHOGENS STRIKE BACK!

Some pathogens can avoid elimination by the body's innate immune system. Pathogens that can succeed in doing this are more dangerous (virulent) than other pathogens. An infection by a highly virulent pathogen is more likely to develop into disease than an infection by a less virulent pathogen. Among the various strategies seen in bacteria that protect them from attacks by the body immune system are the following:

- ***Evading recognition as 'non-self'***

Some bacterial pathogens cover their outer surfaces with molecules that are naturally produced by the human cells. The presence of these molecules means that these pathogens are mistakenly identified by innate immune cells as 'self' and so do not come under immune attack. Bacteria that conceal their 'non-self' identity include *Treponema pallidum*, the cause of syphilis, a sexually transmitted disease (see figure 7.53). These bacteria bind a glycoprotein that is found in human connective tissue to their surfaces. This coating of fibronectin conceals the foreign antigens of the bacteria so that they are identified by the body's immune cells as self.

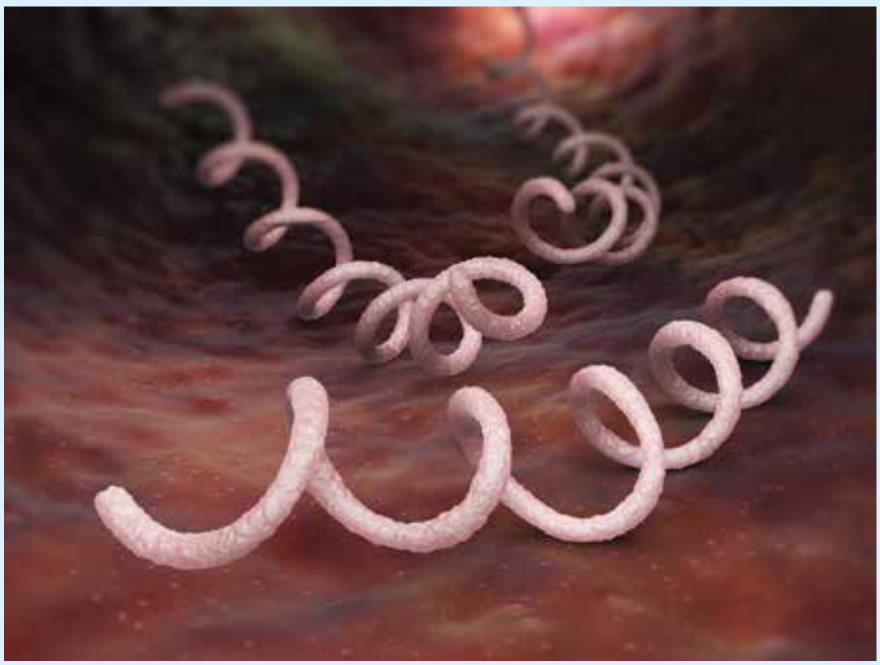


**FIGURE 7.53** Model of *Treponema pallidum*, the causative agent of syphilis. Refer back to chapter 6, page 238, and identify the shape of this bacterium. How does it avoid immune attack?

- ***Avoiding phagocytosis***

The extracellular capsule of some bacterial species (refer back to chapter 6, page 243) contains molecules that enable these bacteria to avoid engulfment by phagocytic cells (see figure 7.54). *Streptococcus pyogenes* is one such bacterial species that avoids phagocytosis because of the presence of a polysaccharide capsule surrounding each bacterial cell. Another bacterial species that has a polysaccharide capsule that confers resistance to phagocytosis is *Klebsiella pneumoniae*, the cause of one form of pneumonia.

*(continued)*

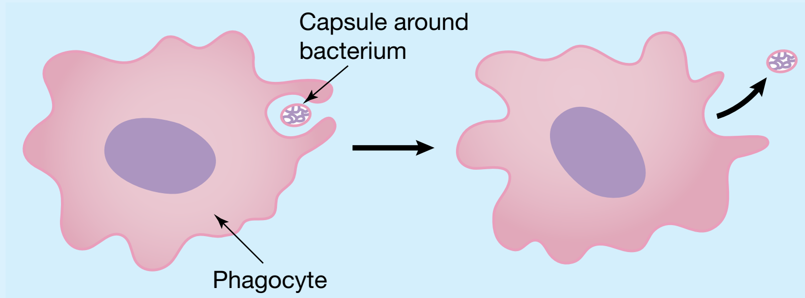



**FIGURE 7.54** Diagram showing how an encapsulated bacterial cell is not engulfed but is rejected by a phagocytic cell.

- ***Interfering with phagocytosis***

Some bacteria are engulfed by phagocytic cells and become enclosed in a phagosome. However, that is where the process of phagocytosis stops. These bacteria prevent the fusion of lysosomes with the phagosome so that phagocytosis stops. The bacteria then live and reproduce within the phagocytic cell that engulfed them, safely hidden from immune responses. Disease-causing bacteria that can successfully interfere with phagocytosis include *Legionella pneumophila* bacteria, the cause of Legionnaires' disease and *Mycobacterium tuberculosis*, the cause of tuberculosis.

- ***Killing the body's immune cells***

Some bacteria produce **exoenzymes**, which attack the immune cells of the body. Exoenzymes are enzymes that are released to the external environment of the cell that produces them. Among the exoenzymes produced by pathogenic bacteria of the genera *Streptoccoccus* and *Staphylococcus* are **leukocidins** (*leuko-* = white; *-cide* = to kill), which destroy white blood cells of the body. These white blood cells include the phagocytic cells of the body's second line of defence and the lymphocytes of the body's third line of defence. Leukocidins break down the lysosomes in the white blood cells, and the chemical contents they release cause the death of these cells and damage other cells in the region. As would be expected, *Streptoccoccus* and *Staphylococcus* bacteria cause serious and life-threatening diseases, including streptococcal toxic shock syndrome.

## KEY IDEAS

- Dendritic cells are the major cells that present antigens to naïve helper T cells.
- Antigens in the external environment that gain entry to the body are captured by dendritic cells, prepared for presentation and transported to lymph nodes.
- Naïve helper T cells are activated by binding to matching antigens on the surface of dendritic cells.
- Naïve helper T cells can only recognise presented antigens that are linked to MHC class II molecules.
- Effector helper T cells play a central role in adaptive immunity by stimulating the actions of both antibody-secreting B cells and cytotoxic T cells.

## QUICK CHECK

**29** Briefly explain the term 'antigen presentation'.

**30** Identify the following statements as true or false:

- **a** Antigen presentation by dendritic cells occurs in lymph nodes.
- **b** Dendritic cells acquire antigens by phagocytosis.
- **c** The target cells for antigen presentation are antibody-producing B cells.
- **d** The loss of helper T cells would be expected to create a life-threatening situation.
- **e** The antigens presented by dendritic cells are peptide fragments of degraded pathogen proteins.

## BIOLOGIST AT WORK

### Dr Marie-Liesse Asselin-Labat — medical researcher and laboratory head

**FIGURE 7.55** (Image courtesy of the Walter and Eliza Hall Institute.)

I grew up in France and, while I was finishing high school, I met a medical researcher who had studied pharmacy. At the time I didn't think I wanted to study medicine or science, but her work sounded very interesting so I decided to do the same.

My pharmacy degree gave me the chance to interact with hospitals and the pharmaceutical industry, which I really enjoyed. It also made me realise that basic research was key to the development of better treatments for patients, so I went on to do my PhD in Paris studying the molecular mechanisms involved in the response of leukaemic cells to blood cell hormones.

Although my husband and I were both born in France, we were eager to live and work in an English-speaking country. Around the same time, my research interest changed — I wanted to concentrate more on understanding the events responsible for the development of cancers in the context of the whole organ. So I applied to do my postdoctoral studies in the Stem Cells and Cancer Division at the Walter and Eliza Hall Institute of Medical Research, working on breast cancer.

There isn't really a 'normal' day for me. As a laboratory head, I have a lot of different responsibilities. Of course I still do experiments in the lab, but I also supervise a small team of scientists. I split my week between desk work, such as writing grant applications and reading scientific papers, and working in the lab. My daily interaction with the younger students is inspiring for me — being able to help develop their passion for science is a real pleasure.

When I first started at the institute I worked on breast cancer. I was part of the team that first discovered breast stem cells and then went on to discover a possible cellular link between female hormones and the development of breast cancer. I was honoured to receive the inaugural Lawrence Creative Prize from the Centenary Institute in 2011 for my work.

I now focus my research on lung cancer — I want to apply the theories we developed about breast cancer to other diseases. Lung cancer still has very poor outcomes for people with the disease, and I am determined to keep working on problems that affect people. It is very important to me to try to understand how diseases such as breast and lung cancer develop.

But I am more than just a scientist — I have three young boys, aged eight, six and three. While I was on maternity leave I used the time to think about what I wanted to do next. I was determined to further develop my career as a scientist. I was fortunate to have support from the institute to facilitate my return to work. It is not always easy, but in the right environment you really can balance a full-time job with having a family.

# BIOCHALLENGE

1 **Pyogenic**, or pus-forming, bacteria have polysaccharide capsules, which means that these pathogens cannot be directly recognised by the **pattern recognition receptors** on phagocytic cells, such as macrophages.

Examples of these bacteria include *Streptococcus pyogenes* (see figure 7.56) and *S. pneumoniae.*

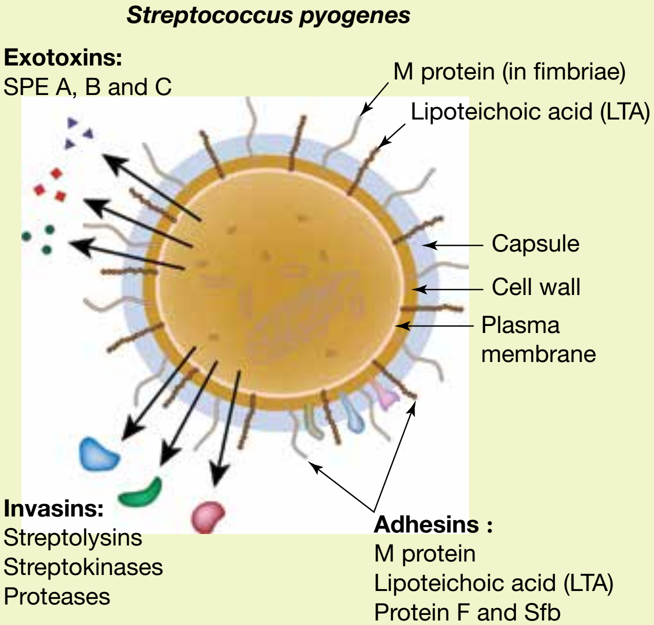



**FIGURE 7.56**

a If these pyogenic bacteria gain entry to the extracellular fluids of the body, could they be eliminated by the immediate cell-based innate immune defences? Briefly explain.
b Which component of the innate immune defence could mount a defence against these bacteria?
c If this defence is successful, how would the bacteria be eliminated?

2 Assume that the innate immune defences are overwhelmed, and the adaptive immune system must come into action against these extracellular pyogenic bacteria.
a Could these bacteria be recognised and eliminated by cytotoxic T cells? Explain.
b Briefly identify the components of the adaptive immune defences that could recognise and eliminate these bacteria.

3 An inherited condition known as Bruton's X-linked agammaglobulinemia is characterised by the absence of immunoglobulin in the blood serum. This is one of several kinds of immunodeficiency diseases and is characterised by the failure of an affected child, typically male, to produce antibodies.
a Identify the probable primary defect in Bruton's agammaglobulinemia.
b Would you expect a male infant with this condition to be susceptible to infection by pyogenic bacteria?

4 Examine figure 7.57, which shows the course of events in an infection by hepatitis A virus (HAV).

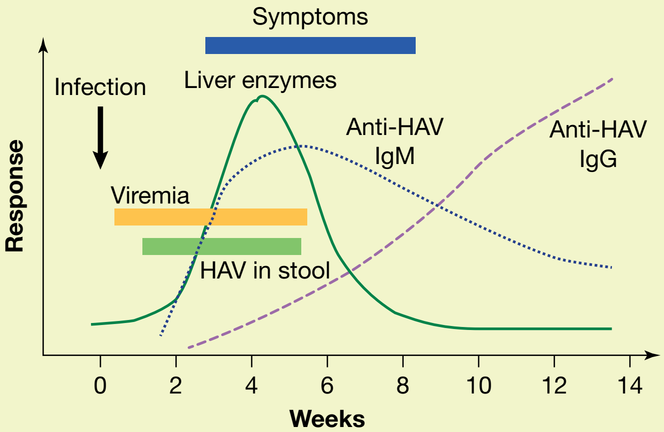



**FIGURE 7.57** Order of events during hepatitis A (HAV) infection. Note that viremia refers to the presence of the virus in the blood.

a What story is told by this graph?
b Could a person with a HAV infection, but who shows no symptoms, transmit the virus to other people?
c What is meant by the label Anti-HAV IgM?
d Which antibody is the first to appear in response to an infection?
e Which antibody appears later, but is more persistent?

Consider two food handlers who could be the source of an outbreak of HAV among clients in their café.

f Could the identification of the HAV antibody levels of each food handler indicate whether or not that person was the likely cause of the HAV outbreak that occurred four weeks previously? Explain your decision.

# Chapter review

study on

Unit 3 · AOS 2 · Topic 2

Responding to antigens

Practice questions

## Key words

acquired immunity
adaptive immunity
antibodies
antigen-binding sites
antigen-presenting cells (APCs)
bone marrow
chronic inflammation
cilia
clonal expansion
clonal selection
clonal selection theory
clone
colonoscopy
complement
cytotoxic T cells
degranulation
epithelial tissues
exoenzymes
extracellular
first line of defence
granzyme
helper T cells
immunity
immunoglobulins
immunological memory
inflammation
innate immunity
interferons (INFs)
intracellular
leukocidins
lymph nodes
lymphocytes
lysis
macrophages
membrane-attack complex (MAC)
memory cells
molecular pattern
mucous membranes
mucus
naïve
natural immunity
neutrophils
non-specific immunity
opsonisation
pattern recognition receptor (PRR)
perforin
phagocytes
phagocytic
phagocytosis
phagosome
primary lymphoid organs
sebum
second line of defence
secondary lymphoid organs
somatic mutations
spleen
third line of defence
thymus
ulcerative colitis (UC)
V(D)J recombination

## Questions

**1 Making connections →** Use at least eight chapter key words to construct a concept map. You may choose to add additional words.

**2 Interpreting graphical data →** Examine the following graph showing the time course of three events that form part of one particular defensive response of the immune system.

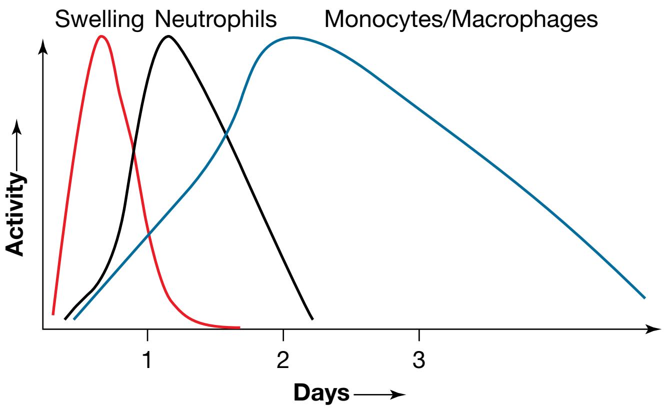



**FIGURE 7.58**

**a** Briefly outline the various events depicted in this graph.
**b** What immune response is shown in this graph?
**c** On what evidence did you make your decision?

**3 Demonstrating understanding →** A pathogen has crossed the mucous membrane of the airways and an immune cell can only identify this pathogen as follows: 'This is a bacterium; it's not a virus.'

**a** Is this an example of a specific or a generic identification?
**b** To which line of defence does this immune cell belong — first, second or third?
**c** What is a possible identity of this immune cell?
**d** Identify a cellular response to this pathogen that might then be expected to occur.
Another immune cell recognises this pathogen as a particular species of Gram-negative bacteria, as follows: 'This is *Vibrio cholera*.'
**e** To which line of defence does this immune cell belong — first, second or third?
**f** Identify a humoral response to this pathogen that might be expected to occur.
**g** What kind of immune cell can produce this response?

**4 Developing valid biological explanations →** Suggest a possible explanation in biological terms for the following observations:

**a** Babies born with an inherited defect in the protein perforin that makes this protein non-functional are at risk of death from viral infection or tumours early in their lives.
**b** A history of recurrent virus infections in a patient suggests a defect in the immune defences carried out by T cells.

c  Photomicrographs of blood smears taken from healthy people are expected to show a majority of red blood cells.
d  A white blood cell in a photomicrograph of a blood smear taken from a healthy person is more likely to be a neutrophil than a basophil.
e  A person feeling unwell often has swollen glands.
f  A person who suffers and recovers from an infectious disease such as measles has lifelong immunity to that disease.
g  Pus may appear during the process of the healing of a skin wound.
h  At birth, a baby has some level of adaptive immunity.

5  **Demonstrating knowledge and understanding →** Read the following statement taken from a research paper:
*'One should be aware that dissecting innate and acquired host defence mechanisms is an artificial approach. In real life the two components of the host response are complementary and synergistic.'* (Source of quotation: van Krevel, R. *et al.* 'Innate Immunity to *Mycobacterium tuberculosis,' Clin. Microbiol. Rev.* April 2002 vol. 15 no. 2 294–309).
a  What key point about immunity is being made in this statement by the authors of this paper?
b  Identify an example of complementary action between cells of innate immunity and cells of adaptive immunity.

6  **Demonstrating knowledge and understanding →** Consider each of the following immune responses and identify each as part of either innate immunity or adaptive immunity or both:
a  inflammatory response to infection
b  activation of the complement cascade to opsonise bacteria
c  movement of mucus from airways by action of cilia
d  recognition and elimination of a virus-infected cell
e  presentation of antigens to naïve immune cells
f  physical barrier to infectious agents
g  secretion of cytokines.

7  **Recognising differences and similarities →** Identify one key difference and one similarity between the members of the following pairs:
a  primary lymphoid organ and secondary lymphoid organ
b  innate immunity and adaptive immunity
c  a cell-mediated immune response and a humoral immune response
d  a B cell and a T cell
e  degranulation and phagocytosis
f  immunoglobulin G (IgG) and immunoglobulin M (IgM).

8  **Interpreting graphical data →** Carefully examine figure 7.59, noting the units on each axis.

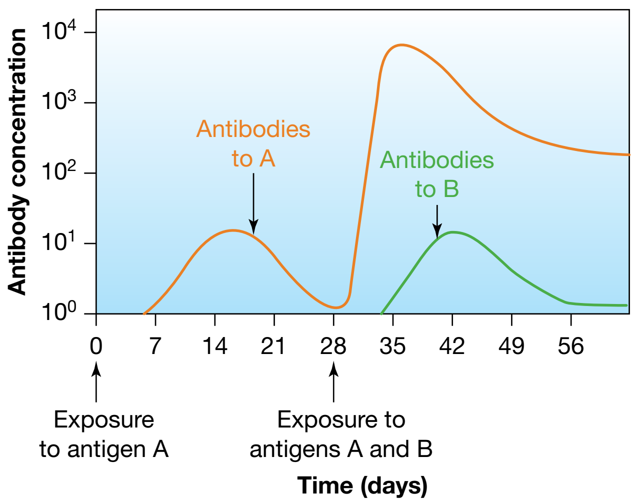



**FIGURE 7.59**

a  What story is told in this graph?
b  What is an antigen? An antibody?
c  Briefly describe what is happening in the period from day 0 to day 5 after a person is exposed to antigen A.
d  In this period, why is there a delay in the appearance of antibodies to antigen A?

At day 28, the person is exposed to two antigens — a new antigen B and the previously met antigen A.
e  What two differences are apparent in the response to antigen A on the second exposure as compared to the first exposure?
f  Explain this difference.
g  True or false: The re-exposure to the antigen results in the maximum production of nearly a thousand times more antibodies.
h  Why was the response to antigen B after day 28 not the same as the response to antigen A over the same period?

9  **Communicating understanding →** A person's innate immune system successfully defended against a viral infection and the person does not develop the viral disease.

Identify how natural killer cells contribute to this successful defence.

10  **Demonstrating knowledge and understanding →** Examine figure 7.60.
a  Complete this diagram by labelling the structures numbered 1, 2, and 3.
b  What is the major antigen-presenting cell?
c  What is an alternative name for a helper T cell: $CD4^+$ cell or $CD8^+$ cell?

**d** Compete the following sentences by choosing from the alternatives supplied:

i Binding of a receptor on a helper T cell to a presented antigen (kills/ activates/ opsonises) the T cell.

ii Another cell that can act as an antigen-presenting cell is a (mast cell/ macrophage/ neutrophil).

iii For presentation, antigens must be linked to (an MHC molecule/ a complement protein/ an antibody).

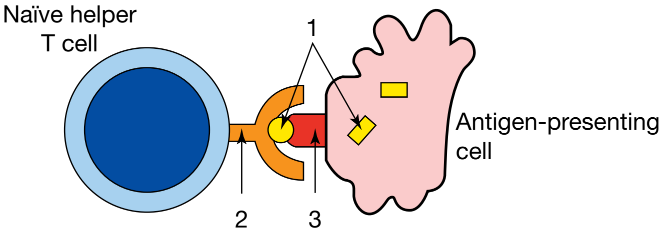



**FIGURE 7.60**

**11 Demonstrating understanding →**

**a** Rejection of a transplanted organ may be *hyperacute*, occurring in the first 24 hours after transplantation. This is caused by pre-existing antibodies in the recipient that bind to antigens on the transplant and results in the activation of the complement system and migration of neutrophils to the transplant site. (Hyperacute rejection can be prevented by pre-screening for these anti-graft antibodies.)

Is hyperacute rejection an immune response of the innate immune system or the adaptive immune system? Explain your choice.

**b** A second form of rejection is termed *acute* rejection and it usually begins after the first week of transplantation. This is caused by a mismatch of HLA antigens and involves the action of T cells. (Acute rejection is minimised by the use of immunosuppressive therapy following transplantation.)

i Is acute rejection an immune response of the innate immune system or the adaptive immune system? Explain your choice.

ii Explain why the acute rejection does not appear until after the first week of transplantation.

**12 Discussion question →** Infectious diseases are currently the world's largest killers of children and young adults, accounting for more than 13 million deaths a year — one in two deaths in developing countries. (Source: H Kuhse *et al. Bioethics: An anthology*. Wiley, 2015). The World Health Organization (WHO) has identified tuberculosis (TB) as a top infectious disease killer worldwide, with 9.6 million people falling ill with TB and 1.5 million dying from the disease in 2014. Of these deaths, over 95 per cent were in low- and middle-income countries.

A new concern is the emergence of a rare form of TB labelled extensively drug-resistant tuberculosis (XDR-TB). XDR-TB is resistant to the first-line drugs used effectively in the standard treatment of TB and is also resistant to at least one of the second-line drugs used to treat multidrug-resistant TB (MDR-TB). Awareness of this threat of XDR-TB was raised in 2006, following an outbreak in South Africa in which 52 of 53 patients with XDR-TB in a rural hospital died.

Several issues are raised by the emergence of drug-resistant infectious diseases, such as XDR-TB, which is spread via airborne particles from coughing, sneezing, shouting or singing by people with the disease. These issues raise the following questions:

**a** What infection controls are appropriate given individual human rights?

**b** Should infection control in a community at risk rely on voluntary cooperation of infected people or should restrictions on their mixing with the general population be considered?

**c** Might any such restrictions act as disincentives for infected people to seek treatment?

In 2006, a person who was diagnosed with XDR-TB refused to remain in hospital and discharged herself. Five days later, this person was traced and forcibly hospitalised.

**d** Discuss whether you consider this limitation of the patient's human rights by health authorities to be appropriate in the interests of public health.

CHAPTER

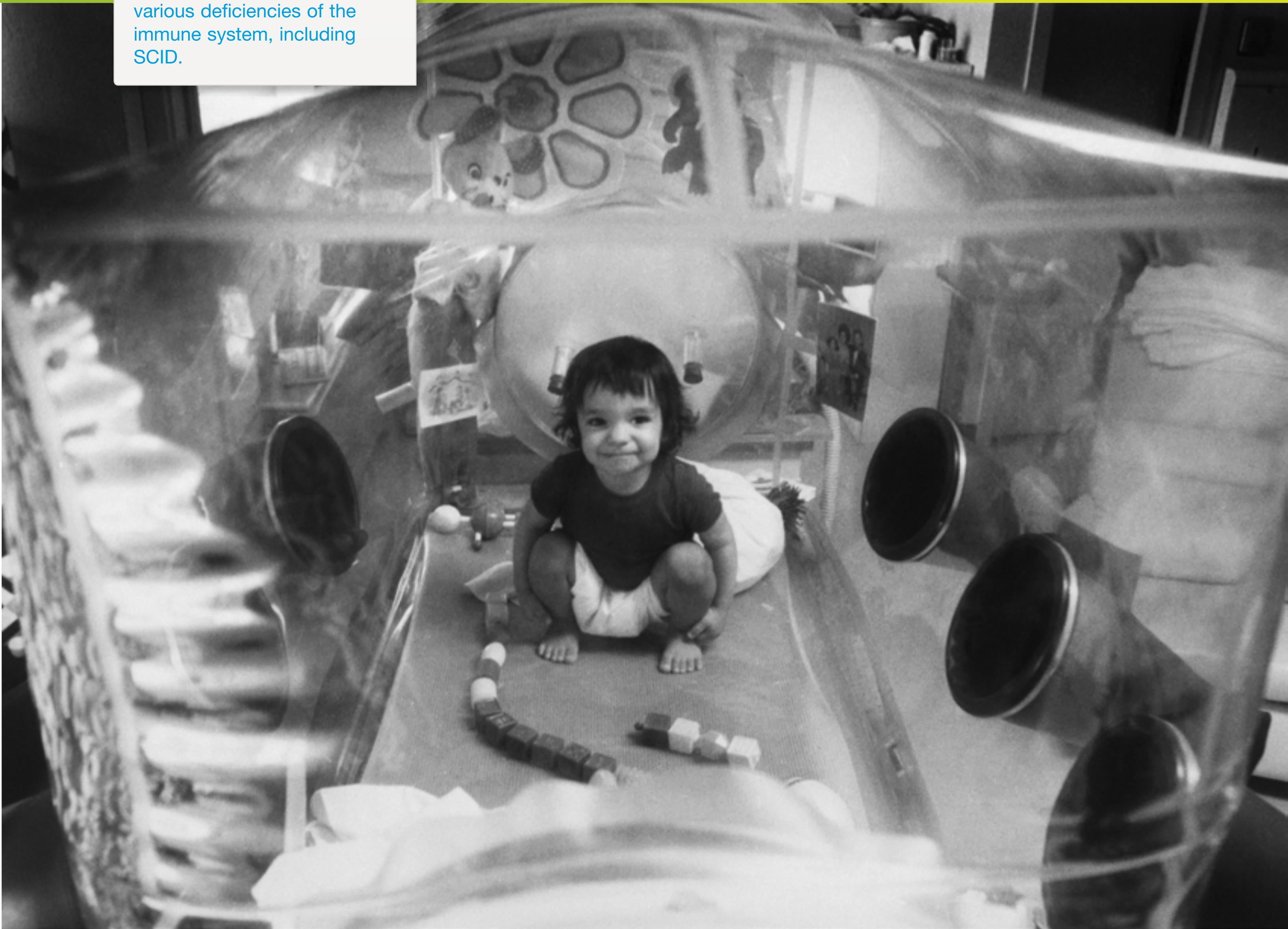
# 8 Immunity: types and deficiencies

**FIGURE 8.1** David Vetter, born in 1971 and perhaps better known as 'the boy in the bubble', had an inherited disorder of his immune system known as **severe combined immunodeficiency** (SCID). At the time of his birth, nothing could be done other than place him in a sterile isolation chamber. Any objects taken into the bubble had to be sterilised. Among topics covered in this chapter are various deficiencies of the immune system, including SCID.

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- gain knowledge of the various forms of immunity that protect against infectious diseases
- understand the link between immunisation levels and herd immunity
- recognise that diseases can arise as a result of various defects of the immune system
- give examples of diseases arising from defects of the immune system
- gain knowledge and understanding of the use of monoclonal antibodies in the treatment of cancers.



# Gardasil: a protective vaccine

Cervical cancer is responsible for the deaths of about 250 Australian women each year. One significant risk factor associated with cervical cancer is infection with particular types (strains) of the common **human papillomavirus (HPV)**. Another risk factor is smoking. HPV is very contagious and is a common sexually transmitted disease.

There are more than 100 types of HPV (see figure 8.2), and they target cells of the skin and the mucous membranes in the various areas of the body. For example:

- more than 60 types of HPV can cause warts in the skin cells of the hands and feet
- HPV types 6 and 11 cause genital warts
- HPV types 16, 18, 31, 33 and 45 cause cervical cancers. However, the high-risk HPV types 16 and 18 are responsible for about 70 per cent of cervical cancers. HPVs also cause some cancers of the anus, vagina and penis; about 95 per cent of the cases of anal cancer are caused by HPV type 16.

**ODD FACT**

HPVs are non-enveloped DNA viruses that are sexually transmitted and replicate in human epithelial cells.

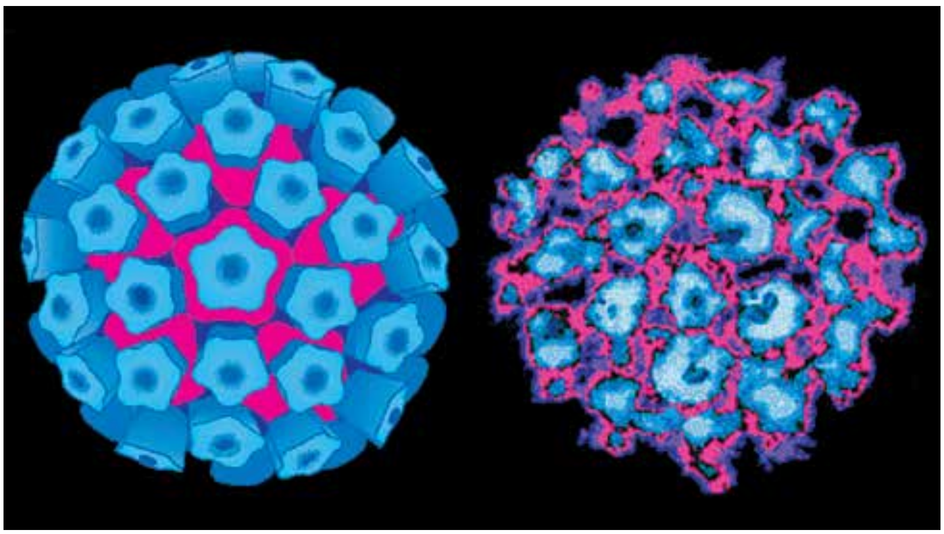


**FIGURE 8.2** Computer artwork (left) and coloured transmission electron micrograph (TEM, right) of human papillomavirus (HPV). Protein coats, or capsids, (red) enclose the genetic material, DNA. The capsids are studded with surface proteins (blue).

The cervix is the lower part, or neck, of the uterus and opens into the vagina.

The Pap test, introduced in Australia in 1991 in the National Cervical Screening Program, can detect cervical cell abnormalities that may lead to cancer. It is recommended that women have a Pap test every two years until age 70, commencing a year or two after they become sexually active. If an abnormality is detected, appropriate treatment is available and further development into cancer may be avoided. Detecting pre-cancerous cervical cells using the Pap test is of great value, but an even better situation is preventing cervical cancer. The development of a vaccine against four viruses, HIV types 6, 11, 16 and 18, was such a preventative measure.

A vaccine against human papillomavirus is now available (see figure 8.3a). It is called Gardasil and was developed by Professor Ian Frazer (see figure 8.3b), now at the Translational Research Institute (TRI) in Brisbane. Gardasil is a recombinant vaccine that originally protected against HPV types 6, 11, 16 and 18. Vaccination with Gardasil is carried out before people have been exposed to these HPV types.

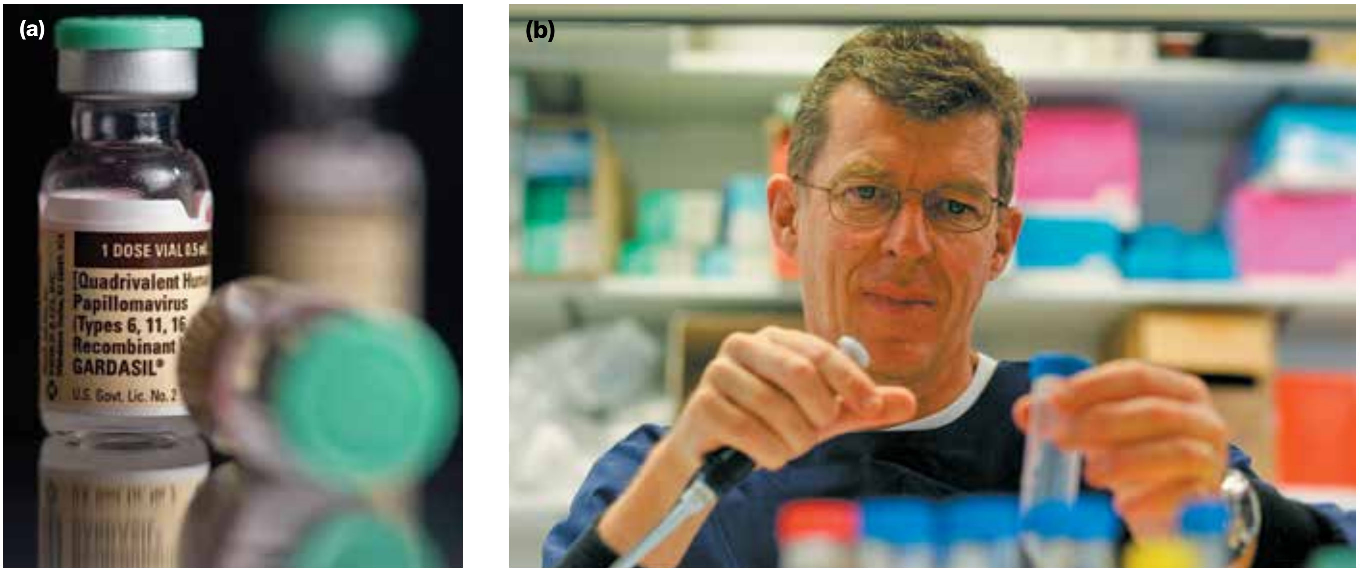



**FIGURE 8.3** **(a)** Gardasil, a vaccine for use against the human papillomavirus. **(b)** Professor Ian Frazer, recognised as Australian of the Year 2006 for his work in developing a vaccine against the papillomavirus. The vaccine will assist the prevention of cervical cancer in more than half a million women worldwide each year. (Image (b) courtesy of Patrick Hamilton, Newspix.)

Critically, in girls and young women, Gardasil helps protect against the high-risk HPV types 16 and 18, which cause 70 per cent of the cases of cervical cancers, and it reduces Pap smear abnormalities that may develop into cancers by 90 per cent. In addition, the vaccine helps protect against 70 per cent of the cases of vaginal cancers. In both females and males, the vaccine helps protect against about 80 per cent of cases of anal cancers and 90 per cent of cases of genital warts. Gardasil does not protect against all cervical cancers nor is it a substitute for cervical screening. For this reason, women should continue to have regular Pap tests.

A further advance was Gardasil 9, approved for use in late 2014, which helps protect against HPV types 6, 11, 16, 18, 31, 33, 45, 52 and 58.

A national cervical cancer immunisation campaign using Gardasil was introduced in Australia for females in April 2007 and was extended to males in February 2013. Australia is the first country to introduce such a program. Through this program, Gardasil vaccine is provided free in schools to all males and females aged 12–13 years. Since the introduction of the HPV immunisation campaign, the incidence of HPV-related infection has fallen. Worldwide, Gardasil is used in 121 countries.

In the next section, we will identify and explore the various kinds of immunity that can protect against infectious diseases.

**ODD FACT**

Many HPV infections are silent, showing no clinical symptoms. People with asymptomatic HPV infections can transmit the virus to their sexual partners.

## Specific immunity

Antibodies are our most powerful protection against infectious disease — each antibody is specifically tailored to counter one particular pathogen. When we have the specific antibodies against a particular pathogen, we are said to 'be immune to' or 'to have immunity against' the disease caused by that pathogen. The antibodies will, in fact, be against specific antigens of the pathogen concerned or antigens of its toxin. So, we can gain immunity to infection by acquiring the specific antibodies against the antigens of the pathogen responsible. Immunity based on specific antibodies can be termed specific immunity.

Specific immunity can be achieved in various ways, depending on:

1. The source of the antibodies

    ACTIVE immunity = antibodies are produced by a person's own adaptive immune system, or

    PASSIVE immunity = antibodies are acquired by a person from an external source.

2. Means of gaining immunity

    NATURAL = immunity that occurs naturally, without deliberate intervention, or

    ARTIFICIAL = immunity that is induced through a deliberate intervention.

Figure 8.4 shows a summary of the different types of immunity.

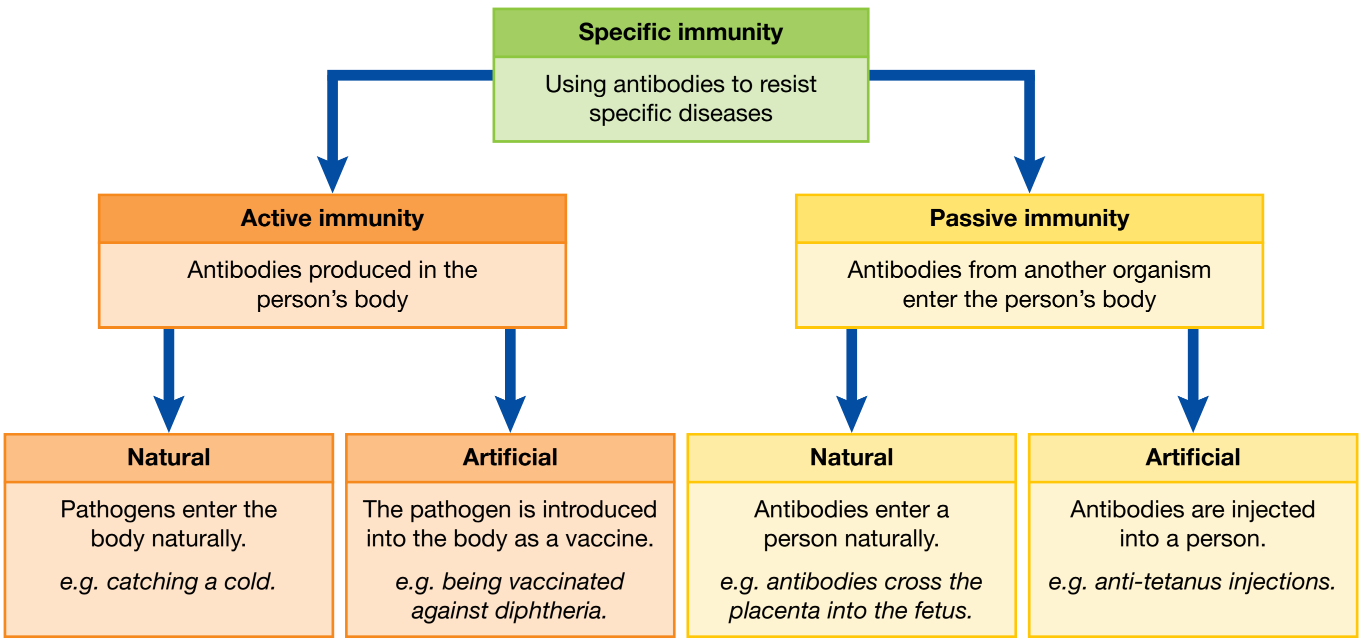



**FIGURE 8.4** Diagram showing the various types of specific immunity. Immunity can be either active or passive and can be either natural or artificial.

## Active immunity

Active immunity involves the production of antibodies by a person in response to exposure to a particular antigen. In addition, B-memory cells and T cells are produced and react quickly if another encounter occurs with the same organism.

Active immunity can be acquired in two ways: naturally or artificially (induced).

### Natural active immunity

When a person comes into contact with a particular disease-causing organism for the first time, no antibodies against the pathogen (the antigen) are present. It takes several days for the appropriate plasma cells and antibodies to form (refer back to figure 7.31 on page 309), and during that time the person begins to show symptoms of the disease.

Within several days, antibodies have formed with identical sites that bind to the pathogen antigens. As the number of antibodies increases, the infecting pathogens begin to be destroyed and the person starts to recover. If sufficient antibodies are made to destroy all the infecting pathogens, the person recovers completely.

This type of immunity is called natural active immunity. It is called *natural* because it is acquired naturally after an infection without any artificial intervention; it is called *active* because the antibodies are made by the immune system of the infected person.

In some cases, the level of infection is sufficient to activate the immune system and yet there may be no outward visible sign that the person has a disease. Such an infection is called a subclinical infection and results in antibody formation in the same way as an obvious infection. Both situations result in immunity against further infection by the same pathogen.

## Artificial active immunity

Artificial active immunity involves the deliberate and artificial introduction of a disabled pathogen or its toxin to the body. Most commonly, this is done by the injection of a particular vaccine that causes the adaptive immune system to produce antibodies to the introduced antigen (see figure 8.5).

**FIGURE 8.5** Administering a vaccine by injection.

**studyon**

| | |
|---|---|
| Unit 3 | **Vaccination** |
| AOS 2 | Summary screen and practice questions |
| Topic 3 | |
| Concept 2 | |

**Vaccines are used to activate the immune system to produce antibodies against specific disease-causing organisms without actually causing the disease.** This is possible because pathogens can be killed or weakened so that they are no longer able to cause disease. The antigens present on dead or treated pathogens in a vaccine are identical to those on the harmful pathogen, and the immune response of the body produces the same antibodies as occur with the harmful pathogen. Organisms that are treated but not killed during the preparation of a vaccine are said to be **attenuated**; their disease-causing capability has been removed, and they can no longer replicate in human cells.

Various means are used to produce vaccines with no disease-causing capability. All these various types of vaccines are able to provoke the adaptive immune system to produce antibodies specific to the antigens of the pathogens or their toxins (see table 8.1).

Vaccines are of different types; for example, some vaccines:

- contain living attenuated pathogens that do not cause disease but can still reproduce. This type of vaccine stimulates a much higher and longer lasting level of antibodies as compared to vaccines that use inactivated or killed pathogens.

studyon

Unit 3
AOS 2
Topic 3
Concept 2

**Do more**
Memory cells and vaccination

studyon

Unit 3
AOS 2
Topic 3
Concept 2

**See more**
Memory cells and vaccination

- contain killed bacteria or inactivated viruses (you can't kill viruses because they are non-living). The treatment to kill bacteria or to inactivate viruses commonly involves the use of heat or chemicals.
- contain one or more sub-units of a pathogen that act as antigen(s) stimulating antibody formation.
- contain bacterial toxins that have been treated to form harmless **toxoids**. This type of vaccine is used for those pathogens that cause disease by secreting poisonous toxins (refer back to chapter 6, page 239). Just like toxins, toxoids act as antigens but they are unable to cause disease.

**TABLE 8.1** Various types of vaccine are formed either by attenuating or inactivating pathogens or their toxins so that they are not capable of causing disease but can still result in specific antibody formation to antigens present on the pathogens or their toxins.

| Type of vaccine | Examples |
|---|---|
| live attenuated vaccines | measles, mumps, rubella (MMR)<br>yellow fever<br>polio (Sabin vaccine) |
| inactivated or killed vaccines | polio (Salk vaccine) (IPV)<br>rabies |
| inactivated toxin of bacteria (called a toxoid) | diphtheria, tetanus<br>(part of DTPa vaccine) |
| sub-units of bacteria or viruses | hepatitis B (hepB)<br>haemophilus influenza type b (Hib)<br>human papillomavirus (Gardasil) |

**ODD FACT**

Rabies was the first virus to be attenuated in a laboratory to produce a vaccine for humans.

Vaccines are sometimes given as two (or three) injections at shorter intervals followed by a booster after a longer interval (see figure 8.6). In general, killed or inactivated vaccines produce a weaker immune response compared to the response from using live attenuated vaccines, and the immunity lasts for a shorter period. As a result, killed or inactivated vaccines have to be administered more than once. The box on page 345 shows the immunisation schedule for babies in Australia.

When a vaccine is first injected into a person, the immune system shows a **primary antibody response**. A second injection of vaccine produces a **secondary antibody response**. The antibody is specific to the treated microorganism used in the vaccine so, if the person comes into contact with the live organisms at some future date, memory cells and antibodies will be ready to act and the person is immune to infection.

Immunisation = vaccination = injection of dead or attenuated pathogens

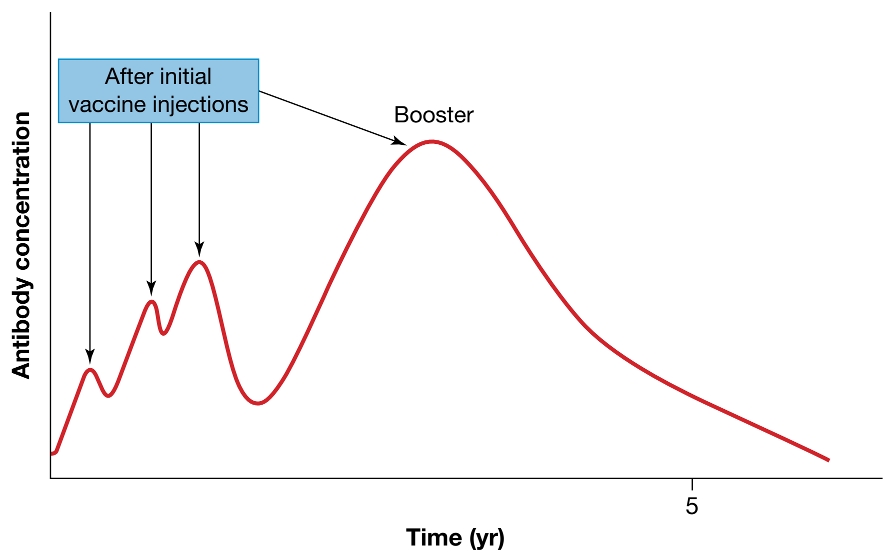



**FIGURE 8.6** Diagram showing the antibody levels in response to administration of repeated doses of a vaccine (in this case Hib vaccine). Note how the concentration of antibodies gradually increases in response to the first three doses of vaccine, followed by a major increase in response to the booster injection. Note also that the antibody level declines over time.

**FIGURE 8.7** Prior to the 1980s, Australians travelling abroad had to carry vaccination booklets to prove they had been immunised against smallpox and other infectious diseases including cholera. (Image courtesy of Marjory Martin.)

Refer back to figure 8.6 on the bottom of previous page. Note that the antibody concentration developed in artificial active immunity gradually declines. This is why a person who has been immunised against some diseases, such as tetanus or diphtheria, needs to have a booster shot about every 10 years.

## Mass vaccinations have eliminated diseases

1. **Eradication of smallpox:** Before the 1980s, smallpox, a disease caused by the variola virus, was widespread in many countries. Australians travelling overseas were required to carry a vaccination booklet (see figure 8.7) showing that they had been vaccinated against various diseases, including smallpox.

Smallpox is a devastating disease (see figure 8.8). The World Health Organization (WHO) began an extensive vaccination program against smallpox in 1959 and, by December 1979, the WHO announced that the disease had been eradicated worldwide. Because of the widespread vaccination against smallpox, the infecting virus could no longer find appropriate hosts who were not immune. Since 1977, no naturally transmitted cases of smallpox have been recorded anywhere in the world.

**FIGURE 8.8** Person suffering from smallpox, the disease caused by the variola virus.

Two laboratories in the world hold stocks of smallpox virus: the US Centers for Disease Control and Prevention, and the Russian State Research Center of Virology and Biotechnology. Some groups argue that these smallpox stocks should be destroyed in order to prevent any possibility of their accidental release. Other groups argue that scientific opportunities will be lost if all smallpox stocks are destroyed and that they are needed for research into the development of antiviral drugs. At the assembly meeting of the World Health Organization in Geneva in 2014, consensus could not be reached and, as at the end of 2015, a decision about the fate of these smallpox stocks has not been reached.

### ODD FACT

People travelling to some parts of South America or Africa should be vaccinated against yellow fever, a viral infection that is transmitted via the bite of an infected *Aedes aegypti* mosquito.

2. **On the road to the eradication of poliomyelitis (polio):** Poliovirus is spread by person-to-person contact through the exchange of nasal or oral secretions, or by contact with faecal-contaminated material. The virus enters the mouth and replicates in the gut. In most cases, polio infection results in a mild illness, but in a very small number of infected people, the poliovirus migrates from the gut, travels through the bloodstream and attacks nerve cells, resulting in paralysis (refer back to figure 5.24 on page 178).

Polio has been eradicated from Australia and almost all other countries through vaccination programs that have given people immunity to polio with the result that the virus cannot find hosts for multiplication. However, pockets of polio infection remain in rural regions of Pakistan and Afghanistan.

The vaccination programs that are leading towards the worldwide elimination of polio are based on two different vaccines: the Salk vaccine and the Sabin vaccine (see table 8.2 below). In Australia, the oral Sabin vaccine was used from 1966 until it was replaced in November 2005 by the Salk vaccine. In developing countries, the Sabin vaccine is used. Can you suggest why?

**ODD FACT**

The last case of locally transmitted poliovirus in Australia was that of a 22-year-old man in Victoria in 1986. The last confirmed case of polio was that of a 22-year-old student who returned from Pakistan in 2007 with the disease.

**TABLE 8.2** Comparison of the Salk and Sabin vaccines against poliovirus. Point 5 for the Salk vaccine is not a problem in Australia since no wild virulent polioviruses exist here. However, this point is a problem in areas where poliovirus is still endemic.

| Salk vaccine | Sabin vaccine |
|---|---|
| 1. chemically inactivated virus developed by Jonas Salk | 'live' attenuated virus developed by Albert Sabin |
| 2. more expensive to produce | cheaper to produce |
| 3. first used widely in 1955 | first available in 1962 |
| 4. administered by intra-muscular injection — requires trained staff and sterile equipment | administered by mouth (orally) — very easy to deliver |
| 5. induces antibody production in blood but not in the gut mucosa; if virulent viruses are ingested later they can replicate in the gut and can be a source of infection for unvaccinated people | induces antibody production in both the gut mucosa and the blood; if virulent polioviruses are ingested later, they will be eliminated by the mucosal antibodies |
| 6. cannot mutate to a virulent form | can mutate and revert to a virulent form |

## Vaccines have reduced the regional incidence of many infectious diseases

The incidence of a number of other diseases — the bacterial diseases diphtheria (caused by *Corynebacterium diphtheriae*), tetanus (*Clostridium tetani*) and pertussis (also known as whooping cough and caused by *Bordetella pertussis*), and the viral diseases measles, mumps and rubella (also called German measles) — has also been significantly decreased by the use of vaccines, in particular, the immunisation programs for babies and children and others (see box on page 345).

If a pregnant woman is infected with the rubella virus, the development of the fetus may be affected. Defects in sight, hearing or other characteristics may occur. This is why the National Health and Medical Research Council (NHMRC) recommends vaccination against rubella for teenage girls. The immunity they acquire will remain with them in adulthood.

## Passive immunity

Antibodies produced in one person and introduced into another person can provide immunity to that person. When this occurs, the individual receiving the antibodies has passive immunity — it is called passive because the antibodies have not been made by the user.

The advantage of passive immunity is that it gives *immediate* protection to the person receiving the antibodies. However, introduced antibodies decline relatively quickly and do not provide long-lasting immunity to the receiver. Figure 8.9 shows the rapid decline in the concentration of antibodies introduced to a person.

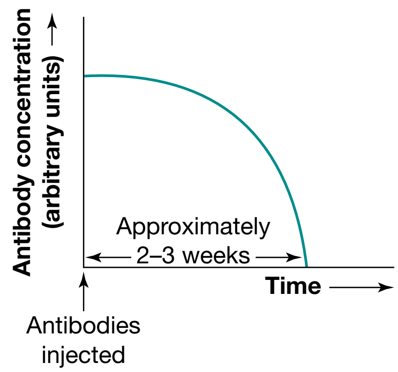



**FIGURE 8.9** Introduced antibodies give immediate protection against infection but their protective action lasts for a relatively short time. This is an example of passive immunity.

Passive immunity can be acquired in two ways: naturally or artificially (induced). Let's look at these.

### Natural passive immunity

A developing fetus receives maternal antibodies across the placenta from its mother. These antibodies (immunoglobulin G [IgG]) provide important protection for the fetus and the newborn baby because the baby's immune system does not mature fully until after birth.

A baby continues to acquire antibodies through the mother's milk. Colostrum is the thick yellowish milk the mother produces for the first two or three days after a baby's birth, and this is particularly rich in antibodies. One of the reasons breastfeeding is recommended is because of the protection against infection that the mother's antibodies give the baby (see figure 8.10a).

Figure 8.10b shows a baby's antibody profile, pre-birth and post-birth.

**(a)**

**(b)**

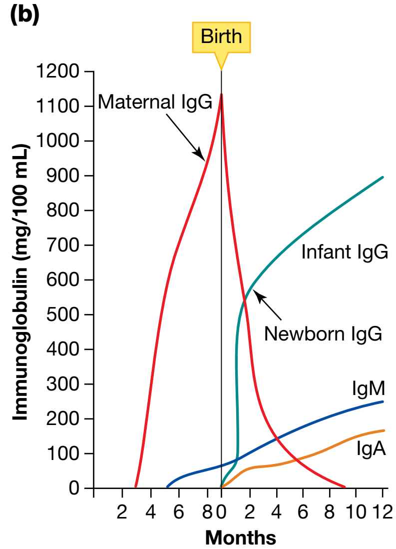



**FIGURE 8.10** **(a)** Breastfeeding by a mother provides her baby with a source of antibodies. **(b)** The immune system does not complete development until sometime after birth. A fetus obtains IgG antibodies from the mother before birth. The maternal IgG antibodies disappear from the baby's blood after birth. At birth, a baby has already started to make its own IgM antibodies. By 12 months of age, an infant produces about 60 per cent of its adult IgG level, 75 per cent of its adult IgM level and 20 per cent of its adult IgA level.

### Artificial passive immunity

Assume a member of your family develops infectious hepatitis (hepatitis A). You are at risk of being infected. If you immediately receive an injection of antibodies specific to hepatitis A, the infection may be avoided. The antibodies used in such injections are obtained from **blood** collected from voluntary donors by the Australian Red Cross Society. The antibodies, or immunoglobulins, are extracted from the blood plasma collected from people known to have had hepatitis A. The plasma contains many different antibodies including those against hepatitis A. Note, however, that the immune defence provided by 'donated' antibodies is only temporary.

Another situation in which a person receives antibodies artificially is in the treatment of snakebite. When a person receives antivenom because they have been bitten by a venomous snake, they are receiving antibodies that act specifically against the snake venom and so they are given immediate protection. The antivenom combines with antigens of the venom so that the venom is no longer free to damage body cells.

In Australia, antivenoms used in the treatment of snakebite are manufactured by Sequiris (formerly BioCSL Limited). Venom used in the production of antivenom is obtained by 'milking snakes' (see figure 8.11).

studyon

Unit 3 | AOS 2 | Topic 3 | Concept 1

**Active and passive immunity**
Summary screen and practice questions

**FIGURE 8.11** 'Milking' a snake to obtain a sample of venom that will be used in the production of antivenom. In Australia, the Australian Reptile Park is the source of venom from terrestrial snakes and funnel-web spiders that is used by Sequiris (formerly BioCSL Limited) to produce antivenoms.

Snake antivenom is produced in Percheron horses that receive injections of venom (see figure 8.12). Why aren't the horses killed? Initially, only a very small dose of venom is injected into the horse. The dose is non-lethal so that the horse survives and produces antibodies that are specific to the venom. A slightly higher dose of venom is then injected into the horse, which responds by producing higher levels of antibodies. At an appropriate time, blood samples are taken from the horse, the antibodies are extracted from the serum, are purified and are used in vaccines for people and other animals that are bitten by snakes. Different antivenoms are produced for various species of snake.

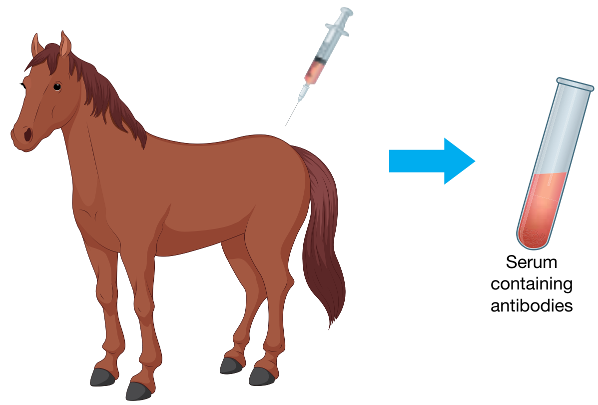



**FIGURE 8.12** Horses are repeatedly injected with small non-lethal doses of venom from a particular snake species. In response to this antigen, a horse's immune system produces antibodies against the venom. (This is an example of artificial active immunity.) These antibodies, or antivenom, are isolated and can then be injected into a person who has suffered a bite from the same species of snake. (This is an example of artificial passive immunity.)

Taking blood from horses is like taking blood from people. The horses replace their blood cells and they continue to make more antivenom (antibodies). More blood can be taken from the horses after an appropriate period. The horses are like antibody-making machines that help save human lives.

## SCHEDULE OF IMMUNISATION FOR AUSTRALIAN BABIES

Table 8.3 below shows the National Immunisation Program Schedule in Australia, effective from 20 April 2015, for babies from birth to 18 months. Rotaviruses are the leading cause of severe diarrhoea among young children. Some minor variations exist between Australian states.

**TABLE 8.3** **National Immunisation Program Schedule for babies in Australia, effective from 20 April 2015. Codes in the Vaccine column identify the disease(s) to which the vaccines give immunity; for example, the DTPa-HBV-Hib-IPV vaccine covers six infectious diseases.**

| Age | Diseases | Vaccine |
|---|---|---|
| birth | • hepatitis B | hepB |
| 2 months | • hepatitis B, diphtheria, tetanus, pertussis, *Haemophilus influenzae* type b and inactivated polio<br>• pneumococcal infections<br>• rotavirus | hepB-DTPa-Hib-IPV<br>13vPCV |
| 4 months | • hepatitis B, diphtheria, tetanus, pertussis, *Haemophilus influenzae* type b and inactivated polio<br>• pneumococcal infections<br>• rotavirus | hepB-DTPa-Hib-IPV<br>13vPCV |
| 6 months | • hepatitis B, diphtheria, tetanus, pertussis, *Haemophilus influenzae* type b and inactivated polio<br>• pneumococcal infections<br>• rotavirus | hepB-DTPa-Hib-IPV<br>13vPCV |
| 12 months | • *Haemophilus influenzae* type b and meningococcal infections<br>• measles, mumps and rubella | Hib-MenC<br>MMR |
| 18 months | • measles, mumps, rubella and varicella | MMRV |

13vPCV = 13-valent pneumococcal conjugate vaccine; 23vPPV = 23-valent pneumococcal polysaccharide vaccine
***Source:*** Australian Government Department of Health, 2016

Several other programs are delivered as part of the National Immunisation Program.

Examples include:

- school programs for children aged 10 to 15 years: HPV vaccine (all adolescents aged between 12 and 13 years), chicken pox (varicella) and diphtheria, tetanus and pertussis (dTpa)
- various 'at-risk' groups including:
  - Aboriginal and Torres Strait Islanders 15 years and older: influenza (flu) vaccine (annual) and pneumococcal (23vPPV)
  - pregnant women: influenza (flu) vaccine
  - people 65 years and older: influenza (flu) vaccine (annual).

## KEY IDEAS

- Natural immunity occurs through contact with a disease-causing agent, when the contact is not deliberate.
- Artificial immunity develops only through deliberate interventions, such as vaccinations.
- Active immunity occurs when people produce specific antibodies in response to antigens to which they are exposed.
- Passive immunity is acquired by a person through the transfer to that person of antibodies from an external source.
- Vaccination entails the exposure of a person to an attenuated, dead or inactivated pathogen that induces active antibody formation in the person receiving the vaccine.

## QUICK CHECK

1 What is the difference between:
  a active immunity and passive immunity
  b natural immunity and artificial (induced) immunity?
2 Identify the following as an example of either active or passive immunity:
  a a person receives antivenom after suffering a snake bite
  b a fetus receives maternal IgG across the placenta
  c a baby girl receives the MMR vaccine
  d a young boy is recovering from the flu.
3 Consider the four examples immediately above and identify each as being an example of either natural or artificial immunity.
4 Which gives longer lasting protection: antibodies that have been actively acquired or antibodies that have been passively acquired?
5 Refer to the immunisation schedule for Australian babies:
  a What is the first vaccine given?
  b What vaccine is administered at 18 months?

# Herd immunity

As individuals, we can be protected against an infectious disease by being immunised against it — this is an example of *direct* protection that we gain by developing antibodies against the disease in response to the injection of an antigen of the pathogen concerned.

In addition, indirect protection from infectious disease can exist, but only at the level of a population. This indirect protection is termed **herd immunity**, also known as community immunity. **Herd immunity is the indirect protection of populations from infection where that protection is created by the presence of immune individuals in the population and the protection is received by unvaccinated individuals**. The immune people in a population are those who acquired their immunity either artificially through vaccination or naturally by having recovered naturally from the disease.

studyon
Unit 3 AOS 2 Topic 3 Concept 3
**Herd immunity**
Summary screen and practice questions

However:

1. Herd immunity operates only when a high proportion of the population has immunity to a particular disease.
2. Herd immunity applies only to those infectious diseases that are contagious, that is, diseases that are spread directly from person to person, such as influenza, measles, mumps, pertussis (whooping cough) and meningococcal disease.

Herd immunity is important for vulnerable members of a community who either cannot be vaccinated or for whom vaccination is ineffective. Such individuals include newborn babies, people whose immune systems are defective because of an inherited disorder (such as SCID) or an acquired condition (such as acquired immunodeficiency syndrome (AIDS)), elderly people, and people undergoing chemotherapy for cancer.

What level of immunity is needed in a population for herd immunity to protect individuals who are not immune against contagious diseases? In a population, the higher the proportion of immune individuals to an infectious disease, such as measles, the lower the chance that an unimmunised person will come into contact with an infectious individual. In other words, the risk of infection among vulnerable individuals lessens as the proportion of immune individuals in a population increases. High levels of immunised individuals in a population have been achieved through mass immunisation programs, such as through the *Immunise Australia Program*.

If enough people are immune, an infectious disease will not get the chance to become established and spread. If the pathogen cannot find a susceptible host, the infectious disease that it causes will, over time, gradually disappear from a population.

Look at figure 8.13a. Assume that the two infected people in a community are showing the symptoms of measles. Measles is transmitted from person to person mainly by respiratory droplets, and infected individuals are contagious from about four days before until four days after the characteristic measles rash appears. What will happen in this community? Even before the two infected people showed any symptoms, they were transmitting the virus to people with whom they came in contact, and those people will, in turn, also transmit the virus, and so on. There is nothing to stop the spread of the disease. The outcome in an unimmunised community is shown in figure 8.13b. Almost everyone in the community has become infected and has developed measles.

Note that where a zero or low proportion of immunised people is present in a community, a chain of infection from person to person can easily be established and a contagious disease will spread rapidly.

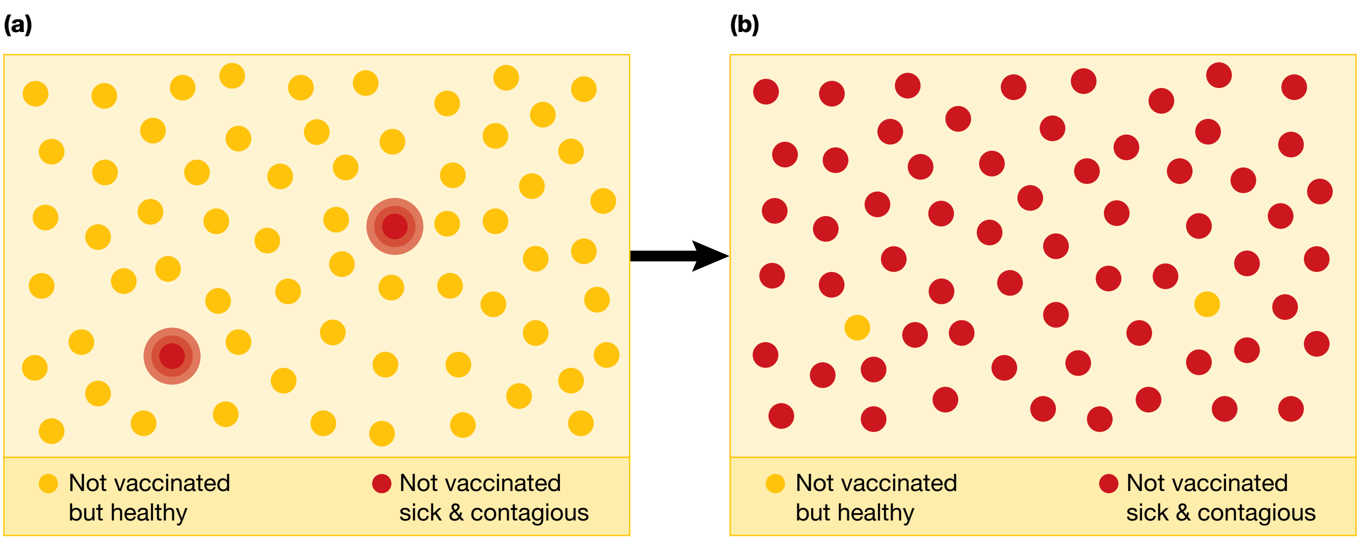



**FIGURE 8.13** **(a)** Diagrammatic representation of an unimmunised population in which two people are sick with the symptoms of a contagious disease, such as measles. What will happen in this community? **(b)** The disease spreads rapidly by person-to-person contact to almost all members of the community.

In contrast, where most people in a community are immune to a disease, the chain of infection from person to person is disrupted so that the infection cannot spread. Look at figure 8.14, which shows that, with so many people immune to an infectious disease, the disease will never get a chance to establish itself and spread throughout the community. This is herd immunity in action. Unimmunised individuals are, in the main, protected by a ring of immune members of the community.

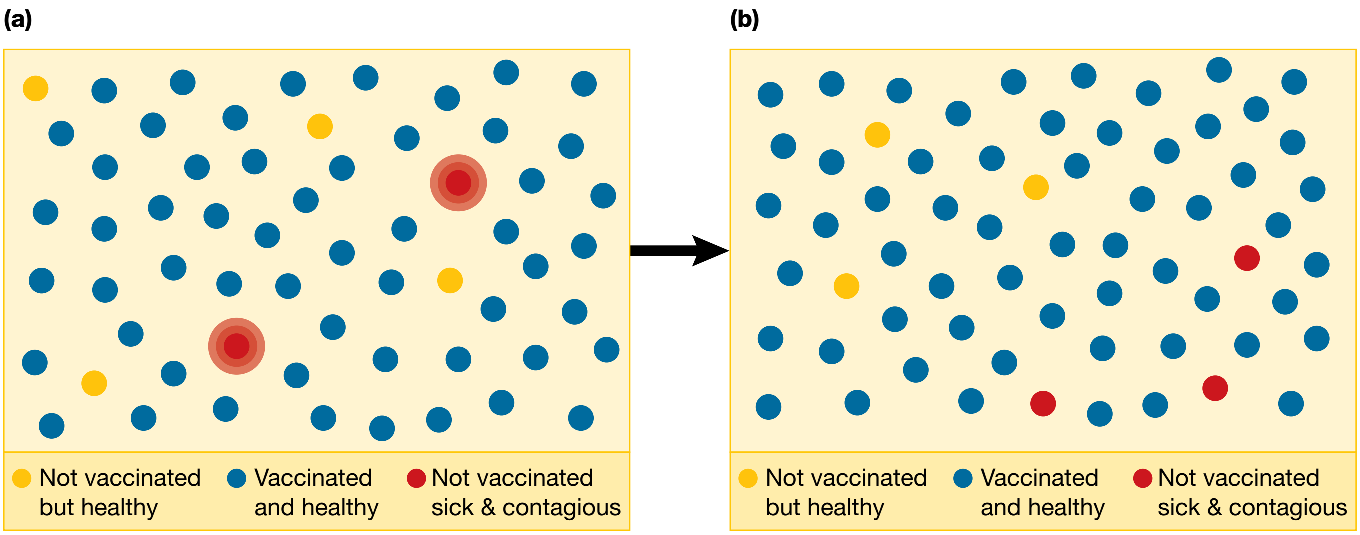



**FIGURE 8.14** **(a)** Diagrammatic representation of a community with a high proportion of immunised people as a result of a mass vaccination program. The larger red dots show two unvaccinated infected people in the community who have the potential to transmit the disease to other members of the community. **(b)** Same population a short time after the infection was introduced into the community. Note that the disease has been contained and cannot spread further. The unimmunised people in this population have, with one exception, gained the protection from the disease because of herd immunity.

**ODD FACT**

It is estimated that for herd immunity to protect people who are not vaccinated against measles, 19 out of every 20 people need to be vaccinated.

Mass vaccination programs have generated herd immunity and have successfully prevented the spread of infectious diseases. In several cases, mass vaccinations have led to the elimination of diseases, either worldwide or from a region, as for example, smallpox and polio. Opposition to vaccination poses a challenge to herd immunity by allowing preventable diseases to persist in or to reappear in communities because of a decline in vaccination rates.

Does mass vaccination help in reducing the incidence of contagious infectious diseases? One example of the power of mass vaccination and the herd immunity that flows from it may be seen in spot data on the incidence of measles in the United States. Before the routine use of the measles vaccine (starting in 1963) and, later, the MMR vaccine (starting in in 1971), there were about 500 000 cases of measles each year, with about 500 deaths. In 2004, a record low number of cases of measles was reported — just 37.

In Australia, contagious diseases have been contained and largely eliminated through mass vaccinations. Figure 8.15 shows the incidence of meningococcal disease in Australia in the period 1991 to 2014. Note the steady decline in the number of cases since the introduction of the pneumococcal vaccine in 2002. However, given the growth in international travel, the reintroduction of some diseases to Australia from offshore where these diseases are still endemic is always a possibility.

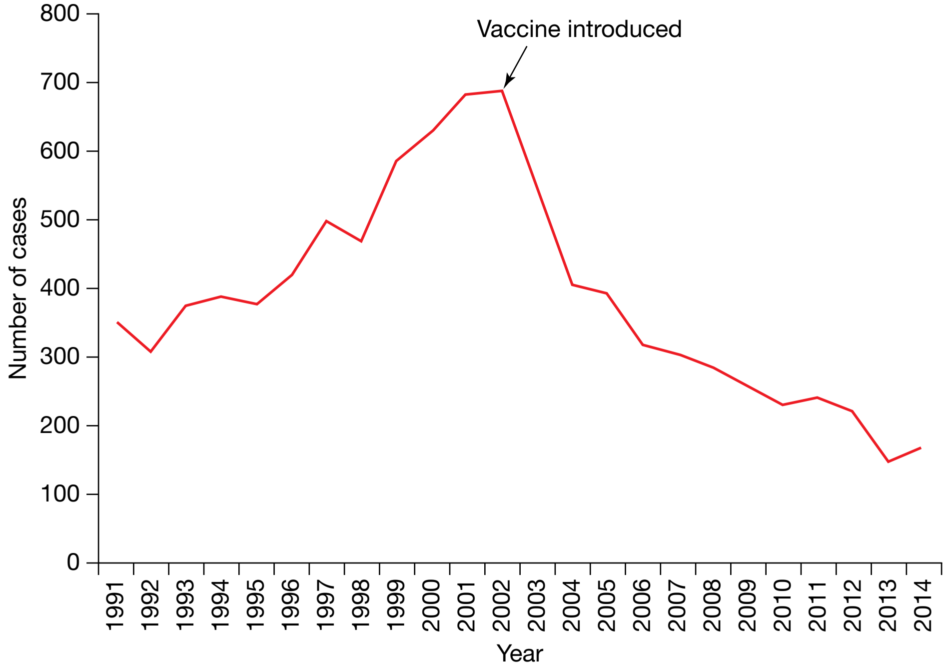



**FIGURE 8.15** Incidence of invasive meningococcal disease in Australia in the period 1991 to 2014. (*Source:* Lahra, M.M. and Enriquez, R.P. 2014, *Annual Report of the Australian Meningococcal Surveillance Program*, Department of Health.)

## KEY IDEAS

- **Herd immunity is an indirect form of protection against infectious contagious diseases that exists in populations containing a high proportion of immune people.**
- **The protection created by herd immunity applies to unimmunised members of the population.**
- **Some people within a population who depend on herd immunity are those who cannot be vaccinated because of age or those with malfunctioning immune systems and for whom vaccinations would be ineffective.**
- **Herd immunity is put at risk when immunisation rates fall because of opposition to vaccination.**

## QUICK CHECK

6 Does herd immunity exist in all populations? Explain.
7 Identify two ways in which a person can become immune to a disease.
8 Give an example of:
 a a vulnerable member of the community who cannot benefit from immunisation
 b a disease that has been eliminated worldwide by mass vaccination.
9 A person stated: 'Meningococcal disease has almost been eliminated. No need to get vaccinated against this disease.' Is this a valid statement? Explain.

# The boy in the bubble

On rare occasions, a baby is born with a severely defective immune system that cannot defend against infection. In September 1971, such a baby boy was born at a hospital in Texas in the United States. This baby, named David, was the second son born to the Vetter family.

For details of X-linked inheritance, check out *Nature of Biology Book 1 Fifth Edition*, chapter 16.

David's older brother had died from an inherited X-linked disorder termed **severe combined immunodeficiency** (SCID) when he was just seven months old. There are several forms of SCID, but the one that caused the death of David's older brother was an inherited X-linked trait. Because of this, the Vetter family knew that any other male baby born to them had a one in two chance of inheriting the defective allele and developing SCID.

The role of T cells and B cells in immunity was described in chapter 7 (see page 317).

SCID is termed 'severe' because an affected male's immune system is greatly weakened, making him highly susceptible to microbial infections. It is termed 'combined' because the disorder affects two types of white blood cells that are key components of the adaptive immune system; one cell type, known as T cells, is non-functional, and the other cell type, known as B cells, is defective. The incidence of SCID of all types in Australia is about 1 in every 65 000 live births.

Baby David Vetter had inherited the defective X-linked allele that produced the SCID disorder. In the 1970s, the only management option available was to place such a baby in a sterile isolator to prevent exposure to infectious microbes. Very soon after he was born, David was placed in a sterile plastic isolator or 'bubble' in the hospital. He lived in the sterile isolator in hospital (refer back to figure 8.1) or in an isolator at his home, apart from a few short periods when he wore a special spacesuit designed for him by the US Space Agency, NASA, that enabled him to go outside.

**eBook*plus***

**Weblink**
David Vetter and SCID disorder

In the 1980s, the option of a bone marrow transplant for David was raised. If successful, the transplant would allow him to live outside an isolator. Although his sister, Katherine, was not a perfect tissue type match, it was decided to proceed with the bone marrow transplant using Katherine as donor. The transplant was carried out but, unfortunately, it was later discovered that the transplanted tissue contained a dormant virus (Epstein-Barr virus, EBV). Once in David's system, this virus became active and he developed Burkett's lymphoma, a cancer that is associated with the Epstein-Barr virus. Twelve-year-old David died in February 1984.

**eBook*plus***

**Interactivity**
Pedigree for X-linked dominant trait
**int-6372**

Today, the standard treatment for SCID is bone marrow transplants (pre-screened for the presence of viruses). This is the only current treatment that may result in a permanent cure. In the future, gene therapy might become a preferred treatment since promising results have been obtained in a clinical trial of gene therapy in boys with X-linked SCID using a new gene delivery vector. (Reference: *N Engl J Med*. 2014 volume 371, pages 1407–17.)

Gene therapy is discussed further in chapter 13.

In the next section, we will examine some of the defects of the immune system.

# Defects of the immune system

The immune system normally operates smoothly and defends the human body against infection. However, this is not always the case. The immune system may malfunction by producing exaggerated responses or responses that do not switch off. On the other hand, the immune system may be suppressed so that its responses are ineffective in defending against infection, as was the case for David Vetter.

Exaggerated immune responses can result in:

- autoimmune diseases
- allergies.

Another example of an inappropriate response is seen in ulcerative colitis. In this condition, the inflammatory response does not switch off after the infection has been controlled. Refer back to chapter 7, pages 278 and 279.

Suppressed immune responses result in an increased susceptibility to infectious diseases and cancers, as seen in:

- acquired immunodeficiency syndrome (AIDS)
- severe combined immunodeficiency (SCID) (see figure 8.1 and the text above).

In the following sections, details of various malfunctions of the immune system will be described and examples of associated diseases will be given.

## Autoimmune diseases

A key feature of the immune system is its ability to distinguish between 'self' and 'non-self'. Normally, a person's immune system does not attack tissues whose cells carry that person's own self HLA markers.

Sometimes, however, this 'self' recognition fails, and a person's own immune system attacks and destroys their own body cells, tissues or organs, resulting in so-called **autoimmune diseases**. In autoimmune diseases, the body produces **autoantibodies**, that is antibodies that attack the body's own cells. Instead of defending the body against foreign pathogens, the immune system attacks the person's own body cells.

When the immune system produces autoantibodies, these antibodies can lock onto self antigens on the person's own body cells, resulting in an immune attack on those cells. What causes the body to mistakenly produce antibodies against its own cells is not known, but several factors including viral infections, genetic factors, hormones and drugs have been suggested as possible triggers for this malfunction.

Some autoimmune diseases affect one type of cell or one organ of the body, while other autoimmune diseases affect several body systems.

Autoimmune diseases that affect one group of cells or one organ include:

- **Multiple sclerosis (MS)** — Multiple sclerosis is considered to be an autoimmune disease that results from autoantibodies attacking the myelin that forms an insulating sheath around nerve fibres in the brain and spinal cord. (The myelin is a single cell that wraps around the axon of a nerve fibre.) When myelin is destroyed, nerve fibres are exposed, and the transmission of nerve impulses along the affected fibres is slowed or blocked (see figure 8.16). Depending on which part of the nervous system is affected, symptoms of MS may include numbness, lack of coordination, unsteady gait, slurred speech, vision impairment and fatigue.

**ODD FACT**

More than 80 different autoimmune diseases have been identified. Some autoimmune diseases are common, such as type 1 diabetes in young people, and rheumatoid arthritis in older people. It is estimated that 1 in 20 people in Australia has an autoimmune disease.

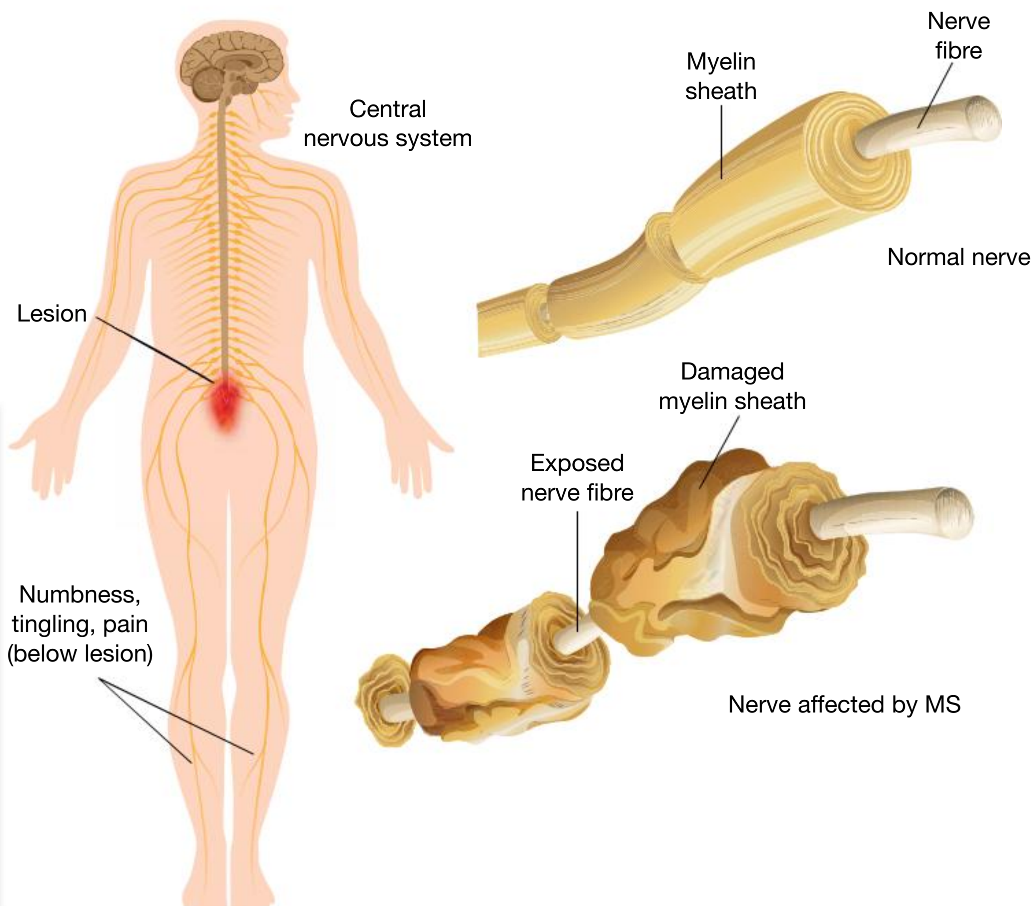



**FIGURE 8.16** Multiple sclerosis (MS) is an autoimmune disease in which autoantibodies destroy the myelin sheath of nerve fibres, producing a range of symptoms depending on the nerve fibres affected. In this figure, the nerves affected are in the lower spinal cord so that the symptoms are expressed in the body region below this damaged area.

studyon

| Unit 3 | Autoimmunity |
| --- | --- |
| AOS 2 | Summary screen and practice questions |
| Topic 3 | |
| Concept 4 | |

MS is most commonly diagnosed in people from early adulthood to middle age (20–40 years of age). More than 23 000 Australians have MS and about three quarters of these people are women. The cause of this malfunction of the immune system is unknown. It has been suggested that exposure to infectious agents such as viruses could be involved, possibly setting off an autoimmune disease directed against myelin. This could occur via a process known as molecular mimicry. This occurs when there is a similarity between proteins expressed by infectious agents, against which an immune attack is appropriately directed, and myelin proteins, against which the immune attack is mistakenly continued. Much research has attempted to find unique features in the immune systems of patients with MS that make them more likely to have ongoing disease. To date, no clear-cut difference has been found, apart from some possible differences in the way that the various proteins might be presented to the immune cells via molecules known as human leucocyte antigens (HLA). Furthermore, despite intensive investigation, no single infectious cause has been identified and it is possible that exposure to a variety of environmental factors can trigger the disease. In both the northern and southern hemispheres, a latitudinal gradient exists in the prevalence of MS, with more cases identified in colder regions. In Australia, the prevalence of MS in Tasmania is several times greater than that in northern Queensland, providing some of the strongest evidence for an environmental influence on the development of MS.

Nevertheless, MS also has an obvious hereditary component, as indicated by the observation that if one identical twin has MS there is a 25–30 per cent chance that the other twin will also get the disease, whereas the risk is only 3–5 per cent among non-identical twins. Consequently, there is active research attempting to identify genes that alter the risk of the disease.

- **Type 1 diabetes** (juvenile onset diabetes) — Type 1 diabetes is most commonly the result of an attack by a person's immune system in which autoantibodies destroy the insulin-producing (islet) cells of the pancreas. As these cells are progressively lost, the insulin-producing capability of the body is destroyed so that the regulation of blood glucose levels can no longer occur.

An autoimmune disease that affects many tissues throughout the body is:

- **Systemic lupus erythematosus (SLE)** — SLE, often just called lupus, is a multisystem autoimmune disease that affects almost all organs of the body, including heart, kidneys and brain. Common symptoms of SLE include sensitivity to sunlight, rashes (typically on the face), fever, fatigue, and decreased circulation to the fingers and toes. One research study found more than 100 different autoantibodies present in the blood of SLE patients. Figure 8.17 shows the blood sample of a person who has tested positive for SLE.

**ODD FACT**

The autoantibodies produced by people with SLE attack many targets, even the DNA of a person's own cell nuclei.


**FIGURE 8.17** A blood sample from a person with systemic lupus erythematosus (SLE) shows a positive response: the presence of mixed antibodies.

## Allergies: when harmless becomes harmful

An **allergy** is present when a person's immune system reacts abnormally to substances in the environment that are harmless to most people. This reaction is an **allergic response**, or a hypersensitive reaction. Allergic responses involve cells of both the innate and adaptive immune systems. Substances to which the person reacts are termed **allergens**. An allergen is

studyon

Unit 3
AOS 2
Topic 3
Concept 6

**Allergens and allergic responses**
Summary screen and practice questions

studyon

Unit 3
AOS 2
Topic 3
Concept 6

**See more**
Allergens and allergic responses

any antigen that causes the immune system to produce an abnormal and inappropriate overreaction when a person is exposed to it.

Many allergens are small, highly soluble proteins present on the surfaces of dry particles, such as pollen grains and dust. Other sources of allergens include food items, pharmaceutical drugs and plant products. Table 8.4 lists some common allergens. These substances are allergens only for those people who have been sensitised to these substances and have an allergic reaction when exposed to them. In contrast, for most people, these substances are harmless. Peanuts are not normally potential killers!

**TABLE 8.4** Some common allergens that produce allergic responses in sensitised people but that are harmless for everyone else.

| Category | Sources of allergens |
|---|---|
| foods | peanuts, shellfish, eggs, soy, milk and dairy products wheat, tree nuts |
| pharmaceuticals | penicillin, aspirin |
| plant and plant products | latex, pollens, sap of rhus tree |
| animal products | shed pet skin cells (dander), excreta of dust mites insect stings (bees and wasps) |
| other | fungal spores |

A sensitised person may make direct contact with a particular allergen by:

- inhaling the allergen (as for example, grass pollen or dust mite excreta)
- ingesting it (as for example, peanuts or prawns)
- touching it (as for example, latex gloves or leaves of the rhus tree)
- when it is injected (bee sting or drug).

The rhus tree (*Toxicodendron succedaneum*) (see figure 8.18) produces a toxic irritant oil. Contact with any part of the tree, and particularly the sap, can trigger an allergic response that is displayed as severe dermatitis with redness, itching and blisters, and localised swelling of the face, arms and legs.

**FIGURE 8.18** Foliage of the rhus tree (*Toxicodendron succedaneum*) showing its striking autumn colours. Some people have an allergic reaction if their skin makes direct contact with the sap of this tree.

## PERSONAL STORY

### Trent — allergic to peanuts and dairy products

Hi, I'm Trent.

When I was seven, I wrote:

Hi my name is Trent and I am allergic to dairy and peanuts. It is a little bit hard with allergies because I get lots of hayfever. I can't eat certain foods; everyone else in my class eats chocolate and I don't.

Good stuff about allergies is that my Mummy gives me treats instead of the stuff I can't have and I trust Mum and Dad to give me the right food.

I have had allergies for a long time. It is like a normal life. At school I love sport, art, maths and writing stories. Having allergies doesn't mean you can't do stuff like other people because I can still play football really, really well.

### Now I am thirteen.

Some things have changed since then and some things have stayed the same. I still have a food allergy, although now it is only peanuts I need to avoid. I am still anaphylactic (most severe form of allergic reaction) to peanuts and peanut products. This is considerably easier to manage than being allergic to dairy products. I have a lot more freedom now with lots of foods that were prohibited before. I do still have to be always on the lookout for peanuts. Doctors have said that my allergy to peanuts will probably last forever. The medical profession continues to research to find the cause, better management and a cure to allergies and in particular anaphylaxis. I do believe one day that they will understand more about the causes. I have learnt to cope though, and I still enjoy things such as swimming and cricket. In this article, I'm going to name some things that have changed, and some things that haven't.

### Things that haven't changed:

I have to carry an EpiPen with me everywhere I go. This could save my life in the event that I accidentally eat peanuts and go into anaphylactic shock, a condition that causes my throat to swell and my airways to close. An EpiPen contains adrenalin. I can inject myself now or, if I go into anaphylactic shock, one of my family or friends can use the EpiPen. Trust is a big thing with an allergy as severe as mine. I am very thankful that my friends and family all watch out for me. I often need to ask about the ingredients in food at social outings.

I do still get lots of hayfever, which can be very annoying at times but is more manageable than a dairy allergy.

I continue to enjoy a very active life including sport, maths and writing stories and I still enjoy lots of things that normal people do as well. I have always believed, whether you have an allergy or not, nothing should stop you from doing what you want to do.

### Things that have changed:

Because I can have dairy products now, I have a lot more choice in food especially, and can join in with what other people are doing. I can eat what others are eating. I can now eat things such as pizza, chocolate, ice-cream, donuts, cheese, milkshakes and commercially prepared biscuits and cakes. Easter is a lot more fun now that I can eat milk chocolate like most other people. Being able to tolerate dairy has made a huge difference to my life.

Given that I am now a teenager, the doctors are monitoring my growth and bone density to ensure my bones are strong and not affected by past years that included a lack of dairy and therefore reduced calcium.

Sometimes I start thinking, 'Why am I the one with the allergies?' I remember then that I am very lucky because others have health issues or disabilities worse than mine. My allergy to peanuts, providing I am always cautious, is quite easy to handle, and I've got to keep remembering that.

Many foods such as these chocolates contain peanuts. Other foods containing peanuts may include pastries and breads, sauces, and African, Asian and Mexican foods.

## Testing for allergens

Skin tests are commonly used to identify the allergens that cause allergic reactions in a person. Figure 8.19 outlines the procedure. A small drop of each suspected allergen is placed on the skin and, using a sterile needle, a small prick is made into the skin. (Allergy testing can also be carried out using a blood test.)

If a person has been sensitised to a particular allergen, a local allergic reaction will occur. A positive reaction shows as a localised swollen red area around the site of introduction of the allergen, with the swelling typically appearing within 20 minutes.

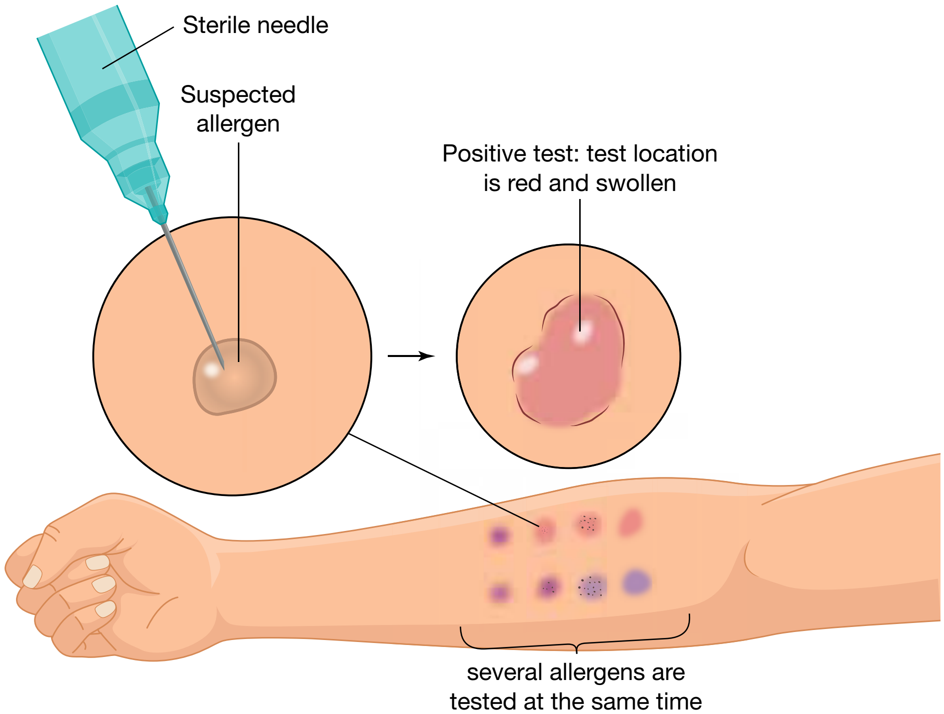



**FIGURE 8.19** Skin test for allergens. A range of presumed allergens is spotted onto a person's arm. For each spot, a sterile needle is used to make a tiny prick into the skin. The confirmation that a substance is an allergen for that person is shown in the appearance of redness and swelling in the test area.

## The allergic reaction

Key players in an allergic reaction include:

- mast cells
- antibodies: immunoglobulin E (IgE).

There are several stages in an allergic reaction, such as the hayfever caused by exposure to pollen. First, a person must be sensitised to a potential allergen, otherwise an allergic reaction will not occur.

### Sensitisation to an allergen

Potential allergens, such as an antigen on airborne pollen cells, are inhaled and become lodged on the mucous membrane lining the airways. Cells of the immune system identify this antigen as non-self, and an immune response is activated. The immune response also includes advance preparations for future exposures to the pollen allergen. This advance preparation involves the production of specific antibodies against the particular allergen. These are **immunoglobulin E (IgE) antibodies**, which are produced by plasma cells in the tissue around the site of entry of the allergen.

The IgE antibodies attach to surface receptors of **mast cells** that are located in the linings throughout the body, including the airways. **Mast cells that are coated with IgE antibodies are said to be primed** (see figure 8.20a). The presence of primed mast cells means that the person is now sensitised to that particular antigen so that the next exposure will cause an allergic response.

Once a person is sensitised, an allergic reaction will occur when that person is exposed to the allergen concerned.

This involves the following steps:

***Step 1*** **Activation of mast cells:**

- When the next exposure to the particular allergen occurs, the IgE antibodies on the primed mast cells recognise the allergen and bind to it (see figure 8.20b).
- The binding of the allergen to the IgE antibodies on the mast cells activates them and they degranulate, releasing their contents of chemical mediators, including histamine (see figure 8.20c).

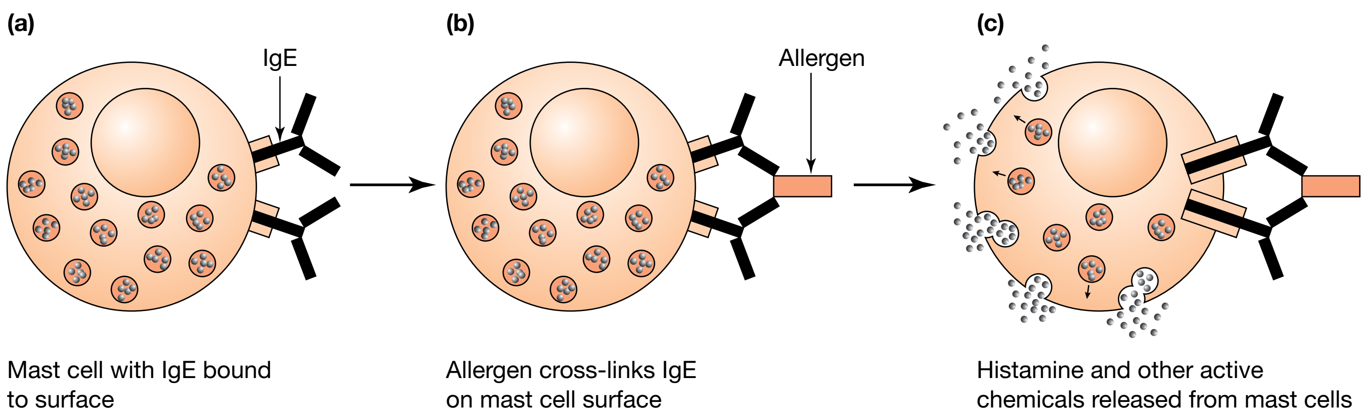



**FIGURE 8.20** **(a)** Sensitisation to an allergen involves priming mast cells by binding IgE antibodies against the specific allergen to their surfaces. **(b)** The allergic reaction begins when the specific allergen binds to and links IgE molecules on the mast cell surface. **(c)** The binding of the allergen sends a signal that results in degranulation. Degranulation releases active chemicals, in particular, histamine, from mast cells.

- The release of histamine results in an increased blood flow to the region (causing redness), increased permeability of blood vessels (causing swelling and pain), sneezing, coughing, nasal congestion, itchy and watery eyes and runny nose. This rapid allergic response to pollen is commonly called hay-fever (see figure 8.21, over page).

***Step 2*** **Migration of more immune cells producing an inflammatory response:** Mast cells produce chemical messengers that signal various circulating immune cells to migrate from the bloodstream into the affected tissue to join the immune attack on the allergen. The influx of more immune cells to the site results in the secretion of more chemicals that sustain the inflammation, cause tissue damage and recruit yet more immune cells. This last phase of an allergic response can continue for hours or even longer.

**FIGURE 8.21** A young man, previously exposed to pollen and sensitised to its antigens, suffers an allergic hayfever reaction when re-exposed to the same antigen.

Most allergic reactions, although very uncomfortable, are mild to moderate, such as redness and swelling from a mosquito bite, or the sneezing, runny nose and watery eyes of hayfever, and do not cause major problems. These allergic reactions typically affect a specific area of the body such as the eyes (red, itchy and watery), the nose and sinuses (runny) and the skin (rash and hives) (see figure 8.22).

**FIGURE 8.22** Allergic reaction appearing as hives (urticaria) on the skin of a person.

## Allergic reactions can be life-threatening

In rare cases, however, people may develop a very severe and life-threatening allergic reaction with a rapid onset that involves many parts of the body, including the cardiovascular, the respiratory and the central nervous systems. This severe and rapid reaction is termed **anaphylaxis** and it requires immediate medical attention. The self-administration of epinephrine (also called adrenalin) using an autoinjector, marketed under trade names such as EpiPen® (see figure 8.23) and Anapen®, can be used in cases of anaphylaxis. Epinephrine counteracts the effects of histamine.

**ODD FACT**

The first documented case of anaphylaxis is believed to be that of the Pharaoh Menes in 2640 BC in Egypt, who died as a result of a severe allergic reaction to a wasp sting.


**FIGURE 8.23** An autoinjector containing epinephrine, also known as adrenalin, with the trade name EpiPen®.

**KEY IDEAS**

- Autoimmune diseases and allergies result from malfunctions of the immune system.
- Autoimmune diseases occur when the body produces autoantibodies that attack the body's own tissues.
- Allergies are damaging immune responses to substances that a person has become sensitised to, but which are harmless to most people.
- Sensitisation to an allergen involves the production of immunoglobulin E (IgE) antibodies that attach to surface receptors on mast cells.

**QUICK CHECK**

**10** Name an autoimmune disease and identify the target of the autoantibodies involved.

**11** What is the difference between:

  **a** an autoantibody and a self antibody

  **b** an allergy and an allergen?

**12** What is meant by the term 'primed' in relation to a mast cell?

**13** What autoimmune disease results from the destruction, by autoantibodies, of insulin-producing cells of the pancreas?

**14** What is the current treatment for SCID?

## Acquired immunodeficiency syndrome (AIDS)

Acquired immunodeficiency syndrome (AIDS) is a disease that results from an uncontrolled and untreated infection by the human immunodeficiency virus (HIV). The HIV targets the helper T cells of the immune system, disabling the body's immune defences and hijacking these cells to produce more and more copies of the virus. The loss of helper T cells effectively cripples the immune system.

Transmission of HIV is through direct contact with body fluids, such as may occur during sexual intercourse, or through blood-to-blood contact when infected needles are shared for intravenous injections.

studyon

Unit 3 | AOS 2 | Topic 3 | Concept 5

**Immune deficiency diseases**
Summary screen and practice questions

HIV can also be passed from a HIV-infected woman to her baby during pregnancy by passage across the placenta, or at childbirth during a vaginal delivery, or by breastfeeding. However, this is a rare occurrence in Australia with an average of just one case per year over the period 2003–2013.

## The human immunodeficiency virus (HIV)

HIV is a retrovirus that carries its genetic information in the form of two fragments of single-stranded RNA. Figure 8.24 shows a drawing of a model of a HIV particle.

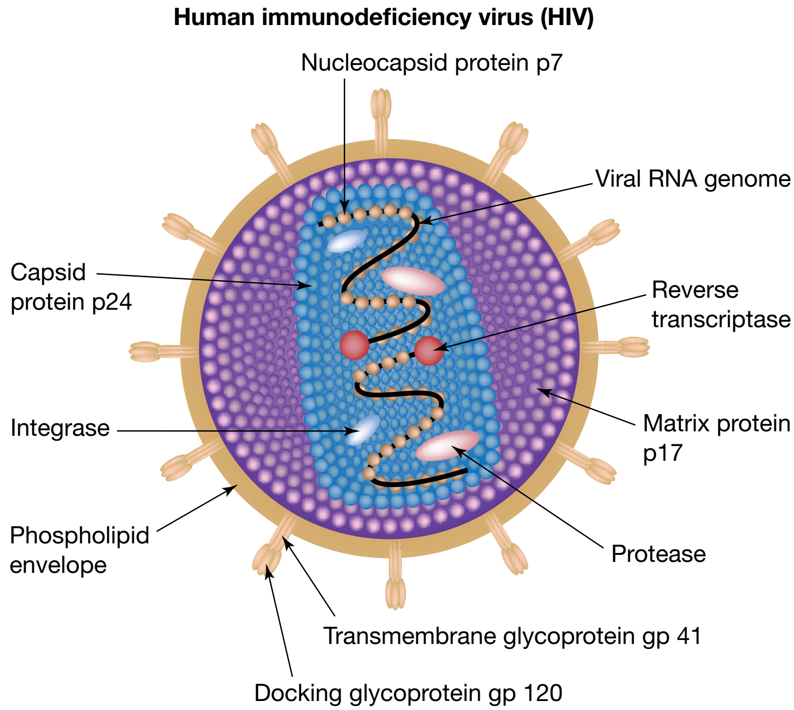
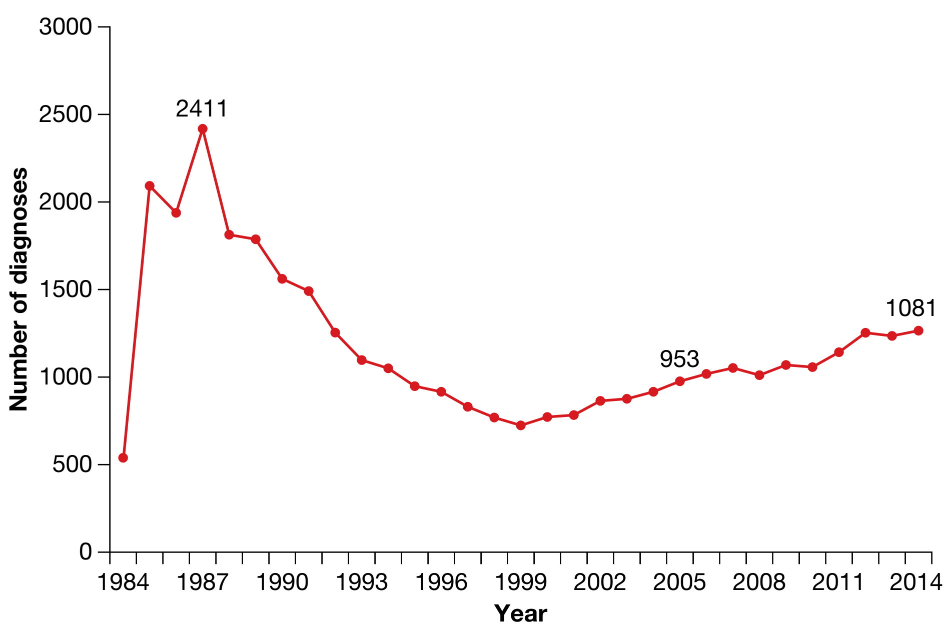



**FIGURE 8.24** Drawing of a transverse section through a model of the human immunodeficiency virus (HIV). Each HIV particle (diameter of 100 nm, or 0.1 μm) is enclosed within a lipid bilayer envelope in which viral glycoproteins are incorporated. The genetic material of the HIV consists of two molecules of RNA that are coated in protein, forming a nucleocapsid. The virus also has several enzymes that are essential for its replication, including reverse transcriptase and integrase.

The incidence of newly diagnosed cases of AIDS in Australia over the period 1983 to 2014 is shown in figure 8.25.




*Excludes people previously diagnosed overseas

**FIGURE 8.25** Graph showing newly diagnosed* HIV infections in Australia by year for the period 1983 to 2014. The majority of new cases (70%) in 2014 resulted from sexual contact between men. (Source of data: Figure 19, *HIV, viral hepatitis and sexually transmissible infections in Australia. Annual Surveillance Report 2015 HIV Supplement*. The Kirby Institute, UNSW, NSW.)

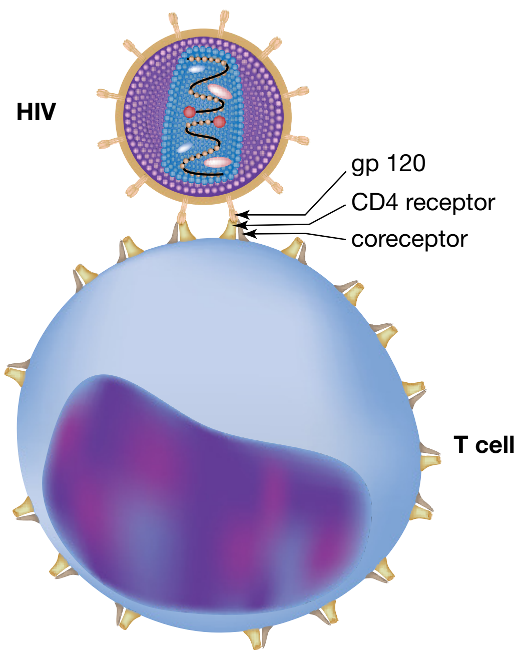



**FIGURE 8.26** HIV binds to CD4 receptors on the plasma membrane of a T-helper cell. Once the virus binds, it takes control of the T-helper cell.

Currently, there is no cure for a HIV infection, so it is a lifelong condition. However, the levels of the virus in the body (viral load) can be controlled using anti-retroviral therapy. This therapy involves a combination of drugs that block specific stages in the process of HIV multiplication in cells.

## HIV targets T-helper cells

Like all viruses, HIV can only replicate in a specific host cell. The host cells for HIV are the T-helper cells that carry CD4 receptors on their outer surface. To gain entry to these cells, the docking glycoproteins gp 120 of the virus must first bind to CD4 receptors (marker proteins) on the surface of the T-helper cells (see figure 8.26).

Once the viral RNA and proteins gain entry to a T-helper cell, the virus takes control of the host cell. Rather than performing its normal functions, the T-helper cell becomes a virus factory, producing multiple copies of the viral genetic material (RNA) and all the viral proteins, both enzymes and structural proteins, needed to make new viruses. Virus-infected T-helper cells are recognised as abnormal by cytotoxic T cells that eliminate them.

In a healthy adult, the T-helper cell count in the blood ranges from 500 to 1200 cells per microlitre (μL). In a person with a HIV infection, the T-helper cell count may drop to less than 200 cells per μL. When the CD4 cell count of a person falls below 200 cells per μL, that person is identified as having a stage III HIV infection, or AIDS. If untreated, the CD4 cell count will continue to fall.

Figure 8.27 shows the decline in the count of T-helper cells (CD4 cells) and the increase in viral load in a person with an untreated HIV infection that develops into AIDS. The loss of T-helper cells effectively disables the immune defences of such people, making them susceptible to many diseases, known as opportunistic diseases.

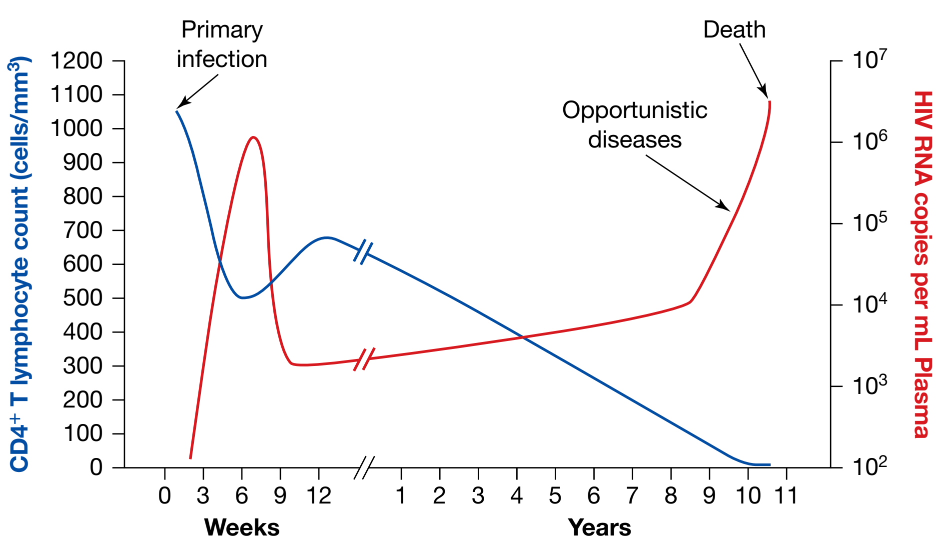



**FIGURE 8.27** Changes in viral load in an untreated HIV infection that develops into AIDS. Data shown are the numbers of copies of HIV RNA per millilitre of blood plasma (red curve) and the numbers of T-helper cells per microlitre of blood (blue curve).

The gradual depletion of the T-helper cell population means that the immune system ceases to protect against pathogens. The immune systems of people with untreated AIDS are suppressed, and these people are at risk of multiple infections. These infections are termed 'opportunistic' because the pathogens take advantage of the weakened immune system. Opportunistic infections can affect many organs, and include fungal diseases such as cryptococcal meningitis, bacterial infections such as tuberculosis, and viral diseases such as Kaposi's sarcoma.

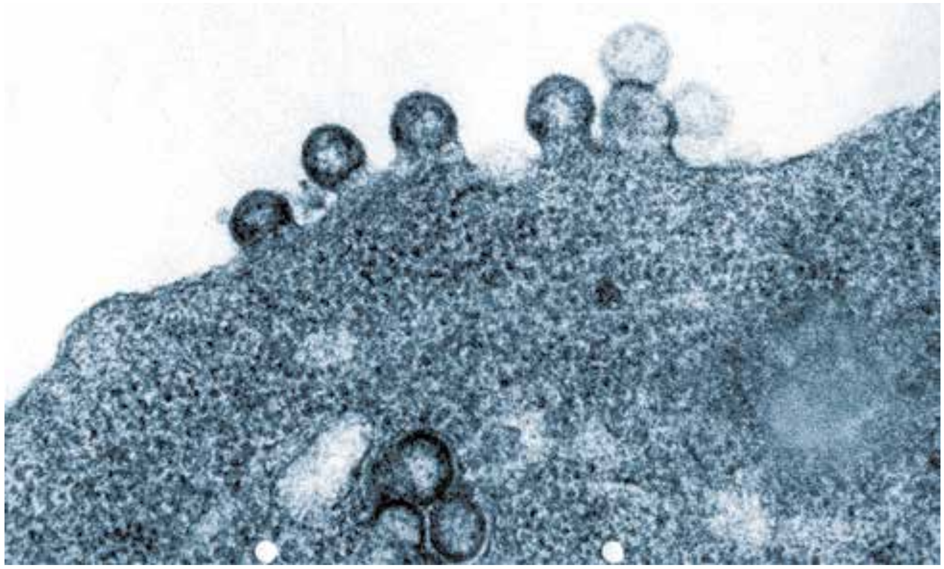


**FIGURE 8.28** Electron micrograph showing HIV particles exiting an infected host cell by a process known as budding. In this process, the virus becomes coated with a portion of the plasma membrane of the T-helper cell. (Image courtesy of Professor Johnson Mak (currently at Deakin University) when he was Head of Virology at the Macfarlane Burnet Institute.)

The identification of the HIV retrovirus and detailed knowledge of its replication process have enabled researchers to develop inhibitors that target different stages of the viral replication cycle. One such inhibitor that has been developed is a protease inhibitor. Another drug is an inhibitor of reverse transcriptase.

Viruses in some HIV-infected T-helper cells actively multiply and bud from the host cells. In other infected T-helper cells, the virus lies dormant, hiding away after inserting one or more copies of its cDNA into the genome of the host cells. These dormant viruses can resume active multiplication at any time.

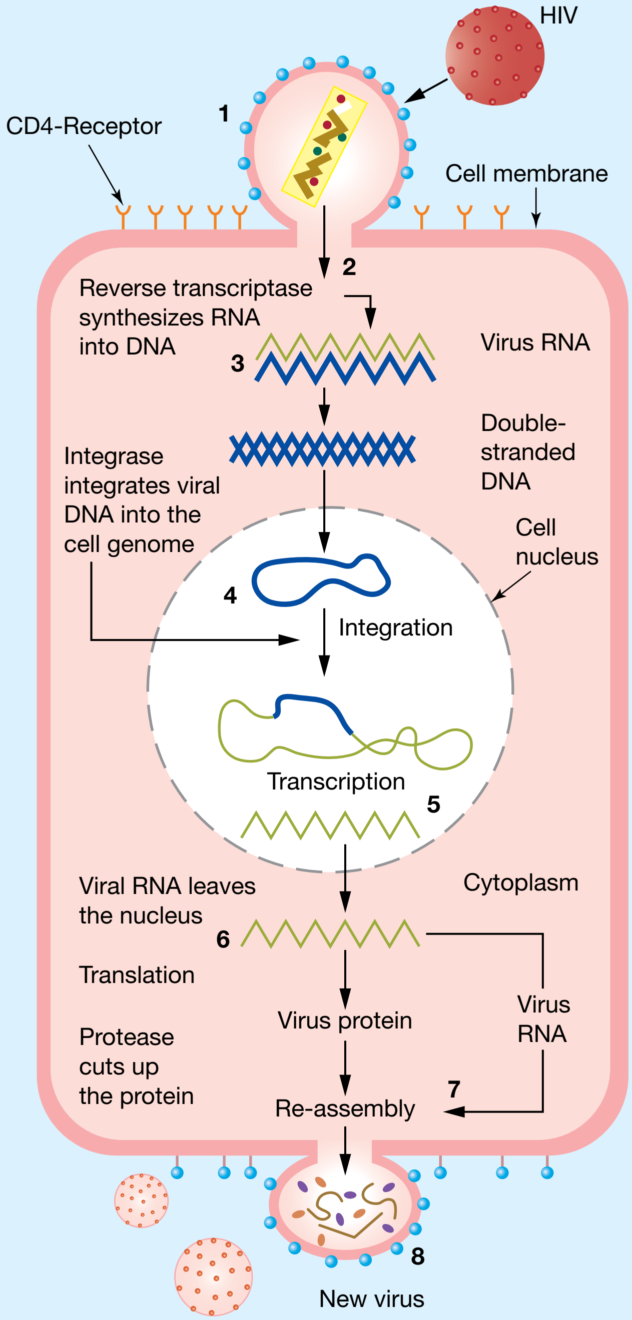
## HOW HIV MULTIPLIES IN ITS HOST CELLS

HIV targets T-helper cells and takes charge of them, using them to make copies of the viral RNA and the viral proteins. Figure 8.29 below is a diagram showing the various steps in the replication (multiplication) of HIV in its T-helper host cell. This diagram shows the replication of one viral particle. However, note that, in a person with untreated AIDS, the viral load can reach up to tens of millions of viral particles per millilitre of blood.

The steps in making more HIV particles are detailed below as follows:

1. **Fusion:** The viral particle binds to CD4 receptors on a T-helper cell and fuses with the plasma membrane of this cell.
2. **Entry:** The viral RNA and the viral reverse transcriptase and other proteins enter the cytoplasm of the host cell.

*(continued)*




3. **Making a DNA copy of the viral RNA:** The viral enzyme, reverse transcriptase, uses the viral RNA as a template to make copies of DNA that are complementary to the viral RNA. This DNA is then made double-stranded.
4. **Insertion:** The double-stranded DNA moves into the nucleus of the T-helper cell where it inserts into the DNA of the host cell. The viral enzyme, integrase, is needed for this insertion.
5. **Transcription of viral DNA insert:** The T-helper cell stops its own activities and transcribes multiple copies of viral RNA, using the viral DNA insert as the template.
6. **Translation of viral RNA:** Copies of viral RNA move into the cytoplasm where the viral RNA is translated to produce the viral proteins (the enzymes, reverse transcriptase, integrase, and protease, and the structural proteins of the capsid and the spikes on the envelope). The ribosomes of the host cell are used for this translation.
7. **Assembly:** New viral particles are assembled in the cytoplasm.
8. **Enveloping and Release:** The new viral particles (virions) are released from the T-helper cell by budding (refer back to figure 8.28). As each virus buds out, it becomes coated with some of the plasma membrane of the T-helper cell, which forms an envelope around the virus. Some viral proteins embed in this envelope.

Note: When the viral RNA is translated, it produces one long polypeptide chain that includes the individual viral proteins. A viral enzyme, the protease, cuts these individual proteins from the chain so they form mature functional proteins.

**FIGURE 8.29** A simplified diagram showing HIV replication. The numbers in this diagram identify the steps as described in the text.

## KEY IDEAS

- Acquired immunodeficiency syndrome (AIDS) results in an uncontrolled infection by the **human immunodeficiency virus (HIV)**.
- HIV targets T-helper cells, also known as CD4+ cells, and uses them as factories for making more virus particles.
- The loss of healthy T-helper cells disables the immune system so that it cannot fight infection.
- Another example of a disease that suppresses the immune system is severe combined immunodeficiency (SCID) disease, which can be inherited.

## QUICK CHECK

**15** What is the difference between the following pairs of terms:
  **a** an acquired disease and an inherited disease
  **b** HIV and AIDS?
**16** What is the genetic material of HIV?
**17** For a person who develops AIDS, what happened over time to their:
  **a** CD4+, or T-helper cell, count
  **b** viral load?
**18** Why does the immune system become suppressed in AIDS?

# Monoclonal antibodies

studyon
Unit 3 | AOS 2 | Topic 3 | Concept 7
**Monoclonal antibodies and cancer**
Summary screen and practice questions

**Monoclonal antibodies (MAbs)** are a relatively new class of drug that can be used in the treatment of cancers. MAbs are specially designed sets of antibodies with every antibody in the set binding to the same antigen (protein marker).

In 1984, Georges Köhler and César Milstein shared the Nobel Prize in Medicine or Physiology for their achievement in creating the **hybridoma technique** for producing unlimited amounts of monoclonal antibodies, which has enabled their clinical use (see the box on page 366 about making monoclonal antibodies in the laboratory).

Let's look at some of the applications, both established and under trial, of monoclonal antibodies in cancer treatment. The information given below does not represent medical advice. The source of professional advice for any individuals diagnosed with cancer must come from their local and specialist medical practitioners.

## Monoclonal antibodies in treating cancer

Researchers can design antibodies that target particular antigens on cancer cells, and they can make multiple copies of these antibodies in the laboratory. Table 8.5 shows some of the monoclonal antibodies that have been designed for use in the treatment of various cancers. Monoclonal antibodies are most commonly not used alone — they are used in combination with chemotherapy and/or radiotherapy. In some cases, the use of a particular MAb may be useful only for a particular subset of cancer patients, as, for example, the use of Herceptin for those breast cancer patients who are HER2-positive (see below).

**TABLE 8.5** Some monoclonal antibodies used in the treatment of particular cancers. The two names of each MAb are its generic name and, in brackets, its trade name. Note that the generic names all have a *–mab* suffix. The trade name is owned by the firm that markets the product.

| | Monoclonal antibody | Examples of type(s) of cancer treated with Mabs | Mode of action |
|---|---|---|---|
| 1 | **bevacizumab (Avastin®)** | various, including advanced colorectal cancers; lung, brain and breast cancers | blocks growth of new blood vessels to cancer |
| 2 | **alemtuzumab (Campath®)** | chronic lymphocytic leukaemia (CLL) | attaches to a surface protein on cancer cells and signals immune cells to eliminate cancer cells |
| 3 | **trastuzumab (Herceptin®)** | some breast and stomach cancers | blocks signals for cancer cells to divide |
| 4 | **brentuximab (Adcetris®)** | Hodgkin lymphoma | carries an attached chemotherapy drug to cancer |
| 5 | **ibritumomab (Zevalin®)** | non-Hodgkin lymphoma | carries an attached radioisotope (Y-90) to cancer |

## Modes of action of MAbs in cancer treatments

Monoclonal antibodies used in cancer treatments can be divided into naked MAbs and conjugated MAbs. Naked MAbs do not have any other molecules joined to them, while conjugated MAbs have an additional group attached. Both types of MAbs are used in the treating of cancers.

### *A. Use of naked monoclonal antibodies in treating cancers*

**1. Stopping the growth of new blood vessels to cancers:** A solid cancer, such as a malignant tumour, cannot grow beyond a size of about two millimetres in diameter without being supplied with new blood vessels. These tumours release growth factors, such as VEGF (vascular endothelial growth factor), that diffuse to nearby blood vessels and signal them to sprout new blood vessels (see figure. 8.30). The new blood vessels grow into the tumour, transporting oxygen and nutrients to the cancer cells, enabling further growth of the tumour.

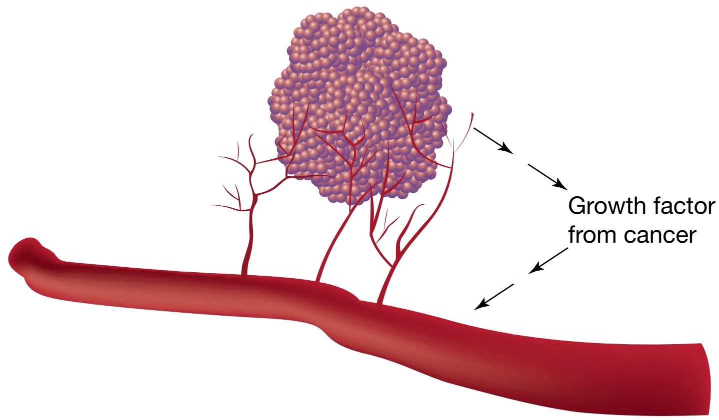



**FIGURE 8.30** Blood vessels sprouting towards a solid cancer in response to growth factor signals from cancer cells. What would happen if this signal was blocked from reaching the target endothelial cells of the blood vessel?

Tumours are solid masses of cells. Some tumours are harmless (benign), while others are malignant (cancerous).

Without additional blood supplies, a solid tumour will not increase beyond about 300 cells.

Monoclonal antibodies, such as bevacizumab (Avastin®), block the growth of new blood vessels to malignant tumours. They do this by binding to a growth factor (VEGF) released by cancer cells. This binding stops the communication between the tumour and nearby blood vessels. Because the blood vessels do not receive the signal, they do not sprout new vessels. Without an increased blood supply to provide oxygen and nutrients to its cells and remove wastes from its cells, a malignant tumour cannot continue growing.

**2. Signalling immune cells to attack cancers:** Some cancers are not very 'visible' to immune cells. Some MAbs bind to antigens on cancer cells (see figure 8.31) and act as markers that attract immune cells to attack the cancer cells.

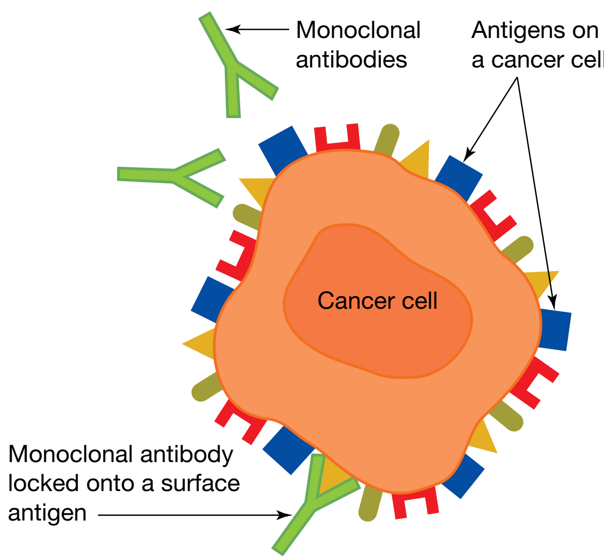



**FIGURE 8.31** Monoclonal antibodies bind to one specific protein (antigen) on a cancer cell.

**ODD FACT**

In one healthy breast cell, about 20 000 HER2 receptors are expressed. In a HER2-positive cancerous breast cell, these receptors are overexpressed, numbering about 2 million per cell.

One monoclonal antibody used for this purpose is alemtuzmab (Campath®). Campath may be used in the treatment of chronic leukaemia, a cancer of one class of white blood cells. Campath binds to the CD52 antigen on certain white blood cells. The attached antibody acts as a signal that attracts immune cells to the cancer, and the immune cells begin eliminating these white blood cells, including the leukemic cells.

**3. Blocking signals for cell division:** When growth factors bind to receptors on cancer cells, this produces a signal for the cancer cells to divide. In healthy breast tissue, surface receptors receive signals from growth factors for normal cell replacement. These receptors are known as HER2 (human epidermal growth factor receptor 2). About 20 per cent of breast cancers are classified as HER2-positive breast cancers. In these cancer cells, the receptor is overexpressed (see odd fact), resulting in increased signalling that directs the cancer cells to divide uncontrollably. The monoclonal antibody trastuzumab (Herceptin®) binds to these receptors, blocking them from receiving signals from growth factors (see figure 8.32). This slows or prevents growth of the cancer.

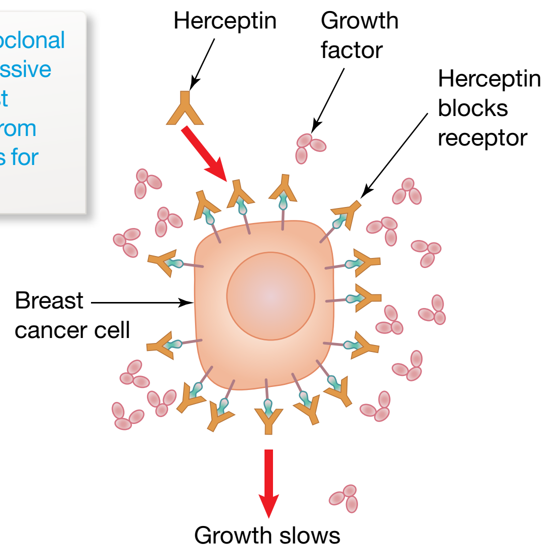



**FIGURE 8.32** Molecules of the monoclonal antibody Herceptin® bind to the excessive numbers of HER2 receptors on breast cancer cells, blocking the receptors from growth factors and preventing signals for uncontrolled cell division.

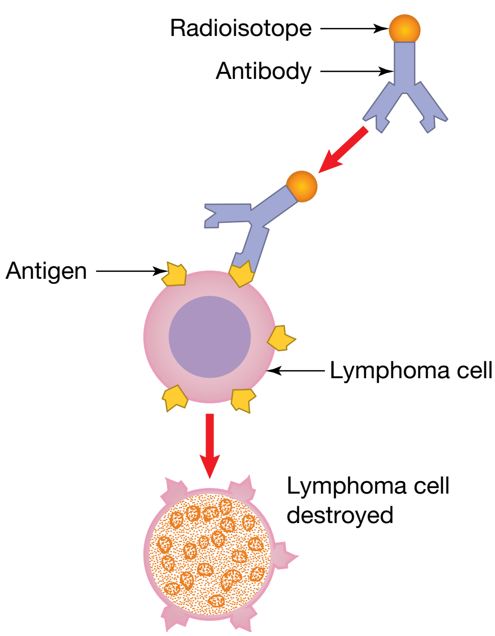



**FIGURE 8.33** A monoclonal antibody joined to a radioactive particle can bind to a cancer cell, provided it has the right target antigen on its surface. One radioactive isotope used for this purpose is yttrium-90, which is joined to the monoclonal antibody called Zevalin. Another radioactive isotope used in this mode of action includes iodine-131.

### *B. Conjugated monoclonal antibodies*

Conjugated monoclonal antibodies are MAbs joined to a second molecule, such as a chemotherapy drug or a radioisotope particle.

**1. Delivering anticancer drugs to cancer cells:** Some monoclonal antibodies, such as brentuximab (Adcetris®), can be joined to molecules of a chemotherapy drug and then deliver the cancer-killing drug directly to the target antigen on particular cancer cells. In the case of Adcetris, its target antigen is the CD30 protein on lymphoma cells, and it can be used to deliver a chemotherapy drug in some cases of Hodgkin lymphoma.

**2. Delivering radioisotopes to cancer cells:** Some monoclonal antibodies can be joined to a particle of a radioisotope and deliver it to cancer cells. Once the antibody links to its target antigen on the cancer cell, the radioisotope emits radiation that can destroy the cancer cell (see figure 8.33). One such MAb is ibritumomab (Zevalin®), which has as its target antigen the CD20 protein on the surface of B lymphocytes. Zevalin can be used in the treatment of some types of non-Hodgkin lymphoma.

## MAKING MONOCLONAL ANTIBODIES IN THE LABORATORY

Figure 8.34 shows the process for making monoclonal antibodies (MAbs) against a particular antigen, in this case, antigen X.

1. A mouse is injected with antigen X. This activates the production of its B cells, which produce antibodies against antigen X. To increase the concentration of these antibodies, repeat injections followed by a booster may be given.
2. The spleen of the mouse is removed, placed in a culture medium and its cells are separated. This produces a mixture of B cells, only some of which can form antibodies against antigen X.
3. Mouse tumour cells (myeloma cells) that can constantly divide are added to the separated B cells. Some B cells fuse with tumour cells to form new cells called hybridomas.
4. The unfused cells die, leaving only hybridoma cells.
5. **Individual hybridoma cells are cultured in a new medium — one cell per well — and allowed to divide repeatedly. This is the cloning step during which multiple identical copies of each individual hybridoma cell are produced.**

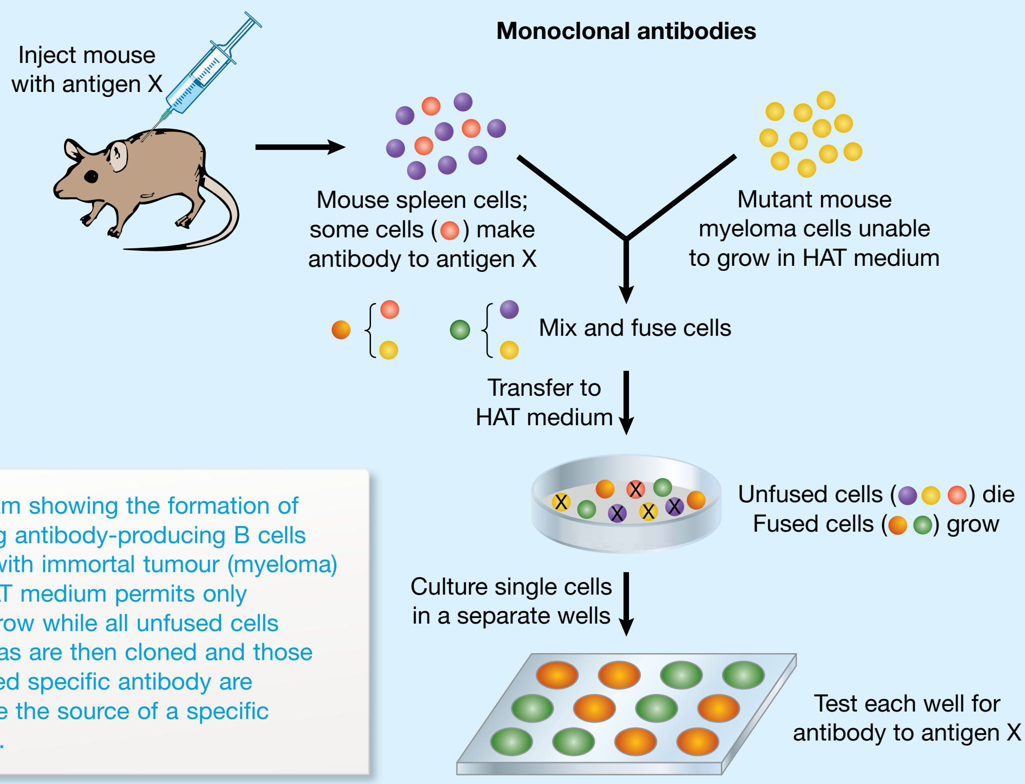



**FIGURE 8.34** Diagram showing the formation of hybridomas by fusing antibody-producing B cells from mouse spleen with immortal tumour (myeloma) cells. The special HAT medium permits only hybridoma cells to grow while all unfused cells die. Single hybridomas are then cloned and those producing the required specific antibody are selected and become the source of a specific monoclonal antibody.

After the steps shown in the figure above, the following occurs:

6. Each individual clone is screened for the presence of the required antibody so that clones of cells that produce antibodies against antigen X are identified.
7. The selected clones can be grown indefinitely in mass culture, and the required antibodies against antigen X can be harvested as required from the culture medium. These antibodies are monoclonal antibodies (MAbs) — 'monoclonal' because the antibodies come from clones of one parent cell, and, importantly, they are antibodies specific for one known antigen.

This process may look complex, but it is really neat. B cells cannot divide, but by fusing a B cell with a tumour cell that can divide indefinitely, a non-stop factory with an antibody production line is created. This was the basis for the 1984 Nobel Prize awarded to Köhler and Milstein.

## KEY IDEAS

- Monoclonal antibodies (MAbs) are produced by a single clone of a cell and consist of identical antibody molecules that bind to the same antigen.
- Monoclonal antibodies are being used in the treatment of various cancers.
- Various monoclonal antibodies block the division of cancer cells through different modes of action.
- Some monoclonal antibodies take action as ‘naked’ antibodies, while others take action as conjugate antibodies.

## QUICK CHECK

**19** Give an example of a MAb that can block the growth of new blood vessels into a solid mass of cancer cells.

**20** What is the expected outcome of preventing the growth of new blood vessels into a solid mass of cancer cells?

**21** A monoclonal antibody is joined to a radioactive isotope. What kind of MAb is this?

**22** **a** Briefly describe how Herceptin can slow or stop the growth of HER2–positive breast cancer cells.

**b** Could another MAb, such as Zevalin, be used for the same purpose? Briefly explain.

## BIOLOGISTS AT WORK

### The Walter and Eliza Hall Institute of Medical Research

In 2015, the Walter and Eliza Hall Institute (see figure 8.35) celebrated ‘100 years of discoveries for humanity’ — remembering the many major discoveries emanating from the institute that have made a tangible difference to human health.

Institute director Professor Doug Hilton (see figure 8.36) says one of the first significant discoveries at the institute, snakebite antivenoms, was made in the 1920s. ‘These antivenoms are still used, nearly 100 years later,’ he says. ‘The capacity to grow influenza in eggs and protect the community against the flu by generating vaccines every year was discovered here in the 1940s. In the 1960s, Professor Don Metcalf discovered colony stimulating factors (CSFs), which are used today to boost the immune system during chemotherapy and have benefited more than 20 million cancer patients. Discoveries about the malaria parasite in the 1990s have underpinned potential new vaccines and treatments that may help the global efforts to combat this devastating disease.’


**FIGURE 8.35** Since 1915, Australia’s oldest medical research institute, the Walter and Eliza Hall Institute, has been making research discoveries to improve our community’s health. The institute is home to more than 1000 scientists, students and support staff. (The photographs in figures 8.35 through 8.39 are courtesy of the Walter and Eliza Hall Institute.)

*(continued)*


**FIGURE 8.36** Since 2009, blood cell researcher Professor Doug Hilton has led the Walter and Eliza Hall Institute. He is the sixth director, following in the footsteps of Professor Suzanne Cory, Sir Gustav Nossal and Sir Frank Macfarlane Burnet.

'There have been discoveries like that every 5 or 10 years that I think have meant a lot to the health and wellbeing of all humanity, and will continue to make a difference in the lives of my children and my grandchildren,' he says.

The institute is home to more than 1000 scientists, students and support staff who are working to understand, prevent and treat diseases including:

- cancers of the blood, bowel, brain, breast, lung, pancreas and stomach
- chronic immune disorders such as rheumatoid arthritis, lupus, inflammatory bowel disease, coeliac disease and diabetes
- infectious diseases such as malaria, toxoplasmosis, hepatitis B, HIV and tuberculosis.

The search for new drugs at the institute is being accelerated by the use of automated technology (see figure 8.37).


**FIGURE 8.37** The search for new drugs at the institute is being accelerated by the use of automated technology. Chemical biology research assistant Ms Melanie de Silva is using a liquid handling robot to accurately and rapidly prepare thousands of minute samples for analysis. Each sample may be as small as one microlitre in volume.

People from a diverse range of backgrounds are needed for this research, so the institute employs staff from disciplines such as biochemistry, molecular biology, genetics, medicine, computer science, mathematics and chemistry. Scientists, students and technicians all play important roles in our research, which occurs in 'wet labs' — the traditional rooms where you find chemical solutions, cells and analytical equipment — and 'dry labs' — areas that look like office spaces, where computers are used to integrate and make sense of the complex data generated by biological studies.

The institute's research effort is also supported by animal technicians, photographers, administrators, accountants, engineers, information technology support staff and event organisers. If you are interested in contributing to advances in medical research, there are myriad opportunities for you.

Professor Hilton says a priority for the institute is communicating its research to the wider community. 'One of the important ways we do this is through the expertise of our two biomedical animators, Drew Berry and Etsuko Uno,' he says. 'They bring our science to life on the big screen, and their animations have been used as educational tools, and exhibited in museums and galleries around the world — including a recent addition to the institute facade, a six-storey illuminated curtain where animations are visible to passers-by.'

### Stem cells and cancer

Stem cells are often lauded for their potential uses in treating human disease. However, there is still much to understand about how stem cells are controlled and maintained, and the properties and signals that allow them to develop into any cell type in the body (see figure 8.38).

Professor Hilton says the institute's research is extending understanding of how stem cells develop and function in different organs of our body, how problems in stem cell function can lead to diseases, and how stem cells can be used to treat diseases.

**FIGURE 8.38** Dr Shalin Naik leads a research team studying how the cells of the immune system are formed from blood stem cells. Understanding how immune cells develop may one day lead to new approaches to treating infections or diseases caused by a faulty immune system, such as lupus or immunodeficiencies.

'One exciting question is how we can reprogram more mature cells to "revert" to stem cells,' Professor Hilton says. 'Blood stem cells were the first example of the successful use of stem cells in medicine to regenerate tissue. However, their low numbers limit more extensive use. We hope our research will redress this problem.'

Stem cell biology also forms a large part of the cancer research program at the institute. Breast cancer researchers have made headway in understanding the cells at the root of breast cancers. Their discovery that stem cell 'daughters', called luminal progenitor cells, are the likely culprit in causing breast cancer was a great advance in our understanding of this disease.

'The inventive approach that Professors Jane Visvader and Geoff Lindeman have taken in identifying the cells at the origin of breast cancers has helped us to take a giant step in our understanding of breast cancer and unravel the link between sustained exposure to female hormones and increased breast cancer risk,' Professor Hilton says. 'The lessons learnt from studying stem and progenitor cells in breast tissue are now being extended to understanding lung cancer and other lung diseases.'

### Understanding immunity

Immunology has been a major focus at the institute for more than 60 years, since Nobel laureate and former institute director Sir Frank Macfarlane Burnet began an immunology research program that culminated in many major discoveries on how the immune system functions.

In 2011, the institute celebrated the fiftieth anniversary of the discovery of the function of the thymus by Professor Jacques Miller. Professor Miller identified that the thymus was critical for the development and function of the immune system. This discovery revolutionised our understanding of the immune system, infection and disease, and led Professor Miller to make several other key discoveries, including identifying two critical components of the immune system — T cells and B cells.

Since then, researchers have continued to make inroads into understanding the function of the immune system, including autoimmune diseases, in which the body's immune cells attack its own tissues. The term 'autoimmune' was coined at the institute by Professor Ian Mackay, who first proposed the idea that disease might be caused by the immune system reacting against its own tissues. Institute researchers continue to work in the field and are developing new, more targeted treatments for these devastating diseases.

One example is Dr Jason Tye-Din, who has been studying coeliac disease, a condition in which immune cells react to gluten in our diet and trigger an immune response that damages many different organs. Coeliac disease is incurable, and puts people at risk of health problems ranging from digestive upsets and iron deficiency through to osteoporosis, other autoimmune diseases, infection and some types of cancer. One in 70 Australians has coeliac disease, but more than 80 per cent of these people have not been diagnosed. Dr Tye-Din's research is leading to better ways to diagnose coeliac disease and distinguish it from other forms of gluten intolerance. Dr Tye-Din has also contributed to the development and testing of the world's first potential vaccine for coeliac disease, Nexvax2®, which is currently in clinical trials.

Researchers at the institute are also continuing to build on the discovery spearheaded by Professor Don Metcalf in the 1960s, in which he identified blood hormones called colony stimulating factors (CSFs). Now, Professor Ian Wicks is examining the role of CSFs in rheumatoid arthritis, an autoimmune disease. 'Understanding how CSFs can be manipulated to treat chronic autoimmune diseases is an exciting field that is already leading to new agents entering clinical trials,' Professor Wicks says. 'As a rheumatologist, I am excited that our research could have opened a whole new field of treatment to help the almost 50 per cent of patients who do not respond adequately to currently available therapies.'

*(continued)*

### Eliminating infection

Many researchers at the institute (see figure 8.39) are devoted to understanding and combating diseases, such as malaria, hepatitis B, tuberculosis and HIV, that affect billions of people globally, particularly in the developing world.

**FIGURE 8.39** PhD student Mr Sofonias Tessema works with laboratory head Dr Alyssa Barry to understand how children in Papua New Guinea develop immunity to the malaria parasite, *Plasmodium falciparum*. This may inform better approaches to developing a malaria vaccine that could help billions of people around the world.

Professor Alan Cowman has had a major impact on infectious disease research in the field of malaria, having spent the past 35 years studying the malaria parasite.

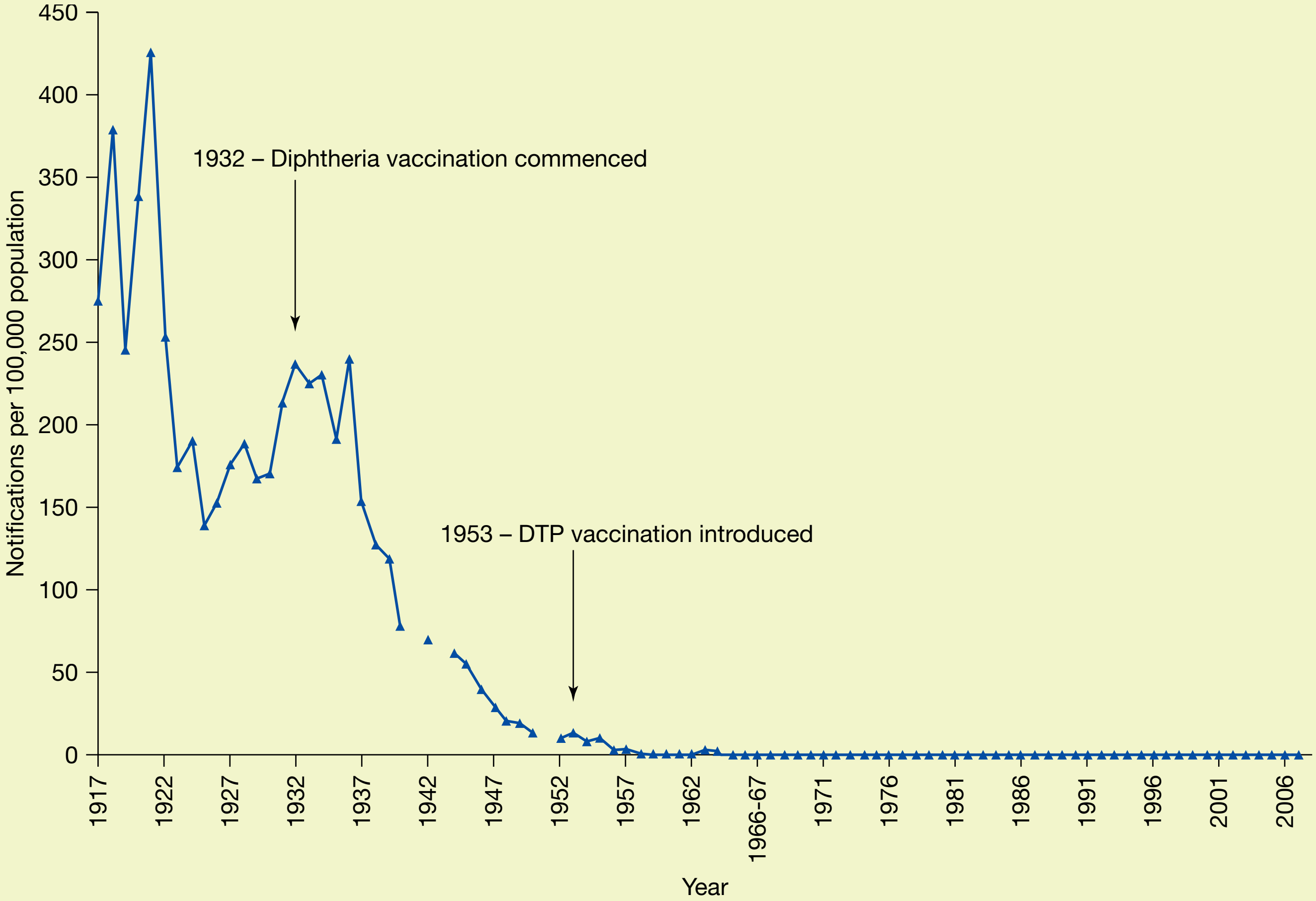
Research led by Professor Cowman has revealed significant detail about how the malaria parasite invades and remodels human red blood cells to multiply and establish infection while escaping the immune responses deployed against it. Collectively, this knowledge is being used to identify vaccine and drug candidates against malaria, including the first potential malaria vaccine that uses a whole, genetically modified malaria parasite to protect against infection, currently in phase 1 clinical trials in the United States.

Institute researchers are also working to understand how viruses such as HIV and hepatitis B, and the bacterium that causes tuberculosis, can overwhelm the immune system, leading to the establishment of chronic infections that are lifelong and incurable. Despite tremendous efforts, long-lived immune responses to some of these viruses are ineffective.

Professor Marc Pellegrini is an infectious diseases clinician who also leads a research group at the institute. His research has revealed how these microbes sabotage the body's normal processes for clearing infections, allowing the microbe to remain within cells. This has led to a clinical trial of a new medication that can promote complete clearance of hepatitis B in laboratory models. It is hoped this may one day lead to a cure for diseases such as HIV, hepatitis B and antibiotic-resistant tuberculosis, which have globally significant health and economic burdens.

# BIOCHALLENGE

1 Examine the graph below, which shows notifications of diphtheria in Australia in the period from 1917 to 2006. (*Source:* Figure 6.1.1, *Communicable Diseases Intelligence Volume 34*, Department of Health, 2010.)




Diphtheria is an acute systemic disease caused by the toxin of the bacterium *Corynebacterium diphtheriae.* Before vaccination programs were established, diphtheria was a dreaded childhood disease. Diphtheria can involve a membranous inflammation of the upper respiratory tract that may cause airway obstruction. Other organs damaged by the toxin include the heart muscle, the nervous system and kidneys. Diphtheria spreads from person to person, most commonly through respiratory droplets from coughing or sneezing.

a Suggest why diphtheria was selected to be the subject of a mass vaccination program.

b The number of notifications is given per 100 000 of a population. A student commented: ‘That’s confusing! Why not just give the numbers of cases?’

Do you agree or disagree with this student? Briefly explain.

c Declines in notifications before immunisation programs began were due to improved living conditions, smaller families and less overcrowding, which reduced exposure to the infectious bacteria. Can you identify this period on the graph?

d If several cases of diphtheria reappeared in an Australian city, would this be expected to lead to a new diphtheria epidemic? Explain.

2 In Australia, the National Notifiable Disease Surveillance System (NNDSS), set up in 1990, monitors the incidence of 65 communicable diseases using data received from each state and territory. Timely fortnightly reports are produced as well as annual summaries.

a Suggest the benefits of having a national surveillance system for communicable diseases.

Communicable diseases are grouped, and table 8.6 shows the total number of cases reported in Australia in 2013.

**TABLE 8.6** Number of cases per communicable disease group reported in Australia in 2013

| Disease group | Number of cases |
|---|---|
| sexually transmissible | 100 949 |
| vaccine-preventable | 59 630 |
| gastro-intestinal | 32 536 |
| blood-borne | 17 919 |
| vector-borne | 10 831 |
| other bacterial diseases | 1932 |
| zoonoses | 634 |
| quarantinable | 3 |

(Source of data: Table 3, *Australia's notifiable disease status, 2013:* Annual Report, CDI Vol 39, No 3, 2015, Department of Health.)

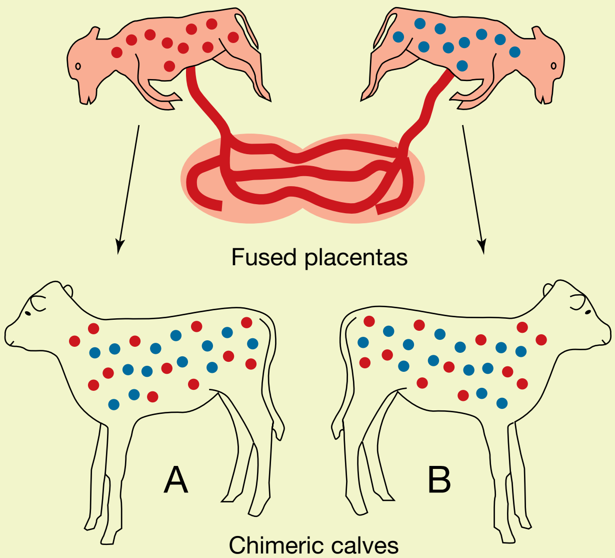
- **b** What is meant by a vaccine-preventable disease?
- **c** Do these data mean that vaccination programs are not protecting vaccinated individuals?
- **d** What is meant by each of the following terms:
  - i a vector-borne disease
  - ii zoonoses (check back to chapter 6, page 230)?

**3** The most serious epidemic of diphtheria in industrialised countries occurred in the 1990s in the newly independent states formed after the break-up of the Soviet Union. In total, between 1991 and 1997, about 140 000 cases of diphtheria were reported, with more than 4000 deaths, with most deaths occurring among adults.

Comment on how each of the factors listed below might have contributed to these epidemics:

- i decreased childhood vaccinations during the 1980s
- ii failure from 1986 to give boosters to children at the time of first entering school
- iii a large population of susceptible adults, many of whom were not vaccinated when the first immunisation programs were introduced in the Soviet Union in the 1950s
- iv mass movements of populations and a decline in socioeconomic standards.

**4** Non-identical twin calves are genetically equivalent to non-twin siblings. They share some proteins but many are different. Sometimes, the two placentas of twin calves fuse and chimeric calves develop. A chimera is an animal that contains cells from two or more genetically different individuals. Stem cells from each calf travel in the bloodstream, enter the body of the other calf and colonise the bone marrow of the other calf. In effect, the result is that both calves have their own cells as well as cells identical to those of their non-identical twin.




- **a** What conclusion could you reasonably make about the ability of fetal cells to distinguish between 'self' and 'non-self' cells?
- **b** Skin from calf A was grafted onto calf B. Explain whether you would expect rejection of the skin graft or not.
- **c** Explain whether you would expect the same result if skin from calf B was grafted onto calf A.
- **d** Tests were carried out on the calves when they were six months old. Would you expect the antibodies in each calf to be identical to those in the other calf? Explain your answer.

# Chapter review

study on
Unit 3 | AOS 2 | Topic 3
Immunity
Practice questions

## Key words

allergens
allergic response
allergy
anaphylaxis
attenuated
autoantibodies
autoimmune diseases
blood
herd immunity
human immunodeficiency virus (HIV)
human papillomavirus (HPV)
hybridoma technique
immunoglobulin E (IgE) antibodies
mast cells
monoclonal antibodies (MAbs)
primary antibody response
secondary antibody response
severe combined immunodeficiency
toxoids
vaccines

## Questions

1 **Making connections** → Use at least eight chapter key words to construct a concept map. You may choose to add additional words.

2 **Analysing information and drawing conclusions** → Di George's syndrome is a disease of the immune system in which B cells are produced but not T cells.
- **a** Explain whether a person with Di George's syndrome would be able to produce antibodies.
- **b** Explain to what extent you think a person with Di George's syndrome is likely to accept or reject a tissue transplant.

3 **Analysing information and applying understanding** → The 'Mantoux' test is designed to determine whether a person has antibodies against the bacillus responsible for tuberculosis. When the test is carried out, a small amount of 'test' solution is injected under the skin on the arm. If the test is positive, a red area appears around the site of injection.
- **a** What material must be in the 'test' solution to react with antibodies in the blood?
- **b** What do you think causes the redness around the injection site?
- **c** Explain the various ways in which a person might have obtained antibodies against tuberculosis.

4 **Demonstrating knowledge and understanding** → Describe the key differences in the anticancer action of the following monoclonal antibodies: bevacizumab and trastuzumab.

5 **Analysing information and applying understanding** → Refer to the box on page 345, which identifies the immunisation schedule for Australian babies. The injections of hepB-DTPa-Hib-IPV that are given at four months and six months are called 'booster shots'. Given that the baby has been already immunised at two months with the hepB-DTPa-Hib-IPV vaccine, explain why booster shots are given.

6 **Communicating knowledge and understanding** → The anti-retroviral therapy used to control a HIV infection involves a cocktail of various drug types, each with different inhibitory actions.

These drugs fall into several classes including:
- i fusion inhibitors
- ii integrase strand transfer inhibitors
- iii reverse transcriptase inhibitors
- iv protease inhibitors.

Refer back to Pages 360–362, which shows how HIV replicates in its target cells: CD4 T-helper cells.

For each class of drug, explain why the drug would be expected to effectively interfere with HIV replication and, as a result, reduce the viral load of a HIV-infected person.

7 **Developing valid biological explanations** → Suggest a reasonable explanation for the following observations:
- **a** From the age of about three months to one year, babies are more susceptible to diseases than after one year of age.
- **b** A person with an uncontrolled HIV infection has a much lower than normal level of T-helper cells.
- **c** The delivery of a chemotherapy drug to a particular target cancer is more direct when the drug has an attached monoclonal antibody.
- **d** Continued growth of a small tumour depends on its establishing a blood supply.
- **e** Herd immunity is effective only if most people in the population are vaccinated.

8 **Applying knowledge in a new context** → The thyroid gland has an organ-specific antigen, thyroglobulin.

In a few people, thyroglobulin can stimulate an immune response, resulting in a disease called Hashimoto's thyroiditis. In this disease, inflammation of the thyroid gland occurs and the thyroid tissue is destroyed.

Explain the cause of this disease.

**9 Demonstrating knowledge and understanding →** Examine the following diagram.

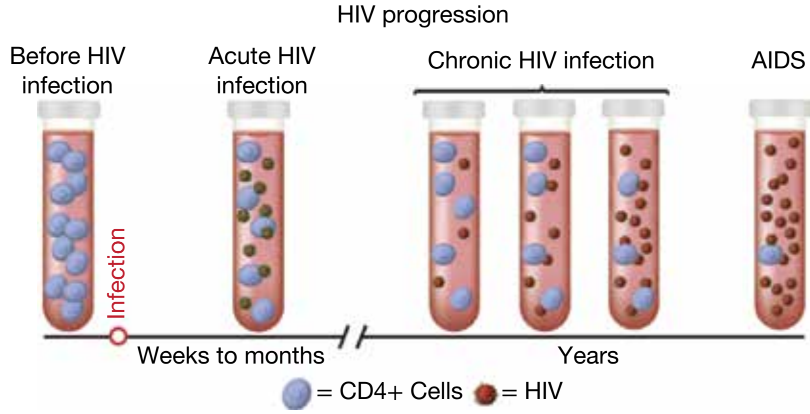



**FIGURE 8.40** Diagram showing the change over time in CD4+ cells and HIV particles in blood samples taken from a person over time.

- **a** What is another name for CD4+ cells?
- **b** What happens to CD4+ cells over time in an untreated HIV infection?
- **c** What happened to the HIV over time?
- **d** Explain what relationship, if any, exists between these two changes?
- **e** Would this person have developed AIDS?
- **f** Briefly explain how the administration of anti-retroviral therapy would be expected to change this picture.

**10 Demonstrating knowledge and understanding →** Following an organ transplantation, the immune system must be suppressed so that it does not attack a transplanted organ and cause rejection. Immunosuppressive drugs are given for this purpose.

To monitor the effectiveness of the immunosuppressive therapy, the CD4+ cell count in the blood of the transplant recipient is checked.

- **a** Why focus on monitoring CD4+ cells?
- **b** What result would indicate that the drug is not working well and that the transplant is at risk of rejection?
- **c** Considering a kidney transplant and a cornea transplant, which is more susceptible to rejection by the immune system of the recipient? Explain your choice.

**11 Communicating knowledge and understanding →** Identify the immune events that have occurred in the following people:

- **a** a person who has been sensitised to the pollen of a particular plant
- **b** a person who has lost the ability to produce the hormone insulin
- **c** a person who has been protected from the bite of a redback spider by administration of an antivenom
- **d** a person who has received the Gardasil vaccine.

**12 Evaluating information →**


**FIGURE 8.41** The Victorian 'No Jab No Play' legislation came into effect on 1 January 2016. (Image courtesy of the State Government of Victoria.)

On 1 January 2016 in the state of Victoria, the '*No Jab, No Play*' legislation came into effect. This requires that parents and guardians enrolling children in kindergarten or child care (long day care, occasional care or family day care) must provide evidence that their children meet one of the following criteria with respect to immunisation: the children in question are (i) fully immunised for their age, or (ii) on a vaccination catch-up program, or (iii) are unable to be fully immunised for medical reasons.

**a** What is the biological rationale for this action?

**b** Give an example of a medical condition that would qualify for exemption under the new laws.

**c** Does this legislation remove parental choice and enforce immunisation without parental consent?

**13 Recognising differences and similarities →** Identify one key difference and one similarity between the members of the following pairs:

**a** active immunity and passive immunity

**b** natural immunity and artificial immunity

**c** dead pathogen and attenuated pathogen

**d** a toxin and a toxoid

**e** allergy and allergen

**f** naked monoclonal antibody and conjugated monoclonal antibodies.

**14 Evaluating alternatives →** In November 2005, an inactivated polio vaccine (IPV, or Salk vaccine) replaced the oral (Sabin) vaccine in Australia. The advantage of using IPV is that it cannot cause vaccine-associated paralytic poliomyelitis. Mutations can occur with the 'live' attenuated oral polio vaccine, producing a rare virulent form that can cause paralytic polio. In developing countries, however, the oral vaccine is the vaccine in use.

Refer back to table 8.2 on page 342 and, taking into account the different features of these two polio vaccines, suggest why the oral vaccine is used in developing countries.

**15 Interpreting data in graphical form →** Examine the following graphs, which show data on cervical cancer in Australia.

Refer to figure 8.42a and answer the following questions:

**a** What general trend is apparent in the incidence of cervical cancer in the period from 1982 to 2012?

**b** What general trend is apparent in the mortality rate for cervical cancer in the period from 1968 to 2012?

**c** Based on a female population of 6 million in Australia in 1968 and using the mortality rate of 7.7 per 100 000, how many deaths from cervical cancer would have occurred in that year?

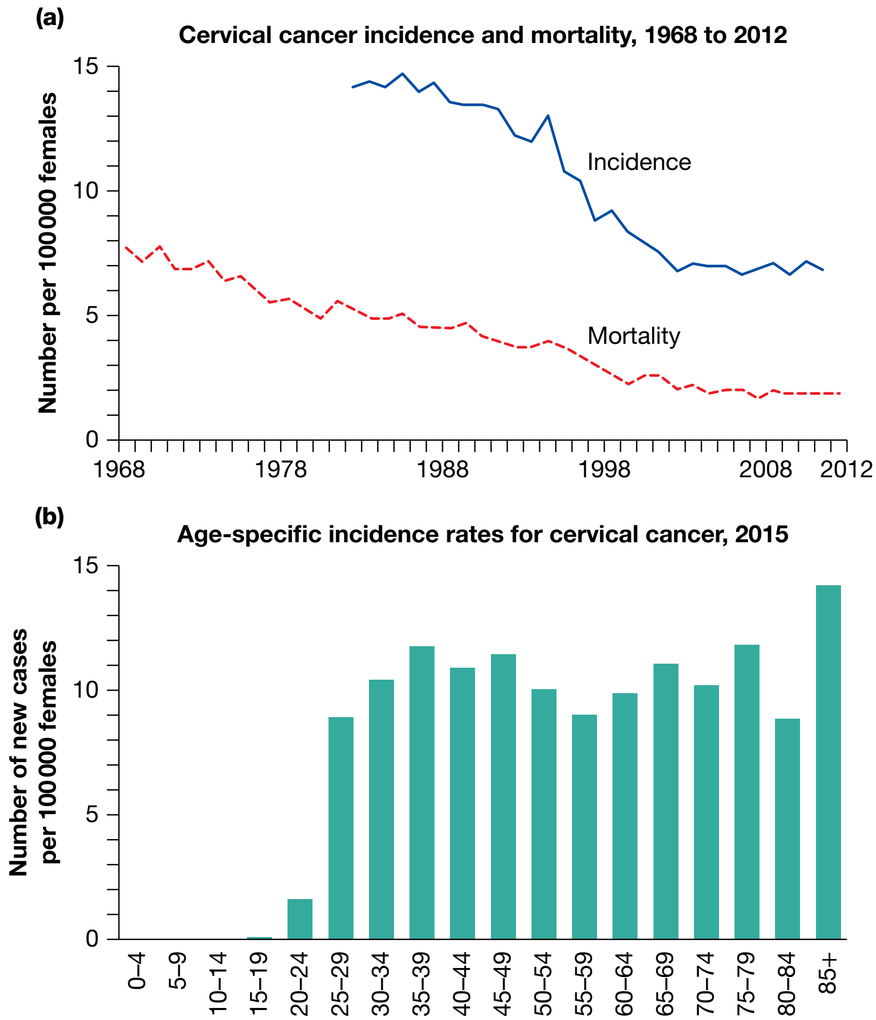



**FIGURE 8.42** **(a)** Incidence of cervical cancer (1982–2012) and mortality (1968–2012) in Australia. **(b)** Incidence of cervical cancer by age group in 2015. (Source of the data in both (a) and (b): Australian Institute of Health and Welfare.)

**d** The female population in Australia in 2016 is about 12 000 000. Using the mortality rate from cervical cancer of 1.8 per 100 000 females, how many annual deaths from this cause would be predicted over the next few years?

Refer to figure 8.42b and answer the following questions:

**e** Which age group shows the most marked 'jump' in incidence compared to the preceding age group?

**f** Cervical cancer immunisation was introduced in Australia in 2007 for 13-year-old females. Would the impact of this immunisation program be expected to have an immediate impact on cervical cancer incidence?

**16** **Discussion question** ➔ Consider the following statements:

1. *'Instead of vaccinating babies and children, we should let them get the disease and when they recover from it, they will have active natural immunity. Surely this is better than vaccinating them and giving them artificial active immunity.'*
2. *'Babies are tiny, vulnerable creatures. Why expose them to vaccines as early as soon after birth or just a few months old. Wouldn't it be better to vaccinate them when they are older and stronger?'*

Discuss these statements with your classmates and determine their biological validity.

CHAPTER

# 9 Changes in genetic make-up of populations

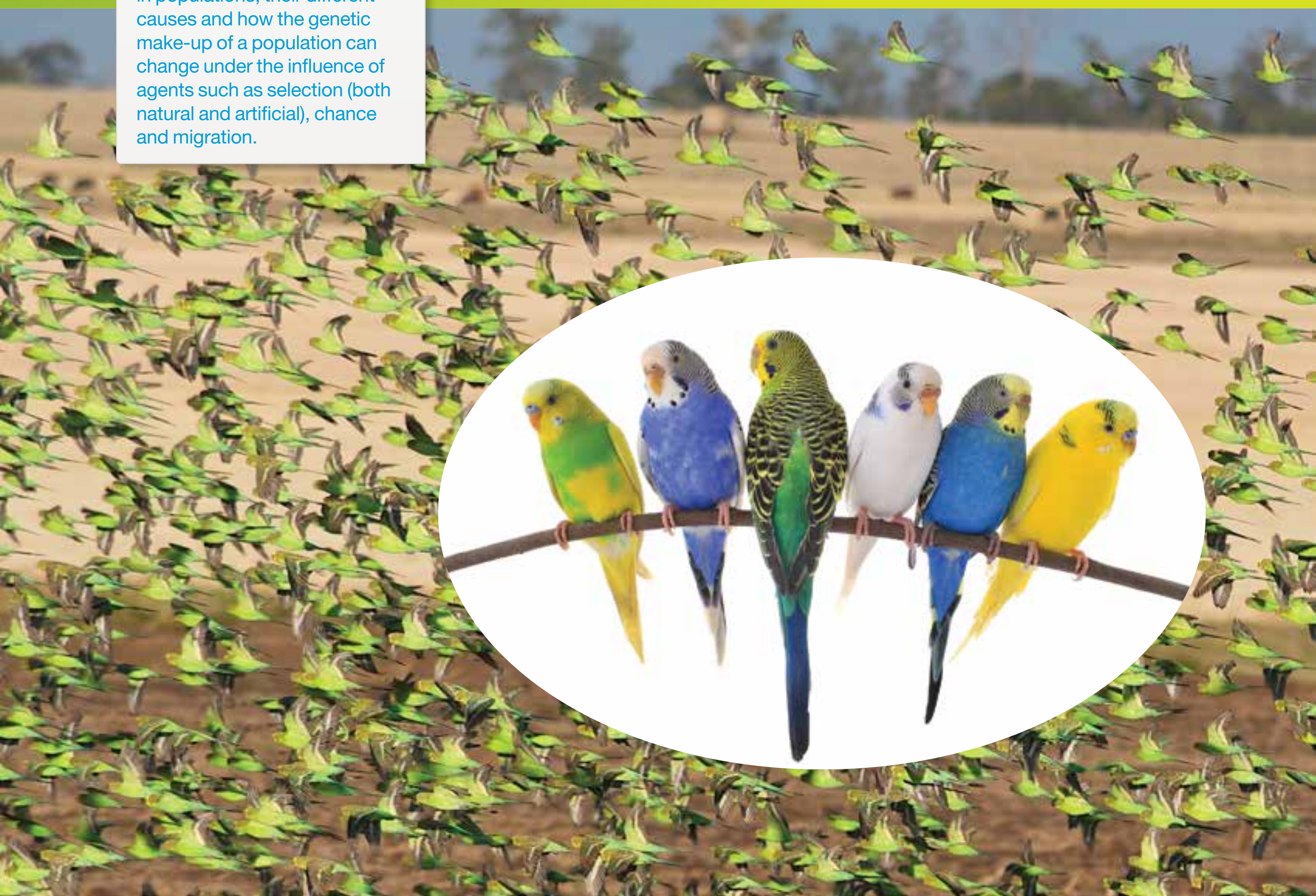


**FIGURE 9.1** **(a)** A flock of wild budgerigars in flight in inland Australia. Budgerigar populations in the wild do not vary in colour, with all birds being the so-called light green colour. **(b)** Other colour variants in budgerigars have arisen by mutation from time to time in captive stock. Much of the variation seen in populations of plants and animals has a genetic basis, as is the case for these budgerigar varieties. In this chapter, we will examine inherited variations in populations, their different causes and how the genetic make-up of a population can change under the influence of agents such as selection (both natural and artificial), chance and migration.

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- understand the concept of the gene pool of a population
- appreciate the range of genic and chromosomal variants that can exist in a population
- identify the factors that can act to change the allele frequencies in a population's gene pool
- recognise that selection pressure on phenotypes varies with changes in environmental conditions
- gain knowledge and understanding of how gene pools can be manipulated through selective breeding programs and the consequences for genetic variation
- become aware of the process of allopatric speciation
- appreciate the value of genetic diversity in a population and the importance of its preservation.

**ODD FACT**

Milk from dairy goats, dairy sheep and even buffalo is in demand in Australia for use in the manufacture of specialist cheeses. A market for goats' milk also exists for people with allergies to cows' milk.

# Can you drink milk?

A glass of cold milk — plain or flavoured — is a refreshing drink for many people (see figure 9.2). In the year 2014–2015 in Australia, sales of drinking milk, principally cows' milk, totalled 2481 million litres. Unflavoured fresh milk comes as regular (at least 3.2% fat), reduced fat (no more than 1.5% fat) and 'no fat' or skim milk (contains a maximum of 0.15% fat).

**FIGURE 9.2** Some of the many fresh milks on sale — regular, reduced fat and low fat, and plain and flavoured.

Like other baby mammals, human babies can be sustained on a diet exclusively of breast milk for a significant period after birth. In the case of human babies, the World Health Organization identifies this as a six-month period after which complementary foods are added to a baby's diet.

Milk contains lactose, a disaccharide sugar, at concentrations of about 5 grams per 100 mL. Lactose must be broken down to its monosaccharide components, glucose and galactose, before it can be absorbed from the gut. This reaction is catalysed by the enzyme **lactase**, which is released from cells lining the small intestine:

$$\text{lactose} \xrightarrow{\textit{lactase}} \text{glucose} + \text{galactose}$$

In humans, the lactase enzyme is encoded by the **LCT** gene on the number-2 chromosome.

Human infants up to about three to four years of age can digest lactose. However, after that time, the majority (about 70%) of the world's population start to lose their lactase enzyme activity and, by about seven or eight years of age, cannot digest lactose. This situation is seen in many people whose ancestry is African, Asian or southern European. For such people, drinking fresh milk can result in abdominal cramps, bloating and watery diarrhoea. Such people are said to be **lactase nonpersistent** (sometimes termed lactose intolerant). Lactase nonpersistence is a recessive trait.

Reminder: The ending for an enzyme name is -***ase***, for example, lactase. The ending for a sugar is -***ose***, hence glucose, galactose and lactose.

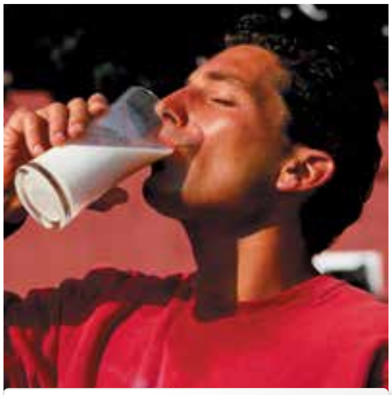


**FIGURE 9.3** A glass of milk being enjoyed by a lactase persistent adult.

**ODD FACT**

Cats and dogs are commonly lactase nonpersistent. Pet food manufacturers produce a special milk for these family pets that is low in lactose, such as Whiskas® Milk Plus, a lactose-reduced milk for cats.

In contrast, the majority of people whose ancestry is from northern Europe retain their lactase enzyme activity into adulthood and are said to be **lactase persistent**. For lactase persistent people, milk can be a source of nutrition throughout their lives (see figure 9.3). Lactase persistence is an autosomal dominant trait.

When several discrete inherited variations are present in the members of a population, that population is said to be **polymorphic** for the trait concerned. The existence of the lactase persistent/nonpersistent variants in human populations is one example of a polymorphism.

## Distribution of lactase persistence

All young non-human mammals, under natural conditions, do not drink milk again after being weaned. For these mammals, no disadvantage is associated with the absence of lactase enzyme activity when they are adults. A similar situation also existed in human prehistory. The ancestral state in human populations was lactase nonpersistence, meaning that all members of early modern human populations were lactase nonpersistent. The appearance of lactase persistence in human populations is a more recent occurrence.

Figure 9.4 shows current regional levels of nonpersistence of lactase enzyme activity (lactose intolerance). Note that in some regions, such as east Asia, the population contains a high percentage of lactase nonpersistent individuals. Note that indigenous populations in Australia and North America are lactase nonpersistent. What does this suggest about their traditional diets?

Table 9.1 provides data for various population groups for lactase persistence/nonpersistence. Note, in particular, the contrasting situation between some African indigenous groups and those from North America and Australia.

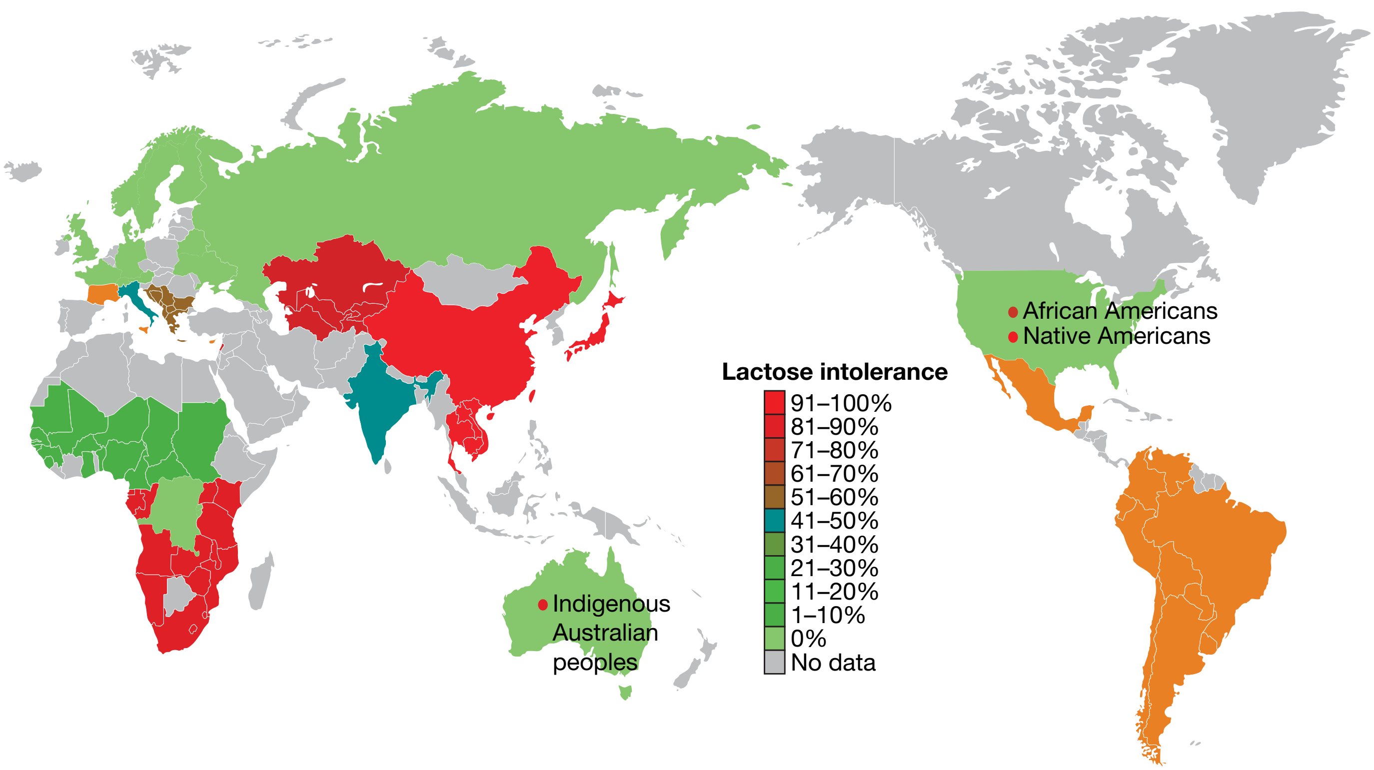



**FIGURE 9.4** Map showing the levels of lactose intolerance (lactase nonpersistence) in various regions of the world. The green-shaded areas indicate regions where the majority of people are lactose tolerant (lactase persistent) and the red-shaded areas are regions where the majority of people are lactose intolerant (lactase nonpersistent). Note that within the Australian population, the Indigenous peoples are lactose intolerant.

**TABLE 9.1** Percentages of lactase nonpersistent individuals in samples of various human population groups. Of the African ethnic groups — Igbo, Yoruba, Tutsi and Fulani — some come from regions where cattle cannot survive because of tsetse fly infestation. Can you identify the ethnic groups from these regions?

| Population sample | Lactase nonpersistence (%) |
|---|---|
| Native Americans | 100 |
| Thai | 99 |
| Igbo (West Africa) | 99 |
| Yoruba (West Africa) | 99 |
| Bantu (Central to Southern Africa) | 99 |
| Chinese | 88 |
| Indigenous Australian peoples | 85 |
| African Americans | 70 |
| Bedouin (North Africa to the Middle East) | 25 |
| Tutsi (Central Africa) | 20 |
| Fulani (West Africa) | 18 |
| French | 17 |
| White Americans | 15 |
| Swiss | 10 |
| Swedish | 6 |
| Australians (European) | 4 |
| Danes | 4 |
| Dutch | 1 |

## Spread of lactase persistence

Perhaps about 10 000 years ago, a mutation occurred in a hunter-gatherer population that produced the allele ***C*** for persistence of lactase enzyme activity into adulthood. When lactase persistence first appeared in human populations, this mutation gave no benefit.

**Soon afterwards, a cultural change occurred in some human populations that would give a selective advantage to people with lactase persistence**. This cultural change was the shift from a hunter-gatherer lifestyle to that of farming.

This transition began in the Middle East about 11 000 years ago and then spread to Europe. Farming with animals — cows, sheep or goats — provided not only meat, but also a source of milk (see figure 9.5). At first, raw milk could not be tolerated by the members of early farming communities because they were all lactase nonpersistent. Archaeological evidence of early cheesemaking indicates that the early farmers may have made fermented milk products such as cheese. The fermentation process involved in making cheese, and other dairy products such as yogurt, breaks down most of the lactose, and so these dairy products could have been consumed by lactase nonpersistent people.

Perhaps about 7000 years ago, the lactase persistent mutation began to spread through farming communities. **Agricultural herding of animals that produced milk for human consumption became an agent of positive selection for individuals with the mutant allele for lactase persistence.** A mutant allele that had previously been of no value became an asset in the new environment of dairy farming. Raw milk provided an additional source of nutrition for lactase persistent members of the population. Groups of lactase persistent farmers with their herds began to spread across Europe.

**ODD FACT**

One speculation is that a large proportion of Europeans may be descendants of the early lactase persistent dairy farmers in Europe.

**FIGURE 9.5** Goats are one example of a herd animal that can provide milk. Here we see a goat being milked.

The spread of dairy farming across Europe is associated with the spread of the ***C*** allele for lactase persistence. In populations that subsisted on dairying, individuals with even a single copy of the ***C*** allele were able to consume milk and milk products into adulthood, giving them an additional source of nutrition that was not available to lactase nonpersistent people, genotype ***cc***, in the same population. Within a dairy-farming environment, lactase persistent individuals have a **selective advantage** over those individuals who are lactase nonpersistent.

In populations where dairy farming was practised, lactase persistence increased in frequency over successive generations, producing the high levels seen today in Europe and in some African ethnic groups. This increased frequency is evidence of the selective advantage conferred by the presence of the lactase persistence allele, either in a single copy in the heterozygous condition (***Cc***) or as two copies in the homozygous condition (***CC***). A selective advantage confers greater genetic fitness on people possessing the lactase persistence allele. Greater genetic fitness is seen in the greater survival rate to reproductive age and a greater reproductive success as expressed in the contribution of genes to the next generation (see figure 9.6).

Refer back to table 9.1. The link between dairy farming and a high level of lactase persistence in a population is also seen in some African ethnic groups, such as the Tutsi people and the Fulani people, whose traditional diet includes fresh milk from their herds of cattle or goats. This link is also seen in nomadic Bedouin ethnic groups whose camels provide both transport and a source of milk.

So, the pattern that emerged is:

- populations that have not had fresh milk as part of their historical diet have high proportions of lactase nonpersistent individuals, and the presence of the lactase persistence allele is of no benefit in a non-dairying environment
- populations that have had fresh milk as part of their historical diet have high proportions of lactase persistent individuals because the lactase persistence allele conferred a selective advantage in a dairy farming and herding environment.

**FIGURE 9.6** Who is fitter (in a genetic sense) — the Olympic weightlifter who is sterile or the relative 'weakling'? Genetic fitness relates to the contribution of genes to the next generation. So, the organism that hides and survives to reproduce is fitter in a genetic sense than the organism that fights and dies before reproducing.

Populations have a pool of inherited (genetic) variation. A change in the environmental conditions in which a population lives can act as an **agent of selection**. The action of an agent of selection can produce differential survival and reproduction rates in members of a population, depending on differences in their inherited physical, physiological or behavioural characteristics. Where members of a population, because of their particular genetic make-up or genotype, have a higher survival and reproductive rate compared to other members of the population, they are said to have a selective advantage.

The lactase persistence/nonpersistence variation that exists in human populations today is just one example of the inherited variation that exists in plant and animal populations. In the next section, we will explore other examples of inherited variations in populations.

### KEY IDEAS

- **Human populations from various regions show a genetic polymorphism in terms of the persistence/nonpersistence of lactase activity.**
- **During their childhoods, individuals who are lactase nonpersistent lose the ability to digest milk, while those who are lactase persistent retain this ability into adulthood.**
- **Lactase nonpersistence is the ancestral state, and lactase persistence arose as a result of a later mutation.**
- **The emergence of dairy herding was an agent of positive selection for lactase persistence.**

### QUICK CHECK

1. What prediction may be made about the relative proportions of lactase persistence and nonpersistence in a population that, early in its history, developed dairy farming?
2. Identify the following statements as true or false:
   - a The more common state in the human global population is lactase nonpersistence.
   - b Lactase persistence is a result of a mutation that disables the gene that encodes the lactase enzyme.
   - c The frequency of lactase persistent individuals in the Tutsi population is greater than in a Chinese population.
   - d Lactase persistence is a dominant trait.
   - e The lactase persistence allele (***C***) gives any person possessing it a selective advantage.

# Inherited variation in populations

The raw material of evolution is the presence in populations of inherited (genetic) variation. **Charles Darwin recognised that if evolutionary change was to occur in a population, it could only do so if inherited variation was present in that population.** Inherited variation may be expressed as structural, physiological, biochemical or behavioural traits.

Inherited variation in a population may have:

- a genic basis: that is, variation that arises from the action of one or more genes
  - action of a single gene
  - action of several genes (polygenes).

- a chromosomal basis: that is, variation that arises as a result of changes in chromosomal make-up
  - gain of entire set(s) of chromosomes (polyploidy)
  - gain or loss of a single chromosome (aneuploidy), where a gain produces a chromosomal trisomy, and a loss produces a monosomy.

Let us look first at variations in population that arise from the action of a single gene. You have already met one of these inherited variations, namely, the lactase persistence/nonpersistence states in people, which are due to the action of alleles of a single gene.

## Variation due to a single gene

The relationships between the three alleles of the ABO gene are detailed in *Nature of Biology Book 1 Fifth Edition*, page 557.

Inherited traits that exist as two or three discrete (non-overlapping) variants in a population are typically controlled by a single gene; these single genes often have just two common alleles. When the relationship between the expression of these alleles is one of simple dominance and recessiveness, two variants of the trait can exist, as, for example, long and short fur length in cats, and black or liver coat colour in dogs. When the relationship between the two alleles is one of co-dominance, three variants will exist, as, for example, red, pink and white petal colour. In the case of the ABO blood type gene, four ABO blood type variants (A, B, AB and O) are produced because this gene has three alleles.

Let's look first at examples of inherited discrete variations that are controlled by a single gene.

*Variation in people* (see figure 9.7a):

- lactase activity — persistent and nonpersistent
- ABO blood types — A, B, AB and O
- hair colour — red and non-red (non-red covers all other colours)
- haemoglobin types — Hb A and Hb S
- tasting ability — able to taste PTC and unable to taste PTC.

**eBook*plus***

**eLesson**
Co-dominance
**eles-2461**

*Variation in cats* (see figure 9.7b):

- pattern — striped (tabby) and non-striped
- colour — black and ginger; grey and cream
- skeletal structure — tailed and no tail (Manx).

*Variation in domestic dogs* (see figure 9.7c; refer also to figures 9.50 and 9.64 on pages 426 and 440):

- colour — black and liver
- hair length — long and short.

*Variation in insects*

- response to pesticides — resistant and sensitive
- wing pattern — spotted and solid (see figure 9.7d).

*Variation in plants* (see figures 9.7e and 9.7f)

- flower colour (many species) — red, pink and white, etc.
- growth habit — prostrate (ground hugging) and erect.

Inherited traits that differ in their expression, such as all the traits shown in figure 9.7, are called **polymorphisms** (*poly-* = many, *morphe* = shape, form). These polymorphic traits are encoded by one gene that has two or more alleles. When more than one variant of a trait is seen in a population, the population can be said to be polymorphic. In contrast, when all members are identical in terms of a particular trait, the population is said to be **monomorphic** (*mono-* = alone, single).

In each case, members of a population can be separated into a small number of discrete non-overlapping classes that can be readily identified, and the number in each class can be quantified.

In this section, the focus is on inherited variants that are discrete or non-overlapping; in other words, the variation is **discontinuous**, with a clear demarcation between the different classes. (It is easy to separate red from pink and white.)


**FIGURE 9.7** Examples of discrete inherited variations controlled by single genes in several different species. **(a)** At left: in people, the A, B, AB and O blood groups are variations that result from various combinations of the three alleles of the ABO gene; at right: the ability to taste bitter substances, such as the chemical phenylthiocarbamide (PTC) as well as the chemically related isothiocyanates (ITC) present in vegetables such as cabbage and broccoli, is under the control of a single gene, with some people able to detect the bitterness while others are unable to detect any bitterness. (Right-hand image in (a) courtesy of Judith Kinnear.)

*(continued)*




**FIGURE 9.7** *(continued)* **(b)** Single gene variants in cats include fur colour (black and orange), colour density (black and grey), white spotting (extensive and limited) and fur length (short and long haired). (Images in (b) courtesy of Judith Kinnear.) **(c)** At left and centre: single gene variants in dogs showing variation in white spotting pattern (Irish pattern and piebald), coat colour (yellow and black) and coat length (long and short haired); at right: short-haired dogs showing colour variation (black and yellow). (Right-hand image in (c) courtesy of Judith Kinnear.) **(d)** Wing pattern variation in the peppered moth, *Biston betularia*, typica variety and solid or melanic variety. **(e)** Petal colour variation in the harlequin flower, *Sparaxis tricolor*. (Images in (e) courtesy of Judith Kinnear.) **(f)** Variation in growth habits as seen in juniper species. *Juniperus chinensis* has an erect growth habit while *J. procumbens* has a prostrate growth habit (at front). In several plant species, this variation has been shown to be due to the action of a single gene. (Image (f) courtesy of Judith Kinnear.)

Let's now look at inherited variation under the control of many genes (polygenes), producing variation that tends to be continuous.

## Variation due to polygenes

Inherited variation between adult members of a population can involve subtle differences, such as slight variations in body dimensions. Red-winged blackbirds, for example, show variation in the lengths of their beaks, bodies, tails and wings. Figure 9.8a shows an image of a red-winged blackbird.

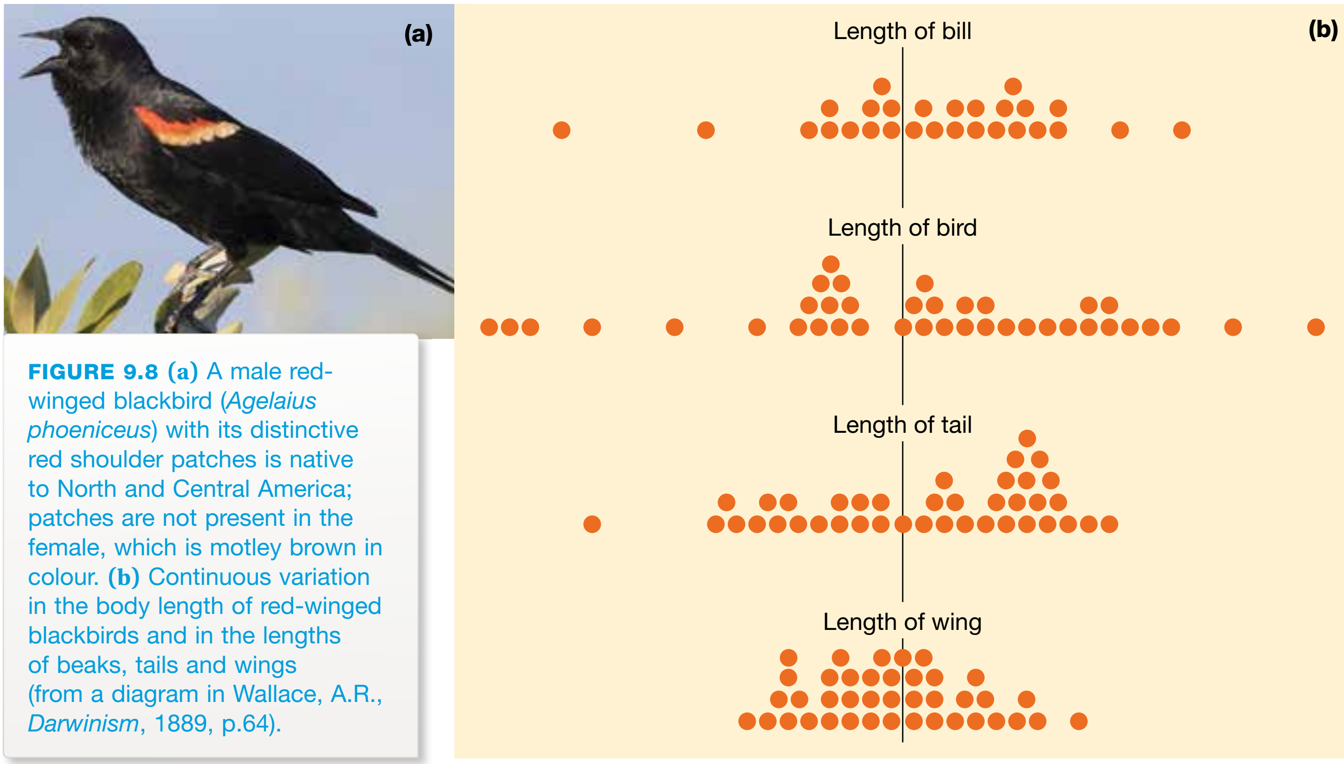



**FIGURE 9.8** **(a)** A male red-winged blackbird (*Agelaius phoeniceus*) with its distinctive red shoulder patches is native to North and Central America; patches are not present in the female, which is motley brown in colour. **(b)** Continuous variation in the body length of red-winged blackbirds and in the lengths of beaks, tails and wings (from a diagram in Wallace, A.R., *Darwinism*, 1889, p.64).

Polygenic inheritance is discussed in *Nature of Biology Book 1 Fifth Edition*, pages 561–563.

In one group of birds, commonly called Darwin's finches, subtle differences in the height and width dimensions of their beaks have been recorded. In this case, the variation is the result of the action of just one gene, the **BMP4** master gene. Shifts in the timing, the level of expression and where this gene is expressed generate a range of different beak shapes (see chapter 11, pages 533–534. This is a different way of generating variants compared to that resulting from the action of polygenes.

Variation of the type shown in figure 9.8b is **continuous**. Other examples of continuous variation include fat content of cows' milk, mass of pea seeds, and many human traits including adult height, body mass index, waist-hip ratio, blood lipid levels and, in females, age at first menstrual bleeding (menarche). Variation of this continuous type is often controlled by a large number of **polygenes**, each with a small but additive effect. The greater the number of polygenes, the greater the number of possible variants. (Many of these traits are also influenced by environmental factors including nutrition, level of activity, and disease incidence.)

For traits controlled by polygenes, the variation seen in members of a population does not fall into a few non-overlapping classes. Instead, the variants are distributed across a continuous range. Figure 9.9 shows the contrast between the continuous distribution of variants for adult male height, a polygenic trait, and the discrete distribution of the same group of people for lactase persistence and nonpersistence, traits controlled by a single gene.

## Chromosomal variation: polyploidy

Polyploidy refers to a condition in which an organism has more than two matched sets of chromosomes. When just the two matched sets of chromosomes are present, an organism (and its cells) are said to be **diploid**. The presence of one or more additional sets of chromosomes to the base diploid condition (2 sets) produces organisms that are termed triploid (3 sets), tetraploid (4 sets), hexaploid (6 sets) and even dodecaploid (12 sets).

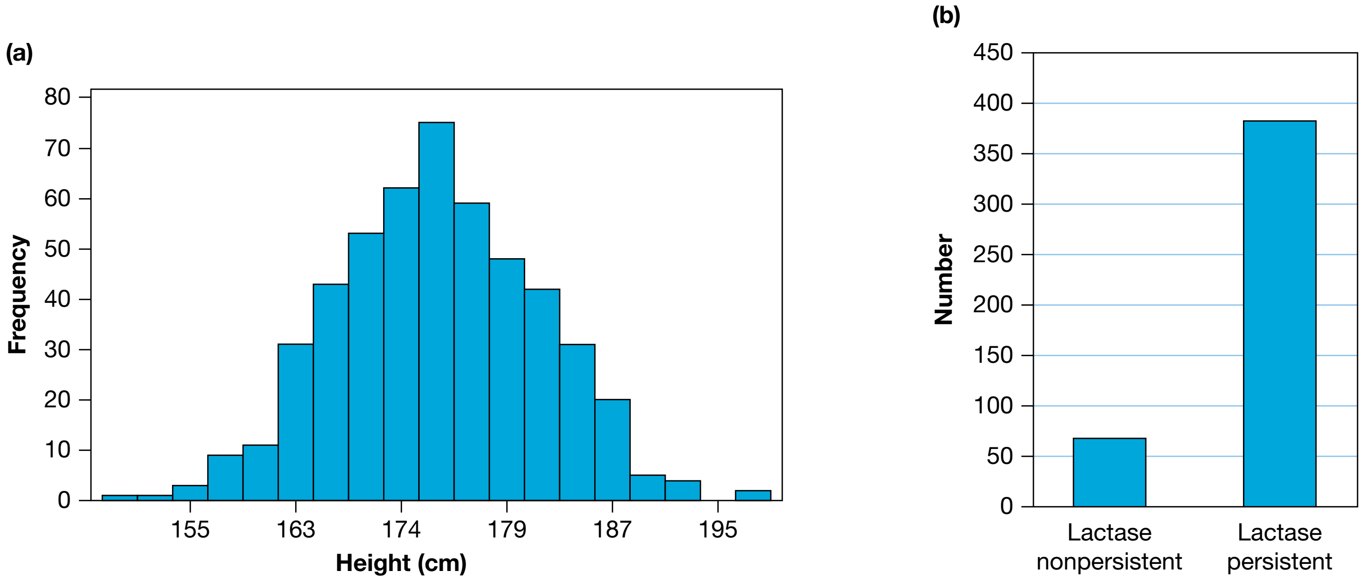



**FIGURE 9.9 (a)** Continuous distribution of heights in a sample of adult males. **(b)** Discontinuous distribution of lactase persistence/nonpersistence status of the same group. Which distribution shows variation controlled by one gene with two alleles with a simple dominant/recessive relationship?

Polyploidy is very common in flowering plants and ferns. As well as occurring naturally, polyploid plants can also be artificially created. This is deliberately done because polyploid flowering plants tend to be bigger than their diploid relatives and have larger flowers. This has occurred in several ornamental gardens, such as tetraploid snapdragons (*Antirrhinum majus*) $4\times = 64$ (see figure 9.10).

**FIGURE 9.10** Tetraploid ($4\times = 64$) snapdragons are bigger, showier and hardier than their diploid ($2n = 32$) close relatives.

Polyploidy is much rarer in animals. Among polyploid animals are the various species of tetraploid salmon; for example, the Chinook salmon (*Ororhynchus tshawytscha*) has a chromosome number of $4\times = 68$. Goldfish (*Carassius auratus*) and carp (*Cyprinus carpio*) are both tetraploid, each with a chromosome number of about 104.

Table 9.2 shows some of the polyploid flowering plants and their numbers of chromosome sets. Note that various cordgrass species in the genus *Spartina* have different levels of polyploidy, up to dodecaploid ($12\times = 120$).

**TABLE 9.2** Polyploidy in flowering plants.

| Plant | Chromosome number | Description |
|---|---|---|
| watermelon, *Citrullus lanatus* | $2\times = 22$, $3\times = 33$, and $4\times = 44$ | diploid, triploid and tetraploid |
| banana, *Musa acuminata* | $3\times = 33$ | triploid |
| durum wheat, *Triticum turgidum* | $4\times = 28$ | tetraploid |
| emmer wheat, *Triticum dicoccum* | $4\times = 28$ | tetraploid |
| cotton, *Gossypium hirsutum* | $4\times = 52$ | tetraploid |
| potato, *Solanum tuberosum* | $4\times = 48$ | tetraploid |
| kiwifruit, *Actinidia deliciosa* | $6\times = 174$ | hexaploid |
| bread wheat, *Triticum aestivum* | $6\times = 42$ | hexaploid |
| oat, *Avena sativa* | $6\times = 42$ | hexaploid |
| cordgrasses, *Spartina* spp. | $4\times = 40$, $6\times = 60$, and $12\times = 120$ | tetra-, hexa- and dodecaploid |

In some cases, a polyploid individual will have additional complete sets of chromosomes from its own species — this is a situation termed **autopolyploidy**, as seen in snapdragons, bananas and watermelon. The polyploidy condition produces plants with larger cells and larger organs, such as flowers, fruits, leaves and even chloroplast counts in leaf cells.

In other cases, the additional chromosome sets in a polyploidy may come from another species. In this case, the resulting organism is called an **allopolyploid**. Many allopolyploids show 'hybrid vigour', indicating that they are more robust and display superior qualities as compared to the parents from which they were derived.

Bread wheat (*Triticum aestivum*) is an allopolyploid ($6\times = 42$) that arose spontaneously in about 7500 BC. Research has shown that its sets of chromosomes came from three different species, not at once, but in a pathway involving several crosses between different species (hybridisations) (see figure 9.11).

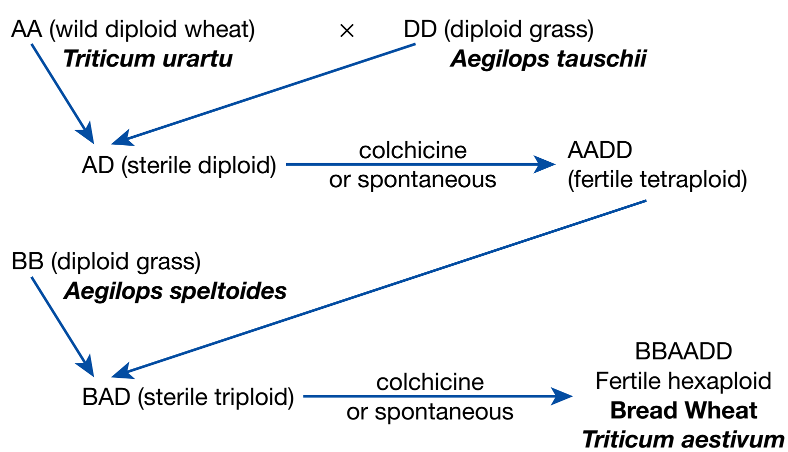



**FIGURE 9.11** Formation of hexaploid bread wheat, *Triticum aestivum* ($6\times = 42$) from three antecedent species, including a wild diploid wheat, *Triticum urartu*, and two wild species of grasses.

### How are polyploids formed?

Figure 9.12 summarises the various ways in which polyploid plants can be produced, both autopolyploid and allopolyploids. The red circle denoted by A indicates the basic (haploid) chromosome set of the diploid species AA, while the blue square indicates the basic (haploid) chromosome set of diploid species BB.

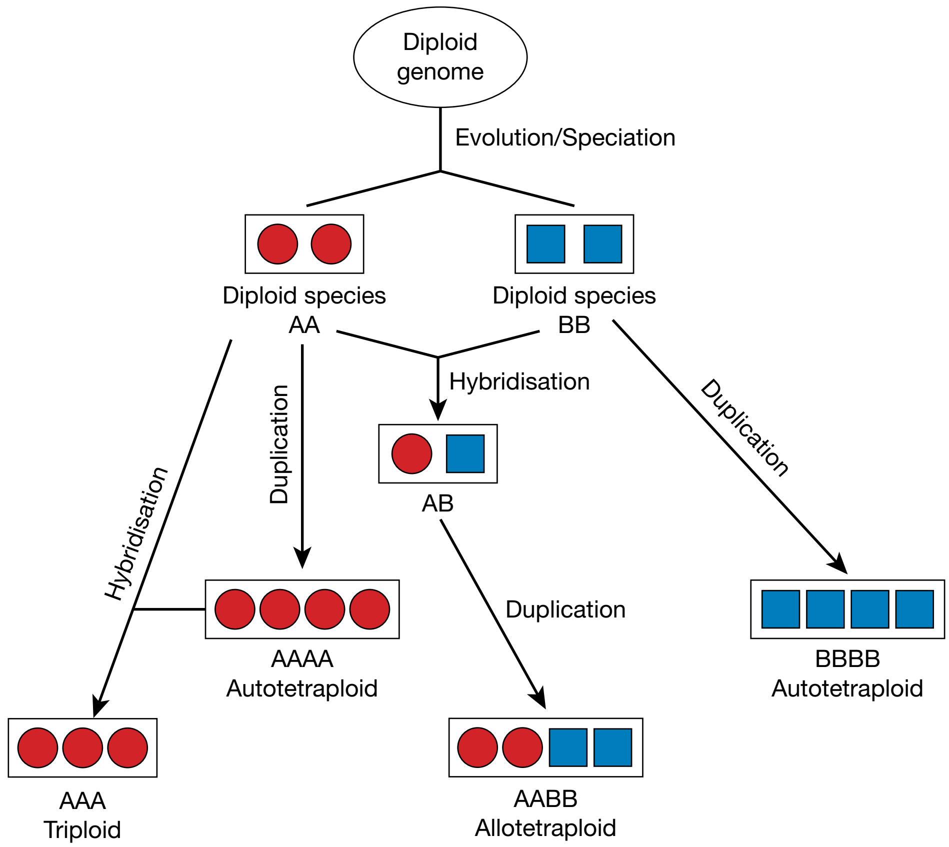



**FIGURE 9.12** Various strategies to make polyploidy plants.

The various processes involved in producing polyploidy organisms include:

- duplication that involves chemical treatment to produce a doubling of the chromosome number; this is most commonly achieved through colchicine treatment
- **hybridisation** that involves crossing parents belonging to two different species, as for example:
  - crossing two different diploid species to produce a diploid hybrid progeny
  - crossing a tetraploid and a diploid to produce triploid progeny.

Polyploidy has changed populations by creating new chromosomal combinations and producing new species that display new variations.

## Chromosomal variation: aneuploidy

An **aneuploid** organism is one that has a chromosome number that differs by a small number from the standard chromosome number for its species. The difference is typically one or two chromosomes more than the standard number or one or two chromosomes less.

**Loss of chromosome(s):** The loss of one of a pair of chromosomes or both members of a pair of chromosomes would be expected to create a lethal condition in diploid organisms because of the loss of the function of the genes carried on those chromosomes. However, polyploid organisms can tolerate the loss of one or even two members of a pair of chromosomes because of the presence of other chromosomes.

**Gain of chromosome(s):** The gain of a chromosome would not be expected to benefit the organism concerned because of the chromosomal imbalance it would produce. In humans, for example, the only autosomal trisomy that survives beyond infancy is Down syndrome, with three copies of the number-21 chromosome. All other autosomal trisomies either result in spontaneous miscarriages or, in the case of trisomy 13 and trisomy 18, the birth of a baby with severe abnormalities, both physical and mental. In both trisomy 13 and trisomy 18, the expected life span is very short, measured in days or weeks.

However, one plant shows an interesting range of variation for fruit shape resulting from chromosomal trisomy events. Thorn apple (*Datura stramonium*), an introduced plant, is a declared noxious weed in many Australian states. The thorn apple is an annual herb that grows to about one metre in height and has large white, bell-shaped flowers. Its fruits, about half the size of an apple, are covered by thorns.

Thorn apples have twelve pairs of chromosomes ($2n = 24$) and show variation in fruit shape. The variations in fruit shape are chromosomal in origin. Each different shape is due to the presence of one specific additional chromosome. Some of the different fruit shapes are shown in figure 9.13, and the additional chromosome that causes the particular shape is identified by a letter.

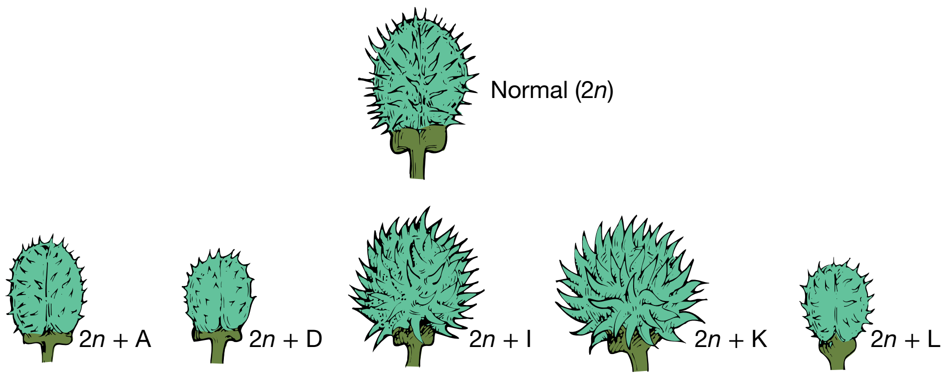



**FIGURE 9.13** Some variations seen in fruit shapes in thorn apple. The fruit of the normal diploid thorn apple is shown for comparison. In the case of an additional copy of the D chromosome, how has the fruit shape been affected?

## HOW TO MAKE A SEEDLESS WATERMELON

Seedless watermelons are the fruits of a sterile triploid ($3\times = 33$) hybrid plant. As such, these watermelon plants cannot reproduce, and viable seeds are not present in their fruits (see figure 9.14). Seedless watermelons were first marketed in the 1990s and they now dominate the market. No-one wants to bite into a watermelon with large hard black seeds!

**FIGURE 9.14** Seedless watermelons are triploid ($3\times = 33$). (Image courtesy of Judith Kinnear.)

To make a seedless triploid watermelon, two plants are needed:

1. A standard diploid ($2n = 22$) watermelon plant to be a source of pollen.
   This pollen is formed by the usual process of meiosis so that the pollen is haploid, with a single set of 11 chromosomes.
2. A tetraploid ($4\times = 44$) plant to be a source of eggs. Tetraploid watermelons are autopolyploids that are produced artificially by colchicine treatment. These eggs are formed by the normal process of meiosis, and each egg has half the number of chromosomes as the tetraploid parent. Each egg of course has 22 chromosomes, the diploid number, and they comprise two sets of watermelon chromosomes.

To make a triploid watermelon plant, artificial pollination is required. Pollen grains, each with a single set of 11 chromosomes, are dusted on the stigma of a female flower of the tetraploid plant. In the ovary of this plant are diploid eggs, each with two sets of chromosomes, 22 chromosomes in total. The resulting fertilisation produces seeds, each with three sets of chromosomes, totalling 11 + 22 = 33 chromosomes (see figure 9.15).

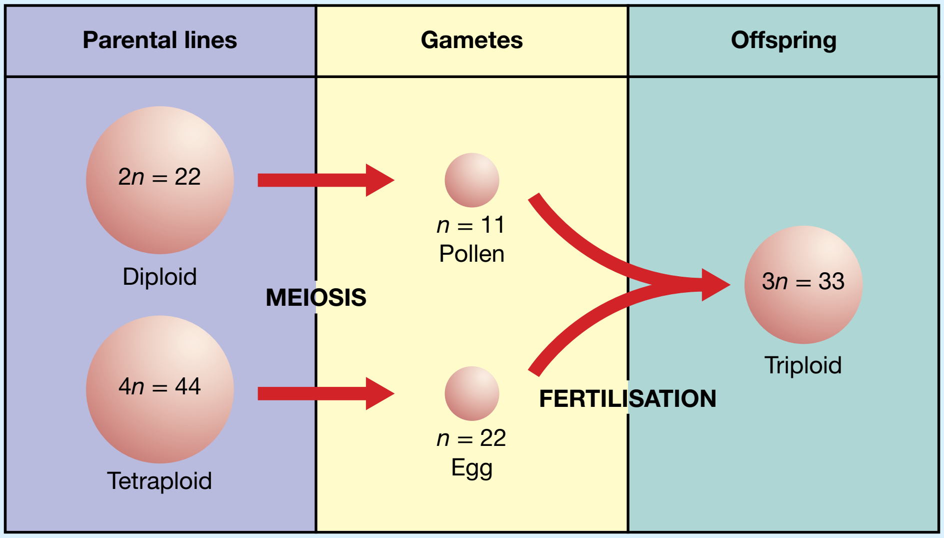



**FIGURE 9.15** Diagram showing a simplified version of the process required to produce triploid watermelons with seedless fruit. The tetraploid ($4x = 44$) parent watermelon is an autopolyploid that is produced artificially using colchicine treatment.

When three copies of each chromosome are present, as in a triploid watermelon, the process of meiosis cannot proceed in an orderly manner. The seedlings that germinate from these triploid seeds cannot produce viable eggs and pollen, and their fruits have small soft white traces instead of large hard black seeds.

Each new generation of triploid watermelon plants must be created anew by seed companies. These companies maintain diploid and tetraploid parental lines of watermelons, which they must pollinate by hand to produce more triploid seeds.

## KEY IDEAS

- Variation can exist among members of a population in the wild and under domesticated conditions.
- Inherited variation in a trait may result from the action of different alleles of one gene or from the action of polygenes.
- Variation in a trait may be classified as continuous or discontinuous.
- Polyploidy, either spontaneous or artificially created, results in new chromosomal combinations and new species.
- Aneuploid changes in chromosome numbers are generally poorly, if at all, tolerated in animals but can be tolerated in polyploidy plants.

## QUICK CHECK

3 An inherited trait in a population shows two variations only. How many genes are likely to be involved?

4 A population of cattle varies in coat colour, being red, roan or white. Is this variation continuous or discontinuous?

5 What is the difference between the members of the following pairs:
   a aneuploidy and polyploidy
   b continuous and discontinuous variation
   c autopolyploid and allopolyploid organisms
   d tetraploid and hexaploid?

6 Identify the following statements as true or false:
   a Polyploidy is far more common in plants than in animals.
   b Salmon is an example of a polyploidy animal.
   c A triploid plant would be expected to be fertile.
   d Hybrids are the result of crossing two different species.

# Genes in populations

A **population** refers to an interbreeding group of organisms of the same species living in the same region at the same time. So, we can talk about the present population of lions (*Panthera leo*) on the Serengeti Plains or the human population of Milan during the period of the Black Death. A population may be as small as a group of one species on a remote island or isolated on a mountain top or it may be as large as the gene pool of one species on a continent.

Each individual in a population has a total set of genetic information known as its genotype. For example, in sheep, the gene that controls wool colour has the alleles white (***W***) and black (***w***), with white being dominant to black. Sheep genotypes at this gene locus are expressed as a combination of two alleles: for example, either ***WW*** or ***Ww*** for a white sheep, and ***ww*** for a black sheep (see figure 9.16a).

## Populations have gene pools

An individual organism has a genotype that identifies the alleles present at each gene locus. A population has a **gene pool** that **consists of the sum of the genotypes of all the members of that population**. Gene pools are described by the frequencies (proportions) of the alleles of each gene present.

To understand allele frequencies, look at figure 9.16b and work through the following information, which will give us the frequencies of the alleles for the gene for wool colour in sheep:

1. There are ten sheep, each with two alleles, so the total number of alleles is $10 \times 2 = 20$.
2. Using the genotypes of the sheep as shown, we can count the number of each allele:

   Number of ***w*** alleles = 6, and number of ***W*** alleles = 14.

3. So, the frequency of the ***w*** allele = 6/20 = 3/10 = 0.3, and the frequency of the ***W*** allele = 14/20 = 7/10 = 0.7.

The frequency of the allele that determines the dominant phenotype is typically denoted by the letter 'p' and that of the allele that determines the recessive phenotype is denoted by the letter 'q'.

So, for this limited gene pool, we can write: freq (***W***) = p = 0.7 and freq (***w***) = q = 0.3.

**studyon**

Unit 4 | AOS 1 | Topic 1 | Concept 1

**Gene pool**
Summary screen and practice questions

**studyon**

Unit 4 | AOS 1 | Topic 1 | Concept 1

**See more**
Exploring a gene pool

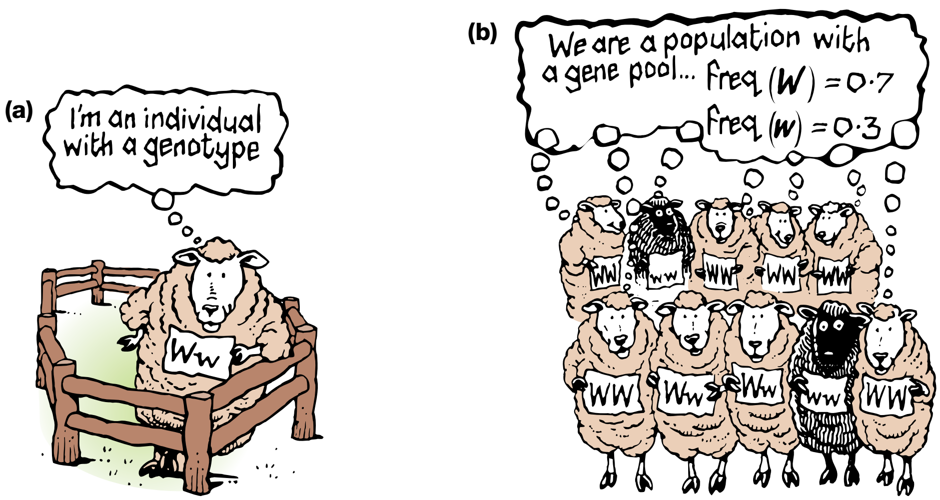



**FIGURE 9.16** **(a)** Genotype of individual sheep. **(b)** Gene pool of sheep population.

**eBook*plus***

**eLesson**
Genotypes
**eles-2459**

**eBook*plus***

**Interactivity**
Genotypes
**int-6370**

**eBook*plus***

**eLesson**
Phenotypes
**eles-2460**

**eBook*plus***

**Interactivity**
Generating the phenotypes
**int-6371**

Note that an allele frequency in a population can have any value from 0 to 1. When the frequency of an allele is equal to 1, all organisms in that population are homozygous at that particular gene locus and show the same phenotype for the trait concerned. When this happens, the allele is said to be **fixed** and, in this case, the frequency of that allele is 1 and that of the alternative allele must be 0. When an allele becomes fixed in a population, the genetic diversity of the gene pool is reduced by the loss of the alternative allele.

Figure 9.16b involves some sleight of hand. What is it? You cannot simply look at a white sheep in a mixed flock and identify its genotype with certainty — it could be ***WW*** or it could be ***Ww***. (No problem with black sheep as they can only be ***ww***.) In reality, in a population where two alleles exist with a simple dominant/recessive relationship, it is not possible to do a direct count of the two alleles such as we have just done above. However, a neat equation overcomes this difficulty. Refer to the box titled *Using the Hardy–Weinberg equation* on page 394.

## Allele frequencies tend to stay constant

In 1908, Godfrey Hardy (1877–1947), a British mathematician working at Cambridge, and Wilhelm Weinberg (1862–1937), a German physician, working independently, identified an important feature of populations. Hardy and Weinberg recognised that allele frequencies in populations tend to remain constant, generation after generation, provided a particular set of conditions is met. This concept is known as the **Hardy–Weinberg principle**.

**The conditions that apply in a population if allele frequencies are to remain constant over generations are as follows:**

- The population must be large.
- Members of the population must mate at random; **random mating** contrasts with non-random or assortative mating in which matings are restricted to those between organisms of like phenotypes, such as blue × blue and green × green. (In random mating, all kinds of crosses are equally likely, such as blue × blue, and blue × green.)
- All matings are equally fertile, producing equal numbers of viable offspring.
- The population must be closed, with no migration either into or out of the population.

Figure 9.17a shows a population in H–W equilibrium in which the allele frequencies for a particular gene remain constant, generation after generation. The ***B*** allele stays at a frequency of about 0.7 (that is, freq (***B***) = 0.7) in the gene pool, and the ***b*** allele at a frequency of about 0.3 (that is, freq (***b***) = 0.3). Figure 9.17b shows a population in which the allele frequencies are changing over generations.

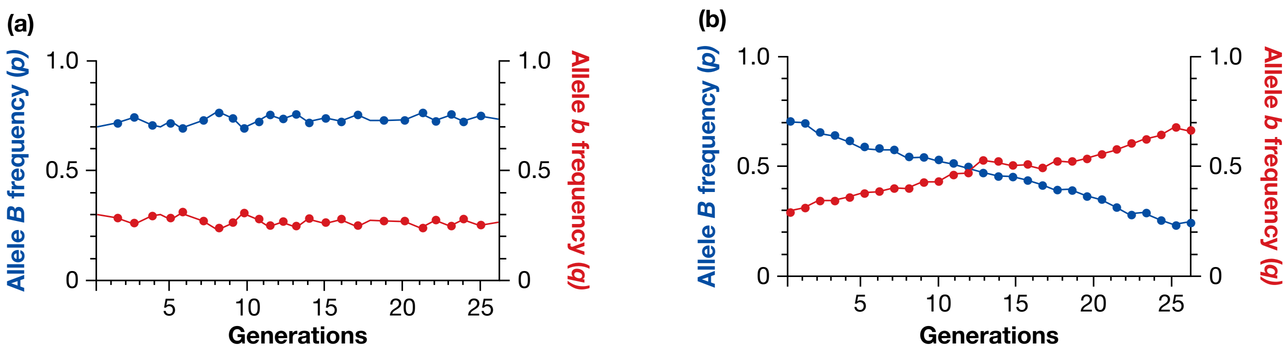



**FIGURE 9.17** Frequency of the ***B*** and ***b*** alleles at a particular gene locus over several generations in two populations (**a** and **b**) under different environmental conditions. Which population is in H–W equilibrium?

## USING THE HARDY–WEINBERG EQUATION

In this box, we are looking at variation in a population determined by two alleles of a particular gene, and where the relationship between the alleles is one of simple dominance/recessiveness.

The Hardy–Weinberg equation is: $p^2 + 2pq + q^2 = 1$.

1. What do the terms in this equation denote?
2. How can we estimate the values of the various terms?

1. **Identifying the terms:**
   We can take as a given that p and q denote the frequencies in the gene pool of the two alleles of a particular gene. In particular, p denotes the frequency of the allele producing the dominant trait, and q denotes the frequency of the allele controlling the recessive trait. We also known that $p + q = 1$.

Check out figure 9.18, which shows what the terms in the Hardy–Weinberg equation denote.

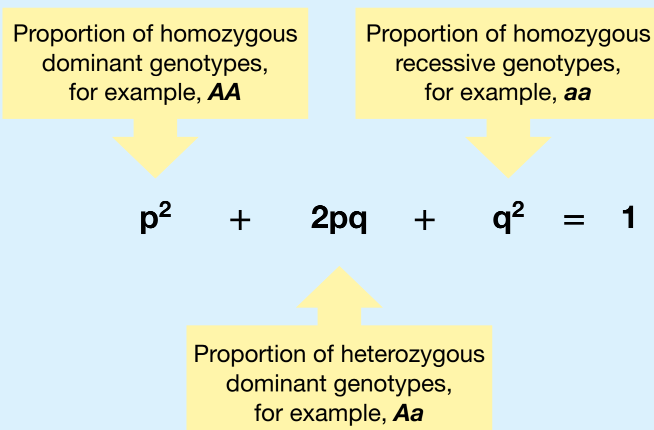



**FIGURE 9.18** The Hardy–Weinberg equation can be used to estimate the allele frequencies in a population and, from these values, to estimate probable proportions of unknown genotypes in a population.

2. **Estimating the value of terms:**
   The key starting value is $q^2$. This is the proportion of homozygous recessive individuals in the population under consideration. The homozygous recessive members of the population can be recognised because of their distinctive phenotype, and so they can be counted, making it possible to obtain a value for $q^2$.
   - Then, knowing $q^2$, it is easy to obtain the value of q.
   - Then, since $p + q = 1$, it is easy to obtain the value of p.
   - Finally, knowing both p and q, it is possible to calculate estimates of $q^2$ and 2pq.

   Now, you have all the details of the population, both the estimated allele frequencies, and the proportions of the three genotypes in the population.

3. **Applying the principles:**
   Consider a population (n = 100) that is tested for lactase persistence/nonpersistence. The results show that 36 individuals are lactase persistent and 64 are lactase nonpersistent. Lactase persistence, ***C***, is dominant to nonpersistence, ***c***.

   We are told that the 36 individuals in this population sample are lactase persistent, but how many of these are homozygous ***CC*** and how many are heterozygous ***Cc***?

   We can now write:

   Given the frequency of ***cc*** individuals $= q^2 = 64/100 = 0.64$

   Therefore, $q = \sqrt{0.64} = 0.8$

   Then, since $p + q = 1$, $p = 0.2$

   We already know that the frequency of individuals with the ***cc*** genotype $= q^2 = 0.64$, that is 64 per cent.

   Now that we have the allele frequencies, we can estimate the frequencies (proportions) of the two remaining genotypes in this sample population:
   - Frequency of individuals with the ***CC*** genotype $= p^2 = 0.2 \times 0.2 = 0.04$, that is 4 per cent are homozygous persistent.
   - Frequency of individuals with the ***Cc*** genotype $= 2pq = 2 \times 0.2 \times 0.8 = 0.32$, that is 32 per cent are heterozygous persistent.

   Note that the majority of individuals who are lactase persistent are heterozygous ***Cc*** and carry the nonpersistent allele.

   So, we can show the lactase status of our sample population as a piechart:

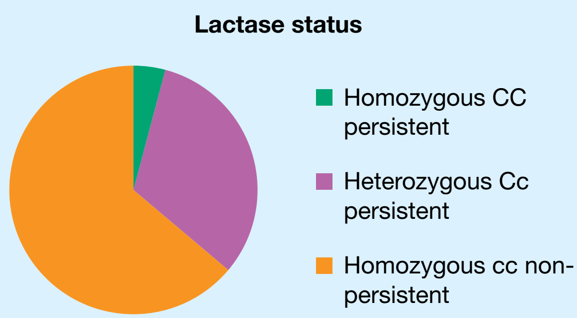



**FIGURE 9.19**

## KEY IDEAS

- The genetic information in an individual is a genotype, while the genetic information in a population is a gene pool.
- A gene pool is described in terms of the allele frequencies (proportions) of each gene.
- Allele frequencies in populations in Hardy–Weinberg equilibrium tend to remain constant generation after generation.
- A population in Hardy–Weinberg equilibrium will be expected to be large, closed and showing random matings, with all matings equally fertile.

## QUICK CHECK

7 Identify what is key between a genotype and a gene pool.

8 A population consists of 50 fish.
  - a What is the total number of alleles for one autosomal gene in this population?
  - b How many different alleles could be present in one fish?

9 Two large populations (M and N) of the same mammalian species contain some members that are pink-eyed and some that are yellow-eyed. Eye colour is under the control of an autosomal gene. Organisms in population M tend to mate with organisms with the same eye colour. However, this mating preference does not occur in population N. One of these populations is in Hardy–Weinberg equilibrium. Explain which population (M or N) this is likely to be.

10 List two conditions that apply to a H–W population.

11 True or false? A closed population is one in which migration, both in and out, is absent.

# Change agents in populations

studyon

Unit 4 | AOS 1 | Topic 1 | Concept 8

**Outcome of changes in allele frequencies**
Summary screen and practice questions

The allele frequencies in a population in Hardy–Weinberg equilibrium tend to remain unchanged, generation after generation. However, in reality, many populations show changes in allele frequencies in their gene pools over time. What agents can bring changes in allele frequencies of these populations?

Agents that can cause allele frequencies to change over time include:

- selection
- gene flow or migration
- chance, as seen in genetic drift, bottlenecks and founder effects.

Let's look at each of these change agents in turn.

## Natural selection as an agent of change

Selection that acts on a population can be of two kinds, depending on whether the population that is being acted upon is living under natural conditions or is a captive population. The two kinds of selection are natural selection and artificial selection. In the following section, we will explore natural selection. Artificial selection is discussed later in this chapter (see page 423) in the context of selective breeding.

The action of selection on populations in the wild is termed **natural selection**. We can define **natural selection as the process by which new heritable traits, whether morphological, physiological or behavioural, evolve and persist in a population**.

Natural selection is brought about by a range of **selecting agents** that act on populations. Examples include:

- physical agents, such as climate change or food shortages
- biological agents, such as infectious diseases, predation or competition
- chemical agents, such as a pollutants in soil or water.

studyon

Unit 4 | AOS 1 | Topic 1 | Concept 5

**Change agent: natural selection**
Summary screen and practice questions

studyon

Unit 4 | AOS 1 | Topic 1 | Concept 5

**See more**
Change agents and natural selection

In the wild, members of a population must compete with each other for access to living space, energy supplies and mating partners in their habitat. Members of a population are also exposed to competition from other species, to predation, to parasites and to disease-causing microorganisms. Variants within a population react differentially to these pressures, with some being more successful than others. **Natural selection occurs when any selecting agent acts on a population in the wild and produces differences in the survival and reproduction rates of variants in that population**.

**Differential reproduction** occurs when one variant (phenotype) in a population produces more viable offspring than other variants, and so makes a greater contribution to the gene pool of the next generation. The action of natural selection over successive generations results in an increase in the frequency of this phenotypic variety in the population relative to other varieties. In other words, the frequencies of alleles present that are associated with this phenotype will increase in the gene pool of the population as the frequencies of alternative alleles decrease.

Darwin and Wallace (see chapter 10, p. 444, The Darwin–Wallace view) recognised natural selection as a causative agent of evolution, based on a number of simple observations and resulting deductions, as follows:

*Observation 1*
In the wild, the numbers of offspring produced by plants and animals over their lifetimes is greater than the number of parents. For example, a single *Eucalyptus* tree produces an enormous number of seeds every year and one female rabbit produces a large number of kits over her lifetime.

*Deduction 1*
A struggle for survival occurs.

*Observation 2*
Over time, the size of natural populations tends to remain fairly constant, although population size may fluctuate depending on environmental factors such as drought, disease and plentiful food supplies.

*Deduction 2*
In this struggle, some variants in a population have a greater chance of survival than others. These variants are favoured under the conditions in a particular environment and are reproductively more successful.

*Observation 3*
Variation exists in populations of plants and animals; that is, no two organisms in a population are identical, and some variations are inherited.

*Deduction 3*
Inherited traits present in surviving parents are passed on to their offspring so that the genetic composition of populations can change over time.

The effects of an agent of natural selection acting on a population depend on the time over which it acts. Over a few generations, natural selection can cause a variant that is at a selective advantage to become more common in a population. When the relative numbers of different variants in a population change over a few generations as a result of natural selection, the outcome is **intra-specific evolution** or evolution within a species. Over longer periods, natural selection can result in the formation of subspecies (see figure 9.31 on page 409); and, over geological time, natural selection can result in the formation of new species.

## Selection acts on phenotypes

Members of a polymorphic population living *under a particular set of environmental conditions* are exposed to a range of selecting agents and the selection pressures that they produce. Under these circumstances, the various phenotypes (variants) may show different survival and reproductive rates. **A phenotype that makes the greater contribution to the gene pool in the next generation has a higher fitness value and is said to be 'at a selective advantage'.** Phenotypes that make lesser contributions to the gene pool of the next generation are 'less fit' or are said to be 'selected against'.

**An observed fitness value for a phenotype is not fixed, but applies under a particular set of environmental conditions**. Table 9.3 shows examples of how different agents of natural selection can act on populations and give a selective advantage to one variant over another under a particular set of environmental conditions.


**FIGURE 9.20** Natural selection acting on inherited variations in height can result in differential survival of variants. Who will survive — shorter or taller plants?

**TABLE 9.3** Action of natural selection in populations with members showing inherited variations. Which plants are at a selective advantage — the faster or the slower growers? In the absence of viral disease, what happens to the selective advantage of resistant rabbits?

| Variants present in population | Agent of natural selection | Outcome of natural selection |
|---|---|---|
| plants with a noxious chemical and plants lacking chemicals | grazing by herbivores | chemical-producing plants are more likely to survive and reproduce than those lacking chemicals (their leaves and reproductive structures (flowers and fruits) are more likely to be eaten by herbivores) |
| plants with faster growth rate and plants with slower growth rates | crowding and limited space | faster growing seedlings are more likely to survive and reproduce than slower growers (that are shaded from light and die) |
| viral-resistant rabbits and non-resistant rabbits | presence of virus in environment | resistant rabbits are more likely to survive and reproduce than non-resistant rabbits (that will be incapacitated or killed by viral disease) |

Consider the following example of natural selection in action on a plant population that consists of a small number of copper-tolerant plants and a majority of copper-sensitive plants. Under normal conditions, this inherited variation does not affect the survival or numbers of offspring produced by each plant variant. What happens if the environment changes and the plant population is exposed to the pressure of high concentrations of copper salts in the soil because of seepage from wastes at a copper mine?

The copper-sensitive plants come under strong selection pressure and will be harmed or killed, but the copper-tolerant plants will survive and reproduce at relatively higher rates. Over time, if high levels of copper remain in the soil, the population will eventually consist of copper-tolerant plants only. *Under this particular set of environmental conditions*, the copper-tolerant plants have a selective advantage relative to the copper-sensitive plants in the population. Table 9.4 summarises this situation.

**ODD FACT**

A study of copper-tolerant plant species growing on sites of copper mine waste found that some copper-tolerant species excluded copper from their shoots by immobilising it in their roots. Neat!

**TABLE 9.4** Possible results in a plant population of continuous exposure to copper.

| Outcome of introduction of copper pollution | |
|---|---|
| variations present in plant population | copper-resistant and copper-sensitive plants |
| environmental change | introduction of copper pollution |
| agent of natural selection | environmental change |
| plants with selective advantage | copper-resistant plants |
| plants with selective disadvantage | copper-sensitive plants |
| outcome of natural selection over several generations | copper-resistant plants become more common |

Now consider the following example of natural selection in action on an insect population.

Figure 9.21a shows a large random mating insect population that includes a small proportion of members that are resistant to a particular pesticide (these insects are shown in red), with the majority being sensitive to the pesticide. Assume that this variation is due to the action of one gene with the alleles ***r*** (sensitive) and ***R*** (resistant). In an environment where there is no pesticide, both the resistant and the sensitive phenotypes survive and reproduce equally. The allele frequencies, as well as the genotype and phenotype frequencies, would be expected to stay constant generation after generation. Why?

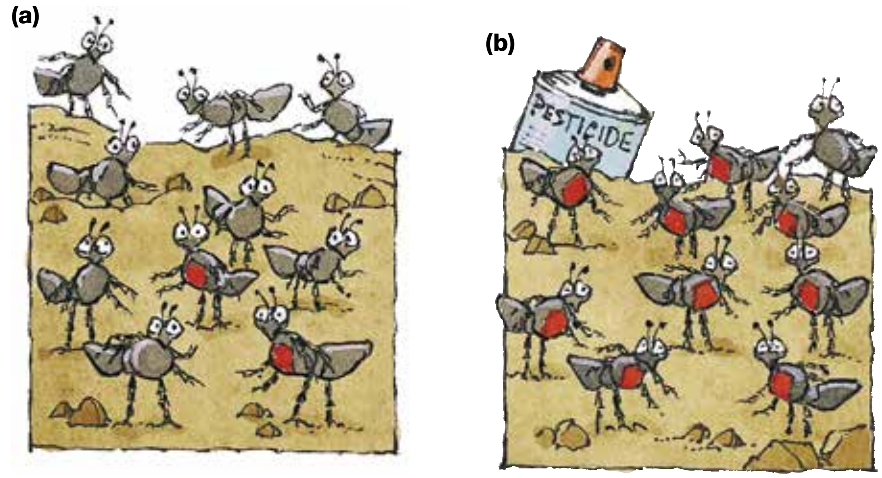



**FIGURE 9.21** **(a)** Sample of insect population in pesticide-free environment. The sensitive phenotype (shown as plain grey) is more common than the resistant phenotype (shown as red). **(b)** Sample of insect population after several generations in the presence of pesticide. The resistant phenotype (red) has become more common. Have the allele frequencies changed?

However, assume that the insects' habitat is cleared for farming and pesticide is introduced into the environment. Under these conditions, many of the sensitive insects are killed by the pesticide spray. The resistant insects are now at a selective advantage and **will survive to form the majority of parents to produce the next generation**. What is the agent of selection?

After several generations, with pesticide in the environment, the population changes so that resistant insects form the majority of the population (see figure 9.21b). The pesticide does not cause the resistant phenotype. Resistant insects were present in the original population, and, in the presence of the pesticide, their proportion increased through the action of selection.

## Genetic fitness varies in different environments

In human history, important selecting agents affecting populations have included widespread infectious diseases that are major causes of death. Malaria is one such disease (see figure 9.22). For people living in areas of the world affected by the malarial parasites, their fitness depends on whether they are able to produce haemoglobin S. The parasite cannot survive in red blood cells containing haemoglobin S, which is due to the $\boldsymbol{H^S}$ allele. However, the malarial parasite thrives in normal red blood cells, which contain haemoglobin A, produced through the action of the $\boldsymbol{H^A}$ allele, so this group is at risk of death from malaria. On the other hand, when red blood cells contain only haemoglobin S, this causes the disease sickle-cell anaemia (see table 9.5).

**FIGURE 9.22** Coloured scanning electron micrograph (SEM) image of red blood cells. The lower cells are infected with malarial parasites as indicated by the yellowish swollen lumps on the cells, which eventually rupture and spread the infection. The lumps cause the cells to adhere in the brain and other tissues, producing symptoms of malaria, such as recurrent fevers. The normal cells, above, lack these lumps because a disrupting gene in the parasite has reduced their ability to adhere.

Table 9.5 shows how genetic fitness can vary in different environments. Note that the $\boldsymbol{H^AH^S}$ individual is at a selective advantage relative to the others in a malaria-affected environment, but loses that advantage in a malaria-free environment.

**TABLE 9.5** The fitness of a phenotype depends on the environment. Is the fitness of people with normal red blood cells containing only haemoglobin A similar in the two environments?

| Genotype | Phenotype | Relative fitness |
|---|---|---|
| *In malaria-affected environments:* | | |
| $\boldsymbol{H^AH^A}$ | usual red blood cells with haemoglobin A only | can die from malaria *reduced fitness* |
| $\boldsymbol{H^AH^S}$ | cells capable of sickling with haemoglobins A and S | minor effects of sickling but resistant to malaria *most fit* |
| $\boldsymbol{H^SH^S}$ | severe sickle-cell anaemia with haemoglobin S only | resistant to malaria but can die from sickle-cell anaemia *least fit* |
| *In malaria-free environments:* | | |
| $\boldsymbol{H^AH^A}$ | usual red blood cells | no effects of sickling *most fit* |
| $\boldsymbol{H^AH^S}$ | cells capable of sickling | minor effects of sickling *reduced fitness* |
| $\boldsymbol{H^SH^S}$ | severe sickle-cell anaemia | can die from sickle-cell anaemia *least fit* |

## Levels of selection

The levels of selection against phenotypes in a population can vary.

- **Complete selection** against a phenotype occurs when any organism with a given phenotype cannot reproduce because of death before reproductive age is reached or because of sterility.
- **Partial selection** against a phenotype occurs when matings involving that phenotype produce on average fewer viable and fertile offspring relative to other matings.

## Case studies in selection

The following section includes details of two examples of natural selection in an Australian setting.

### Case study 1: Selection in cattle tick populations

The cattle tick (*Boophilus microplus*) is a major pest in northern Australia (see figure 9.23). It carries the parasite that causes tick fever in cattle.

**FIGURE 9.23** Cattle showing a heavy infestation of ticks. (Image courtesy of CSIRO Publishing.)

The first attempt to control ticks used arsenic-based cattle dips. At first, this killed many ticks. By 1936, the tick populations were showing resistance to this poison so arsenic use gradually disappeared. By 1946, the insecticide DDT was introduced to control cattle ticks and, at first, this reduced the tick populations. By the early 1960s, however, the tick populations had become resistant to DDT insecticide. The next chemicals brought into the attack on cattle ticks were organophosphate pesticides. Soon after these were introduced, the first resistant ticks were detected and, by the 1980s, resistance to many of these compounds had developed in tick populations. How?

Resistance to each chemical pesticide involves a different gene. Before a particular pesticide was widely used, the frequency in the cattle tick gene pool of the various alleles that confer resistance to these pesticides was low. Once a pesticide became widely used, however, it acted as a selecting agent that produced an increase in the survival and reproduction rates of resistant ticks relative to those of nonresistant ticks. As a result of this selection pressure, the proportion of resistant ticks increased and the frequency of the resistant allele in the gene pool of the population increased. This recurred in turn with different alleles as each new pesticide was introduced. A pesticide does not produce the resistant allele. The resistant allele is already present in the gene pool of the tick population. So, the pesticide positively selects resistant members of the tick population, with the result that resistant members have a greater genetic fitness than the nonresistant variety and, over a few generations, will dominate the tick population.

Other control measures for cattle tick infestation are now being used, including strict controls on the movement of cattle from tick-infested regions to tick-free areas.

studyon

Unit 4

AOS 1

Topic 1

Concept 5

**See more**
Ticks

## Case study 2: 1080 resistance in native mammals

In the southwestern corner of Western Australia, several species of native plants produce sodium fluoroacetate, a poison known commercially as 1080 (ten-eighty). These plants include species within the genera *Gastrolobium* and *Oxylobium*(see figure 9.24). These genera are members of the pea family (Papilioaceae) and typically grow as perennial shrubs.

**FIGURE 9.24** One of Western Australia's native toxic plants, sandplain poison *(Gastrolobium microcarpum)*, that produces fluoroacetate poison, a substance that can cause the death of introduced farm stock.

Populations of the common brushtail possum, *Trichosurus vulpecula*, include some living in regions where these toxic species grow. When their response to fluoroacetate is measured, these possums are found to be resistant to the poison. Other possum populations live outside the range of fluoroacetate-producing plants. Not surprisingly, when these possums are tested for resistance to fluoroacetate, they show symptoms of poisoning at low concentrations. Data on the resistance of native mammals to 1080 poison are shown in table 9.6.

The presence of fluoroacetate in the diet of marsupials living within the range of the toxic plants has **acted as an agent of selection so that, over time, resistant possums have higher survival and reproduction rates than nonresistant possums**. Because of this increased genetic fitness of resistant possums in the presence of fluoroacetate, the allele frequencies in the possum gene pool have changed over several generations so that the resistant allele has increased in frequency.

**ODD FACT**

When myxomatosis virus was first released via mosquito vectors to control rabbits in 1950, the death rate for rabbits in the Corowa region of New South Wales was more than 99 per cent but by the second year had dropped to 85 per cent. Why?

**TABLE 9.6** Approximate mean $LD_{50}$ for fluoroacetate poison for different populations living in environments either with exposure to the toxin or without known exposure. $LD_{50}$ is a measure of toxicity: the higher the number, the more resistance to the poison.

| Animal group | Prior exposure to 1080 $LD_{50}$ (mg/kg) | No prior exposure to 1080 $LD_{50}$ (mg/kg) |
|---|---|---|
| introduced herbivores: rabbit, goat | n/a | 0.3 |
| introduced carnivores: fox, cat | n/a | 0.2 |
| native marsupial herbivores: possum, wallaby | 42.0 | 0.3 |
| native marsupial carnivores: quoll, antechinus | 8.3 | 2.7 |

Table adapted from McIlroy, J.C., 1986, *Australian Wildlife Research*, Vol. 13, p. 39.

## KEY IDEAS

- Various phenotypes in a population may differ in their fitness value.
- The fitness value of phenotypes varies according to the environment.
- Selecting agents act on populations and change allele frequencies.

## QUICK CHECK

**12** A herd of wild mountain goats moves permanently to a higher region and over many generations shows an increase in average body size. If this change is a result of selection, is it natural or artificial? Explain.
**13** What is meant by 'fitness' in a genetic sense?
**14** In a plant population, some phenotypes survive periods of drought better than other phenotypes. What might be expected if this population was exposed to drought conditions for several generations?
**15** Would complete selection have a greater effect against a dominant trait in a population or a recessive trait? Explain.

## Gene flow as an agent of change

**study on**

Unit 4 | AOS 1 | Topic 1 | Concept 7

**Change agent: gene flow**
Summary screen and practice questions

In the first half of the twentieth century, the allele for beta thalassaemia was very rare in the populations of Australian capital cities. Today, however, this allele is common. What has caused this change in allele frequency?

**Gene flow**, also known as **migration**, can change the allele frequencies of a population. Unlike selection, which requires several generations to have an effect, **changes due to migration can occur very quickly**.

Regional human populations have differences between their gene pools that reflect the long-term effects of natural selection in the prevailing environmental conditions, as well chance events and gene mutations. The gene pool of western Europeans has a higher frequency of the allele for cystic fibrosis compared to the gene pool of southern Europeans. The gene pool of that population has a relatively higher frequency of the allele for beta thalassaemia. The **immigration** of groups of people from one region to a new population in a different region can introduce new alleles into the gene pool of the new population or can alter its existing allele frequencies. This is what occurred in the post–World War II period when large-scale migration to Australia, including that from southern European countries such as Italy and Greece, resulted in an immediate increase in the frequency of the beta thalassaemia allele in the gene pool of the Australian population. At the same time, this mass immigration produced a decrease in the frequency of the allele for cystic fibrosis in the same gene pool.

**FIGURE 9.25** Photo showing Ron Cooke, his wife and 10 children on Southampton Dock prior to boarding a ship in early 1969 to emigrate to Australia. More than one million Britons emigrated to Australia on an assisted scheme between 1947 and 1982. Immigration or gene flow can have an immediate effect on the gene pool of a receiving population. Emigration may also have an effect on the gene pool of the source population.

**ODD FACT**

In the period from 1945 to 1975, the population of Australia increased from about 7.5 million to 13 million, and this increase included about 3 million migrants and refugees.

Emigration can also change allele frequencies if the emigrant group is not a representative sample of the original population. Imagine a small hypothetical population that comprises mainly blue-furred organisms, both homozygous ***BB*** and heterozygous ***Bb***, and a few homozygous ***bb*** pink-furred organisms. If a group emigrates from that population that is not representative of its existing allele frequencies, the gene pool of that population will immediately change. If, for example, all the pink-furred organisms emigrate, the frequency of the ***b*** allele will decrease. Because the pink-furred phenotype is no longer present in the population, it may appear that the ***b*** allele has disappeared. However, this allele is 'hidden' in the remaining heterozygous ***Bb*** blue-furred organisms and can reappear in the next generation as a result of the mating of two heterozygotes.

Gene flow can increase the diversity of a gene pool. If, for example, the group of pink-furred organisms emigrates to a new population that consists entirely of ***BB*** blue-furred organisms, the gene pool of their new population immediately becomes more diverse.

With the development of mass transport, human migration continues. As human groups move, their new environmental conditions may be very different from those in their place of origin. Phenotypes that were selected for in the 'home' environment may be disadvantageous to the migrants in their new setting, as seen in the migration of fair-skinned Celtic people from Ireland and Scotland to Australia.

In the northern high-latitude home environment of Celtic people, sunlight is not intense, even in summer, and, during winter, the days are short. Fair skin is favoured by natural selection because it allows the absorption of sunlight that is needed for the synthesis of essential supplies of vitamin D. In the Australian environment, with its more intense sunlight, people with fair skin are poorly protected against UV radiation. People of Celtic ancestry are at higher risk of skin cancer than people who have heavily pigmented skin.

The forcible movement of people from west Africa to the United States as part of the slave trade introduced the sickle-cell $H^S$ allele into the American population. However, since malaria was not prevalent in the new environment, people carrying this allele were no longer at a selective advantage. Over time, the incidence of this allele in the African American population has decreased.

Australia is now a multicultural population. Each of the groups that migrated to Australia has brought with it specific alleles that are common in the home regions of each group. The high incidence of cystic fibrosis in Australia is due to the migration here of people from western Europe, where the allele that causes cystic fibrosis is common. Since World War II, migration of people from southern Europe, the Middle East, India and South-East Asia introduced the beta thalassaemia (***t***) allele into the gene pool of the Australian population. This allele gives a selective advantage to the heterozygous ***Tt*** individuals because this genotype protects against malaria. In the Australian setting, however, this heterozygote advantage no longer exists.

studyon

Unit 4 | AOS 1 | Topic 1 | Concept 6

**Change agent: genetic drift**
Summary screen and practice questions

## Chance as an agent of change

**Chance factors can cause allele frequencies in a population to change over time.** In this section we will explore chance factors that impact on the genetic make-up of a population, that is, change the gene pool of a population. These chance events include:

- genetic drift
- bottleneck events
- founder effects.

studyon

Unit 4 | AOS 1 | Topic 1 | Concept 6

**See more**
Genetic drift

### Genetic drift can change gene pools

When chance operates on the allele frequencies in a population, the direction of the change is unpredictable and can vary from one generation to

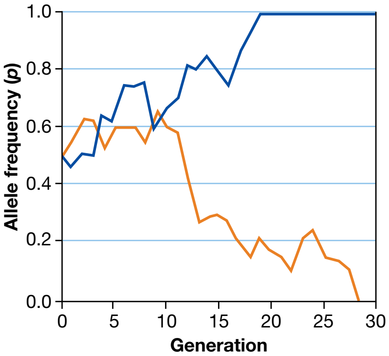



**FIGURE 9.26** Changes in allele frequency in replicate populations due to chance. Is the direction predictable from one generation to the next? What has happened in population number 1 (blue line) by the nineteenth generation? Now check population number 2 (orange line). Contrast what has happened in the period from generation 0 to 7 with what happened in the period from about generation 10 to 17.

the next. The resulting pattern of random change is known as **genetic drift** (see figure 9.26). The change in allele frequency from one generation to the next in genetic drift is random, and this is in marked contrast to the directional change that occurs in natural selection.

**The smaller the population size, the greater the potential impact of genetic drift**. Unlike the action of natural selection, genetic drift does *not* favour one allele over another; both are equally subject to being affected by genetic drift. In a very small population, genetic drift can lead to the decrease, and eventual loss, of favourable alleles from the gene pool. For this reason, when a species is reduced to one or a few small populations, the species is at great risk of extinction. It may become extinct in spite of conservation measures.

## Bottleneck effects can change gene pools

Bottleneck effects come into operation when the size of a population is drastically reduced for at least one generation. This reduction may be the result of a major disaster, such as a widespread bushfire or flood, or the introduction of a new disease to which the population has not previously been exposed. **The few survivors that reproduce to give the next generation may by chance be an unrepresentative sample of the gene pool of the original population**. This is known as the **bottleneck effect**.

An evolutionary bottleneck occurs when some disaster or disease causes a rapid decrease in the size of a large population. The original large population may have included a diverse set of alleles in its gene pool. The small post-disaster population is likely, by chance, to have a much less diverse set of alleles in its gene pool (see figure 9.27).

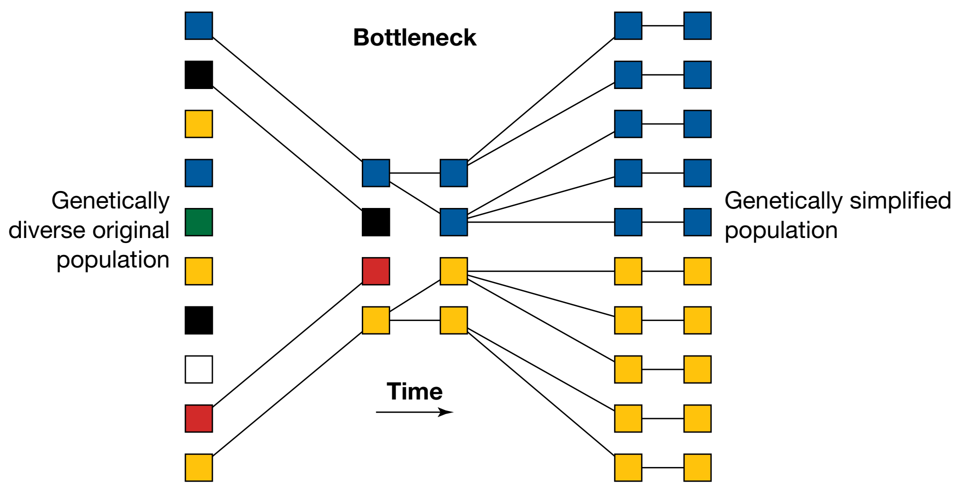



**FIGURE 9.27** Diagram illustrating the concept of an evolutionary bottleneck that occurs when the population size is very sharply reduced by a disaster of some kind. Note that the survivors at the bottleneck have only some of the genetic diversity present in the original population. Note also that, by chance, only some of these survivors contribute to the gene pool of the 'recovery' population that rebuilds after the disaster.

After the disaster has ceased, the population will rebuild. However, the allele frequencies can be very different from those in the original (pre-disaster) population, and the genetic diversity that is the population's insurance policy may be greatly reduced.

## Founder effects can change gene pools

A founder effect occurs when a new colony is started by a few members from a larger population. This small population size means that the gene pool of the new colony is highly likely (i) to have reduced genetic variation and (ii) to be a non-random sample of the original population (see figure 9.28).

To understand the influence of sample size, think about netting samples of fish from a large tank that contains 50 orange fish and 50 black fish. If a sample of only four fish is netted, this sample might comprise fish of one variety only. If the sample size is larger, say 14, the sample will almost certainly include fish of both varieties and be representative of the tank population.

*With a sample of 14, the chance that the fish will all be black is about one in 16 000.*

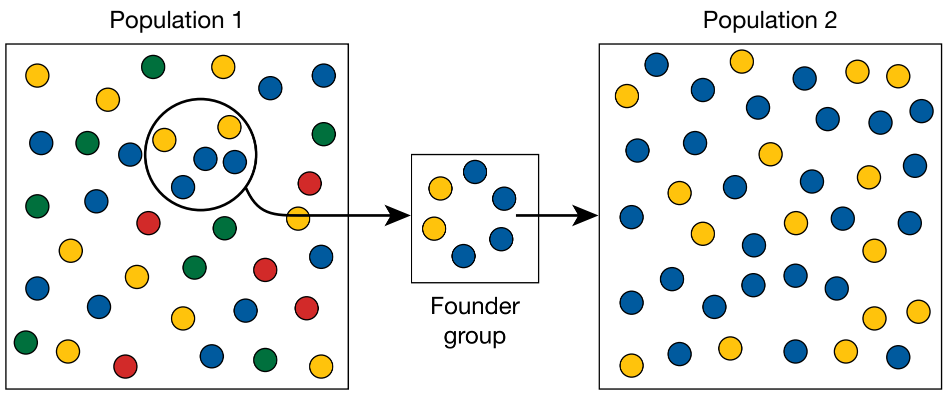



**FIGURE 9.28** Diagram showing the concept of the founder effect. Population 1 is more genetically diverse than the small founder group that leaves it. In consequence, population 2, which develops over time from the founder group, is also less diverse than population 1.

**FIGURE 9.29** A macaroni penguin. Is this a white-faced or a black-faced variant?

Populations of macaroni penguins (*Eudyptes chrysolophus*) live on subantarctic islands and on the Antarctic Peninsula (see figure 9.29). Most have black faces but a small proportion have white faces. On Macquarie Island, however, the population is composed almost entirely of the white-faced variety. How did this occur? It may be by chance that the small **founder population** of Macaroni penguins that first occupied Macquarie Island consisted of the white-faced variety only. So, **when a small unrepresentative or non-random sample of a population leaves to colonise a new region, this is known as the founder effect**.

Chance effects, including founder effect, can be significant in small populations, such as in zoos, where captive breeding programs must be carefully managed to avoid inbreeding and to maximise genetic diversity.

Founder effects may also be the cause of unusually high incidences of particular inherited diseases in a particular geographic area. One person comes into a region and reproduces there, introducing into the local gene pool the allele that determines a particular disease. Years later, or generations later, **the founder effect is made visible in a cluster of a particular disease in a restricted region**. One Australian example can be seen with Huntington's disease (HD).

A medical historian has shown that virtually all the families in Tasmania, and in the south-eastern states of Victoria and South Australia who are affected by HD can trace their family trees back to an English woman who migrated to Tasmania in 1842 with her husband and seven children. This woman, 'Mary' was born in 1806 in Somerset, England, and she inherited the HD allele from her father. Mary and her husband had a further seven children who were born in Tasmania. Nine of Mary's fourteen children developed HD, and all of these produced children.

A similar founder effect lies behind the high prevalence of a form of early-onset Alzheimer disease in the town of Yarumal and surrounding district in Colombia, South America. About 5000 people presently carry a particular mutation that causes this disease that typically shows its effects when people are in their early forties. This early-onset form of Alzheimer disease is a dominant trait. Researchers analysed the sequence of the DNA segment involved in the affected people and found that all these people carried the same mutation. It has been estimated that this particular mutation was introduced to the population about 375 years ago. Further, it has been concluded that this allele was brought into the district by a Spanish conquistador in the early sixteenth century. This conquistador did not remain in the district but, as a result of his sexual activities, his genetic material remained through at least one child whom he fathered with a local woman. This child survived to pass this DNA segment to the next generation. As mentioned, about 5000 people are now known to carry this DNA and will develop the destructive symptoms of early-onset Alzheimer disease.

## BIOLOGIST AT WORK

### Dr Kate Charlton-Robb — researching a population

I am a dolphin biologist and conservation geneticist. My work involves researching populations of dolphins in southern Australia and takes me into the field, into museums across Australia and into laboratories to process and analyse dolphin DNA. I was the principal investigator in the discovery of a new dolphin species that is endemic to southern and south-eastern Australia. I was fortunate enough to name the new dolphin, *Tursiops australis*, with the common name Burrunan dolphin, following aboriginal narrative. Only four new dolphin species have been discovered since the late 1800s. This was not only incredibly exciting but also quite a scientific journey for me.

My research began in 2003 when, as a Bachelor of Science undergraduate student at Monash University, I examined different techniques for obtaining skin samples from live dolphins and extracting and assessing their DNA. I completed my Honours year working with the dolphin population in Port Phillip Bay, Victoria. I used a common biopsy sampling technique that let me take a small piece of skin using a biopsy dart at a distance from the dolphins. This technique is not harmful to the dolphin and, in many cases, the dolphins were inquisitive about the dart floating in the water after the biopsy had been collected. The skin samples allowed me to investigate many things about the dolphins. Using their DNA signatures, previously unknown information about the dolphins can be revealed, including kinship relationships, gender, movement and breeding patterns and importantly the genetic variability and thus the genetic health of the dolphins.

During this initial project we discovered that their mitochondrial DNA differed substantially from that of the other two recognised bottlenose dolphin species, *Tursiops truncatus* (common bottlenose dolphin) and *Tursiops aduncus* (Indo-Pacific bottlenose dolphin). The discovery of unique mtDNA first highlighted that the smaller 'inshore' dolphins in Port Phillip Bay and in the Gippsland Lakes might potentially represent a new dolphin species. In 2006 I received the Dean's Postgraduate Research Scholarship to continue my research at Monash University as a part of my PhD studies. During this time I investigated numerous other DNA regions and, at each of these regions, these dolphins differed from the other bottlenose dolphin species. However, the taxonomic description of a new dolphin species requires that more than one line of evidence be presented (not just DNA), so I investigated the dolphins' skulls, colouration and external morphology (body, beak and flipper length, etc.). We then used numerous statistical tests to compare our findings with those of the other recognised bottlenose dolphin species. Our multiple lines of evidence showed that the 'inshore' dolphins were in fact different enough to represent a new dolphin species.

This work involved many years of dedicated research during which I collaborated with many fascinating researchers, from specialists in ancient DNA to taxonomists, and learned many different fields of science. Along with my collaborators, we put forward a taxonomic paper 'A new dolphin species, the Burrunan dolphin *Tursiops australis* sp. nov., endemic to southern Australian coastal waters'. This paper formally identified, described and named the new dolphin species and was scrutinised by our scientific peers and The International Commission on Zoological Nomenclature (ICZN), which is responsible for the International Code that governs the naming of animals. We were delighted that in 2011 our paper was accepted and published in a leading scientific journal, *PLoS ONE*.

I love my work because I get to be out on the water with these amazing animals, work with museum collections dating back to the early 1900s and use applied genetic technologies. The Burrunan dolphin is now listed as 'Endangered' under Victoria's Flora and Fauna Guarantee Act, and I continue my research on the Burrunan dolphin as Head of Research at the Australian Marine Mammal Conservation Foundation. My main aim for future research is to find out where else these dolphins occur in Australia, to further assess the genetic variability, population demographics and, most importantly, to work to conserve and protect this amazing new species for future generations to enjoy.

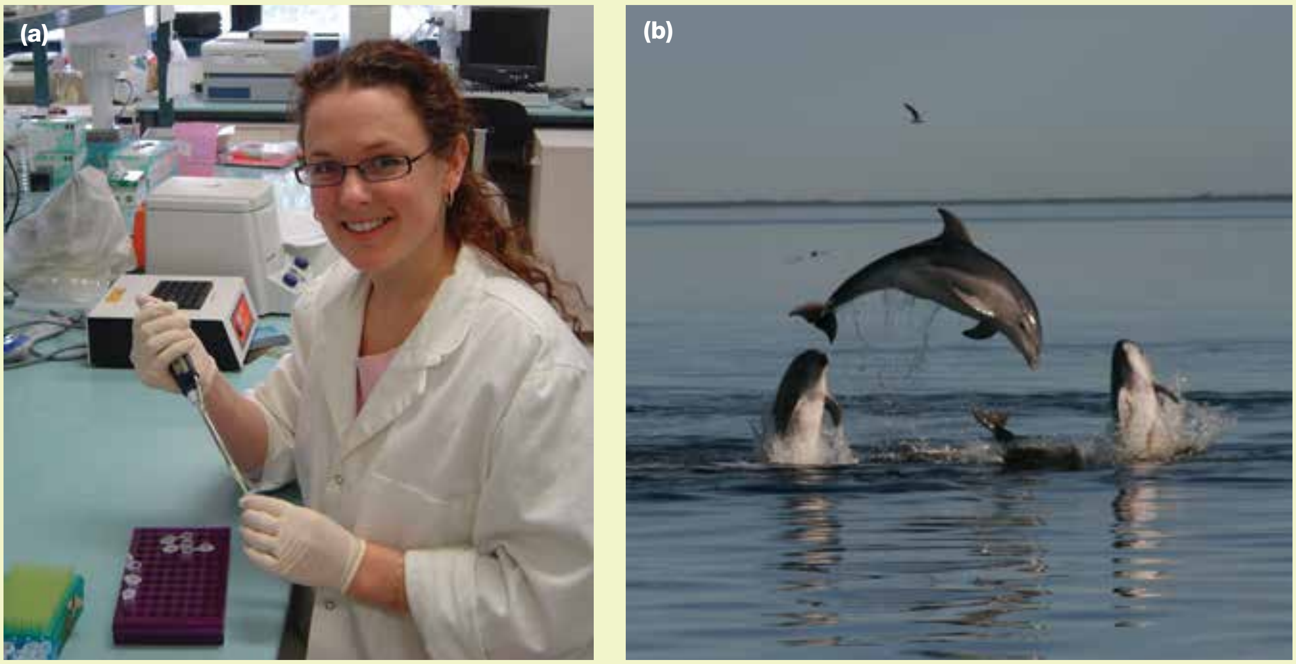



**FIGURE 9.30** **(a)** Kate Charlton-Robb during her PhD at Monash University. **(b)** Burrunan dolphins (*Tursiops australis*) in Port Phillip Bay, which were recognised as a new species in 2011. (Images courtesy of Dr Kate Charlton-Robb.)

## KEY IDEAS

- In addition to selection pressure, other agents of change in populations are gene flow and several chance factors.
- Gene flow can produce an immediate change in the gene pools of populations.
- Chance factors that can alter the gene pool of a population are genetic drift, bottleneck events and founder effects.
- Genetic drift effects on allele frequencies in gene pools are not directional, but are random.
- Gene flow, such as the migration of a group into an existing population, may result in an increase in the diversity of the receiving population.
- Bottleneck events typically result in a decrease in the diversity of the gene pool of the recovery population.
- Founder effects occur when a small sample of one population becomes the foundation of a new population.

## QUICK CHECK

**16** The frequency of a particular allele in the gene pool shows a steady decline over each successive generation. Is this change the result of genetic drift? Briefly explain.

**17** Bottleneck events and founder effects are both chance factors that can act on the gene pool of a population.
- **a** Identify one way in which they are similar.
- **b** Identify a key difference between these two factors.

**18** Identify the following statements as either true or false:
- **a** Genetic drift is expected to produce an increase in the gene pool of those alleles that confer greater genetic fitness.
- **b** Gene flow can produce an immediate increase in the diversity of a gene pool.
- **c** The smaller the population size, the greater the impact of genetic drift.
- **d** Bottleneck effects can occur when a population undergoes a sudden decrease in size.
- **e** A high prevalence of an inherited disease in a restricted region might be explained as being due to a founder effect.

# Speciation

We have seen how the gene pools of large populations living under particular environmental conditions can change in response to selection pressures acting on the different variants present in the population. In small populations, change agents such as genetic drift can also produce changes in their gene pools.

So, populations of the same species living under different environmental conditions are likely to be subject to different selection pressures so that the allele frequencies in their gene pools diverge. In addition, any new alleles added by mutation to these gene pools are likely to be dissimilar.

The Darwin–Wallace theory of evolution by natural selection proposes that populations of the same species living in different locations under different conditions can evolve in different directions. Over many generations, under different selection pressures, these populations can become increasingly different from each other in structure, physiology and behaviour. Eventually they can become so different that they form two distinct species.

The process of formation of new species is termed **speciation**. For living organisms that are sexually reproducing organisms, species are recognised as different when they are not able to interbreed under natural conditions, or, if they interbreed, the offspring are either inviable or, if they survive, are sterile.

Several factors keep different species separated and prevent interbreeding or gene flow. These may come into operation before mating takes place — these are often called pre-mating isolation mechanisms — or they may operate after mating — these are post-mating isolating mechanisms.

**Pre-mating isolation mechanisms** include the following factors that are barriers to finding and securing a mate:

- Isolation in time: one species may be active by day and the other species is active at night — so, they are unlikely to meet.
- Isolation in space: one species may live on mountain tops and the second in valleys — again, they are unlikely to meet.
- Isolation by behaviour: one species does not recognise the signs of sexual readiness in a second species or is unable to perform the correct courtship rituals — so they meet, but nothing comes of it.
- Isolation by anatomy: mating may be attempted but, because of physical differences, it is not successful — so they meet, they try, but the parts don't fit.

**Post-mating isolation mechanisms** operate after the transfer of sperm from male to female and are effective because of chromosomal and chemical imbalances between the different species. Example include:

- Incompatibility of gametes: sperm cannot penetrate the outer coats surrounding the egg of the second species, so fertilisation does not occur.
- Zygote mortality: fertilisation occurs, but the zygote fails to develop.
- Inviability of zygote: zygote develops into an embryo but does not develop beyond that.
- Sterility of hybrid: hybrid offspring survive, but are sterile.

## Subspecies

Before becoming different species, separated populations may be identified as distinct subspecies. At this stage, members of geographically separated populations are recognisably different from each other. However, if the populations come together again, members of the different subspecies groups can still interbreed and produce viable and fertile offspring.

Subspecies are denoted by the addition of a third part to the scientific name. Figure 9.31 shows some of the different subspecies of the crimson rosella (*Platycercus elegans*):

A red-headed subspecies (*Playtycercus elegans elegans*) is found mainly on the south-east coastal regions of Queensland, New South Wales and Victoria. The orange-headed subspecies (*Platycercus elegans subadelaidae*) is found mainly in the Gulf region of South Australia. The yellow-headed subspecies (*Platycercus elegans flaveolus*) is found in the Murray–Darling Basin of New South Wales, Victoria and South Australia.

**studyon**

Unit 4 | AOS 1 | Topic 1 | Concept 9

**Allopatric speciation**
Summary screen and practice questions

**studyon**

Unit 4 | AOS 1 | Topic 1 | Concept 9

**See more**
Allopatric speciation

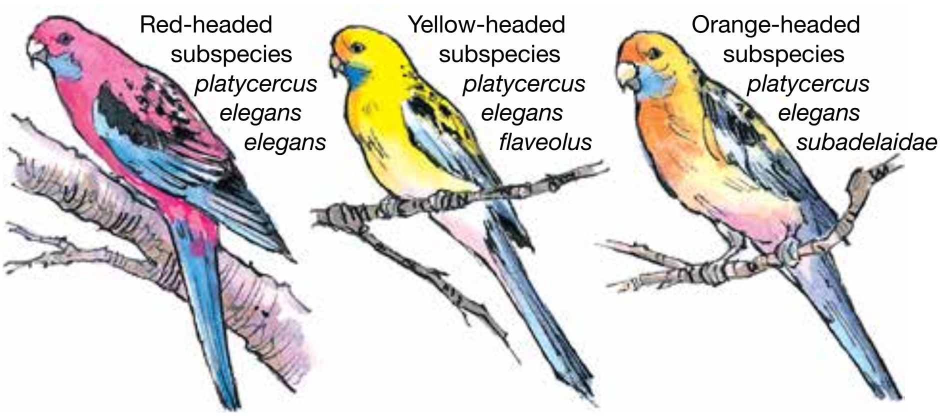
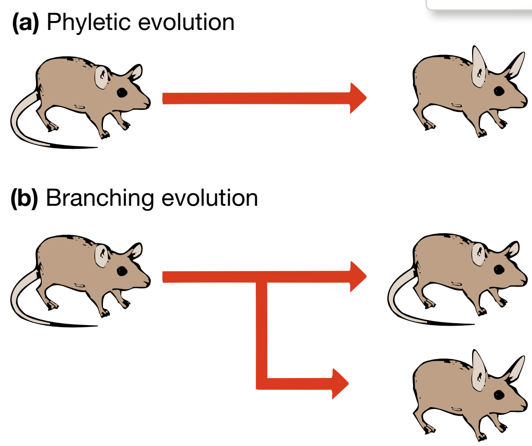



**FIGURE 9.31** Three subspecies of the rosella (*Platycercus elegans*). If colour is important in attracting a mate and in pre-mating rituals, what is likely to happen to these subspecies over time?




Two patterns of speciation are possible:

- phyletic evolution in which one population progressively changes over time to become a new species (see figure 9.32a)
- branching evolution in which a population of one species splits and one part of the population evolves separately to form a new species distinct from the original species (see figure 9.32b).

**FIGURE 9.32** Two different patterns of speciation: **(a)** phyletic evolution; **(b)** branching evolution. Which pattern is more common?

A pattern of branching evolution is more common than phyletic evolution. Branching evolution typically involves:

1. splitting of a smaller group from an original population so that the small group becomes geographically isolated
2. over many generations, the isolated population being subjected to different selection pressures because of different environmental conditions and to other change factors such as genetic drift
3. the isolated population changing over time such that, even if the two populations were to come together again, their members are unable to reproduce successfully — they are now two distinct gene pools and two different species (see figure 9.33).

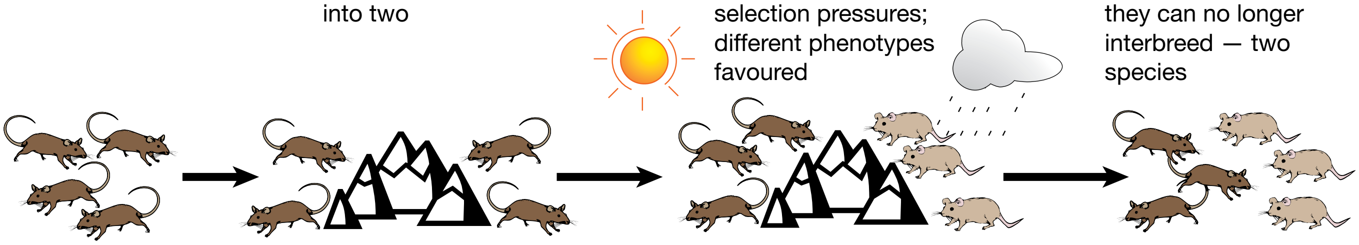



**FIGURE 9.33** Natural selection can, over time, produce two different species from a single ancestral species. This most commonly occurs when a population is split into two isolated groups, a situation called allopatric speciation.

When speciation occurs under circumstances where populations become geographically separated, it is termed **allopatric speciation** (allopatric = 'different homeland'). Factors that can split a population into geographically isolated groups may be:

- quick acting, such as habitat fragmentation owing to clearing or construction
- slow acting, such as the change of a river course
- even slower geological processes, such as the uplift of mountains or rising sea levels.

Another concept is **sympatric speciation** (sympatric = 'same homeland'). This occurs when members of a population living in the same area diverge sufficiently to become two different species. It was thought for some time that speciation could not occur under sympatric conditions. However, more recent evidence indicates the cichlid fish living in Lake Nicaragua are in the process of becoming a new species.

**FIGURE 9.34** Charles V: a person with Hapsburg lip. The transmission of the autosomal dominant Hapsburg lip through many generations is revealed in portraits, allowing the trait to be identified.

## Mutations: source of new alleles

**Generally, the DNA of the genetic material is stable, both in its base sequence and in its chromosomal location, and is passed unchanged from generation to generation.** For example, members of one European family, the House of Hapsburg, have been immortalised in art over several generations. Their portraits show that many members of the family had an excessively developed lower jaw that projected in front of the face (see figure 9.34). This condition, known as prognathism (also called Hapsburg lip), can be followed from Ernest the Lion (1377–1424) through ten generations to Maria Theresa (1717–1780).

However, the DNA of the genetic material can change. This change is very commonly a change in the base sequence of a gene. When this happens, a new form of a gene, called an allele, is created. Such a change is known as a **gene mutation** and it can alter the instructions that are encoded in the DNA.


**FIGURE 9.35** Laila El Garaa of Morocco is a short-statured woman who competed in the 2008 Beijing Paralympics where she won a bronze medal in the women's sports class F40 shot put event. Achondroplasia is due to the action of one allele at the **ACH** gene locus, with small stature being dominant to common stature. While aspects of their skeletal structure are affected, short-statured people (a preferred description) are otherwise unaffected and play standard roles in society.

This may be seen in the unexpected birth of a baby with achondroplasia, a form of dwarfism, to two parents of average height (see figure 9.35). This event is the visible expression of one change in the base sequence of a gene in a gamete (either egg or sperm) of one of the parents. Other gene mutations result from a change in the **location of a gene, such that the DNA moves to a new chromosomal neighbourhood and this move alters its expression**.

Earlier in this chapter, you saw examples of inherited variation in several organisms (refer back to page 383). These phenotypic variations are expressions of just a tiny sample of the different forms (alleles) of each gene that have arisen by mutation of ancestral genes.

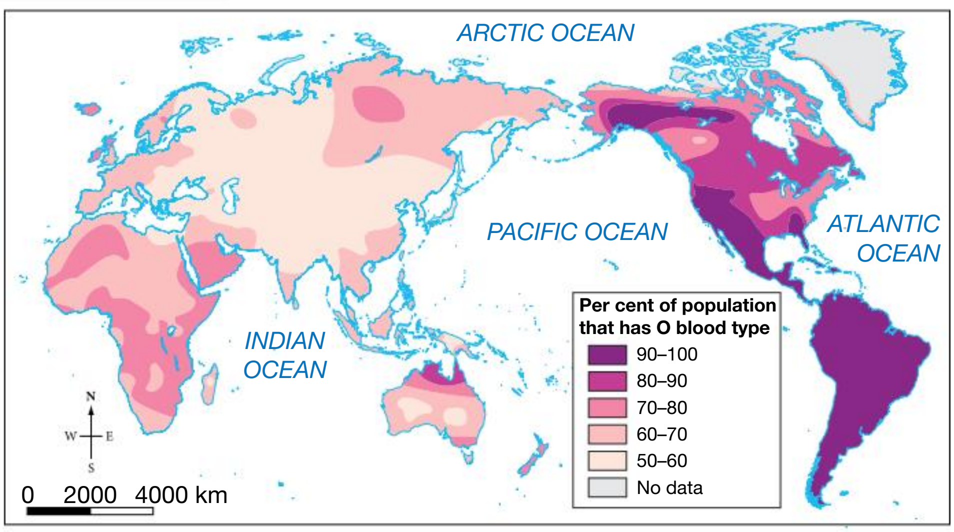
## Spread of mutations in populations

Some mutations spread locally, regionally and even globally. For example, we know that prehistoric modern humans were all lactase nonpersistent. The gene mutated, giving rise to a new allele that conferred lactase persistence and has spread where dairying cultures were established. Archaeological evidence and ancient DNA studies support the inference that in Europe this occurred about 7500 years ago.

The ancestral form of the ABO gene that determines ABO blood types is believed to be the $I^A$ allele that determines group A blood. This allele was already present in the first modern humans, having evolved much earlier in their remote ancestors. Sometime later in human prehistory, other mutations of this gene occurred that gave rise, first to the $I^B$ allele, and later to the $i^o$ allele. These alleles are spread globally. Worldwide, the most common blood type is group O, with more than 60 per cent of the world's population having this type (see figure 9.36).

**eBook*plus***

**Interactivity**
Blood typing
**int-3041**




***Source:*** Anthro Palomar

**FIGURE 9.36** Distribution of blood group O in indigenous populations worldwide. Note the very high percentage in Central and South America.

Table 9.7 shows the current distribution of the ABO blood types in several populations.

**ODD FACT**

Karl Landsteiner was the first to identify three of the four ABO groups when, in 1901, he discovered the A, B and O blood groups. Landsteiner originally called them groups A, B and C, but later renamed group C as group O (= Ohne, German for 'zero' or 'without').

**TABLE 9.7** Distribution of the four ABO blood groups in a number of human populations.

| People group | O | A | B | AB |
|---|---|---|---|---|
| American (Caucasian) | 45 | 41 | 10 | 4 |
| Australian (other) | 49 | 38 | 10 | 3 |
| Bororo (Brazil) | 100 | 0 | 0 | 0 |
| British | 47 | 42 | 8 | 3 |
| Chinese (Beijing) | 29 | 27 | 32 | 13 |
| Danes | 41 | 44 | 11 | 4 |
| Dutch | 45 | 43 | 9 | 3 |
| East Indian (Bangalore) | 40 | 24 | 30 | 6 |
| East Indian (Banjara caste) | 28 | 23 | 37 | 12 |
| Indigenous Australian peoples | 61 | 39 | 0 | 0 |
| Japanese (Ainu) | 17 | 32 | 32 | 18 |
| Maori (New Zealand) | 46 | 54 | 1 | 0 |
| Mayan (Central America) | 98 | 1 | 1 | 1 |
| Navajo (American Southwest) | 73 | 27 | 0 | 0 |
| Pakistani (NW Frontier) | 35 | 25 | 32 | 8 |
| Peruvian (Indians) | 100 | 0 | 0 | 0 |
| Spanish | 38 | 47 | 10 | 5 |
| Sudanese | 62 | 16 | 21 | 0 |
| Thai | 37 | 22 | 33 | 8 |

Data adapted from bloodbank.com, *Asian Journal of Transfusions Science, Journal of Pakistan Medical Association, The Internet Journal of Biological Anthropology,* and *A Contribution to the Physical Anthropology and Population Genetics of Sweden* (Lund: Sweden, 1959).

**FIGURE 9.37** Wall painting from tomb of Nebamun in Thebes, Egypt, about 1350 BC, showing Nebamun accompanied by his wife, Hapshetsut, hunting birds in the marshes of the Nile river. Note the striped (tabby) cat near the reeds that is looking up at the birds.

## Mutations in domestic pets

We can track the appearance of some mutations in the domestic cat (*Felis catus*) from historical records. For example, wall paintings and papyrus manuscripts from ancient Egypt show that the first domesticated cats of Egypt (and indeed the world) were striped (tabby) cats with their stripes arranged in the so-called mackerel tabby pattern. Figure 9.37 shows part of a wall painting, dated more than 3000 years ago, showing such a cat.

Since their domestication and spread from Egypt, many other mutations have spontaneously occurred in cat populations around the world, including:

- a blotched mutation that produces a different pattern of tabby striping that is known to have first appeared in England (see figure 9.38b)
- a non-agouti mutation that produces a solid-coloured cat, such as an all-black cat (see figure 9.38c)
- an X-linked orange mutation that produces ginger cats, both male and female, and female tortoise-shell cats
- a dilute mutation that produces grey, rather than black cats, and cream, rather than orange (ginger) cats
- a Manx mutation in which the cat has either no tail or has a greatly reduced stumpy tail (see figure 9.39a).

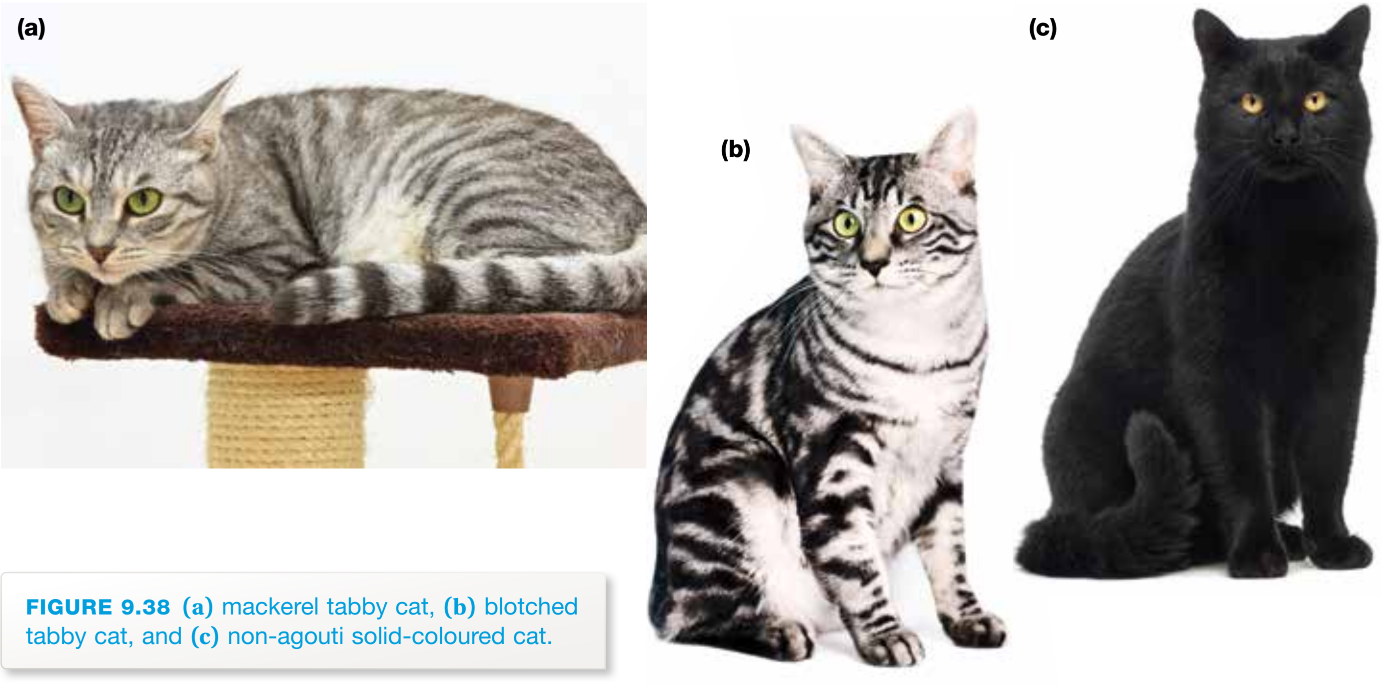


**FIGURE 9.38** **(a)** mackerel tabby cat, **(b)** blotched tabby cat, and **(c)** non-agouti solid-coloured cat.

Mutations continue to occur in cats, as in other species. In 1950, in Cornwall, England, the rex mutation appeared, in which the cats' coats consisted of the soft and curly undercoat with the normal overcoat missing (see figure 9.39b). In 1966, in New York, the 'wirehair' mutation was detected in which the cats' fur is very coarse.

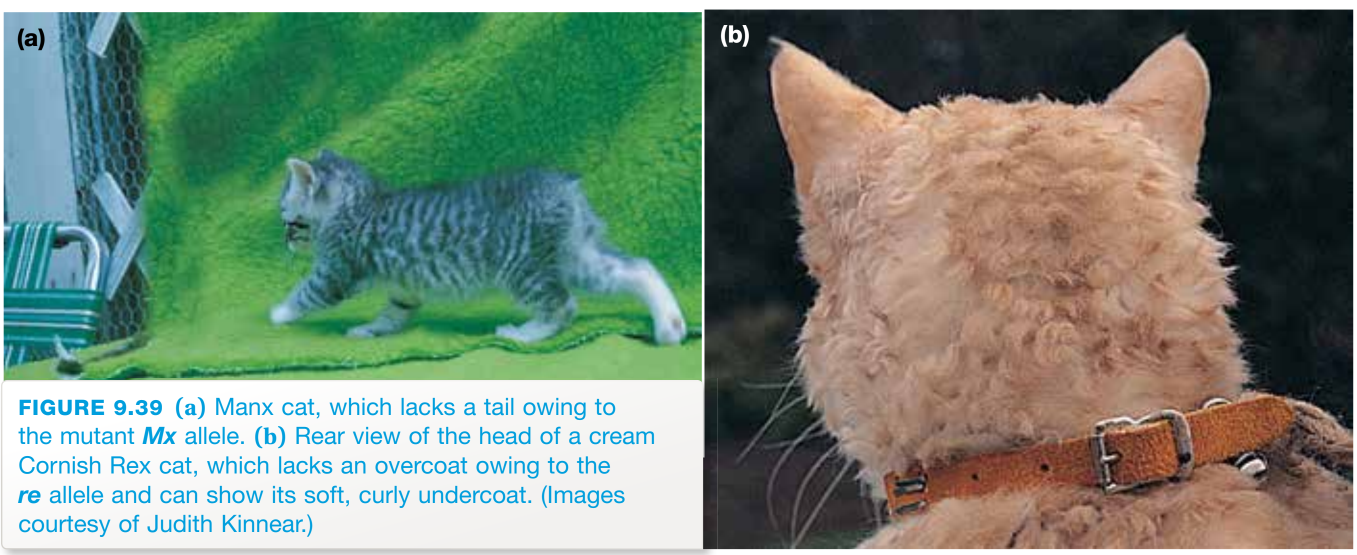


**FIGURE 9.39** **(a)** Manx cat, which lacks a tail owing to the mutant ***Mx*** allele. **(b)** Rear view of the head of a cream Cornish Rex cat, which lacks an overcoat owing to the ***re*** allele and can show its soft, curly undercoat. (Images courtesy of Judith Kinnear.)

The budgerigar, *Melopsittacus undulatus*, is the world's most popular cage bird. In the wild, populations of these birds are predominantly green. However, many gene mutations have occurred in captive breeding populations of these birds (refer back to figure 9.1 on page 377).

The first budgerigars were taken from Australia to England in 1840. For about 30 years, light green was the only colour seen in these pet birds. Breeders found that, generation after generation, these birds bred true. However, in 1872, the first colour mutant appeared in birds in Belgium and Germany; this was the

study on

Unit 4
AOS 1
Topic 1
Concept 2

**Sources of new alleles: mutations**
Summary screen and practice questions

study on

Unit 4
AOS 1
Topic 1
Concept 2

**See more**
Mutation

*dilute* mutation. Soon afterwards, in 1878, the *light-blue* mutation appeared — this caused quite a sensation — green birds now came in a blue variety, and these blue budgerigars initially commanded very high prices. Later, other colour variants appeared, including the *ino* mutation, which produces all yellow and all white colours (1879), the dark mutation (1915), the *cinnamon* mutation (1931), and the *yellow-faced blue* mutation (1934), and, to date, about 30 different colour variants have appeared in captive budgerigar populations as a result of gene mutations.

The continued appearance of mutations reflects the fact that DNA is a large macromolecule and that errors can occur during processes in meiosis, including DNA replication and synapsis and crossing over.

### Fate of new mutations

In order for a gene mutation to be sustained in a population, it is first necessary that the mutation occurs in the germline of an organism so that it may be passed to the next generation. Somatic mutations, that is, mutations in body cells of an organism, disappear with the death of the organism in which the mutation occurred.

If the germline mutation produces an allele that controls a dominant trait, its phenotypic expression will appear in the next generation. If, however, the new allele controls a recessive trait, it may be several generations before the trait is expressed phenotypically in the population. Why? The first appearance requires the mating of two heterozygous carriers of the trait, and if the allele is rare, as would be the case for a new allele, such matings would also be rare.

After a mutation is expressed in a phenotype, it will be the actions of change agents, in particular natural selection, that determine whether the allele is lost from the population or whether it becomes established in the gene pool of a population. Some gene mutations disappear after one or a few generations, because of strong selection pressure acting against them. In addition, in a small population, even advantageous mutations may be lost through the action of genetic drift.

If a new allele created by mutation confers a strong selective advantage in terms of survival and reproduction rate under the prevailing environmental conditions, it is reasonable to assume that this allele has a high probability of becoming established in a population and increasing its frequency in the gene pool over successive generations. This occurred in the case of the lactase persistence allele of the **LCT** gene in populations with a dairying culture, and with the $\boldsymbol{H^S}$ mutant haemoglobin allele in environments where malaria is a major killer.

Does a very high frequency of a particular allele in the gene pool of a population always indicate that the corresponding phenotype has a selective advantage as compared to other variants in a population? Not necessarily. This could be a founder effect. For example, indigenous populations in South America are 100 per cent blood group O. It has been suggested that the founder groups of these populations might all have been group O, genotype $\boldsymbol{i^o\ i^o}$.

**ODD FACT**

In the 1920s, flies were taken up in a balloon to a height of more than 18 kilometres and it was found that gene mutation in these flies occurred greater than five times more rapidly at this height than at sea level. Can you suggest a possible explanation?

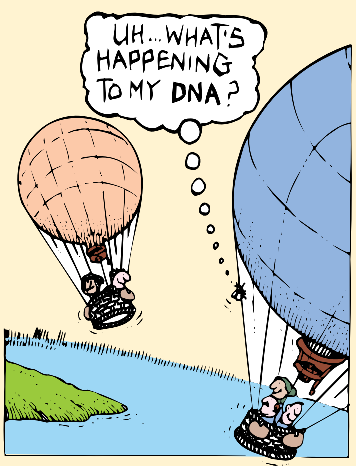



**FIGURE 9.40**

## Causes of mutations

Mutations may be spontaneous or may be induced by known exposure to **mutagenic agents**, such as irradiation (for example, X-rays and gamma rays), some chemicals (for example, benzene, dioxane and mustard gas) and some viruses.

Protection against mutagenic agents is necessary to protect the DNA in both body cells and germline cells. Have you had an X-ray during a dental check-up? The use of a lead apron during this procedure is to protect body tissues from exposure to radiation. Doses received by people working with radiation are monitored by wearing badges that contain radiation-sensitive devices (see figure 9.41).

**FIGURE 9.41** Nuclear decontamination workers removing their radiation protection. Note the protective breathing apparatus and the radiation-sensitive dosimeter that can be seen on the chest one of the workers.

Mutation rates are not equal for all genes and the spontaneous mutation rate for achondroplasia, one form of inherited dwarfism, is much higher than that for some other inherited disorders (see table 9.8). This is in accord with the fact that about 90 per cent of the cases of achondroplasia are sporadic, and are due to mutation. On average, more sporadic cases occur in association with increased paternal age at the time of conception.

**TABLE 9.8** Spontaneous mutation rates at different gene loci in various human population samples. Estimated mutation rates are based on frequencies of births of babies showing these dominant conditions, to unaffected parents. Note the high rate of mutation at the achondroplasia gene locus.

| Mutation | Locus | Mutations per 100 000 gametes |
|---|---|---|
| *retinoblastoma* | number-13 chromosome | |
| United Kingdom | | 1.2 |
| United States | | 1.8 |
| *achondroplasia* | number-4 chromosome | |
| Northern Ireland | | 14.3 |
| Japan | | 12.2 |
| *Huntington's disease* | number-4 chromosome | |
| United States | | 0.5 |

Autosomal dominant conditions that have a serious deleterious effect on an individual are usually the result of a new mutation. Can you suggest why?

studyon

Unit 4
AOS 1
Topic 1
Concept 3

**Gene mutations**
Summary screen and practice questions

studyon

Unit 4
AOS 1
Topic 1
Concept 3

**Do more**
Mutations

# Types of mutations

In this section, we will examine mutations, both gene mutations and chromosomal (block) mutations.

## Gene mutations

Gene mutations are changes in the genetic material or DNA. Several kinds of change in DNA can be identified, such as:
point mutations that include

- base substitution
  - 'silent' mutation
  - nonsense
  - missense
    - conservative
    - non-conservative.
- base insertion
  - frameshift mutation.
- base deletion
  - frameshift mutation.

### Point mutations

Point mutations affect a single base in DNA, either by substitution, insertion or deletion. Figure 9.42, on page 418, shows some of the possible point mutations in the coding sequence of a gene.

These mutations commonly affect the base sequence of a structural gene. However, these mutations may also affect a segment of DNA that regulates the expression of a structural gene. This was the case for the lactase persistence allele. We now know that the mutation that produced the lactase persistent phenotype is a single base substitution in the region upstream of the **LCT** gene. This base substitution enhances the promoter of the **LCT** gene.

### Point mutations: base substitutions

In a substitution, one base in a DNA triplet is replaced by another. Three kinds of base substitutions are recognised: silent, nonsense and missense, as follows:

**1. Silent mutations:** A silent mutation is a base substitution in DNA that does not result in a change in the polypeptide product encoded by this gene. This is because the genetic code is redundant and many amino acids are encoded by several DNA triplets. For example, the amino acid valine (*val*) is encoded by four DNA triplets: CAA, CAG, CAT and CAC

**. . . ATG CAA GGT . . . → . . . ATG CAT GGT . . .**

This DNA substitution mutation will not affect the expression of the gene because the mutated DNA triplet (CAT) still codes for the same amino acid (*val*) as the original DNA triplet (CAA), hence the label 'silent'. Silent mutations most commonly occur in the third base of a DNA triplet.

While these mutations are 'silent' in terms of their polypeptide product, they are of course visible when the DNA sequence is examined.

**2. Nonsense mutations:** A nonsense mutation in DNA is a base substitution that changes a DNA triplet so that, instead of coding for an amino acid, it codes for a STOP signal. The stop or termination signals in DNA are the triplets ATT and ATC. Look at the following substitution mutation:

**. . . ATG TTT GGT . . . → . . . ATG ATT GGT . . .**

Such a mutation will have a serious effect on the polypeptide product of the gene because transcription will stop at the mRNA codon transcribed by the mutated triplet DNA sequence. In turn, this will shorten the polypeptide chain encoded by the gene. The degree of shortening of the polypeptide gene product will depend on the position of the point mutation relative to the start of the coding sequence — the earlier in the DNA sequence, the greater the shortening of the polypeptide.

**3. Missense mutations:** A missense mutation is a base substitution in DNA that results in a single amino acid alteration in the polypeptide product of the gene. Missense mutations may be conservative or non-conservative.

When a missense mutation occurs and a different amino acid is incorporated into a polypeptide chain, the outcome depends on whether the mutated polypeptide can carry out its normal function. This is likely to be the case in a **conservative missense mutation** and such a mutation can be tolerated. A **conservative missense** mutation of DNA results in the substitution of one amino acid by a different amino acid with similar chemical properties, for example:

**. . . ATG CAA GGT . . . ➔ . . . ATG CGA GGT . . .**

This mutation results in the replacement of the amino acid valine (*val*) with alanine (*ala*), both amino acids having hydrophobic side chains.

In other cases, the missense mutation results in the substitution of an amino acid by one that has different chemical properties. Such a mutation is said to be a **non-conservative missense mutation.** In this case, the resulting polypeptide cannot carry out its normal function and such a mutation will have severe clinical effects.

A non-conservative missense mutation of DNA results in the substitution of an amino acid with an amino acid containing a different class of side chain and different chemical properties. For example, the replacement of histidine (*his*), which contains a charged side chain, with tyrosine (*tyr*), which contains a hydrophobic chain:

**. . . ATG GTA GGT . . . ➔ . . . ATT ATA GGT . . .**

One example of a non-conservative missense mutation is haemoglobin M disease. Haemoglobin M results from a missense mutation that results in the substitution of histidine with tyrosine in the alpha chain of haemoglobin. Haemoglobin M can carry oxygen but cannot release it to body cells where it is needed. Another example of a missense mutation results in achondroplasia (refer back to pages 411 and 415). This condition is the result of the replacement of the amino acid glycine (*gly*) with arginine (*arg*) at amino acid number 380 in the receptor protein for a fibroblast growth factor. We can write this replacement as Gly380Arg. This replacement is caused by a single base substitution in the DNA triplet that encodes the 380th amino acid in this protein that mutates as **CCA ➔ GCA**. Just one specific base substitution among hundreds of bases in a segment of DNA has this major effect.

**ODD FACT**

Changes in one of the first two bases of a DNA triplet are more often a source of nonsense and missense mutations than changes in the third base.

## Point mutations: Frameshift mutations

Single base insertions have a major effect on the genes involved because they alter the DNA triplet at the point of the insertion and also affect every DNA triplet following that point. Likewise, single base deletions have the same effects. These single base insertions or deletions are known as **frameshift** mutations.

Frameshift mutations alter the base sequence of a gene so that the message it encodes no longer makes sense. Remember that the genetic information in DNA is organised into groups of three bases or triplets that are then read in order, with one triplet coding for one amino acid.

Consider the sentence: **the boy hit the dog** with each word representing a DNA triplet. In this form, this sentence makes sense.

Now, add an extra letter — this is like a base insertion: **the bxo yhi tth edo g**

The addition of the x causes the sense to be lost at the point of the addition and beyond.

Now delete the first e from the original sentence — this is like a base deletion: **thb oyh itt hed og**

An insertion occurs when one nucleotide base is added to a DNA strand. This type of mutation is more damaging than a point mutation because they cause a frameshift with all the bases from the point of insertion being moved down by one position to make room for the extra base.

A base deletion also causes a frameshift with all bases from the point of the deletion being moved back by one position to compensate for the deletion. In addition, both insertion and deletion may form a stop codon where one should not exist, having drastic consequences by creating a premature cessation of translation of a polypeptide chain. Figure 9.42 shows both of these types of frameshift mutations.

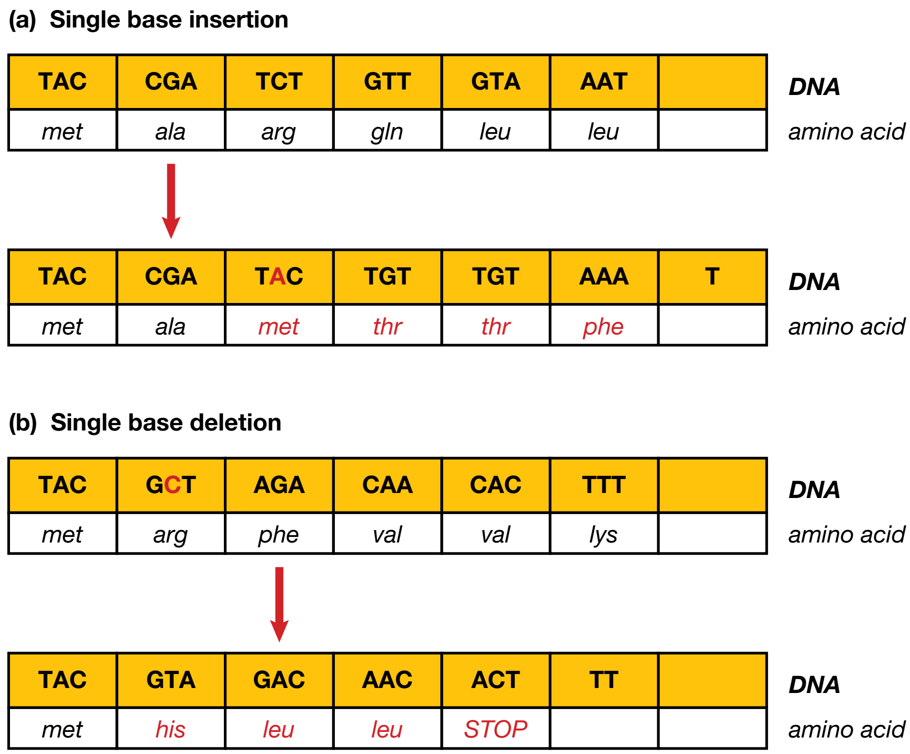



**FIGURE 9.42** Diagram showing the effects of **(a)** a base addition or insertion into a DNA sequence and **(b)** a base deletion. Both mutations cause a significant change in the polypeptide that is encoded by this DNA by changing the amino acids from the point of addition or deletion and beyond as a result of the changes in the DNA triplets.

The point mutation discussed above involves single base changes. However, some mutations, known as trinucleotide repeats involve mutations that repeat a large number of bases. Go to the box at the end of this section to read about these mutations, which are the cause of two major inherited disorders, fragile X syndrome and Huntington's disease (HD).

## Chromosomal or block mutations

Block mutations are chromosomal changes affecting large segments of a chromosome. These block mutations most commonly arise as a result of spontaneous errors in crossing over during meiosis or they may be induced by mutagenic agents, such as X-rays.

**study on**

Unit 4
AOS 1
Topic 1
Concept 4

**Chromosomal mutations**
Summary screen and practice questions

**study on**

Unit 4
AOS 1
Topic 1
Concept 4

**Do more**
Chromosomal changes

**study on**

Unit 4
AOS 1
Topic 1
Concept 4

**See more**
Changes in chromosomes

These block mutations include:

- deletion of part of a chromosome
- duplication or gain of part of a chromosome
- translocation or reciprocal exchange between non-homologous chromosomes
- inversion when a segment of a chromosome rotates through 180 degrees.

Figure 9.43 shows some of these block mutations. In each case, chromosome breakages occur followed by rearrangement and rejoining.

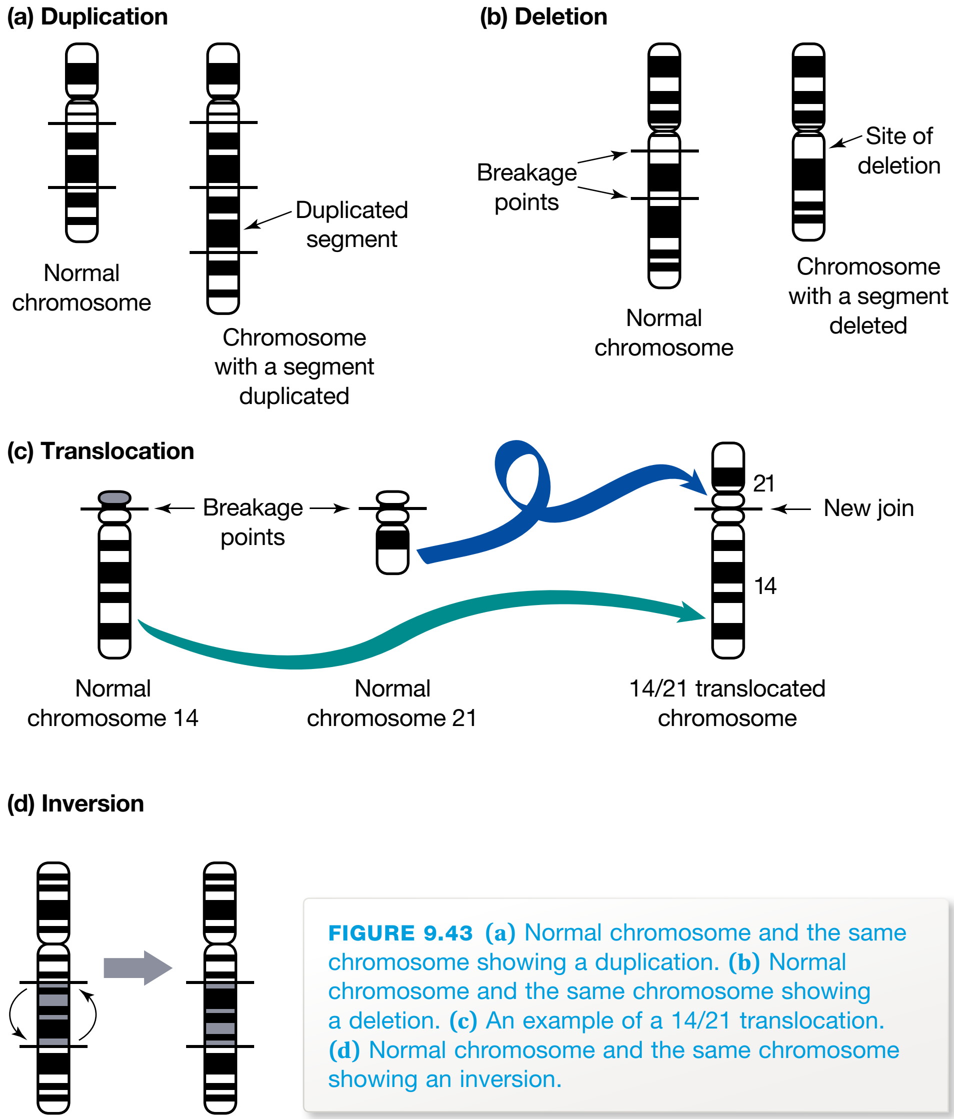



**FIGURE 9.43** **(a)** Normal chromosome and the same chromosome showing a duplication. **(b)** Normal chromosome and the same chromosome showing a deletion. **(c)** An example of a 14/21 translocation. **(d)** Normal chromosome and the same chromosome showing an inversion.

None of these chromosomal block mutations involves a loss of genetic material, but they do alter the location of some genes, relocating them in a new position on a chromosome, giving them different next-door neighbours. Research has shown that relocating can alter the activity or the expression of a gene — such effects are termed **position effects**. In some cases, a gene at a breakpoint may be silenced or its expression may become unstable or, in the case of some reciprocal translocations, abnormalities such as cancers occur. Figure 9.44 shows the details of a reciprocal translocation that results in chronic myeloid leukaemia.

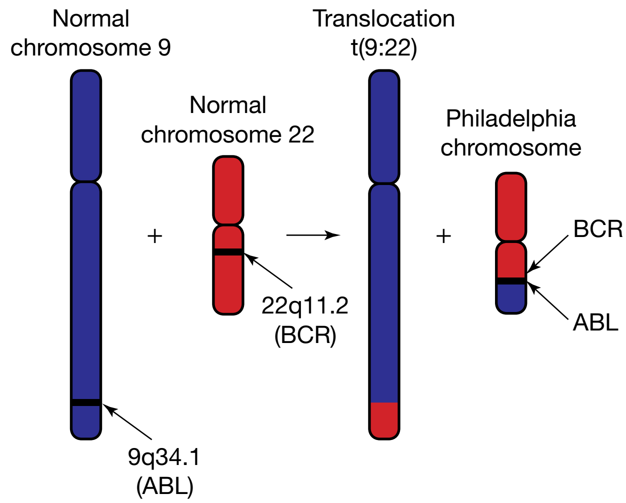



**FIGURE 9.44** Diagram showing the reciprocal exchange that occurs between chromosome 9 and chromosome 22 in chronic myeloid leukaemia. The relocation of the **ABL** and the **BCR** genes in close proximity on the Philadelphia chromosome has been shown to be the cause of chronic myeloid leukaemia.

Figure 9.45 shows the unstable expression of a gene for eye colour in the compound eye of a fruit fly (*Drosophila melanogaster*) after relocation of the gene involved to a new chromosomal location by an inversion. Note that some eye facets are dark where the gene is active but others are pale where the gene is not expressed.

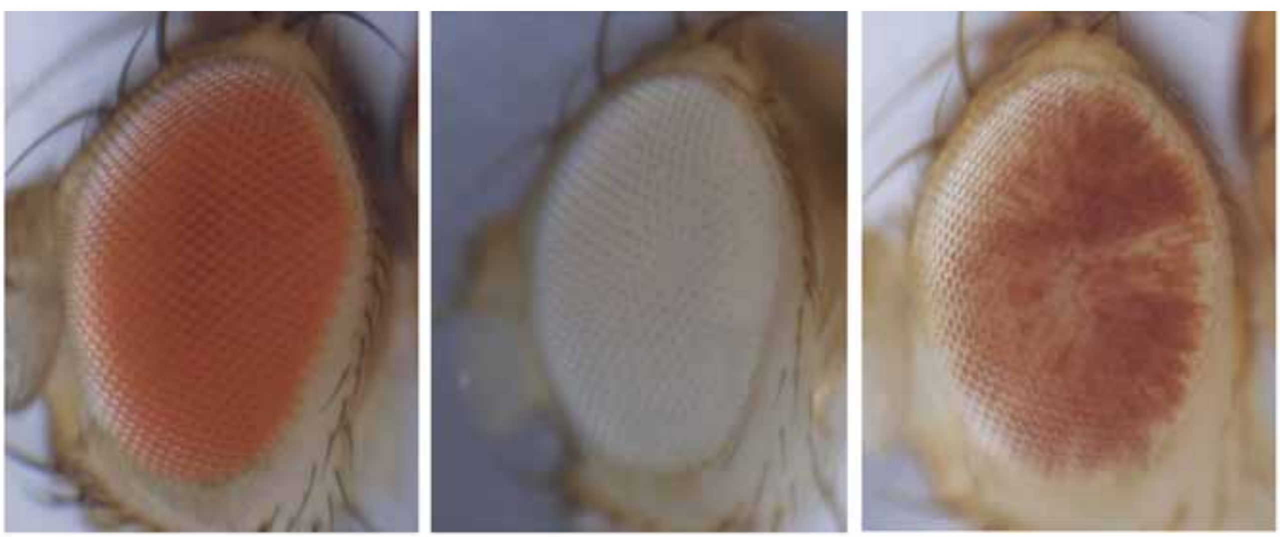


**FIGURE 9.45** Images of the compound eye of the fruit fly *Drosophila melanogaster*, showing the normal red eye (left), the normal white eye (middle) and the variegated eye (at right). The variegation results from the unstable expression of the gene determining red eye colour because of a chromosomal inversion that has relocated it. (Images courtesy of A.M. Bauer and S.C.R. Elgin, Washington University, St Louis.)

Earlier in this chapter (refer back to page 389), you were introduced to chromosomal changes involving whole chromosomes (aneuploidy) or the gain of whole sets of chromosomes (polyploidy). Refer back to page 386.

Table 9.9 summarises some of these aneuploid changes in humans. Note that polyploidy does not occur in humans.

studyon

Unit 4
AOS 1
Topic 1
Concept 4

**See more**
Trisomy 21

**TABLE 9.9** Some examples of chromosome changes and approximate incidence rates. Which syndrome is an example of a trisomy? A monosomy? The XYY condition does not have a clinical name.

| Chromosome change | Resulting syndrome | Approximate incidence rate |
|---|---|---|
| ***Addition: whole chromosome*** | | |
| extra number-21 (47, +21) | Down syndrome | 1/700 live births |
| extra number-18 (47, +18) | Edwards syndrome | 1/3000 live births |
| extra number-13 (47, +13) | Patau syndrome | 1/5000 live births |
| extra sex chromosome (47, XXY) | Klinefelter syndrome | 1/1000 male births |
| extra Y chromosome (47, XYY) | n/a | 1/1000 male births |
| ***Deletion: whole chromosome*** | | |
| missing sex chromosome (46, XO) | Turner syndrome | 1/5000 female births |
| ***Deletion: part chromosome*** | | |
| missing part of number-4 | Wolf-Hirschhorn syndrome | 1/50 000 live births |
| missing part of number-5 | cri-du-chat syndrome | 1/25 000 live births |

## TRINUCLEOTIDE REPEAT EXPANSION MUTATIONS

One class of gene mutations involves the addition or deletion of a large number of short sequences of bases.

Several normal human genes contain multiple copies of a three-base (trinucleotide) sequence, such as **CCG** and **CAG**. One kind of mutation, known as a **trinucleotide repeat expansion (TRE)**, involves additional repeats of these sequences beyond the normal range. These TRE mutations are the cause of two inherited disorders: fragile X syndrome (FMR1) and Huntington's disease (HD) (see table 9.10).

**TABLE 9.10** Examples of trinucleotide repeat expansion mutations

| Gene | Mutant condition | Trinucleotide repeats |
|---|---|---|
| FMR1 | fragile-X syndrome | CCG: 200 to 1000+ repeats |
| HD | Huntington's disease | CAG: 36 to 120 repeats |

Compared with the normal alleles that have a small number of trinucleotide repeats, the mutant alleles that determine these disorders have many more repeats. This can result in a significant increase in the total number of bases compared to the number in the normal allele.

Fragile X syndrome is the most common inherited form of mental retardation in males and occurs at a frequency of about 1 in 1250 male births. Affected males often have long faces, large ears and, after puberty, may have enlarged testes (macro-orchidism). In many cases, this mutation can be visibly seen when the X chromosome is examined using light microscopy.

In 1991, it was recognised that males affected by fragile X syndrome (see figure 9.46) have from 200 to 1000 plus repeats of a CCG trinucleotide within their mutant **FMR1** allele. In contrast, an unaffected person has from 6 to about 50 repeats of the CCG trinuceotide.

Both males and females affected by Huntington's disease have from 36 to over 100 repeats of the CAG trinucleotide in their mutant **HD** allele, while unaffected people usually have from 6 to about 35 copies of this trinucleotide.

Trinucleotide repeat expansion mutations tend to be unstable so that the number of repeats can change from one generation to the next. For example, the number of trinucleotide repeats has been observed to increase when the mutant **HD** allele is transmitted from an affected male parent to his children.

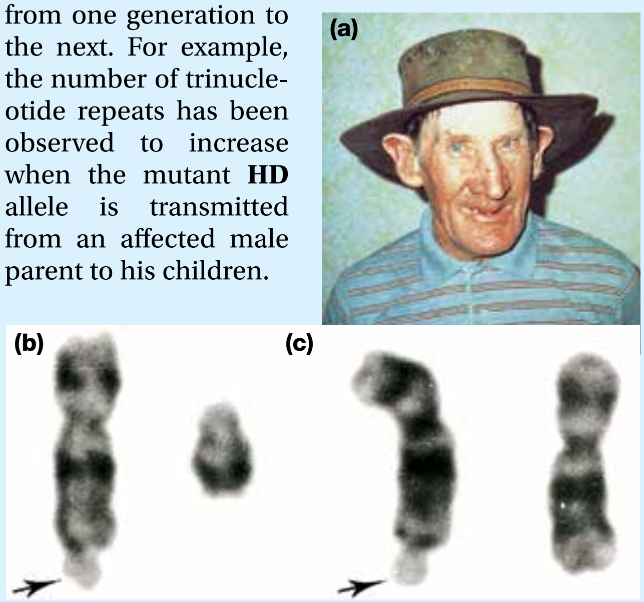



**FIGURE 9.46** **(a)** A male with fragile X syndrome. **(b)** Fragile X (left) and Y chromosomes from an affected male. **(c)** A fragile X (left) and normal X chromosome from a female who is a heterozygous carrier of the fragile X. (Images courtesy of Marjory Martin.)

## KEY IDEAS

- Mutations are the source of new allelic forms of genes.
- Mutations continue to occur in the gene pool of populations.
- Mutations may be either spontaneous or induced by exposure to a known mutagenic agent.
- Spontaneous mutation rates are not identical for all genes.
- Mutations may be classified as point mutations or chromosomal (block) mutations.
- Relocation of genes can produce changes in their expression.

## QUICK CHECK

**19** Give two examples of:
  **a** a point mutation
  **b** a block mutation.

**20** Identify the following statements as either true or false:
  **a** Gene mutation can occur in the DNA of structural genes only.
  **b** The relocation of a gene may cause its expression to be changed.
  **c** Chronic myeloid leukaemia is caused by a reciprocal translocation.
  **d** A frameshift mutation would be expected to have more significant effects on a gene product than a missense mutation.

**21** Identify the difference between the members of the following pairs:
  **a** a gene mutation and a block mutation
  **b** a frameshift mutation and a substitution mutation
  **c** a chromosomal deletion and a chromosomal translocation
  **d** an induced and a spontaneous mutation.

# Manipulating gene pools

## Lambs in spring

Early spring is the lambing season in Australia. In certain sheep breeds, such as merino, a female sheep or ewe typically produces just a single lamb each season. However, these days, it is not uncommon to see merino ewes with twins, sometimes triplets and, in some cases, even quads. What has caused this change? It is human intervention through artificial selection of an allele for increased fecundity by selective breeding with rams and ewes that have that particular allele.

*The Booroola gene is formally known as* **BMPR-1B** *and its alleles are denoted as* ***FecB**$^B$* *and* ***Fec**$^{B+}$* *but, for simplicity, they are shown here as* $B^B$ *and* $B^+$.

The gene concerned is known as the Booroola gene, formally designated **BMPR-1B**. Its action in producing increased fecundity was recognised by the Seears brothers in the merino flock at their 'Booroola' property, near Cooma, New South Wales in the 1950s. The Booroola gene is located on the number-6 sheep chromosome and its rare $B^B$ allele increases ovulation rate while its more common $B^+$ allele has no effect on ovulation rate. In heterozygous $B^BB^+$ ewes, the ovulation rate is increased by a factor of 1.5, while in homozygous $B^BB^B$ ewes, their ovulation rate is increased by a factor of 3.0. (The ovulation rate in homozygous $B^+B^+$ ewes is 1.0.)

**ODD FACT**

The Booroola gene appears to have been introduced by Bengal sheep imported to Australia from India in 1792. Records show that the Booroola merino flock can be traced back to the merinos owned by Rev. Samuel Marsden (1764–1838), who also owned a Bengal flock.

Using selective breeding techniques, the $B^B$ allele has now been incorporated into other sheep breeds that normally produce a single lamb from each mating. This was done by crossing these breeds with sheep known to be homozygous for the $B^B$ allele. The process in which specific animals are used in breeding programs to introduce particular alleles of a gene into the gene pool of a population and to fix the preferred alleles in the homozygous state in members of the population is termed **selective breeding**.

**FIGURE 9.47** Sheep with four lambs. The high ovulation rate of this ewe may be assumed to be due to the presence of a double dose of the $B^B$ allele of the Booroola gene.

Selective breeding is an example of artificial selection. In this procedure, only those animals that display a particular trait in their phenotype or are known carriers of the trait are chosen to reproduce. Selective breeding is an example of how humans can intervene in natural evolutionary processes by artificially manipulating the gene pool of a population. Let's look at some other examples of selective breeding.

## Selective breeding in action

studyon

Unit 4 | AOS 1 | Topic 1 | Concept 10

**Selective breeding** Summary screen and practice questions

From early times, artificial selection was carried out to improve herd quality. Farmers selected the best males for mating with their breeding females. In the case of beef cattle, bulls were chosen for their genetic superiority in terms of desirable market characteristics, such as meat yield and non-fatty carcass. For dairy cattle, desirable market characteristics included milk yield and butterfat content. Merino rams were chosen for the fineness (thinness) of their wool fibre and the yield of greasy fleece. Males were also selected based on other inherited features, including good conformation (form, outline or shape), high fertility based on sperm counts and the absence of any known genetic defects. Through artificial selection, farmers could improve the quality of their herds.

The deliberate selection by a breeder of specific animals to provide the genetic material for the next generation is a process known as selective breeding or artificial selection. This is in contrast to the random breeding that occurs when any male animal in a population has an equal chance of mating with any female.

### Selective breeding in action: sheep

More than 200 recognised breeds of sheep exist, derived from wild sheep or mouflons (*Ovis musimon*) that were domesticated at least 6000 years ago. Mouflon sheep are native to parts of Europe and central Asia and are still living today.

About 75 per cent of the sheep in Australia today are merino sheep, which are prized for their wool quality. The first merinos were imported to Australia by John Macarthur in 1796 — these were Spanish merinos from South Africa. In Australia, several strains (breeds) of merino are recognised, including the

Peppin merino, the South Australian merino and the Saxon merino. Selective breeding (artificial selection) over generations has enhanced particular features of these breeds and their survival in particular climatic conditions. For example, the large South Australian merino, which has been bred to thrive in the arid regions of that state and in similar regions in others, produces greasy wool with strong (thick) fibres. In contrast, the much smaller Saxon merino thrives in wetter regions and has been bred for wool of increasing fineness (thin fibres).

Table 9.11 shows some of the different wool characteristics of several sheep breeds, including some merino breeds.

**TABLE 9.11** Wool characteristics of some sheep breeds. A staple is a collection of a large number of wool fibres. Typically, the fleece grows to a fixed length (like the coats of cats and dogs). In contrast, the fleece of Drysdale sheep continues to grow (like human hair) and so the staple length is given as growth per year. Which breed produces the finest (thinnest) wool?

| | Breed of sheep | | | | | |
|---|---|---|---|---|---|---|
| Trait | Drysdale | Saxon merino | Peppin merino | Strong Wool merino | Polwarth | Corriedale |
| body mass (kg) | 116 | 60 | 132 | 147 | 130 | 104 |
| greasy fleece weight (kg) | 12.0 | 5.0 | 18.6 | 20.0 | 12.0 | 13.0 |
| diameter of fibre (μm) | 40.0 | 17.4 | 22.0 | 23.5 | 24.6 | 28.0 |
| length of staple (mm) | 400 (annual) | 75 | 150 | 95 | 150 | 175 |

Data from the National Wool Museum (Geelong, Victoria).

The value of a wool fleece is determined principally by two traits, namely, clean fleece weight — the heavier the better — and fibre diameter — the thinner the better! These are inherited traits but, because they are polygenic, the rate of change that can be achieved through selective breeding over many generations is slower than with monogenic traits. (Can you suggest why?) For example, selection programs at the Triangle Research Station in New South Wales using several merino strains produced gains in these two important traits. The results of selective breeding, with an equal emphasis on increasing clean fleece weight and decreasing fibre diameter, are shown in table 9.12. Although the gains may appear small, they represent significantly increased profits for woolgrowers.

**ODD FACT**

The first merino sheep brought to the Australian colony in 1796 had fleeces that yielded about 1.5 to 2.0 kilograms of wool per year. Today, fleeces of merino sheep can commonly yield up to 10 kilograms of wool per sheep per year.

**TABLE 9.12** Results of a 10-year selective breeding program on two strains of merino sheep with an equal emphasis on increasing clean fleece weight and decreasing fibre diameter.

| Merino strain | Group | Clean fleece weight (kg) | Fibre diameter (microns) |
|---|---|---|---|
| Fine wool merino | Experimental | 4.0 | 18.3 |
| | Control | 3.5 | 19.7 |
| Medium wool merino | Experimental | 5.0 | 19.3 |
| | Control | 4.7 | 20.8 |

Data from NSW DPI Prime Facts 579, March 2007.

Several sheep breeds have been developed in Australia, including the Poll Dorset, a meat producer, and the White Suffolk, a dual-purpose wool and meat producer. Suffolk sheep have black faces and legs (see figure 9.48a). This breed produces superior, fast-growing lean lambs but growers were not receiving an appropriate market price for these lambs because their wool included some dark fibres. By selective breeding over many generations involving crosses of black-faced Suffolks and white-faced Poll Dorsets, a new breed, the White Suffolk, was produced in Australia. This new breed has all the features of the Suffolk breed except for the black face and legs (see figure 9.48b).


**FIGURE 9.48** Sheep breeds **(a)** Suffolk sheep with their black faces and legs. **(b)** White Suffolk, which has the features of the Suffolk except for the black face and legs. The White Suffolk was derived from a program of selective breeding and artificial selection, and was developed for a range of Australian pastoral conditions.

## Selective breeding: Australian cattle

The many different breeds of cattle are believed to have been derived by artificial selection from a wild species known as the auroch (*Bos taurus*), which was domesticated at least 6500 years ago.

The original cattle brought to Australia were British breeds that were suited to the temperate regions but did not thrive in tropical Australia. Selective breeding has produced a number of breeds for the tropics. These new breeds include the Australian Braford, which was developed in Queensland in the period from 1946 to 1952. The Australian Braford breed combines features of the Brahman, such as the hump, loose skin, short coat and heat and tick resistance, with features of the Hereford, including its colour markings (see figure 9.49). Another cattle breed developed in Australia is the Australian Milking Zebu, which resulted from selected crosses between Asian dairy cattle and European dairy cattle. This breed is a reliable milk producer in tropical regions, it is heat resistant because it is able to sweat and its loose skin makes it resistant to ticks, which it can easily shed.

**ODD FACT**

Aurochs are now extinct. The last member of this species was killed by a Polish poacher in 1627.

**FIGURE 9.49** Australian Braford cattle. Note the folds of loose skin. (Image courtesy of the Australian Braford Society.)

## Selective breeding by fanciers

Selective breeding of animals is not restricted to agricultural stock, such as cattle and sheep. Selective breeding has been used by animal fanciers and hobbyists to produce the great variety of breeds that may be seen in various pets, such as dogs, and in other domestic animals, such as poultry. This selective breeding continues today.

**Dogs:** The Australian National Kennel Council recognises 200 different breeds of dog, from Affenpinschers to Yorkshire Terriers. Figure 9.50 shows three dog breeds that show striking variation in terms of their coats — from almost zero hair in the Chinese Crested Dog to the Puli with its profuse corded coat.


**FIGURE 9.50** Selective breeding of dogs has produced a remarkable range of variation seen in the diversity of breeds recognised by **(a)** the Chinese crested dog, which is almost hairless with only fine tufts of hair on its head, **(b)** shih tzu, a toy-sized variety with long fine hair, and **(c)** puli, with its profuse corded coat hair. Note other structural differences, such as muzzle length. (Images courtesy of Judith Kinnear.)

**Poultry:** Frizzle, Indian game, Houdan, Cochin, Hamburgh, Australorp, Silkie, Andalusian, Dorking, Scots dumpy . . . what are they? These are all breed or varieties of the domestic fowl (*Gallus domestica*). The domestic fowl is believed to have descended from the wild red jungle fowl (*Gallus gallus*), which is native to a region of south Asia extending from India to Indonesia.

Over time, many mutations have arisen in the domestic fowl that have produced features that are not present in their wild type ancestor, the jungle fowl (see figure 9.51a), for example:

- Most domestic fowls have legs covered with scales, but some breeds, such as the Cochin, have feathered legs (see figure 9.51b).
- Most domestic fowls have four toes, but some breeds, such as the Dorking, have five toes (see figure 9.51c).
- Dark lacing on the edges of feathers can be seen in several fowl breeds (see figure 9.51d) in contrast to solid-coloured feathers in other breeds.
- Some fowl breeds have crests on their heads (see figure 9.51d) unlike most other breeds.

These mutations have been preserved in fowls by fanciers through their selective breeding practices.

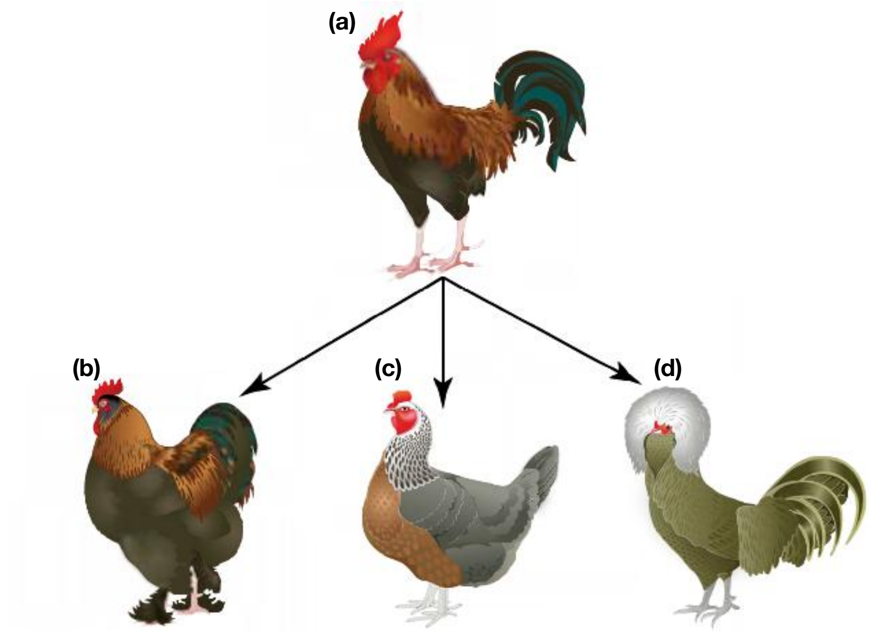



**FIGURE 9.51** **(a)** Red jungle fowl, the presumed ancestor of the domestic fowl. **(b)** Feathered legs in the Cochin breed of domestic fowl. **(c)** The Dorking breed with its five-toed feet. **(d)** The Polish breed of domestic fowl with crest and dark lacing on the feather edges. (Images (b), (c) and (d) courtesy of Judith Kinnear.)

The jungle fowl has a simple comb that is also seen in many domestic fowls, including the Australorp and the Leghorn. However, other comb shapes have arisen by mutation and have been retained through selective breeding. These different comb shapes include the V-shaped comb of the Crevecoeur, the rose comb of the Dominique, the walnut comb of the Silkie and the pea comb of the Brahma (see figure 9.52).

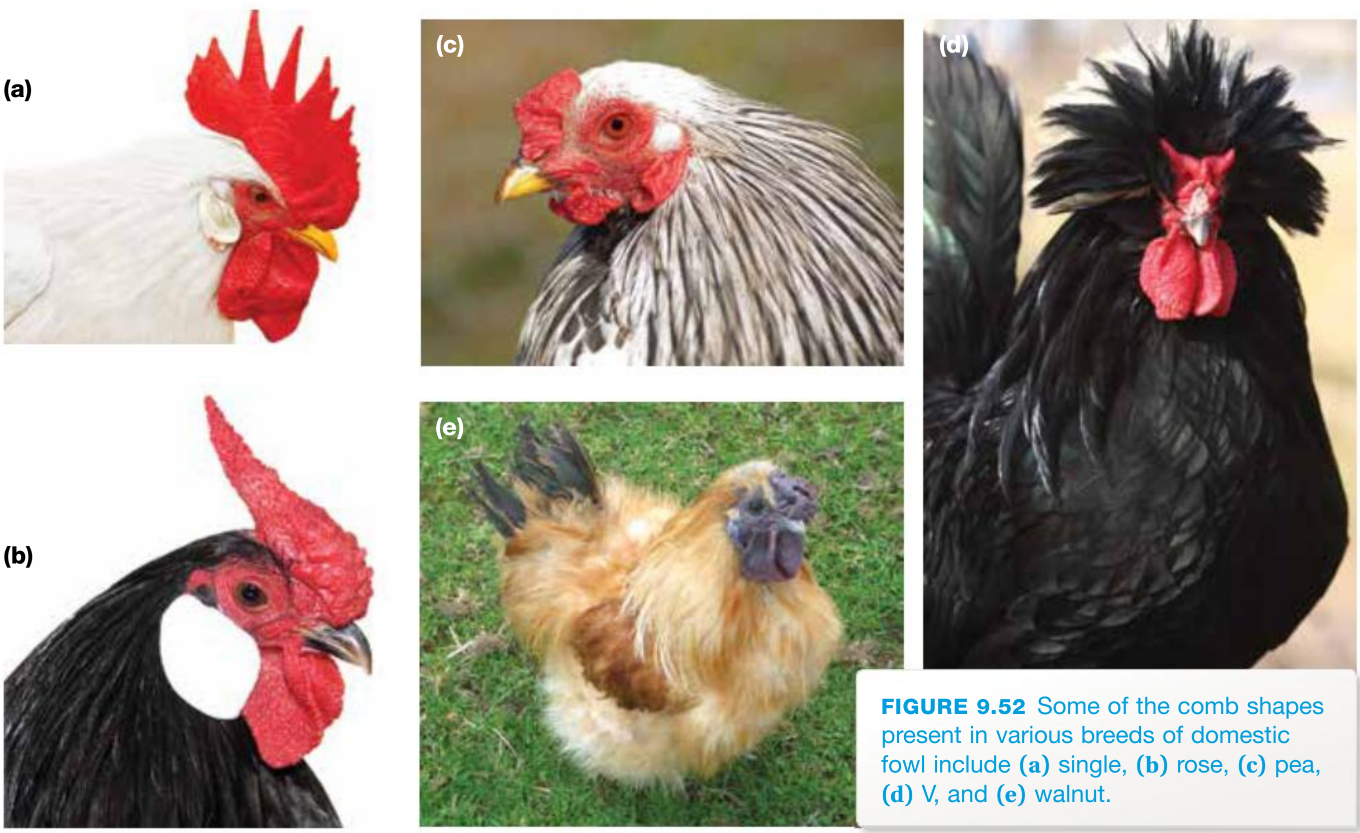



**FIGURE 9.52** Some of the comb shapes present in various breeds of domestic fowl include **(a)** single, **(b)** rose, **(c)** pea, **(d)** V, and **(e)** walnut.

The diversity of breeds of various farm and pet animals is the result of selective breeding. This selective breeding may have been part of formal breeding programs or the breeding activities of animal fanciers. The power of selective breeding is seen, for example, in chickens. Selective breeding over many generations has transformed wild fowls into farmed chickens with very different egg production capabilities. Wild fowls produce a total of only about 12 eggs during their short breeding season each year. In contrast, domesticated chicken varieties lay eggs continuously throughout the year and annually produce very large numbers of eggs.

## Effect of selective breeding on genetic fitness

Selective breeding or **artificial selection** occurs when breeders, fanciers and farmers prefer particular inherited features in their show animals and livestock because of their economic value or aesthetic appeal, and use selective breeding to enhance those features and increase their frequency in the breeding stock. In doing this, breeders manipulate the gene pool of their breeding stock.

Artificial selection is in contrast to natural selection, which favours only those inherited features that enhance survival and reproduction in the wild. Artificial selection can maintain features in a population that are economically important or aesthetically appealing but which are disadvantageous in terms of survival and reproduction. In natural populations in the wild, these features would come under negative selection pressure so that individuals with these features would not be maintained, and the alleles responsible would tend to be lost from the gene pool.

So, inherited characteristics that are the goal of selective breeding programs are not necessarily those features that best equip animals for survival and reproduction under natural conditions in the wild.

Artificial selection in domesticated species, particularly in so-called show varieties, can favour features that are clearly disadvantageous for survival and reproduction and would be positively selected against in the wild. Examples of features maintained only by artificial selection can be seen in domesticated animals such as:

- Jacobin pigeons with a distinctive arrangement of neck feathers that forms a ruff, masking their faces except from immediately in front (see figure 9.53a). Continued selective breeding over many generations has enhanced the size of the feather ruff that surrounds the face of the pigeon. Compare this image with figure 10.9, which shows a Jacobin pigeon in 1835. Because the large ruffs make it difficult for these pigeons to see, many breeders clip the ruff short during the breeding season to assist these birds to mate and ensure that they can feed their hatchlings.
- hairless Sphynx cats (see figure 9.53b) and dogs (see figure 9.50a above). The lack of an insulating fur coat in these pets means that, in cold environments, the homeostatic mechanisms that regulate core body temperature are at risk of failing.
- English bulldogs with greatly shortened muzzles that result in breathing problems (see figure 9.53c).

Other examples of disadvantageous features maintained thorough selective breeding can be seen in the short-legged dog breeds, such as dachshunds. These dogs are at a greatly increased risk of a disc herniation, in which one of the discs between the spine vertebrae is displaced. This creates pressure on the spinal cord, resulting in pain and, in some cases, temporary or permanent nerve damage.

Disadvantageous features such as these are maintained in the gene pools of these populations only through human intervention and selective breeding.

(a)

(b)

(c)

**FIGURE 9.53** **(a)** Jacobin pigeon. Compare this image with figure 10.9b on page 449, which shows a Jacobin pigeon in 1835. Continued selective breeding over many generations has enhanced the size of the feather ruff that surrounds the face. **(b)** Sphynx cat. These cats are essentially hairless but most have some very fine soft hair on their noses and ears. Hairlessness in cats has arisen as a spontaneous mutation several times in different parts of the world. **(c)** English bulldog. The characteristic short muzzle of this breed has several disadvantages, including impaired breathing.

## Technologies in selective breeding

In commercial herds and flocks, new reproductive technologies resulting in selective breeding include:

- artificial insemination
- sex selection through sperm sorting
- multiple ovulation and embryo transfer
- oestrus synchronisation.

### Artificial insemination

Natural breeding in mammals can be modified by **artificial insemination (AI)**. This technique changed herd management by altering the 'how', 'when' and 'where' of breeding.

AI involves collecting **semen** from a selected stud animal and then introducing this semen by artificial means into the reproductive tract of females of the same species. When first developed, AI involved the use of fresh semen only.

The use of AI increases the number of offspring that one stud animal can produce. One ejaculate from a male bull contains sufficient sperm to fertilise ten eggs. This volume of semen can be divided into ten portions and used to artificially inseminate ten cows and so can produce ten offspring. In contrast, in a natural one-to-one mating, that same volume of ejaculate would normally produce just one offspring.

In 1949, a successful technique was developed for freezing semen. 'Successful' meant that, when the frozen semen was thawed, it contained motile sperm that could fertilise eggs. The freezing technique involves adding

**ODD FACT**

To collect sperm from a stud animal, a 'teaser' animal is used to entice the stud to mate. When the stud tries to mount the teaser, his penis is redirected into a warmed tube that is used to collect the semen.

**ODD FACT**

By means of AI, a single ram can 'mate' with 2000 ewes as compared to a maximum of 200 ewes through natural matings.


**FIGURE 9.54** Transport of frozen semen and artificial insemination has geographically extended the reproductive capacity of prize studs.

semen to a special solution with a controlled pH and that comprises a mixture of various chemicals, including glycerol. Samples of this diluted semen (0.25 mL volume) are taken up in 'straws', frozen rapidly and stored in liquid nitrogen at –196 °C. Under these circumstances, semen samples can be stored for many years and still retain their ability to fertilise an egg after thawing.

Through AI technology using frozen semen, physical and temporal barriers to mating are removed. This technology means that one prize stud animal can:

- fertilise many more females than under natural conditions
- fertilise females located hundreds or thousands of kilometres distant from the stud animal because its frozen sperm can be easily transported over great distances (see figure 9.54)
- fertilise female animals and produce offspring long after its death.

A process similar to AI is used by plant breeders with populations of cultivated plants. In plants, the process is termed **artificial pollination** and it is another example of human intervention in the evolutionary process. Unlike AI, artificial pollination has been used for centuries.

## Embryo transfer in livestock (MOET)

AI and the use of frozen semen can greatly increase the contribution of particular bulls and rams to the genetic make-up of herds and flocks. In a similar way, a technique known as **multiple ovulation and embryo transfer (MOET)** allows high-quality cows and ewes to make a much greater than usual contribution to the future generations.

- **Multiple ovulation** refers to a process whereby a female receives injections of the follicle-stimulating hormone (FSH), which stimulates her to super-ovulate, or produce multiple eggs. An injection of gonadotrophin-releasing hormone (GnRH) is also given to make all the eggs mature at the same time.
- **Embryo transfer** refers to the process through which embryos at six to seven days of development are removed from the reproductive tract of a female and transplanted into the tracts of other females of the same species. These females act as surrogate mothers and carry the embryos to term and give birth.

MOET is another example of human intervention in the evolutionary process. In the process of MOET in sheep, for example, a high-quality donor ewe is treated so that she super-ovulates. When her eggs are released, they are fertilised, typically through AI with sperm from a selected ram. The fertilised eggs develop within the ewe's uterus for about six days. At the end of that time, the embryos are flushed from the ewe's uterus. On average, about seven embryos can be collected from a single flush. These embryos are immediately transferred directly into the uteruses of young recipient ewes (see figure 9.55), which will be the surrogate mothers of these embryos. Embryos not transferred to recipient ewes are frozen in liquid nitrogen and stored for later use. (Frozen embryos can be stored indefinitely.)

### ODD FACT

In sheep, AI can be achieved through either an intra-uterine or a cervical technique. In intra-uterine AI, semen is placed directly in the uterus of a ewe through use of a laparoscope, with average pregnancy rates of 70 to 90 per cent. With cervical AI, semen is introduced via the vagina, with average pregnancy rates of 55 to 65 per cent.

**ODD FACT**

MOET has made the movement of genetic stock between countries easier because quarantine regulations are less stringent for frozen embryos than for live animals.

**FIGURE 9.55** Transfer of an embryo into the uterus of a recipient, or surrogate ewe.

**ODD FACT**

The sex of an embryo produced by IVF can be identified at a very early stage of its development before it is transferred to a female. A single cell is taken from the embryo, cultured and the sex chromosome make-up is identified.

The same donor ewe can be used for MOET procedures several times during a breeding season. Over a normal reproductive lifetime, one ewe might produce 30 eggs. Multiple ovulation, however, greatly increases this egg output. The advantages of embryo transfer are that genetically important female lines can be multiplied at much faster rates than can occur through normal reproduction and that valuable embryos can be stored. Under conditions of natural selection, this would not be possible.

## Sex selection through sperm sorting

Under normal circumstances, a sex ratio of about one male to one female is expected in live-born mammals. In the beef industry, male calves are preferred because they have more beef (muscle) on their carcasses at a given age than females. In contrast, in the dairy industry, female calves are necessary for milk production and are preferred (see figure 9.56).

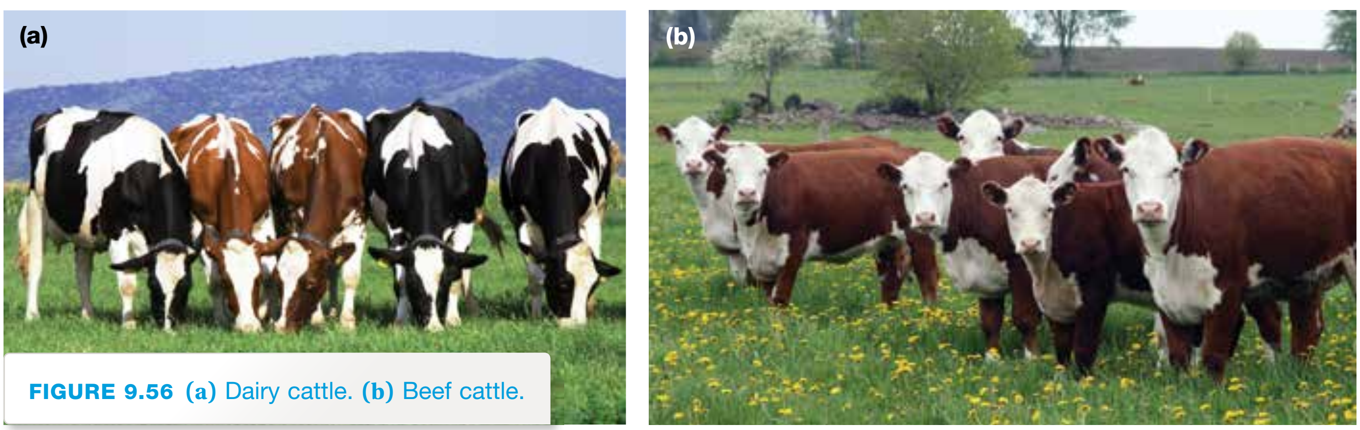



**FIGURE 9.56** **(a)** Dairy cattle. **(b)** Beef cattle.

**Sex selection** is now possible. After semen has been collected from a stud bull, for example, it is possible to treat the semen and separate sperm with X chromosomes from those with Y chromosomes. Sperm cells are first labelled with a harmless fluorescent dye that binds to the DNA. The X chromosome in mammals is larger and contains more DNA than the Y chromosome.

As a result, the sperm with X chromosomes fluoresce more brightly than those with Y chromosomes. After labelling, the sperm are then separated into two groups depending on their fluorescence. The use of this sperm separation technique has allowed sex selection to occur on a large scale (see figure 9.57).


**FIGURE 9.57** A technician at work with a FACStarPLUS flow cytometer used for the rapid and accurate separation of cells. Cells are stained with a fluorescent dye and this instrument uses a laser beam to sort cells according to differences in their fluorescence. If sperm cells were being sorted using fluorescence from a DNA-binding dye, which sperm would fluoresce more — sperm with an X or a Y chromosome?

## Manipulating breeding cycles

It is now possible to **synchronise the time of oestrus** or sexual receptivity of female farm animals, such as cattle and sheep. **Oestrus synchronisation** results in all sexually mature females being in oestrus within a predictable and narrow time frame, with the result that the time of fertilisation in a herd or flock, either by AI or natural mating, can be more efficiently managed. Synchronisation is also necessary for MOET procedures so that the intended embryo donor and the recipient surrogate mothers come into oestrus at the same time.

Advantages of synchronisation include:

- less time (and hence lower labour costs) needed to test animals to see if they are in oestrus
- higher fertilisation rates and birth rates
- more uniform and manageable crops of calves or lambs, since all young are born within a short period
- lower mortality rates because greater oversight of all newborns is possible.

Oestrus synchronisation can be achieved in a number of ways. One method depends on the fact that the hormone progesterone inhibits ovulation by stopping production of another hormone, oestrogen, which is needed to bring female animals into oestrus. By adding an external source of progesterone to female livestock, oestrus production and the associated ovulation are suppressed. When the source of progesterone is simultaneously removed from a group, mature females go into oestrus and ovulate within a short time period. How is this external progesterone delivered?

**ODD FACT**

Release of the hormone oestrogen from ovarian follicles stimulates the release of luteinising hormone (LH) from the pituitary gland. Release of LH in turn causes ovulation and suppresses oestrogen production.

Methods of supplying progesterone to farm livestock include:

- feeding using a dietary supplement
- implants under the skin
- sponges inserted into the vagina
- CIDRs (controlled internal drug-releasing devices) inserted into the vagina.

In the case of CIDR (pronounced cee-dar) use in cattle, the insert is left in place for seven days. When the insert with its supply of progesterone is removed, the level of circulating progesterone drops and oestrus begins within three days.

## Genetic impact of selective breeding

**Genetic variation** refers to the genetic differences present in the members of a population. In the gene pool of a population, the greater the number of different alleles of each gene, the greater the genetic variation in that population.

Selective breeding of domesticated animals, such as cattle and sheep, occurs when the parental input to the next generation is restricted to a small number of animals, selected on the basis of their superiority for a limited number of inherited traits. This process of artificial selection can result in a loss of genetic variation in the gene pool of the herd or flock concerned. The loss of genetic variation may not be noticed immediately and may become apparent only when there is a change in the environmental conditions. If, for example, members of an animal population have never been exposed to a particular disease-causing virus, the loss of an allele that confers resistance to that disease would not be noticed. The value of the allele in question would be realised only after the virus comes into the environment.

Genetic variation can also be lost when plant crops are cultivated from commercially produced seeds that come from just one, or a small number, of the many varieties of that plant species.

Loss of genetic variation can be an unintended consequence of the use of reproductive technologies, such as artificial insemination (AI) combined with the transport of frozen semen, which increases the genetic influence of a small number of stud males in animal herds and flocks. Similarly, the use of multiple ovulation and embryo transfer (MOET) restricts the number of breeding females in animal herds and flocks. Likewise, the use of artificial pollination with a limited number of plants as the pollen source may reduce the genetic variation of a plant population under cultivation.

Genetic variation is lost when habitats are destroyed in a region with the local loss of plant and animal species. Although other populations may exist in other regions, what will be lost is the genetic variation that is unique to those local gene pools.

## Does loss of genetic variation matter?

The death of more than one million Irish peasants from starvation in the period from 1845 to 1849 is a tragic reminder that genetic variation matters. Potatoes were the staple diet of Irish peasants. When an outbreak of late potato blight occurred, caused by the fungus *Phytophthora infestans*, potato tubers turned black and rotted in the ground.

Potatoes are endemic to South America, where hundreds of species grow wild. This food crop was introduced to Europe, first by the Spaniards in 1570 and later, in about 1590, by the English. All potatoes grown in Europe were plants descended from the 'European' potato (*Solanum tuberosum*). When potato blight broke out in Ireland in 1846, this potato variety was genetically uniform for susceptibility to fungal infection so that all the potato tubers rotted. This produced a mass famine that resulted in the deaths of an estimated one million people, either directly from starvation or from starvation-related diseases. About another one million people emigrated, the majority to the United States, but significant numbers came to Australia. This tragic event is commemorated in a sculpture on the Custom House Quay beside Dublin's Liffey River (see figure 9.58).

Natural populations of the wild relatives of plants can be the source of alleles for potentially valuable traits. For example, in the 1940s, resistance to the late potato blight disease was identified in *Solanum demissum*, which grows wild in the Andes and is a close relative of the European potato. Crossbreeding was used to introduce this allele into the commercial varieties of potato, providing them with resistance to blight.

**FIGURE 9.58** Sculpture by Rowan Gillespie commemorating the Great Famine, located on Custom House Quay in Dublin's Docklands. (Image courtesy of Judith Kinnear.)

## Saving genetic variation

As wild populations are destroyed through land clearance and as larger areas are devoted to the cultivation of smaller numbers of commercial crop varieties, the safeguarding of genetic variation in wild populations is critical. In the case of plant varieties, some contribution to this is being achieved through the establishment of seed banks.

### Gene banks and seed banks

In 1992, Western Australia became the first Australian state to establish a genetic diversity seed bank when the *Threatened Flora Seed Centre* was established at the WA Herbarium. In 2011, the centre relocated to state-of-the-art facilities within the WA Conservation Science Centre. Since starting operations in 1993, the centre has built a store of seeds from many threatened plant species. Seeds of rare and endangered plant groups are collected from various populations of the native plant groups. (Why collect from more than one population?) After collection, the seeds are carefully cleaned and counted in a laboratory (see figure 9.59a). The seeds are then dried to reduce their water content, then stored in sealed aluminium foil packs at low temperatures (see figure 9.59b). The centre also engages in research and has reintroduced many threatened species back into the wild.


**FIGURE 9.59** **(a)** Processing and testing of seed in the laboratory prior to storage. **(b)** The freezer. Extracted and dried seed is sealed in laminated aluminium foil packages and frozen at –18 °C for the long term. (Images courtesy of Threatened Flora Seed Centre, Science and Conservation Division Department of Parks and Wildlife, Western Australia.)

Other states have also established seed banks to store seeds from native plants and together they form the Australian Seed Bank Partnership. The Victoria node of this partnership is located in the National Herbarium of Victoria at the Royal Botanic Gardens Victoria in Melbourne.

Other facilities have a major focus on the seeds of crop plants. In March 2014, the Australian Grains Genebank was opened in a new facility in Horsham, in country Victoria. This centre combines the temperate field crops collection from Horsham, the cereals crop collection from Tamworth in New South Wales, and the tropical crops collection from Biloela in Queensland. The gene bank houses seeds from pulses (pea and bean species), cereals, oilseeds and, importantly, from their wild relatives from temperate and tropical regions of Australia. Seeds from indigenous wild crop varieties also form part of this collection. This facility will hold up to 180 000 samples from Australia and around the world. Samples of 1000 seeds are dried to 6 per cent moisture and then stored in vacuum-sealed, triple-foil aluminium pouches at –20 °C. Small seed samples of 100 seeds are distributed free on request to plant breeders, scientists and research students, and are also exchanged with other gene banks internationally. When stocks become low, seeds are grown for regeneration. This is not only a means of replenishing seed stocks, but it also provides an opportunity to collect relevant data on seed performance, such as time to flowering, height, and grain yields.

An important step in the preservation of the genetic variation and biodiversity of the world's major crop plants and their wild relatives occurred in February 2008 when the Svalbard Global Seed Vault (SGSV) was opened on remote Spitsbergen Island close to the North Pole (see figure 9.60). The SGSV facility has large underground storage areas dug into the side of a mountain and these are maintained at a temperature of −18 °C (see figure 9.61). Here, duplicate samples of seeds from seed banks throughout the world can be placed in secure long-term storage at no cost. This storage acts as an insurance against the loss of the valuable genetic variation in food crop plant varieties and their wild relatives.


**FIGURE 9.60** The Norwegian island of Spitsbergen, north of the Arctic Circle, is the location of the Svalbard Global Seed Vault.

At capacity, SGSV can store more than 4 500 000 samples, each containing up to 500 seeds. During the first year of its operation, various national and international groups deposited more than 400 000 unique seed samples. The first seeds from Australia were deposited in the SGSV in early 2011.

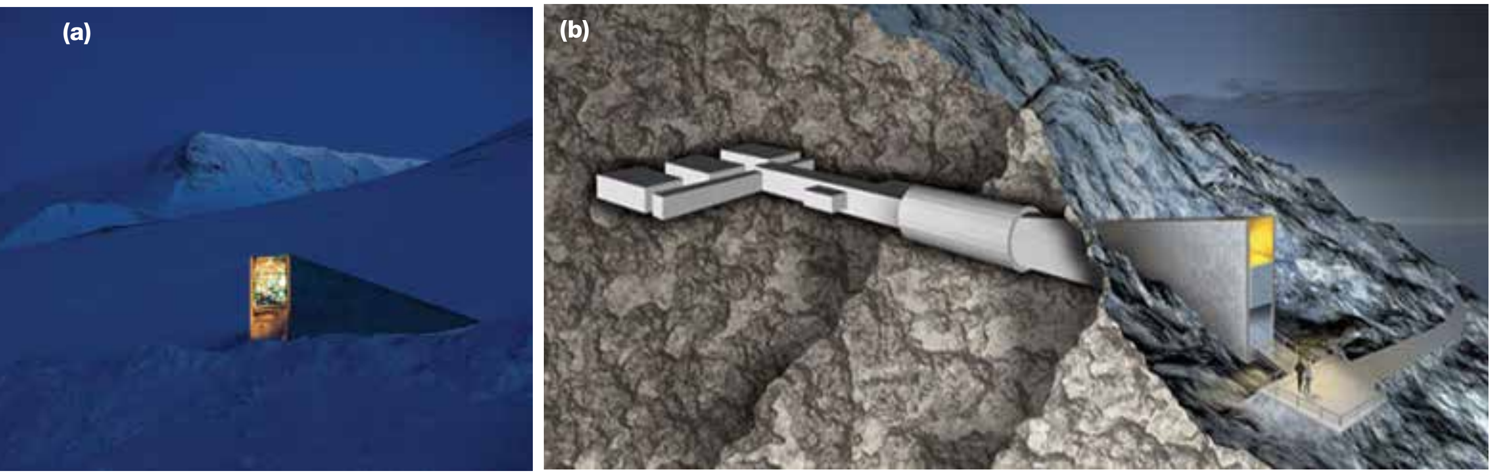



**FIGURE 9.61** **(a)** Entrance to the Svalbard Global Seed Vault on the island of Spitsbergen. **(b)** Diagram showing the three storage areas at the end of a tunnel deep within a mountain. This facility will ensure that the genetic variation in plant crops is preserved for future generations. (Images courtesy of the Global Crop Diversity Trust.)

## BIOLOGISTS AT WORK

*Catherine Cavallo and Sonia Sanchez are PhD students with Monash University and Phillip Island Nature Parks. Together, they study little penguins in their environment, in an effort to better understand and protect marine ecosystems for the future.*

Cathy and Sonia write: 'Marine ecosystems are under increasing stress from climate change and other pressures such as pollution and overfishing. We rely on these ecosystems for food, oxygen production, employment and recreation, so it is critical that we protect them. But first, we have to understand how they function. In marine environments, this is a difficult and expensive task, requiring an innovative approach.'

'To understand what goes on beneath the waves, we monitor little penguins at Phillip Island. As top predators, penguins reflect changes in the food web below them, and are useful indicators of ecosystem health. We use an automated monitoring system to count and weigh penguins as they enter and exit the colony over a specially designed bridge. This tells us how long penguins have to spend hunting to find enough food, and how successful they are. We also check nests during the breeding season, counting eggs and chicks and weighing chicks to monitor their growth rates. Little penguin breeding success is very sensitive to prey availability, so we can learn a lot about the ecosystem by observing how quickly their chicks grow and how many survive to leave the burrow.'

Cathy continues: 'My specific role is to use DNA analysis to understand penguin diet, by identifying prey DNA in their poo. It's a smelly task, but an important one, which gives us invaluable information about the abundance and availability of little penguin prey, and, consequently, health and change in the penguins' environment.

'Several times a year, I get to travel to the Australian Antarctic Division in Tasmania to analyse my samples alongside scientists who work in Antarctica. I also go to conferences to share ideas on how we can protect the planet with our colleagues from all over the world. My PhD project involves my working alongside government (Australian Antarctic Division), tourism (Phillip Island Nature Parks) and media, and means working in some incredible locations.

'As a kid, I always knew I wanted to work with animals, but as I grew I fell deeply in love with the sea, and felt compelled to understand and protect it. Through my research, I feel that I am helping to protect our precious marine environments and species from the growing pressures humans place on them. My study path included a Bachelor of Science (majoring in zoology) and a Master of Science studying sea turtles and climate change, as well as taking every opportunity that was offered to me to increase my research experience and networks.'

Sonia adds: 'Currently, less than 4 per cent of our oceans are protected. Tracking marine top predators at sea can be an important tool to identify important sites to protect within marine protected areas, as they feed in areas where food is available and abundant. I am trying to find out where the feeding hotspots of little penguins from Phillip Island are. I attach two miniaturised devices to their backs, a GPS and an accelerometer, to record their location and behaviour at sea. With this technology I can see where the penguins go and where they find food. By protecting these areas, we protect not only the food source for the penguins, but the wider ecosystem as a whole. I spend many nights waiting for my penguins to return to their burrows, so I can retrieve their tracking devices.

'When I was a kid I used to spend all my summers swimming and exploring at the beach and I guess that is when I decided I wanted to study marine sciences, to protect and understand that sea that made me so happy. I did a Bachelor of Science and a Master in Marine Sciences in Barcelona (Spain), my home city, where I used to work on anchovies and sardines, and spent time living and working with other researchers on an oceanographic vessel.'


**FIGURE 9.62** **(a)** Cathy (at right) and Sonia at work. **(b)** Adult penguin and baby. (Images courtesy of Phillip Island Nature Parks Australia.)

## KEY IDEAS

- Selective breeding is one form of artificial selection used in plant and animal breeding.
- Artificial selection using selective breeding can maintain inherited features in a population that, under natural conditions, would be selected against.
- Several new reproductive technologies are available to assist selective breeding in commercial flocks and herds.
- Reproductive technologies, such as AI, artificially increase the contribution of selected animals to the gene pool of the next generation.
- Use of reproductive technologies can result in reduced genetic variation in a gene pool.
- Loss of genetic variation in a population may result in failure of the population to survive environmental change.
- Seed banks and gene banks are among the facilities that seek to conserve the genetic diversity of plant species.

## QUICK CHECK

**22** Define the following terms:
  **a** artificial insemination
  **b** artificial pollination.

**23** What is the likely effect on the genetic variation in the gene pool of herds where reproduction involves AI using semen from one prize stud?

**24** Identify two consequences for AI when technology allowed semen to be successfully frozen.

**25** What is meant by the abbreviation MOET?

**26** **a** Why did the potato crop in Ireland fail in the 1840s?
  **b** What does this show about genetic variation?

# BIOCHALLENGE

1 Consider the diagrams (A to D) below. Match each diagram to one of the chapter keywords giving a brief explanation for your choice. Note that each coloured dot represents an individual organism.

2 Explain what change, if any, each of these events would produce on the gene pool of the population denoted on the right-hand side of each diagram.

3 The two generations shown in D below have been living under a particular set of environmental conditions. The release of a pollutant changes the environmental conditions so that individuals denoted by the pink dots are at a selective advantage. Draw a possible representation of the generation 3 population.

4 A normal protein has 141 amino acids.
A single gene mutation produces a mutant form of the protein that is 31 amino acids longer.
Describe what kind of mutation could have caused this.

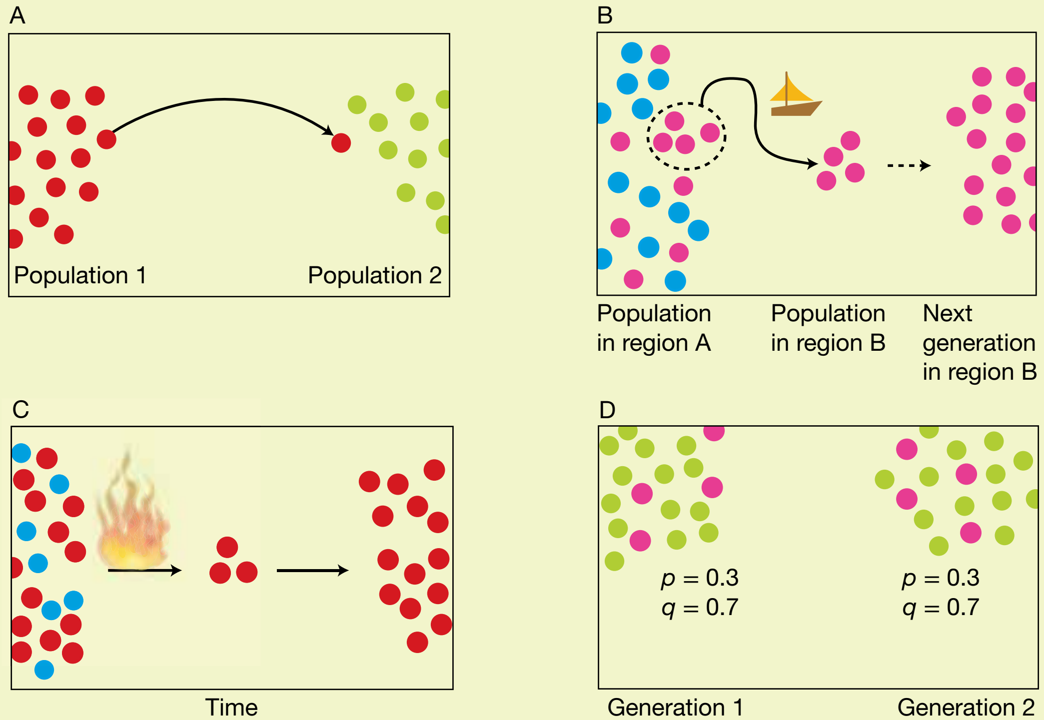



**FIGURE 9.63**

# Chapter review

studyon

Unit 4 | AOS 1 | Topic 1 — Changes in the genetic make-up of a population

Practice questions

## Key words

agent of selection
allopolyploid
aneuploid
artificial insemination (AI)
artificial pollination
autopolyploidy
bottleneck effect
complete selection
conservative missense
conservative missense mutation
continuous
differential reproduction
discontinuous
embryo transfer
fixed
founder effect
founder population
frameshift
gene mutation
gene pool
genetic drift
genetic variation
Hardy–Weinberg principle
hybridisation
immigration
intra-specific evolution
lactase nonpersistent
lactase persistent
LCT
monomorphic
multiple ovulation
multiple ovulation and embryo transfer (MOET)
mutagenic agents
natural selection
non-conservative missense mutation
oestrus
oestrus synchronisation
partial selection
polygenes
polymorphic
polymorphisms
population
position effects
post-mating isolation mechanisms
pre-mating isolation mechanisms
random mating
selecting agents
selective advantage
selective breeding
semen
sex selection
speciation
trinucleotide repeat expansion (TRE)

## Questions

1 **Making connections ➔** Draw a concept map for 'genetic changes in populations' incorporating key words from this chapter. You may add any other concepts that you wish.

2 **Demonstration knowledge and understanding ➔** Identify the difference between the members of the following pairs:
- a gene flow and gene pool
- b genic change and chromosomal change (in the make-up of a population)
- c natural selection and artificial selection
- d germline mutation and somatic mutation.

3 **Demonstrating knowledge and understanding ➔** In an animal species, body colour is inherited with yellow being dominant over purple. Consider two populations of this species: Population A consists of 9 purple and 1 yellow individual; population B consists of 900 purple and 100 yellow individuals.
- a In which population would genetic drift be more likely to lead to loss of the allele controlling the yellow colour? Explain.

  Population C consists of 500 purple and 500 yellow individuals.

  A sample of ten individuals is chosen at random from each of populations B and C. These small groups will be the founder groups of new populations in two different regions.
- b Which founder group is more likely to have less genetic diversity in terms of body colour. Explain.

4 **Demonstrating knowledge and understanding ➔** Consider the following mutations in the coding strand of DNA and identify the kind of mutation that they represent. Assume that the code will be read starting from the first base shown.
- a . . . T T T  T C T  A G G  G T C ➔
  . . . T T T  T **G** T  A G G  G T C
- b . . . T T T  T C T  A G G  G T C ➔
  . . . T T T  T C T  **A** A G  G G T  C
- c . . . T T T  T C T  A G G  G T C ➔
  . . . T T T  C T A  G G G  T C
- d . . . T T T  A T T  G T C  C C T ➔
  . . . T T T  A T **C**  G T C  C C T

5 **Developing valid biological explanations ➔** Suggest a reasonable explanation for the following observations:
- a Indigenous populations in Australia and North America show a high percentage of lactase nonpersistence, but some indigenous populations in Africa exhibit a low percentage lactase nonpersistence.
- b Populations descended from groups that do not have a history of herding dairy animals have only a small proportion of lactase persistent individuals.

6 **Demonstrating knowledge and understanding ➔** An autosomal gene in a population of diploid organisms has eight different alleles. How many different alleles could be present in:
- a one member of this population
- b the gene pool of this population?

7 **Analysing data and applying principles ➔** In an imaginary large population of sexually reproducing

organisms, foot colour is a monogenic trait with green feet (***G***) being dominant to yellow (***g***). Both green-footed and yellow-footed organisms occur in this population. These short-lived organisms normally breed and then die. Assume that all organisms with green feet are prevented from breeding.

**a** Under these conditions, how rapidly will the ***G*** allele be lost from the gene pool of this population?

**b** What are values of *p* and *q* when the ***G*** allele is removed from the gene pool?

Return to the original population with organisms of both foot colours and assume that all yellow-footed organisms are prevented from breeding.

**c** Under these conditions, will the ***g*** allele be lost from the gene pool as quickly as the ***G*** allele? Explain.

**d** Which genotype acts as the hidden reservoir of an allele that controls a recessive trait?

**8 Developing valid explanations →**

**FIGURE 9.64** Two adult dogs, a Great Dane and a Chihuahua.

This figure above shows examples of inherited variation in dogs.

Identify, giving an explanation for your decision, the most likely genetic basis for:

**a** the observed variation in body size between the two dogs

**b** the observed variation (black/yellow) in the base colour of the two dogs (ignore the white spotting).

Consider the proposal that these size differences emerged in wild dog populations by natural selection.

**c** Is this assumption reasonable or not? Explain your choice.

**d** If you decided that the assumption is not reasonable, suggest an alternative explanation.

**9 Analysing information →** Read the following abstract of a research report by Iannuzzi, M.C. *et al.* published in the *American Journal of Human Genetics* vol. 48m, page 227 (1991):

'Cystic fibrosis (CF) is a recessive disease caused by mutations in the CF transmembrane conductance regulator (CFTR) gene. We have identified in exon 7 two frameshift mutations, one caused by a two-nucleotide insertion and the other caused by a one-nucleotide deletion; these mutations–CF1154insTC and CF1213delT, respectively, are predicted to shift the reading frame of the protein and to introduce termination codons at residues 369 and 368.'

**a** What is a frameshift mutation?

**b** Two frameshift mutations are identified in this case of cystic fibrosis. Note that CF means cystic fibrosis and the following number identifies a base position in the DNA of that gene.
What do you think is meant by the shorthand CF1154insTC and CF1213delT?

**c** What effect did these frameshifts have on the protein encoded by this mutated gene?

**10 Demonstrating understanding and communication →** The sheep strike blowfly, *Lucilia cuprina*, is a major pest for sheep farmers. In 1955, the insecticide dieldrin was first introduced to control this pest. The introduction was initially highly successful, but, within two years, this pesticide began to be less effective against these blowflies.

**a** Explain carefully what was happening in this two-year period?

**b** Outline what would be expected to happen if dieldrin continued to be used.

**11 Analysing data and making predictions →** Populations of giraffe (*Giraffa camelopardalis*) from different regions on the African continent show distinctive differences in inherited traits, including colour and pattern. In all, up to nine **subspecies of giraffe** are recognised. Figure 9.65 shows the geographic distributions of the various groups.

**a** What might be expected to happen over a very long period if the populations remain geographically separated?

**b** Suggest a likely origin for the different colours and patterns in the different subspecies.

**c** What test would tell if the giraffe populations are the same or different species?

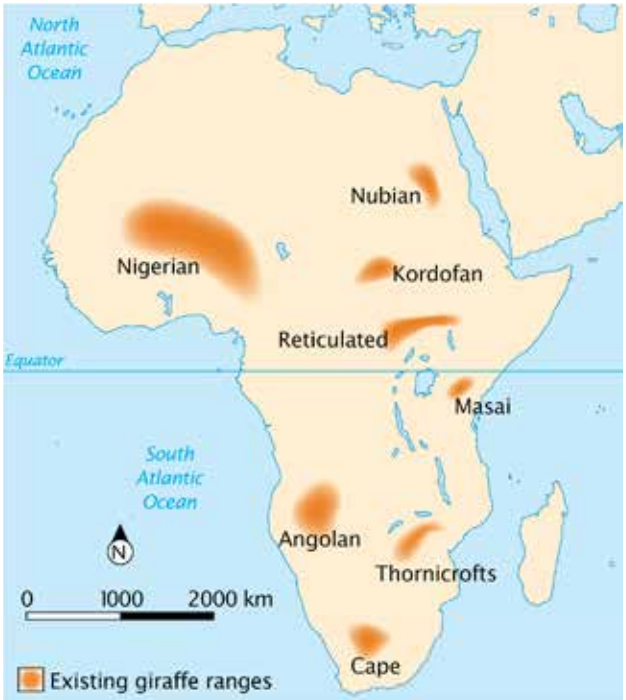



**FIGURE 9.65** (Map provided by MAPgraphics Pty Ltd, Brisbane.)

**12 Demonstrating understanding and communication** → Examine figure 9.66, which shows stylised representations of the gene pool of the same population at three different times. The time periods are sufficiently long for several generations to have passed. The letters denote the alleles, C and c, of one gene.

The images are not in any particular order. Place the boxes in an order of your choice that you think is feasible and tell the story of what has happened in the population over the time period involved, consistent with your chosen sequence of gene pool images.

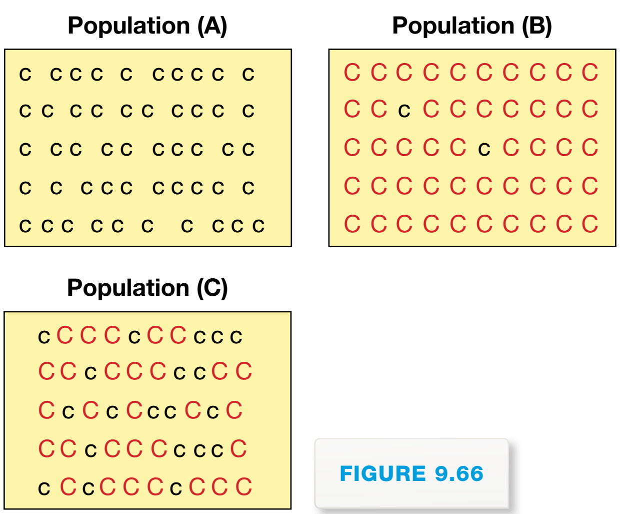



**FIGURE 9.66**

**13 Analysing data and drawing conclusions** → Refer back to table 9.5 on page 399. The areas of Africa, south-west Asia and southern Europe subject to malaria are shown in figure 9.67a. The frequency of the $H^S$ allele that is responsible for sickle-cell anaemia is shown in figure 9.67b.

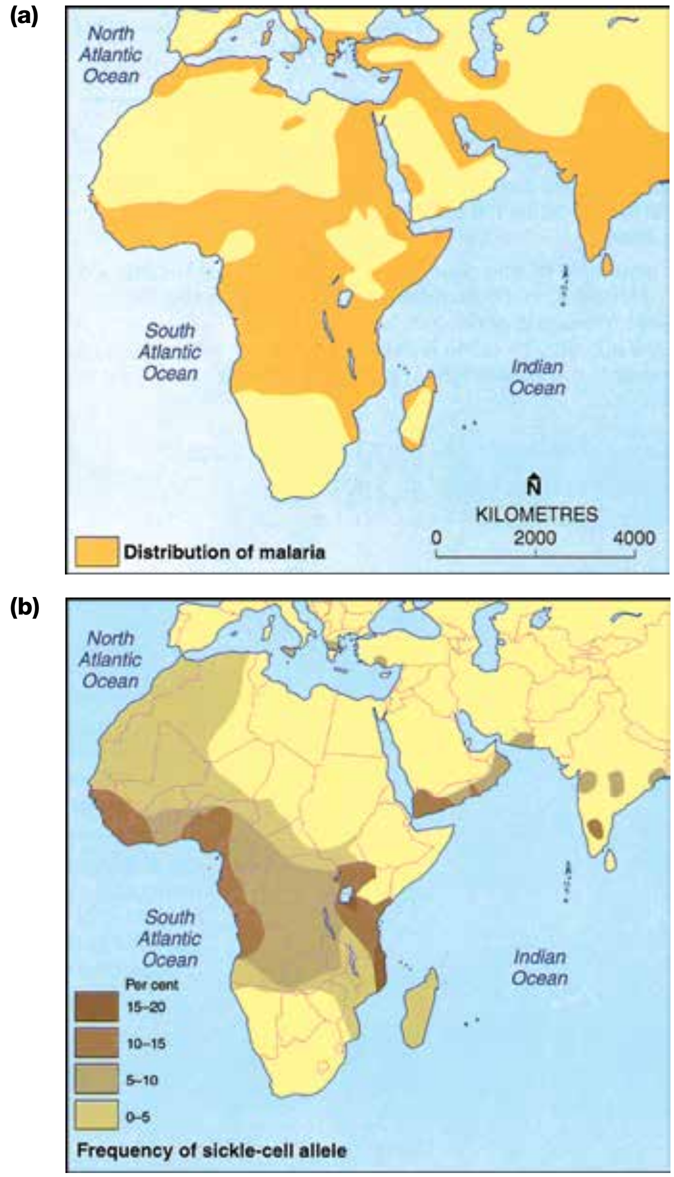



**FIGURE 9.67** **(a)** Map showing malaria-prone areas. **(b)** Map showing frequency of $H^S$ allele. (Maps provided by MAPgraphics Pty Ltd, Brisbane.)

**a** What is the most likely origin of the $H^S$ allele?
**b** Is the high frequency of the $H^S$ allele in areas of Africa best explained by the action of natural selection or by genetic drift? Explain.
**c** If malaria were eradicated from this planet, what changes in the frequency of the $H^S$ allele might be expected to occur over time? Explain.

d The frequency of the $H^S$ allele is not rare in the United States, yet that country is not significantly affected by malaria. Suggest a possible explanation for this observation.

e The $H^S$ allele is extremely rare in Australian populations and has a frequency far less than the frequency found in the United States. Suggest a possible explanation for this observation.

14 **Discussion question ➜**

**Consider the following statement:**

*Selection pressures can change — it is likely that the predation pressure acting on hominins was reduced by cultural development such as toolmaking and the controlled use of fire.*

(Hominins include the early human species *Homo habilis* and a later species, *Homo erectus* (whom you will meet in chapter 12.) Note that cultural practices are not inherited traits.)

a What is meant by the term 'predation pressure'?

b What is meant by 'the controlled use of fire'?

The first humans to exercise the controlled use of fire are believed to be the human species *Homo erectus*.

c Discuss how the ability to exercise the controlled use of fire might contribute to a reduction in predation pressure on populations of *Homo erectus*.

d Identify other benefits that could been generated by the controlled use of fire in these populations.

Simple stone tools first appeared in association with *Homo habilis* and then came more sophisticated tools made by *Homo erectus*. Populations of early human species survived by hunting and gathering.

e Discuss how the development of various tools might have resulted in more efficient and more successful hunting and gathering practices of early humans.

CHAPTER

# 10 Changes in biodiversity over time

**FIGURE 10.1** An excavator working on the fossilised bones of a plant-eating sauropod dinosaur (*Jobaria* sp.), which in life was about 21 metres long. This dinosaur fossil was found in a region of the Sahara Desert in present day Niger, in western Africa. Among topics covered in this chapter are fossils, their types, how they are dated and interpreted, and how the fossil record provides evidence of the changes in Earth's biodiversity over geological time.

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- appreciate the explanatory power of Darwinian evolutionary theory
- understand that Earth's life forms have changed or evolved over geological time
- become familiar with the geological time scale
- gain knowledge of the nature of fossils, their types and how they are formed and dated
- recognise the various lines of evidence of biological change over time
- demonstrate knowledge of the patterns of biological change over geological time.

# Species: unchanging or not?

**ODD FACT**

A medieval view was that birds and fish originated from leaves that fell, either onto land or into water. Leaves falling into water slowly transformed into fish, while leaves falling onto land slowly transformed into birds (see figure 10.2.)

A remarkable diversity of living organisms is found today on planet Earth — an estimated 30 million different species. In addition, if extinct species are also included, the diversity of life on planet Earth is even more striking. How did this diversity come about?

In the early 1800s, the general view held by many scientists, particularly in Britain, was that species were unchanging and that each species was fixed in its structure and characteristics for all time. According to this view, each species was the result of an act of creation — a view known as the **special creation of species**. If this were the case, each of the many millions of different species, living and extinct:

- would have been specially and individually created
- would not have been capable of change
- would not have been able to produce different descendant species.

In contrast, an alternative view was that species were not fixed but were capable of change over successive generations, a view known as the **transmutation of species**. If this were the case, ancestral species could produce descendants with different and new features so that, over time, these ancestral species could give rise to new species.

These contrasting views — transmutation of species versus special creation of species — are not compatible. The view that species were fixed and each was specially created was finally dismissed as a result of the theory of evolution by natural selection developed by Charles Darwin and Alfred Wallace.


**FIGURE 10.2** A representation of one medieval view of the origin of birds and fish.

## The Darwin–Wallace view

Various naturalists in the seventeenth and early eighteenth centuries proposed that species could evolve or be transformed to produce new species, but none of these naturalists identified a testable mechanism to explain how evolution could occur. This changed when Charles Darwin (1809–1882) and Alfred Russel Wallace (1823–1913) (see figure 10.3) proposed their theory of evolution by natural selection.

The power of the Darwin–Wallace theory of evolution by natural selection was that:

- the theory identified a mechanism or a cause of evolution
- this mechanism was testable by observation and experimentation.

Refer back to chapter 9 (page 395) for an outline of the mechanism of natural selection.

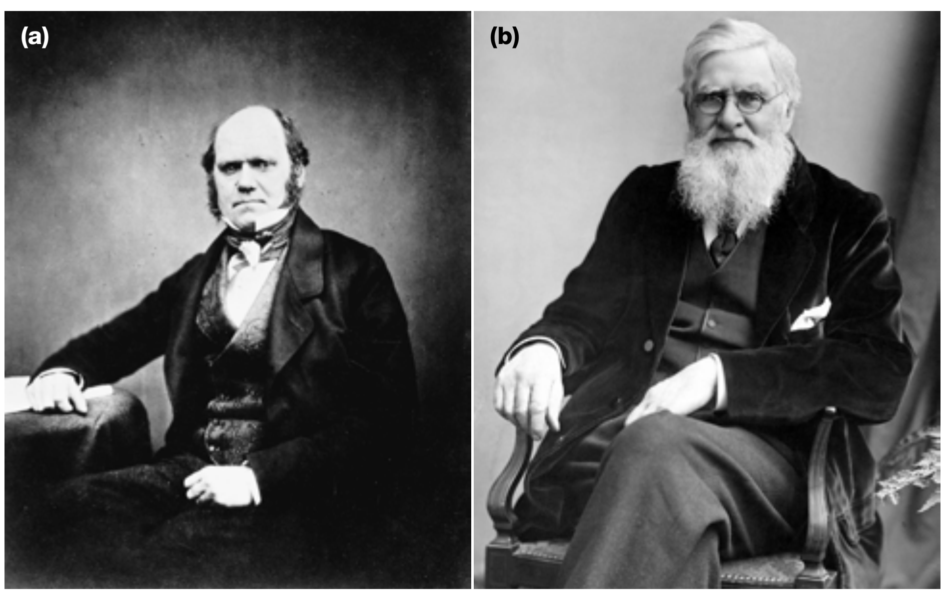



**FIGURE 10.3** **(a)** Charles Darwin (1809–1882) aged 40 in 1849. **(b)** Alfred Russel Wallace (1823–1913).

**ODD FACT**

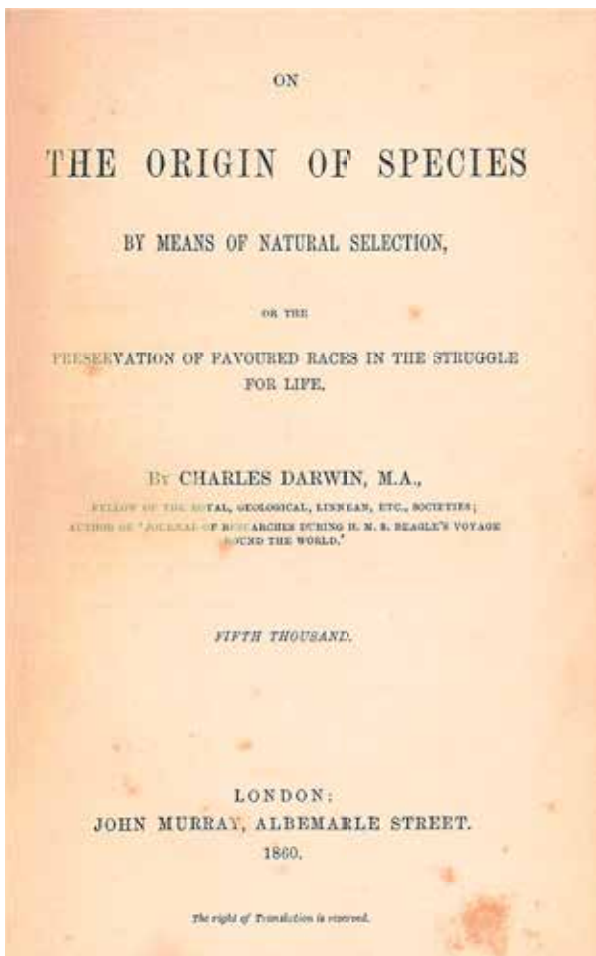
At the time of Darwin's death in 1882, his *The Origin of Species* was into a sixth edition and 24 000 copies had been published. *The Origin of Species* has since been translated into twenty-nine languages including Hindi, Turkish and Flemish. In 1934, the first English Braille edition was published.

The theory of Darwin and Wallace was first presented publicly at a short scientific meeting of the Linnean Society in London on 1 July 1858 and this passed largely unnoticed. Just over a year later, in October 1859, the first edition of Darwin's book entitled *On the Origin of Species by Means of Natural Selection* was published (see figure 10.4). The second and third editions quickly followed in 1860 and 1861 respectively. This book caused quite a stir. It put the theory of evolution by natural selection into the public domain and stimulated widespread and sometimes bitter debate — perhaps like debates today on topics such as abortion and euthanasia. Some prominent people supported Darwin's theory while other equally prominent people opposed this theory.

By arguing that species could evolve by natural selection, Darwin's theory challenged head-on the commonly accepted view that each species was specially created and unchanging. Darwin's theory of evolution also changed for all time the way in which human beings saw their place in the natural world. No longer could they stand apart from that world because, like other species, they too had evolved.

The topic of evolution became a matter of major discussion and debate in newspapers, magazines, churches and public meetings. The box *The Oxford Debate of 1860* on pages 446 to 447 describes a famous public debate that took place at the meeting of the British Association at Oxford in 1860. The key players in this debate were:

- Samuel Wilberforce (1805–1873), then Bishop of Oxford, known as 'Soapy Sam'
- Thomas Henry Huxley (1825–1895), a biologist, known as 'Darwin's bulldog'
- Joseph Dalton Hooker (1817–1911), a botanist and biogeographer.



ON

THE ORIGIN OF SPECIES

BY MEANS OF NATURAL SELECTION,

OR THE

PRESERVATION OF FAVOURED RACES IN THE STRUGGLE FOR LIFE.

BY CHARLES DARWIN, M.A.,

FIFTH THOUSAND.

LONDON:
JOHN MURRAY, ALBEMARLE STREET.
1860.

*The right of Translation is reserved.*

**FIGURE 10.4** The title page of the second edition (published in 1860) of Charles Darwin's *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, more commonly known as *The Origin of Species*. Following the publication of the first edition in late 1859, the ideas that it contained about evolution caused wide public discussion. (Image courtesy of Judith Kinnear.)

## THE OXFORD DEBATE OF 1860

It is Saturday 30 June 1860 at Oxford, England. A debate on Darwin's theory of evolution is about to take place. Excitement is high as it is widely known that Bishop Samuel Wilberforce (see figure 10.5) intends to 'smash Darwin'. The room where the debate is to take place is crowded. Darwin is not present but some of his supporters are, including Thomas Huxley (figure 10.6) and Joseph Hooker ( figure 10.7).

**FIGURE 10.5** Samuel Wilberforce, Bishop of Oxford, outspoken critic of Darwin's theory.

Various accounts of this debate exist and they do not all agree. Some accounts follow. (In the extracts, the use of '. . .' denotes that part of the account has been omitted for brevity. Material appearing [in square brackets] is an addition to assist readers' understanding.)

Hooker describes the debate as follows:

> The excitement was tremendous . . . The numbers [in the audience] . . . were estimated at from 700 to 1000. The President [who] occupied the chair wisely announced that none who had not valid arguments to bring forward on one side or the other would be allowed to address the meeting . . .
>
> The Bishop was up to time, and spoke for full half-an-hour with inimitable spirit, emptiness and unfairness . . . He ridiculed Darwin badly, and Huxley savagely, but all in such dulcet tones, so persuasive a manner . . . that I who had been inclined to blame the President for allowing a discussion that could serve no scientific purpose, now forgave him from the bottom of my heart. Unfortunately the Bishop, hurried along on the current of his own eloquence . . . turned round and addressed Huxley . . . [and] asked whether Huxley was related by his grand-father's or grand-mother's side to an ape.
>
> [Huxley replied:] 'I asserted, and I repeat, that a man has no reason to be ashamed of having an ape for his grandfather. If there were an ancestor whom I should feel shame in recalling, it would be a man, a man of restless and versatile intellect, who, not content with an equivocal success in his own sphere of activity, plunges into scientific questions with which he had no real acquaintance, only to obscure them by an aimless rhetoric, and distract the attention of his hearers from the real point at issue by eloquent digressions, and skilled appeal to religious prejudice.'
>
> . . . the Bishop made no reply . . . Lady Brewster fainted [probably from the heat as the room was crowded to suffocation] and the meeting broke up.

(*Source:* Darwin, Francis (ed.), *The Life and Letters of Charles Darwin*, Vol. II, John Murray, London, 1887, pp. 321–2)

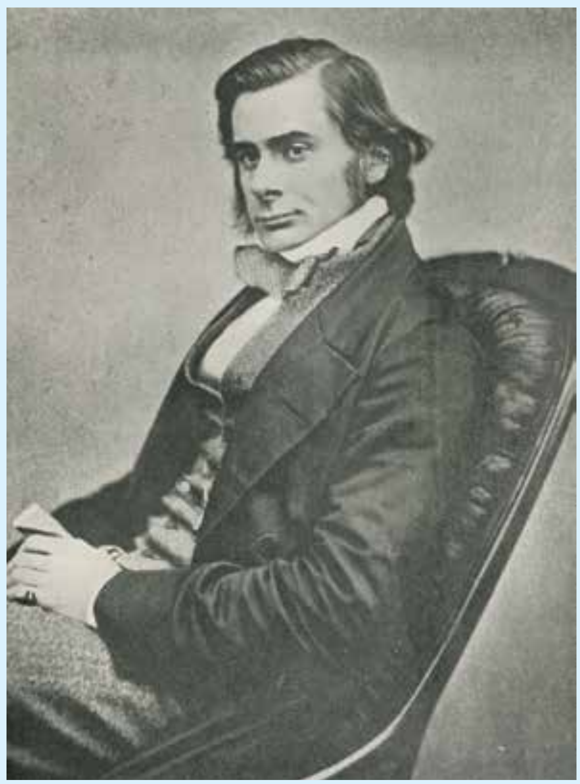


**FIGURE 10.6** Thomas Henry Huxley, supporter of Darwin's theory of evolution.

Another version of Huxley's response to Bishop Wilberforce's question follows:

> The Bishop asks how would I [Huxley] like it if my Mother had been an ape. I answer, putting aside the bad taste of an allusion to a relationship which I have in my own case regarded as calling up the tenderest memories, and having regard to that argument from the 'Godlike gift' of language which the Bishop has put forward as paramount against Mr Darwin's theory, that I would rather I had a parent wanting that 'Godlike gift' than a parent who devoted that great and godlike gift to the perversion of truth or to diverting the minds of an audience from the facts that support a great scientific hypothesis by ridicule, etc.

(*Source:* Huxley, Leonard, *Life and Letters of Sir JohnDalton Hooker*, Vol. II, John Murray, London, 1918, p. 303)

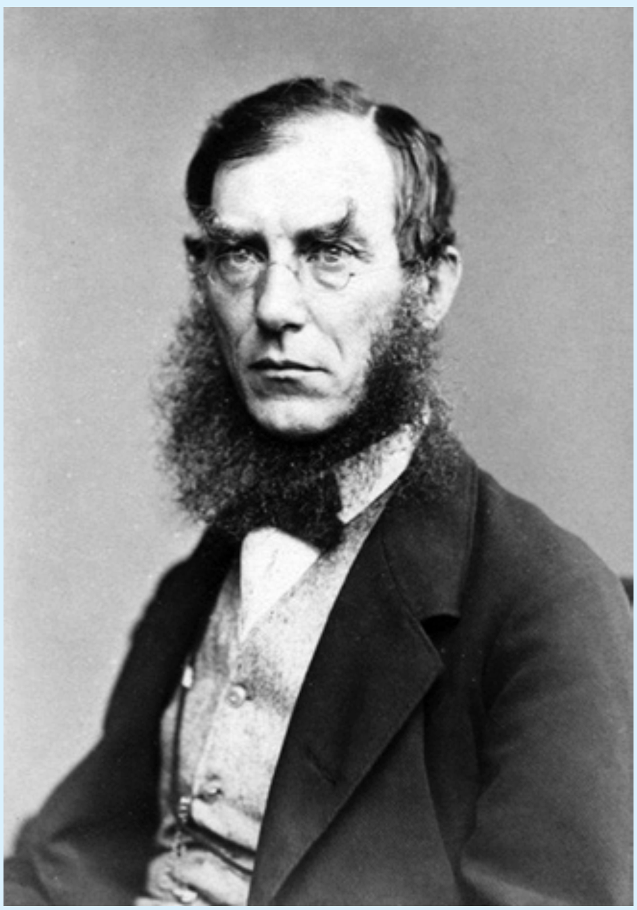


**FIGURE 10.7** Joseph Dalton Hooker, supporter of Darwin's theory of evolution.

Hooker described his own involvement in the debate in a letter dated 2 July 1860 that he wrote to Charles Darwin:

> Well, Sam Oxon [Bishop Wilberforce] got up and spouted for half an hour with inimitable spirit, ugliness and emptiness and unfairness. I saw that he was coached up by Owen [a leading opponent of Darwin's theory] and knew nothing, and he said not a syllable but what was in the Reviews [critical reviews of Darwin's book written by Wilberforce].
>
> . . . my blood boiled . . . so there I was cocked up with Sam [Bishop Wilberforce] at my right elbow and there and then I smashed him [metaphorically] amid rounds of applause. I hit him in the wind at the first shot in ten words taken from his own ugly mouth; and then proceeded to demonstrate in as few more: (1) that he could never have read your book, and (2) that he was absolutely ignorant of the rudiments of Bot[anical] Science.
>
> I said a few more [words] on the subject of my own experience and conversion, and wound up with a very few observations on the relative positions of the old and new hypotheses, and with some words of caution to the audience. Sam was shut up — had not one word to say in reply, and the meeting was dissolved forthwith, leaving you [Darwin] master of the field after 4 hours' battle.

(*Source:* Huxley, Leonard, *Life and Letters of Sir John Dalton Hooker*, Vol. I, John Murray, London, 1918, pp. 526–7)

On the same day, Darwin replied to Hooker by letter, dated 2 July 1860, that includes the following:

> . . . how still more I [Darwin] should have liked to have heard you [Hooker] triumphing over the Bishop. I am astonished at your success and audacity . . . I had no idea you had this power [of counter argument].

(*Source:* Darwin, Francis (ed.), *The Life and Letters of Charles Darwin*, Vol. II, John Murray, London, 1887, p. 323)

The Oxford debate of 1860 is important because, in arguing in favour of Darwin's theory of evolution by natural selection, Darwin's supporters (Huxley and Hooker) openly resisted the religious authority of the day (Wilberforce). This debate led to further serious and public discussion of Darwin's theory, based on scientific arguments.

## Developing a hypothesis about evolution

In developing their thinking about evolution, both Darwin and Wallace were influenced by observations made on independent and separate voyages to remote islands (see the box *Travels of Darwin and Wallace* on page 451). During the voyage of the HMS *Beagle*, Darwin saw a diversity of bird species (finches) on the Galapagos Islands, in the Pacific Ocean. Darwin noted that they reminded him of birds that he had seen in South America. He wrote that it appeared that 'one species had been taken and modified for different ends'. It was observations like this that finally led Darwin to the view that one species could evolve and produce a variety of different species.

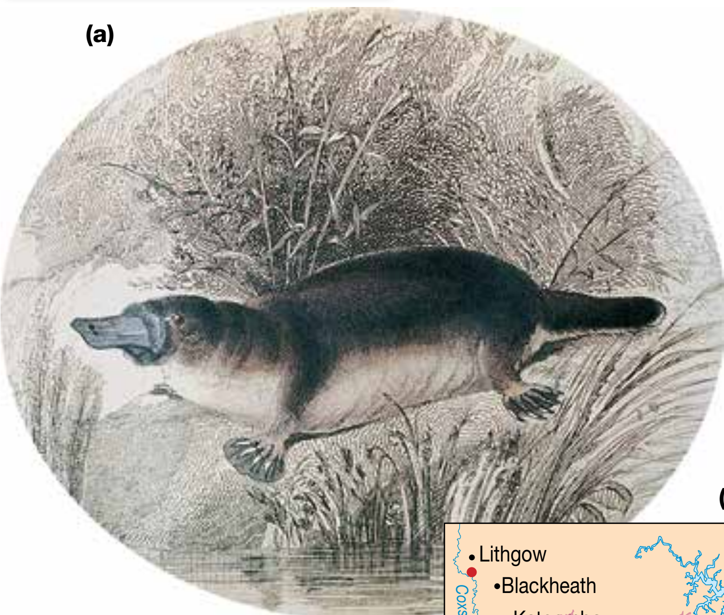
From the Galapagos Islands, the *Beagle* travelled to New Zealand and then to Australia. On 11 January 1836, the *Beagle* arrived in Sydney Harbour. During a period of less than three weeks in Sydney, Darwin visited countryside areas and travelled as far as Bathurst. On the evening of 19 January, Darwin saw platypuses swimming in Cox's River (see figure 10.8). He wrote:

> A little time before this I had been lying on a sunny bank, and was reflecting on the strange character of the animals of this country as compared with the rest of the world. An unbeliever in every thing beyond his own reason might exclaim, 'Two distinct Creators must have been at work; their object, however, has been the same, and certainly the end in each case is complete'.

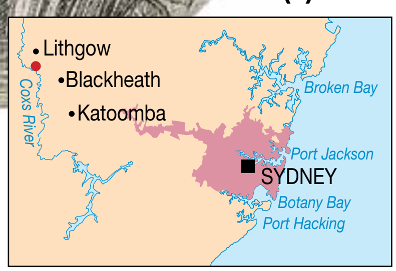
(*Source:* Darwin, Charles, *Journal of Researches into the Geology and Natural History of the Various Countries Visited by HMS Beagle,Under the Command of Captain Fitzroy, RN, from 1832 to 1836*, Henry Colburn, London, 1839, p. 526)

**FIGURE 10.8 (a)** A platypus on a riverbank. This image comes from a lithograph published in 1841. Very different from other mammals, the platypus is an egg-laying monotreme. (Image (a) courtesy of Judith Kinnear.) **(b)** Location of Cox's River where Darwin saw platypuses swimming.

(a)



(b)




Darwin's reference to 'two distinct Creators' emphasised that Australian animals were very different from those found in the rest of the world. At this time (the 1830s), the generally accepted view was that species were unchanging and that each was specially created. Darwin's comment in January 1836 might indicate that, at this time, Darwin held the view that species were specially created.

It is reasonable to suggest that the many observations made by Darwin during the *Beagle* voyage would have raised questions in his mind, such as:

- If species are specially created, why create 12 or more slightly different finch species on the various Galapagos Islands rather than just one finch species?
- Why were Australian species different from those found in the northern hemisphere? Why were platypuses found in Australian creeks instead of water rats such as would be seen in European streams?
- Why did modern species resemble extinct fossil species?
- Why was the Australian vegetation so different from that of Europe? Darwin wrote: *'The trees nearly all belong to one family; and mostly have the surface of their leaves placed in a vertical, instead of as in Europe, a nearly horizontal position; the foliage is scanty, and of a peculiar, pale green tint, without any gloss.'*

(Source: *Journal of Researches*, page 517.)

To what kind of vegetation was Darwin referring?

On his return to England, Darwin continued to struggle with the question of whether species could evolve. If species could change, then what was a possible mechanism by which evolution occurred? By 1838, in his notebooks, Darwin began to outline his view that species were *not* fixed and unchanging and that the evolution of new species from existing species occurred through the mechanism of natural selection.

Darwin spent many years refining his thinking about evolution in light of observations made during the *Beagle* voyage. In addition, Darwin was influenced by information from other sources, such as:

- his interactions with plant and animal breeders, including pigeon fanciers (see figure 10.9)
- *An Essay on the Principles of Populations*, published in 1798 by Thomas Malthus (1766–1834), in which Malthus argued that, because human populations were growing faster than food supplies, this would lead to a struggle for existence.

**ODD FACT**

Charles Darwin joined two London Pigeon Clubs and came to recognise the enormous variation in domesticated pigeons, including the following varieties: the carrier, the tumbler, the runt, the barb, the pouter, the turbit, the Jacobin, the trumpeter, the laughter and the fantail. These varieties were produced by breeders who used artificial selection procedures.

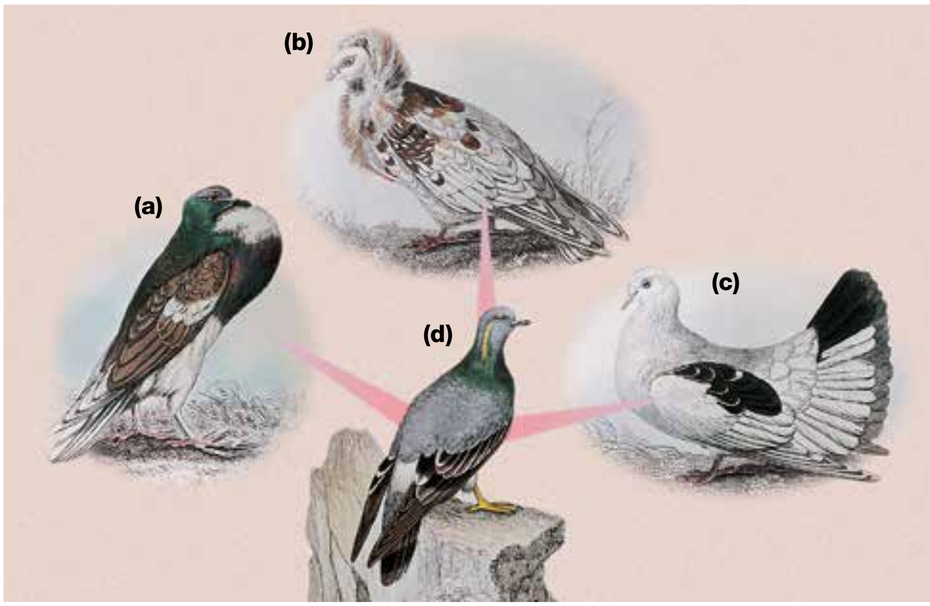



**FIGURE 10.9** Pigeon variants (*Columba livia*): **(a)** pouter, **(b)** Jacobin, and **(c)** fantail — all derived from **(d)** the common rock pigeon, by artificial selection. Darwin was aware that, through artificial selection over several generations, pigeon breeders were able to produce different varieties of pigeon. (Images courtesy of Judith Kinnear.)

Because he wrote so extensively, it is possible to follow changes in Darwin's thinking over time. His writings include a publication about his voyage on the *Beagle* (see figure 10.10a), his notebooks written in 1837 and 1838 (see figure 10.10b) and two lengthy essays written in 1842 and 1844. It is apparent that both before and during his voyage on the *Beagle*, Darwin believed that species were fixed and that each was specially created. By 1837, however, Darwin's notebooks reveal that he had moved to a position of recognising that species were capable of transformation and that new species evolved and were not specially created. In January 1844, Darwin wrote to his friend J.D. Hooker:

> At last gleams of light have come, and I am almost convinced (quite contrary to the opinion I started with) that species are not (it is like confessing a murder) immutable.

(*Source:* Darwin, Charles, letter to J.D. Hooker, 11 January 1844)

(a) (b)

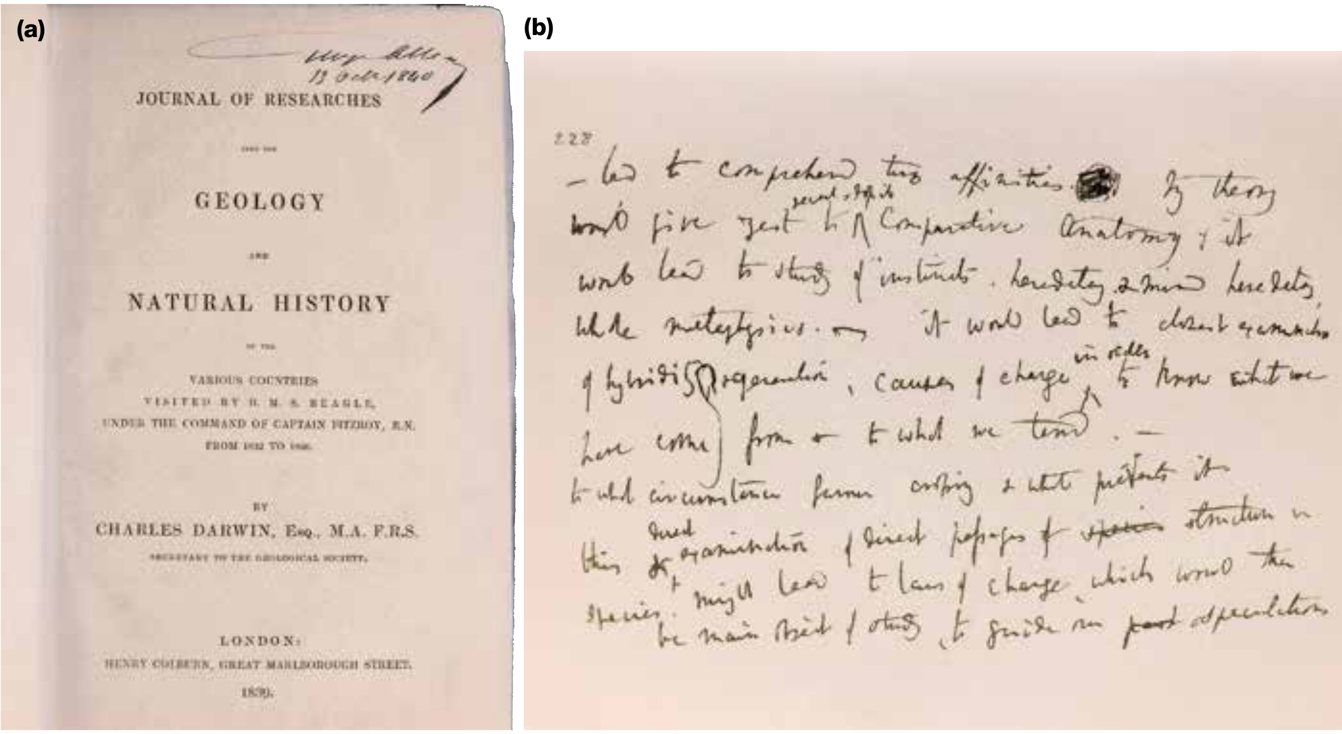



**FIGURE 10.10** **(a)** Front page of Darwin's *Journal of Researches*. This book is now better known as *The Voyage of the Beagle*. (Image courtesy of Judith Kinnear.) **(b)** A page from one of Charles Darwin's notebooks. Here, Darwin writes: '. . . . My theory would give zest to recent and Fossil Comparative Anatomy . . . [it] would lead to causes of change in order to know what we have come from & to what we tend . . . [it] might lead to laws of change, which would then be [the] main object of study, to guide our speculations.'

Darwin did not initially make his views public but shared them with his closest colleagues and recorded them in his notebooks. However, in June 1858, Darwin received a manuscript written by Wallace that contained essentially the same ideas about evolution by natural selection as those Darwin had formulated in his private notes and essays. Because of this, Darwin agreed to make his own ideas public jointly with those of Wallace at the meeting of the Linnean Society in London in 1858. The following year, *The Origin of Species* was published. Subsequent testing of predictions arising from the Darwin–Wallace view of evolution has changed its status from a hypothesis under test to a theory that is supported by evidence from many sources.

**ODD FACT**

In 1870, the journal *Nature* published the following about Darwin's theory: '[It has] so completely taken hold of the scientific mind . . . that our rising men of science may be classed as belonging to this school of thought. Probably since the time of Newton no man has had so great an influence over the development of scientific thought as Mr Darwin.'

## Impact of Darwinian thought

The theory of evolution of species by natural selection was a major contribution to biological thought. No longer were the living things on planet Earth regarded as unchanging and products of special creation that were unrelated to any other kind. Instead, the various kinds of living things on planet Earth were seen as capable of change and related by evolutionary descent. The theory of evolution provides an explanatory framework that can account for many observations, including the diversity of species, past and present. Evolution means that past and modern species are related and that different kinds of organisms living today are descended from various kinds of organisms that lived in the geological past.

Before *The Origin of Species* was published, human beings were viewed as the special product of divine creation and were not seen as part of nature. The study of the human species was seen as outside science. After the text was published, human beings were viewed as part of nature and able to be studied by scientific methods.

## THE TRAVELS OF DARWIN AND WALLACE

### Darwin's voyage on the *Beagle*

In 1831, Charles Darwin was offered the unpaid position as naturalist on HMS *Beagle*, which undertook a five-year voyage around the world. The ship left England on 27 December 1831 and returned on 2 October 1836, after visiting many countries (see figure 10.11). Throughout this journey, Darwin suffered from incurable seasickness. However, he described the voyage as 'by far the most important event in my life and had determined my whole career'.

One major event was a stay on the Galapagos Islands, a cluster of more than one dozen volcanic islands located in the Pacific Ocean, nearly 1000 kilometres west of Ecuador. Darwin saw many remarkable sights here, including giant tortoises and species of iguanas (large lizards), some terrestrial and others that browsed on algae and swam in the sea. Darwin realised that, while the islands had similar plants and animals, each island had particular and distinctive species.

Darwin also discovered that the Galapagos Islands were home to more than a dozen species of finches. These finches showed some similarities to finches he had seen in Chile, in South America. Darwin later wrote:

> '*It was most striking to be surrounded by new birds . . . and yet by innumerable trifling details of structure and even by tones of voice and plumage of the birds to have . . . the hot dry deserts of Northern Chile vividly brought before my eyes*'.

After his return to England in 1836, Darwin spent many years thinking about how species could change. Darwin came to realise that each species of finch was not specially created. Instead, he recognised that each kind of finch on the Galapagos Islands had evolved from a few ancestral finches that reached there from South America.


**FIGURE 10.11** Countries visited by HMS *Beagle*. In the period 16 September to 20 October 1835, the ship was at the Galapagos Islands. During January and February 1836, the *Beagle* stopped at Sydney and Hobart, Australia. (Map provided by MAPgraphics Pty Ltd, Brisbane.)

### Wallace's voyages in South-East Asia

Alfred Wallace also travelled widely and visited many islands. In 1858, Wallace was in the Moluccas thinking about the problem of how species are formed. Wallace wrote:

> '*. . . checks must also act upon animals, and keep down their numbers . . . While vaguely thinking how this would affect any species, there suddenly flashed upon me the idea of the survival of the fittest — that individuals removed by these checks must be, on the whole, inferior to those that survived. Then considering the variations continually occurring in every fresh generation of animals and plants, and the changes of climate, of food, of enemies always in progress, the whole method of species modification became clear to me*'.

So, Wallace concluded that natural selection was a plausible mechanism for evolution of species and he wrote a manuscript outlining his ideas and '*sent it by the next post to Mr Darwin*'. Darwin realised that Wallace had independently reached the same conclusion as he had, namely, that new species could arise as a result of the action of natural selection acting over many generations on populations. They agreed to present their proposal jointly to the scientific community, and this occurred at a meeting of the Linnean Society held in London on 1 July 1858. Darwin was writing a long account of his theory but he quickly produced a shorter version for publication. Darwin's book *On the Origin of Species*, was published in 1859.

## KEY IDEAS

- Two conflicting views existed in the past: the view that each species was unchanging and was specially created; and the view that species were mutable and could evolve to produce new species.
- The Darwin–Wallace theory of evolution by natural selection identified a mechanism for evolution that allows for predictions that can be tested by observation and experiment.
- Darwin's ideas about evolution and the changeability of species were influenced by his voyage on the *Beagle*, including his visit to Australia.

## QUICK CHECK

1 Identify the following as true or false:
   a The theory of evolution by natural selection was developed independently by Darwin and by Wallace.
   b The concept of the special creation of species means that species can change over time.
2 Identify an important observation made by Darwin during his time in Australia.
3 Identify two features of the Darwin–Wallace theory of evolution by natural selection that distinguished this theory from most other views about evolution.

# Time scales in evolution

Time is a critical element in evolution. Estimating the ages of rocks and of fossils that they contain is important in reconstructing the evolution of life on Earth. In the following section we will look at:

- various estimates of the age of the Earth
- how events in the history of Earth are organised on a geologic time scale
- how the ages of rocks and fossils can be determined.

## Estimating the age of Earth

Just how old is the Earth? Answers to this question have been many and varied as shown in the following examples:

- In 1650, the Anglican Archbishop of Armagh, James Ussher (1581–1656), published his calculation of the age of the Earth based on the Bible. By adding the ages of Adam and Eve and their descendants as given in the 'Genesis' chapter, Ussher concluded that the Earth was created on the evening of Sunday 23 October 4004 BC, making it about 6000 years old. This view prevailed until the late eighteenth century.
- In 1785, a Scottish farmer and amateur geologist, James Hutton (1726–1797), published his conclusions about the age of the Earth in his *Theory of the Earth*. Hutton identified that 'countless ages are required to form mountains, rock and soil'. He recognised that repeated cycles of sedimentation, compression into rock, uplift and erosion had occurred, and concluded that the Earth was millions of years old. Hutton recognised that the processes of building layers of sedimentary rock were slow and yet, in spite of this, some of these layers were several kilometres thick. His views were shaped in part by the so-called **unconformity** in the rocks at Siccar Point, east of Edinburgh (see figure 10.12), where rocks of two distinct age groups can be seen.

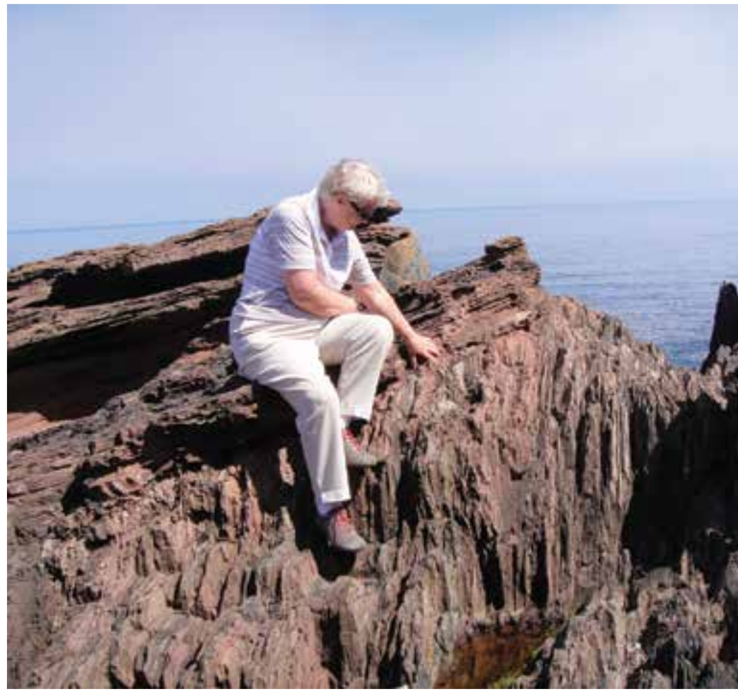


**FIGURE 10.12** James Hutton (1726–1797) recognised two different groups of sedimentary rocks at Siccar Point, in Scotland. Gently tilted layers of red sandstone lie above the almost vertical layers of rock known as greywacke. He recognised that these two different rock groups were laid down at different times. Which is the older group? The hand of the seated person straddles the unconformity or junction between the two different sequences of rocks. (Images courtesy of Judith Kinnear.)

**ODD FACT**

When he left on his *Beagle* voyage, Charles Darwin carried a copy of Volume I of Charles Lyell's *Principles of Geology*. Volume II reached him during the voyage in late 1832. Lyell's book about processes of geological change influenced Darwin's thinking about processes of biological change.

**ODD FACT**

Bright spark! Lord Kelvin, then William Thomson, graduated from Glasgow University in 1834, aged 10.

- In 1830–33, Charles Lyell (1797–1875) published the three-volume *Principles of Geology*. Lyell also recognised that present-day cycles of sedimentation, compression into rock, uplift and erosion must have been repeated many times in the past and that the same geological forces acting today would also have acted in the past. Lyell attempted to identify absolute times for geological events. He observed sedimentary rock strata that were 600 feet (about 183 metres) above sea level and recognised that these sediments were originally below sea level. Lyell assumed a rate of elevation of 2.5 feet (approximately 0.76 metres) per century and calculated that a period of 24 000 years was required for these rocks to reach their present elevation.
- In 1862, Lord Kelvin (1824–1907) calculated the age of the Earth based on an estimated rate of cooling of Earth from an original molten state. His first estimate was 98 million years old. (In 1897, he amended this estimate to 20 to 40 million years.) However, Kelvin's estimates were too low. Why? At that time, no-one knew about the radioactive elements that are a source of internal heat in the Earth's crust and so affect the rate of cooling.
- In 1899, an Irish scientist, John Joly, estimated the age of the Earth from salt content in the oceans. He calculated that it would have taken about 90 million years for the oceans to accumulate the present-day salt levels from inflow of river waters. This method was an underestimate of Earth's age because it failed to account for salt in other locations, such as salt deposits.
- Other attempts to date the Earth were based on estimating rates at which sediments were deposited under the sea and then measuring the total thickness of sedimentary rock strata. Based on average sedimentation rates

of 0.3 metres per 1000 years, one estimate of the Earth's age as being about 1600 million years was obtained in 1910.

- From about 1910 onwards, estimates using simple radiometric methods were made of the age of the Earth. By the mid-twentieth century, modern **radiometric dating** techniques had been developed. In 1956, Clair C. Patterson (1922–1995) used these techniques to calculate an accurate and reliable estimate of the age of the Earth based on the age of an iron meteorite. This meteorite had an estimated age of about 4500 million years. Since meteorites came into existence at the time of formation of the solar system, their age also identifies the age of the Earth. Since then, the age of many more meteorites has been determined and all have been found to have ages between 4400 and 4500 million years. This value has been confirmed by the dating of some moon rocks at 4500 million years old. A more recent dating in 2005 is that the Earth is 4569 million years old, estimated by the age of a meteorite.

### Geologic or deep time

Our understanding of geologic time or deep time is recent. The change in thinking about Earth from being just a few thousand years old to the concept of a very ancient Earth thousands of millions of years old was initially due to the work of the two Scottish thinkers, James Hutton, the farmer and amateur geologist, and Charles Lyell, the pre-eminent geologist of his time.

Based on modern techniques of radiometric dating of lunar rocks and meteorites, we now know that the Earth is about 4500 million years old — an interval that provides sufficient time for evolutionary and geological processes to occur.

## Organising the past: the geologic time scale

Geologic ages almost defy comprehension. We can gain some understanding by representing the age of Earth as a 100-metre track, with the starting line being the present time and the finishing line being the time of formation of Earth more than 4500 million years (Myr) ago. A human life of 72 years on this scale would be an undetectable 0.002 millimetres. Just the tip of a fingernail (one millimetre) over the starting line would take us back 45 000 years. Two steps (two metres) would take us to about 90 million years ago, to the time when Australia was part of the Gondwana supercontinent and dinosaurs dominated life on Earth.

The entire history of humanity (as represented by fossil evidence of members of the genus *Homo*) would be just 6 centimetres from the starting line, and the 200 000-year history of our species, *Homo sapiens*, would be placed less than 5 millimetres from the starting line.

Perhaps the most remarkable fact about this geologic time scale is that the first indirect evidence of life on Earth appeared about 3850 million years ago, but it was not until the **Ediacaran period** about 620 million years ago that the first multicellular animals appeared in the fossil record. **The first members of the genus *Homo* appeared only 2.8 million years ago** and the first modern humans (*H. sapiens*, our species) appeared about 200 000 years ago.

The geologic history of the Earth is divided into various time intervals as a hierarchy that includes eons, eras and periods (see figure 10.13). This geologic time scale developed in the eighteenth and nineteenth centuries when scientists organised sedimentary rock strata in the same region into groups of similar **relative ages** and also recognised similarities in rock strata in different regions because they contained identical fossils. By about 1840, the major divisions of the time scale based on relative ages were established but the **absolute ages** were not identified until during the twentieth century.

**ODD FACT**

In March 2004, a new period was added to the geologic time scale — the first to be added in 120 years. This is the Ediacaran period, which precedes the Cambrian and started 630 Mya. This lower boundary is defined in terms of rocks in the Flinders Ranges of South Australia.

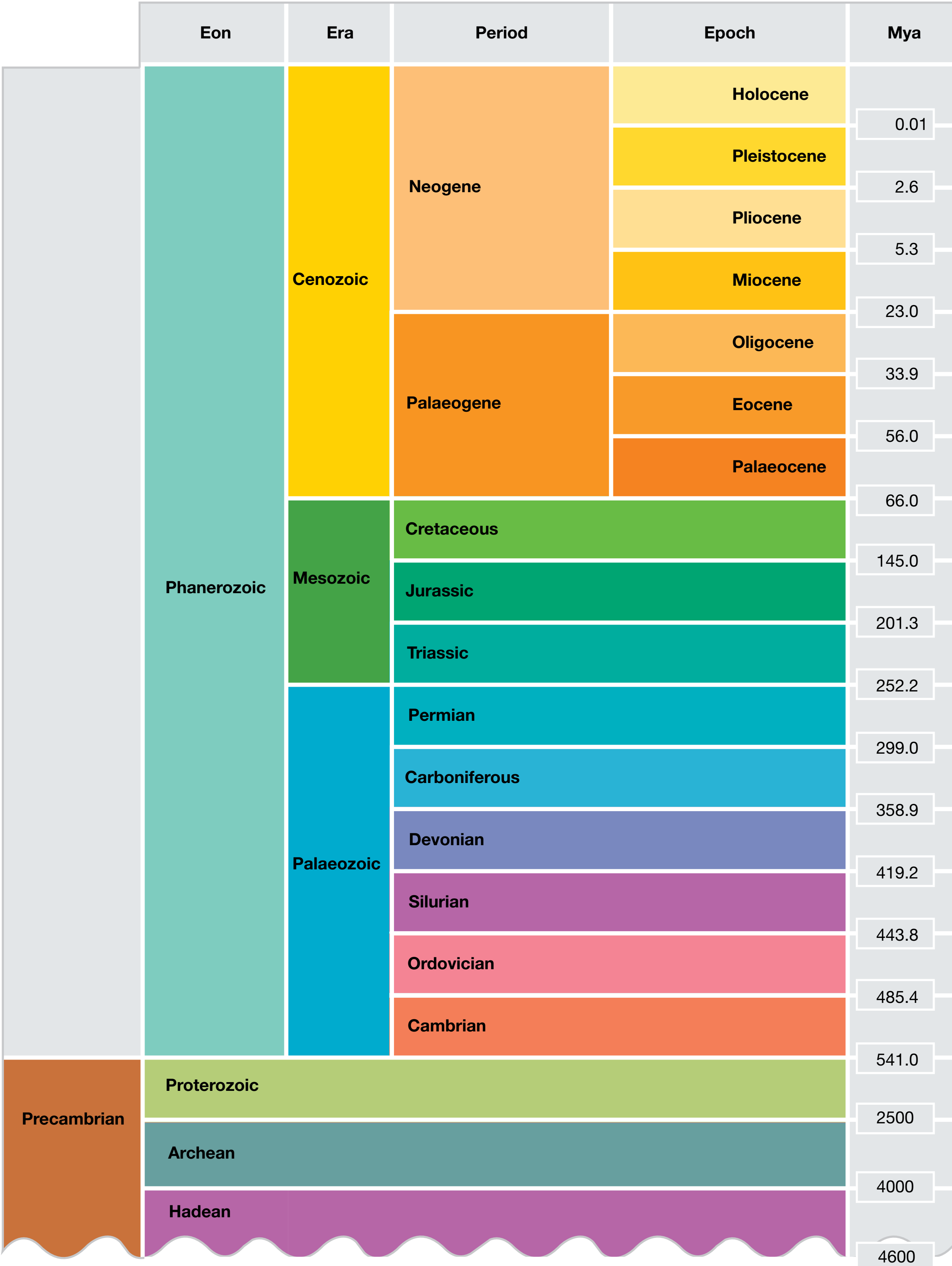



**FIGURE 10.13** Geologic time scale, as published by the American Geological Institute. Dates are in accord with the International Commission on Stratigraphy (2016 version).

# How old is it?

The age of a person can be stated in two ways — by relative age or by absolute age. For example:

- the statement 'Kym is older than Tran who is younger than Shane' gives the relative ages of these people but does *not* give their chronological ages. Relative age allows us to place the people in order of age: Kym then Shane then Tran, from oldest to youngest.
- the statement 'Kym is 34 years, Shane is 18 years and Tran is 16 years old' gives the absolute ages of these people.

Likewise, the ages of fossils or of geological structures, such as a deposit of mudstone or a layer of solidified volcanic ash (tuff), can be given in either relative or absolute terms.

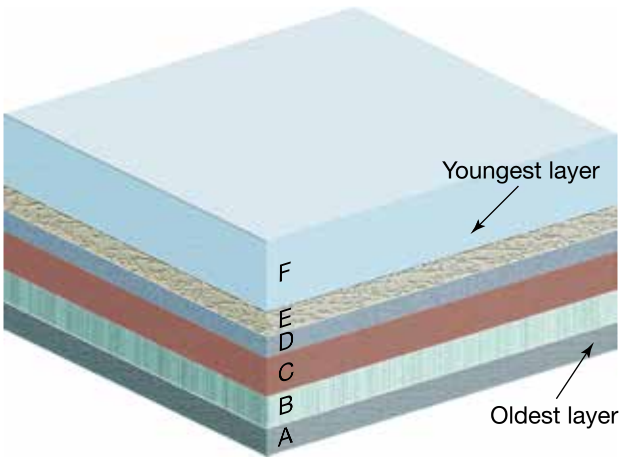
## Relative age: which is older?

Two methods that may be used to identify the relative ages of rocks and fossils are:

- the stratigraphic method, which is based on the principle of superposition
- the use of index fossils, which is based on the law of fossil succession.

Fossils are found in sedimentary rocks, such as a mudstone, sandstone, shale or limestone. The stratigraphic method of dating sedimentary rocks gives the relative ages of the rock strata by using the principle of superposition. This principle simply states that, for rock layers or strata (singular: stratum), the oldest stratum is at the bottom and progressively younger layers lie above it (see figure 10.14). Likewise, fossils found in a lower stratum are older than fossils in any overlying strata in the same sequence of rocks.

**FIGURE 10.14** The principle of superposition states that older layers of rock are at the bottom of a sequence of strata and are overlain by progressively younger layers. Note that, of the six strata (layers) shown, the oldest is layer A and the youngest F.




This method can only be used to identify the relative ages of rock strata in the same sequence. Based on position alone, the relative ages of strata from different sequences in widely separated regions of the world cannot be identified.

However, another method can be used to identify the relative ages of rock strata from different locations throughout the world. This method makes use of the **law of fossil succession**, namely that the fossils found in sedimentary rocks change over time and that the species that were alive at one time differ from those that were alive before or after that time. Because species change over time, the fossils that they leave are distinctive and can be used to identify rocks of the same age, even in different sequences and in different locations. (See figure 10.15 and the discussion on the next page concerning graptolites.)

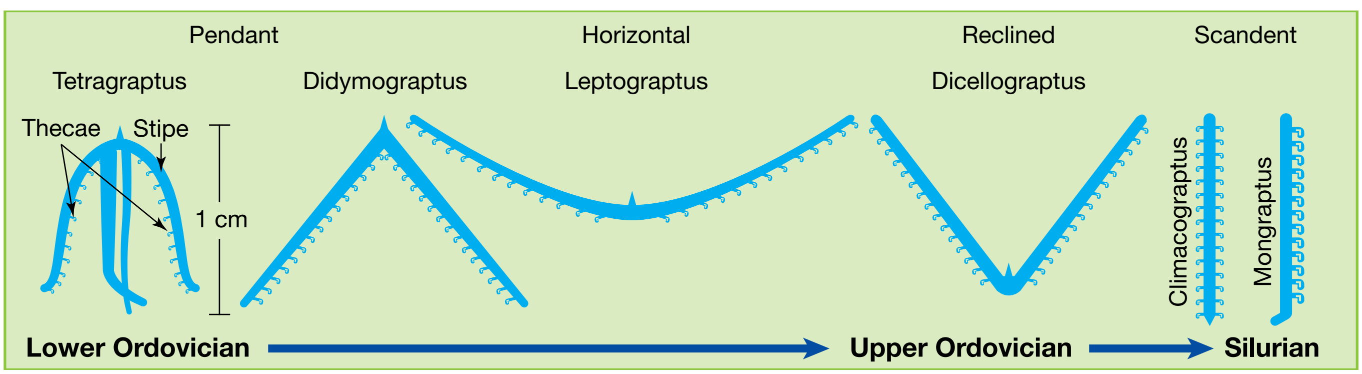



**FIGURE 10.15** Examples of some of the many different species of graptolite that existed during the Ordovician period (485–444 Mya) and into the Silurian period. Note the different arrangements of the branches (stipes) and the cups (thecae) on the branches.

**ODD FACT**

Using the law of fossil succession, William Smith (1769–1839) produced, in 1815, the first large-scale geologic map that showed the different rock strata of England, Wales and part of Scotland. Would you be surprised to know that Smith was a surveyor who worked on the construction of canals in Britain?

The fossils that are most useful for this purpose are so-called **index fossils**. These are geologically short-lived species that have a limited occurrence in the fossil record and so are found only in a restricted depth of sedimentary rock strata, and are widely distributed. The presence of these fossils in rock strata even in widely separated regions of the world can identify these rocks as having the same age (see figure 10.16). Before radiometric dating became available, correlating the fossils present in rocks was the best method available to date rock layers relative to one another.

Graptolites are one important group of index fossils because:

1. several hundred species of graptolites exist, each readily distinguishable by the number of branches (stipes) and the arrangement of the cups (thecae) on the branches (see figure 10.15)
2. they occur in the fossil record from the Cambrian to the Carboniferous
3. each different species existed for a limited time only in the fossil record; that is, graptolites underwent rapid evolution
4. they are widely distributed globally because graptolites formed floating colonies that drifted on the ocean surface.

**study on**

Unit 4 | AOS 1 | Topic 2 | Concept 3

**Relative fossil dating**
Summary screen and practice questions

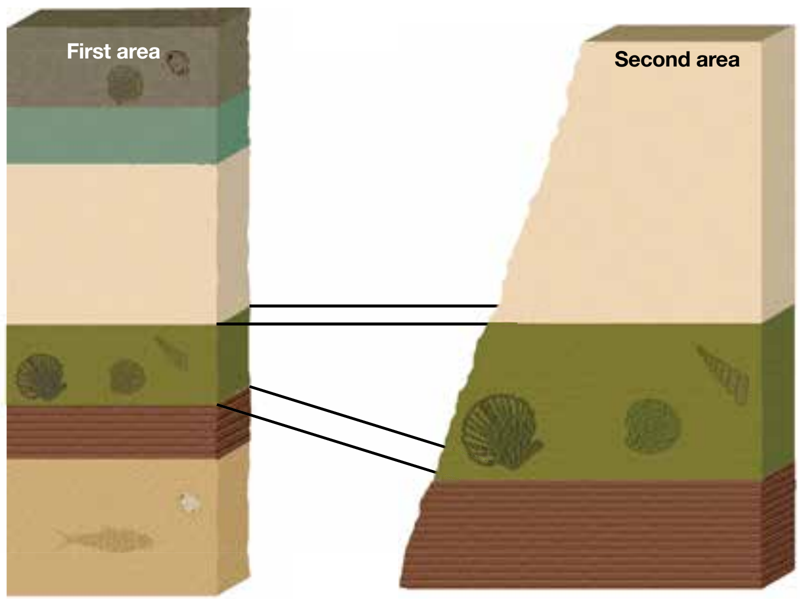



**FIGURE 10.16** The law of fossil succession allows the relative ages of rocks in different parts of the world to be inferred, based on distinctive index fossils.

These two methods give the relative ages of rock layers throughout the world. Rocks of the same age are identified by the name of a geological period, such as Devonian or Jurassic. Likewise, fossils can be dated as the same, or as younger or as older than a particular index fossil.

**study on**

Unit 4 | AOS 1 | Topic 2 | Concept 3

**Do more**
Relative age of rocks

**study on**

Unit 4 | AOS 1 | Topic 2 | Concept 3

**See more**
Relative age of rocks

## Absolute age: using 'rock clocks' to tell age

During the twentieth century, techniques for identifying the absolute ages of rocks were developed. The most important method of estimating the absolute ages of rocks is the **radiometric dating technique**, which is based on the decay of certain radioactive elements. **The radioactive elements concerned are those present in minerals in igneous rocks, and each element decays at a known rate.** Igneous rocks form when molten rock solidifies, either below the Earth's surface (for example, granite, a plutonic rock) or on the Earth's surface (for example, basalt and tuff, which are volcanic rocks).

**Radiometric** dating cannot be applied to sedimentary rocks that are derived from the erosion of pre-existing rocks because the minerals that they contain were formed prior to the sedimentary rocks themselves. Although igneous rocks do not contain fossils, radiometric dating of igneous rocks can give an estimate of the age of fossils.

This can occur, for example, when fossils are embedded in a sedimentary rock layer that is located between layers of solidified volcanic ash (tuff). The age of a fossil can be inferred to lie within the range defined by the ages of the tuff layers above and below the fossil (see figure 10.17).

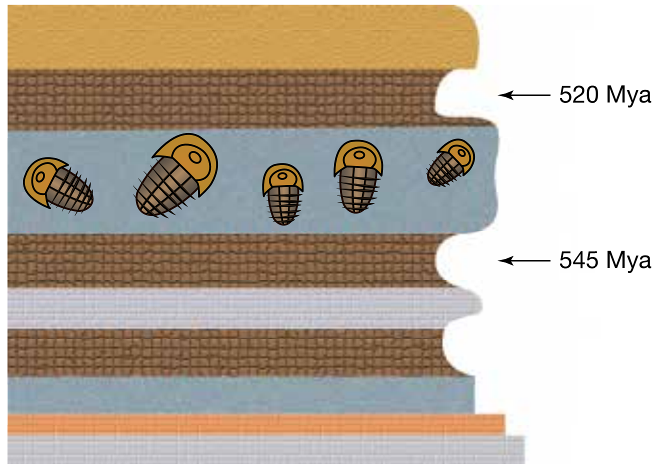



**FIGURE 10.17** Diagram showing fossils in a sedimentary layer located between layers of solidified volcanic ash or tuff (shown in brown). These volcanic ash layers can be given an absolute age using a radiometric dating technique. The age of the fossils can be inferred to lie in the range of 545 to 520 million years ago.

**ODD FACT**

Using the decay of radioactive elements, the age of the oldest rocks yet found on Earth has been estimated at 4300 million years. These rocks come from an ancient bedrock in northern Canada. Rocks of this age are not widespread. Can you suggest why?

The principle of radiometric dating depends on the fact that various elements exist in two or more forms known as isotopes, some of which are radioactive. The radioactive isotopes, which can be called 'parents', spontaneously decay or break down over time to form stable daughter products. The rate of the decay is specific for each radioactive isotope, and **the half-life is the time taken for half the original radioactive isotope to decay** (see figure 10.18). Table 10.1 shows some of the radioactive isotopes used for absolute dating and their stable daughter products. Each technique can be used over a particular age range that depends on the half-life of the parent radioactive isotope.

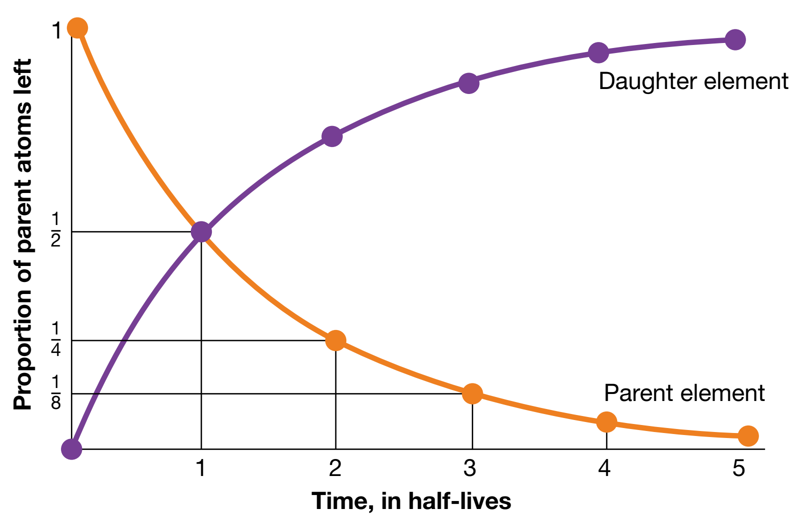



**FIGURE 10.18** Idealised decay of a radioactive parent element over a time period of five half-lives. What percentage of parent is left after one half-life? After two half-lives?

**TABLE 10.1** Some radioactive parent isotopes and the daughter products of decay. The longer the half-life of a radioactive isotope, the older the material that can be dated using a particular radiometric method.

| Radioactive parent isotope | Daughter product | Half-life (years) | Comments |
|---|---|---|---|
| Carbon-14 | Nitrogen-14 | 5 730 | Used for dating organic (carbon-based) remains; useful for ages up to about 60 000 years old |
| Uranium-235 | Lead-207 | 710 000 000 | Used for dating igneous rocks containing uranium-based minerals; useful in the range from 10 million years and older |
| Potassium-40 | Argon-40 | 1 300 000 000 | Used for dating igneous rocks containing K-bearing minerals, such as feldspars; useful in the range from 0.5 Myr and older |
| Rubidium-87 | Strontium-87 | 47 000 000 000 | Can be used to date the most ancient igneous rocks on planet Earth |

A particular radiometric technique can be used to date only material that contains a specific radioactive parent isotope. For example, carbon-14 dating *cannot* be used to estimate the age of zircon crystals ($ZrSiO_4$) in a rock but can be used to date linen fabric from an ancient Egyptian grave. Why?

Let's look at some radiometric dating techniques and their applications.

## Potassium–argon (K–Ar) dating

Absolute dating of rocks using K–Ar dating depends on the decay of radioactive potassium-40 isotope present in minerals found in rocks that form when molten magma cools and solidifies either at depth below the Earth's surface or on the surface. Rocks formed in this way are called igneous rocks (from Latin *ignis* = fire). Examples of these rocks are granite, which solidifies at depth below the Earth's surface, and basalt, which solidifies on the Earth's surface.

Potassium (K) is present in some of the minerals found in many igneous rocks. One common mineral is orthoclase, a potassium aluminium silicate ($KAlSi_3O_8$) that is found in granite (see figure 10.19a). A small fixed proportion (approximately 0.012%) of potassium is in the form of the radioactive isotope potassium-40, while most is in the stable form of potassium-39. Potassium-40 decays at a known constant rate to form the inert gas argon-40, which accumulates within potassium-containing mineral crystals in rock (see figure 10.19b). The older a rock, the longer the time available for potassium-40 to decay and for argon-40 to accumulate. The potassium-39 does not decay and it provides a measure of the original amount of potassium-40 present.

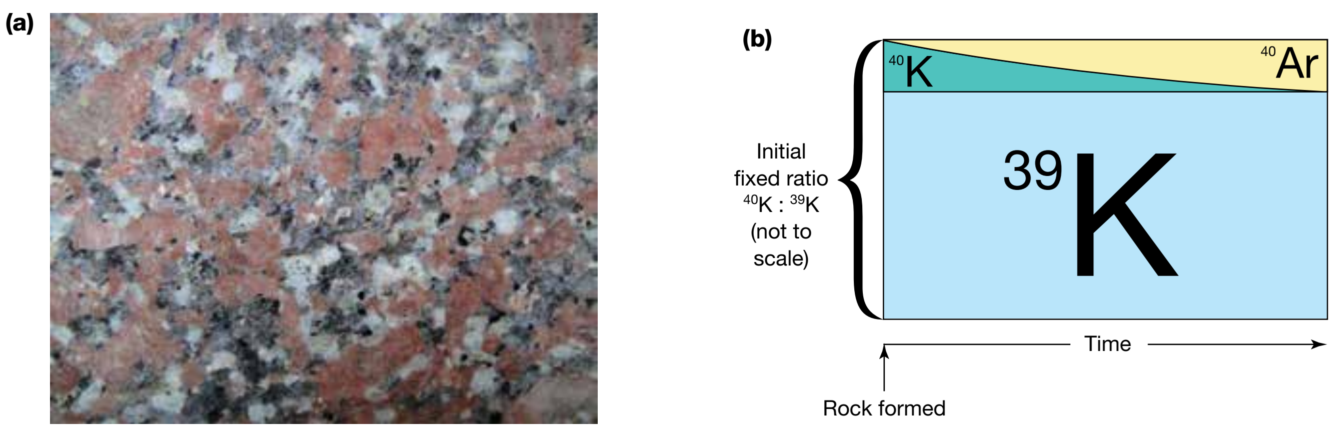



**FIGURE 10.19** **(a)** Granite, an igneous rock, showing its crystalline structure. The pink crystals are the mineral orthoclase, the black are mica and the white are quartz crystals. Of these three minerals, which could be used to date this rock using the potassium–argon (K–Ar) procedure? (Image (a) courtesy of Judith Kinnear.) **(b)** Over time, the amount of potassium-39 remains constant, but the radioactive potassium-40 progressively decays to form argon-40, a gas that is trapped in the rock.

Molten rock (magma), which has temperatures of about 1000 °C, contains no argon-40 because this gas is driven out of the rock at these temperatures. However, once the rock has cooled and solidified, the 'rock clock' starts and the argon-40 begins to accumulate.

Age estimates are done under carefully controlled conditions in which the sample of rock under test is heated in a vacuum, releasing argon-40 atoms, which are measured using a highly sensitive instrument. From this result and measurements of the amount of potassium-39, the absolute age of the rock can be calculated.

### Argon-39/argon-40 dating

A modification of the K–Ar dating technique is the argon-39/argon-40 dating technique. This technique can be done with just one mineral crystal. In contrast, the K–Ar technique described above requires two samples: one to measure the K-39 content and another for a separate measure of the Ar-40 content. In the argon-39/argon-40 technique, the sample to be dated is irradiated with neutron particles and this converts the stable K-39 to Ar-39. The amounts of the two argon isotopes are measured, giving an age estimate of the mineral.

### Carbon-14 dating

All living organisms are built of carbon-containing organic matter, such as proteins (for example, the keratin of hair, nails, hooves and claws, and the collagen of bones) and structural carbohydrates (for example, the cellulose of plant tissues). When an organism is alive, the carbon in its organic matter is a mixture of two isotopes:

- the stable isotope carbon-12 (C-12)
- the radioactive isotope carbon-14 (C-14), which decays to N-14.

In life, the proportion of these two isotopes is constant and matches that in the carbon dioxide in the atmosphere. This proportion remains constant during the life of an organism.

After the death of an organism, however, the proportion of carbon-14 decreases as this isotope decays and is not replaced. As the time after death increases, the ratio of C-14 to C-12 becomes progressively smaller. After one half-life of 5730 years, half the original amount of C-14 present in the organic material at the time of death will have disappeared, and so on for each successive half-life period. By measuring the ratio of C-14 to C-12 in a sample of organic material, an estimate of the time since death of the organism that produced this material can be obtained.

C-14 dating can date bones and also artefacts, such as the wooden handle of a tool and a localised collection of ash and fragments of burnt wood found in a cave. This burnt wood might be from the hearth of some early humans, such as Neanderthals, who sheltered in the cave 35 000 years ago or it might be of recent origin. C-14 dating can estimate the absolute age of this material and distinguish between these two possibilities provided the wood sample is not older than about 60 000 years.

**ODD FACT**

C-14 dating was used to obtain an estimate of the age of the Shroud of Turin since the shroud material was made of flax, a plant material.

**ODD FACT**

Willard Libby (1908–1980) developed carbon dating in 1947 and was awarded the Nobel Prize in Chemistry in 1960.

### Electron-spin resonance (ESR)

Carbon dating is not useful for organic material older than about 60 000 years. However, another technique of dating, known as electron-spin resonance (ESR), is useful for ages of about 50 000 years old to 500 000 years old. This period encompasses much of the evolutionary history of the genus *Homo*.

The ESR dating technique depends on the fact that when objects are buried they are bombarded by natural radiation from the soil. When these objects are composed of minerals, such as flint tools and fossil teeth, this bombardment causes some of the electrons in the minerals to move from the ground state to a higher energy level and some to remain trapped at higher energy levels.

studyon

Unit 4 | AOS 1 | Topic 2 | Concept 4

**Absolute dating techniques**
Summary screen and practice questions

studyon

Unit 4 | AOS 1 | Topic 2 | Concept 4

**Do more**
Absolute age

studyon

Unit 4 | AOS 1 | Topic 2 | Concept 4

**See more**
Absolute age

The rate at which high-energy electrons become trapped is determined by the background radiation; the longer the material has been buried, the greater the accumulation of higher energy electrons.

When material is exposed to heat, fire or bright sunlight, all the electrons return to the ground state so that the 'electron clock' is reset to zero. As a result, ESR can be used to estimate the time since material under investigation was last heated, such as when a flint instrument was burnt in a fire or when a tooth last lay exposed on the ground in sunlight.

In ESR, scientists make direct measurements of:

1. the number of high-energy electrons trapped in the material under investigation
2. the radiation owing to any unstable isotopes in the material itself (why?)
3. the background radiation in the soil in which the sample was buried.

Using these measures, it is possible to calculate a date that indicates the time that has passed since the 'electron clock' was last set to zero.

**KEY IDEAS**

- The age of the Earth has been variously estimated using different procedures.
- Radiometric techniques have provided reliable estimates of the Earth's age.
- The history of Earth is organised into an international geologic time scale.
- Ages can be expressed in either relative or in absolute terms.
- The principle of superposition and the principle of fossil succession provide the relative ages of rocks and fossils.
- Radiometric techniques based on the decay of naturally occurring radioactive isotopes can be used to estimate the absolute ages of rocks.

**QUICK CHECK**

4 List a particular contribution of the following to the debate about the age of the Earth:
  a James Hutton.
  b Charles Lyell.
5 Identify the following statements as true or false:
  a Absolute age gives the age of an object in a number in years.
  b Radiometric methods of dating yield the relative age of an object.
  c Carbon dating could be used to tell the age of Australian coal deposits from the Permian period.
  d When half of a radioactive parent element has decayed, an equivalent amount of daughter element(s) will have been produced.
  e Over two half-lives, the initial amount of a radioactive parent element will be halved.
  f All radioactive elements have equal rates of decay.
6 Explain briefly the principle of superposition.
7 List the following geological periods in order, from oldest to most recent: Cretaceous, Triassic, Cambrian, Carboniferous.

# Changing life forms over time

Imagine that we could travel back in time on planet Earth down the geologic time scale (refer back to figure 10.13 on page 455). Let's borrow a time machine from the *Back to the Future* film and journey to four periods of time on the geologic time scale. The further back in time we travel, the less familiar the scene around us will be.

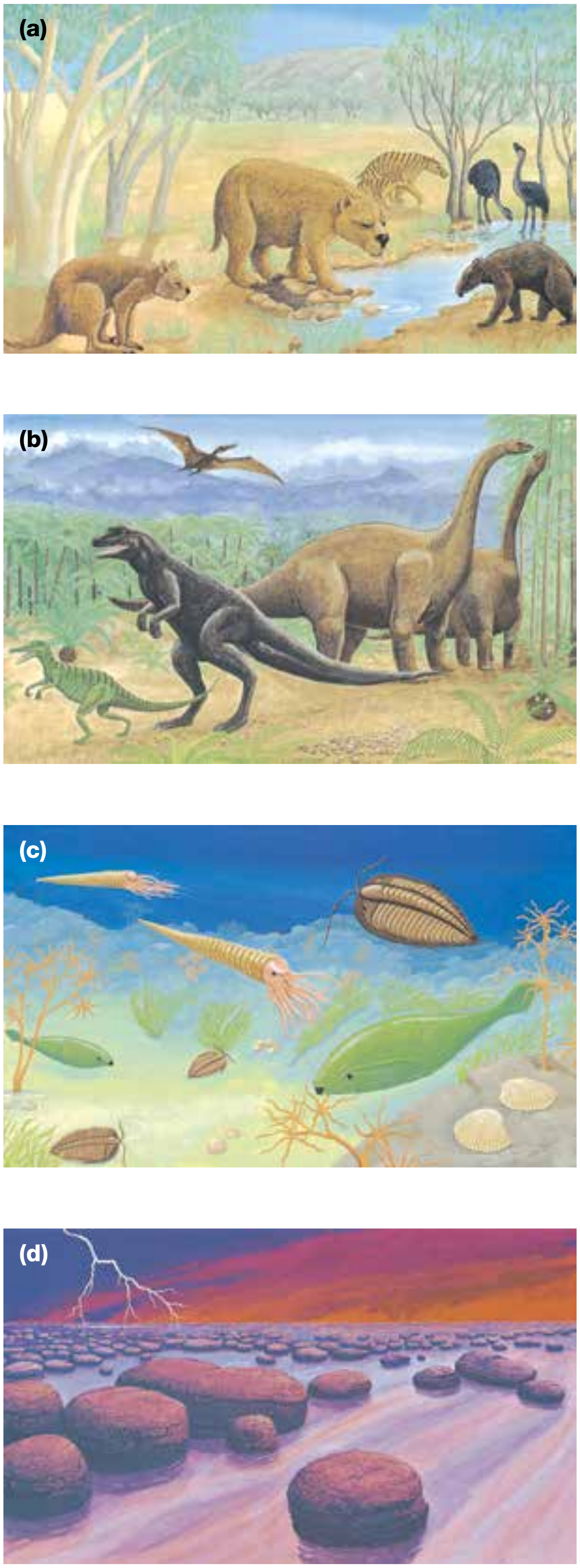



**FIGURE 10.20** Reconstructed views of life on Earth at various times: **(a)** the Pleistocene epoch, **(b)** the Cretaceous period, **(c)** the Ordovician period, and **(d)** the Archaean eon.

### First stop: Pleistocene epoch

Travel back to about 60 000 years in the late Pleistocene epoch in Australia and the view would look fairly familiar. You would find a familiar plant community that included grasses and familiar flowering plants including trees and shrubs such as eucalypts and wattles. Many of the birds and the mammals, both marsupial and monotreme, would also look somewhat familiar. However, you would be surprised to see wombats (*Phascolonus gigas*) that are the size of cows, and kangaroos (*Macropus titan* and *Procoptodon goliah*) that would match the height of the tallest AFL ruckman, and *Diprotodon optatum*, the largest marsupial, with a mass of over 1000 kg (see figure 10.20a, and figure 10.21 below). This was the time of the so-called megafauna (*mega-* = large); not only large mammals, but also giant goannas (*Megalania prisca*) that reached lengths of more than six metres, and giant flightless birds, such as *Genyornis newtoni*, two metres tall. (Go to the box at the end of this section to read about the Australian megafauna.)

### Second stop: Cretaceous period

Travel further back to the mid-Cretaceous period, about 110 Mya, and the dominant terrestrial life forms are no longer mammals, but are dinosaurs, including *Triceratops prorsus* and *Tyrannosaurus rex* (see figure 10.20b). You might be lucky enough to notice fur-covered mammals, but they are small and probably nocturnal. Pterodactyls can be seen in the skies. The seas are the home not only of fish and sharks, but also of large marine reptiles. These marine reptiles, such as pliosaurs and plesiosaurs, swam in the vast Eromanga Sea, which covered much of present-day inland arid Australia from about 90 to 130 million years ago. Central Queensland is a rich source of these fossils, including the as-yet unclassified magnificent 'Penny the plesiosaur' (see figure 10.22 below). The dominant plant life consists mainly of ferns, cycads and conifers, but flowering plants are present and are diversifying.

### Third stop: Ordovician period

Travel further back to the mid-Ordovician period, about 460 Mya, and the land is virtually bare. Only a few small primitive plants are beginning to appear on the land. No terrestrial animals are to be seen; animal life is confined to aquatic habitats (see figure 10.20c). No croaking of frogs, since amphibians appeared only about 360 Mya. No fear of a crocodile attack since the first reptiles do not appear in the fossil record until about 300 Mya. The animal life of the Ordovician is dominated by marine invertebrates. The dominant invertebrates are the trilobites, but also present are other invertebrates including crinoids, molluscs such as brachiopods, coral polyps, cephalopods, and graptolites. Graptolites were tiny marine animals that lived in colonies, most often free-floating. A colony consisted of one or more small branches that supported cup-like structures in which each animal lived (see figure 10.23). If we look carefully, we might see the first vertebrates — the jawless fish. However, there are no sharks, no bony fish, no lobe-finned fish in the Ordovician seas of planet Earth.

**studyon**

| | |
|---|---|
| Unit 4 | **Changes in life forms** |
| AOS 1 | Summary screen and practice questions |
| Topic 2 | |
| Concept 1 | |

## Last stop: Archaean eon

Travel back to the Archaean eon, about 3000 Mya, to a world that appears very strange to our human eyes. The atmosphere would be deadly for us because it is without oxygen (anoxic). The land is devoid of plants and animals. Because the ozone layer was not yet formed, high levels of damaging ultraviolet radiation reach the Earth's surface. The only life forms are prokaryotic microbes, largely unseen. Visible signs (biosignatures) of their existence might be seen in the characteristic marks and patterns left on sediments where microbial mats were once present. Other biosignatures of microbial activity that we would more easily see are dome-shaped structures called **stromatolites** in shallow seas (see figure 10.20d). Stromatolites are layered rocks that build up one layer at a time when microbial colonies of cyanobacteria trap sediments. Stromatolites can be see today in the hypersaline waters in Shark Bay in Western Australia (see figure 10.24).

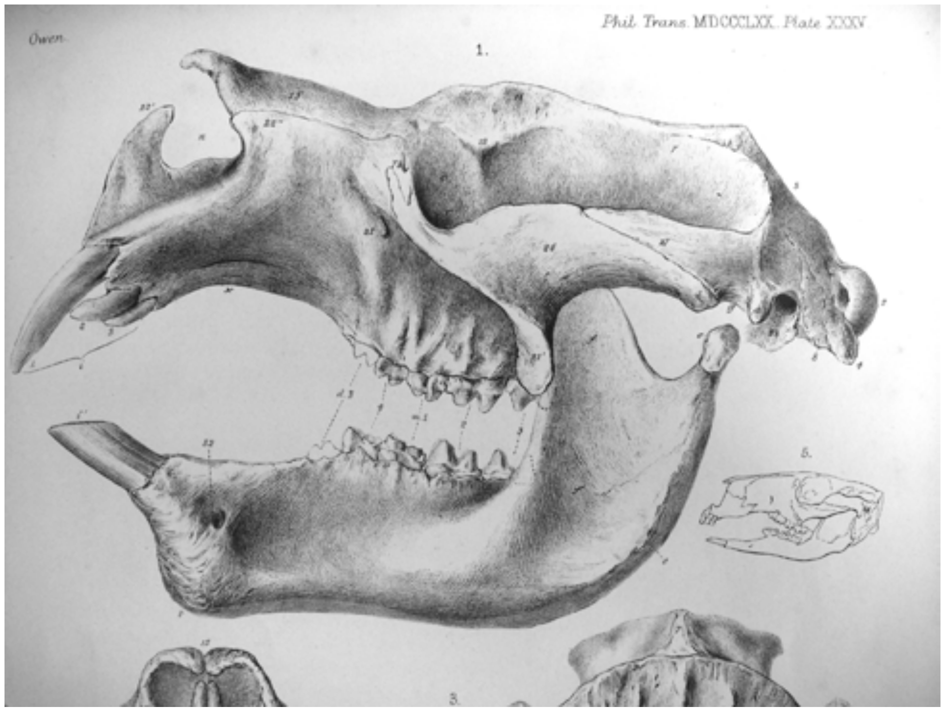



**FIGURE 10.21** First stop — Pleistocene: Skull of *Diprotodon optatum.* The length of the skull is almost one metre. The diprotodons were members of Australia's megafauna. (Image from Plate XXXV of Owen's Fossil Mammals of Australia, 1870.)

**FIGURE 10.22** Second stop — Cretaceous: Fossil bones of the as-yet unclassified 'Penny the plesiosaur' on display at the Kronosaurus Korner in Richmond, central Queensland. Remarkably, the skeleton is almost complete. Note the four large paddle-like limbs that propelled Penny through the Eromanga Sea in the Cretaceous period. (Image courtesy of Judith Kinnear.)

**FIGURE 10.23** Third stop — Ordovician: *Tetragraptus fruticosus*, a graptolite colony that floated in the Ordovician seas. Note the small cups that appear as serrations along the branches of the graptolite. Each cup housed one of the animals that formed part of the living graptolite colony. (The longest branch shown here is about five centimetres long.) (Image courtesy of Judith Kinnear.)


**FIGURE 10.24** **(a)** Fossil stromatolites from the Bitter Springs formation east of Alice Springs. (Specimen from the Houston Museum of Natural Science.) Note the multi-layered structure that is a distinctive biosignature of microbial activity, in particular, cyanobacteria. These layers are built when sediments became trapped by a sticky microbial mat of cyanobacteria. **(b)** Stromatolites as seen today in the hypersaline Hamlyn Pool in Shark Bay, Western Australia.

In our time travels, we saw changes in Earth's biodiversity. Table 10.2 shows approximate times of some major landmarks in these changes. Until about 400 Mya, plants and animals were almost all limited to life in the oceans. Since plants and animals became established on land, Earth's life forms have become more diverse.

**TABLE 10.2** Approximate times of appearance of various life forms on Earth. Note that for most of Earth's history, the only life forms were unicellular organisms — first, prokaryotic microbes, which later were joined by unicellular eukaryotic organisms.

| Approx. time | Emergence of life forms |
|---|---|
| ~ 4500 Mya | The lifeless (abiotic) Earth |
| ~ 3800 Mya | First prokaryotic cells |
| ~ 3500 Mya | First photosynthetic bacteria |
| ~ 2500 Mya | First aerobic (oxygen-using) microbes |
| ~ 2100 Mya | First unicellular eukaryotes |
| ~ 600 Mya | First multicellular eukaryotes |
| ~ 500 Mya | First vertebrates (jawless fish) |
| ~ 479 Mya | First insects |
| ~ 450 Mya | First land plants |
| ~ 365 Mya | First amphibians |
| ~ 360 Mya | First ferns |
| ~ 330 Mya | Earliest reptiles |
| ~ 250 Mya | First conifers |
| ~ 160 Mya | First mammals |
| ~ 130 Mya | First flowering plants |

In the next section, we will explore fossils — the various kinds and how they are formed. Earlier in this chapter (see page 456), the relative and the absolute dating of fossils was discussed.

## AUSTRALIAN MEGAFAUNA

Major Thomas Mitchell (1792–1855), an explorer and surveyor, collected some fossil bones in a cave near Wellington, New South Wales. These bones were sent to England to Sir Richard Owen (1804–1892), the leading vertebrate fossil expert at that time (and an outspoken opponent of Darwin's theory of evolution). From these bones, Owen reconstructed the skeleton of an extinct marsupial mammal, now known as *Diprotodon* (see figure 10.25). This animal was a member of a diverse group of animals known as the Australian megafauna (= 'giant animals'). The term 'megafauna' refers to large extinct vertebrates that lived during the Pleistocene epoch and which grew to sizes larger than their close relatives.

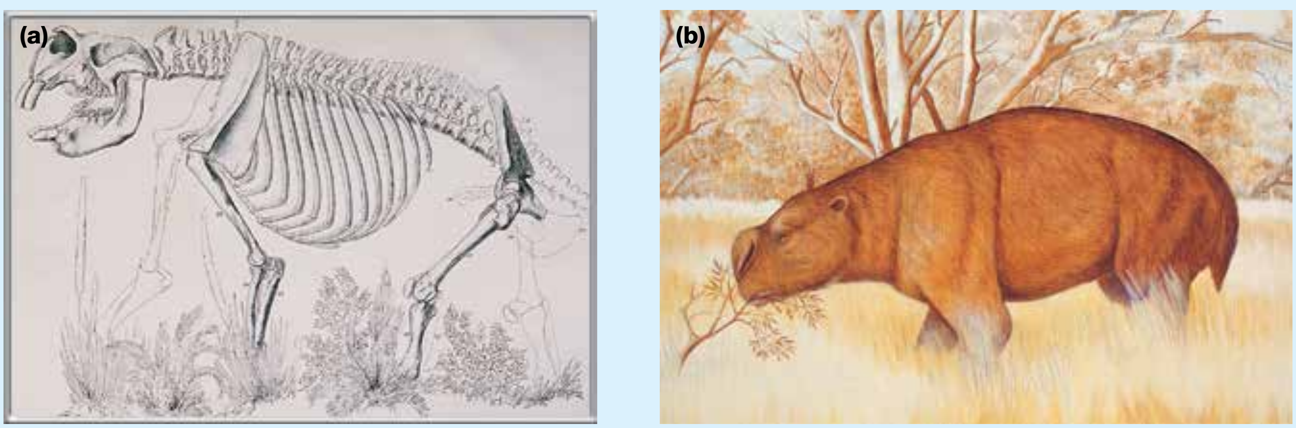



**FIGURE 10.25** **(a)** Skeleton of *Diprotodon optatum*, an extinct marsupial and one of the Australian megafauna. (Image courtesy of Judith Kinnear.) **(b)** Reconstruction of *Diprotodon optatum.*

*(continued)*

The Australian megafauna included animals from different vertebrate classes (see figure 10.26 and table 10.3). Starting about 50 000 years ago, megafauna species began to disappear. The cause of their extinction is not clear and debate continues on this matter.

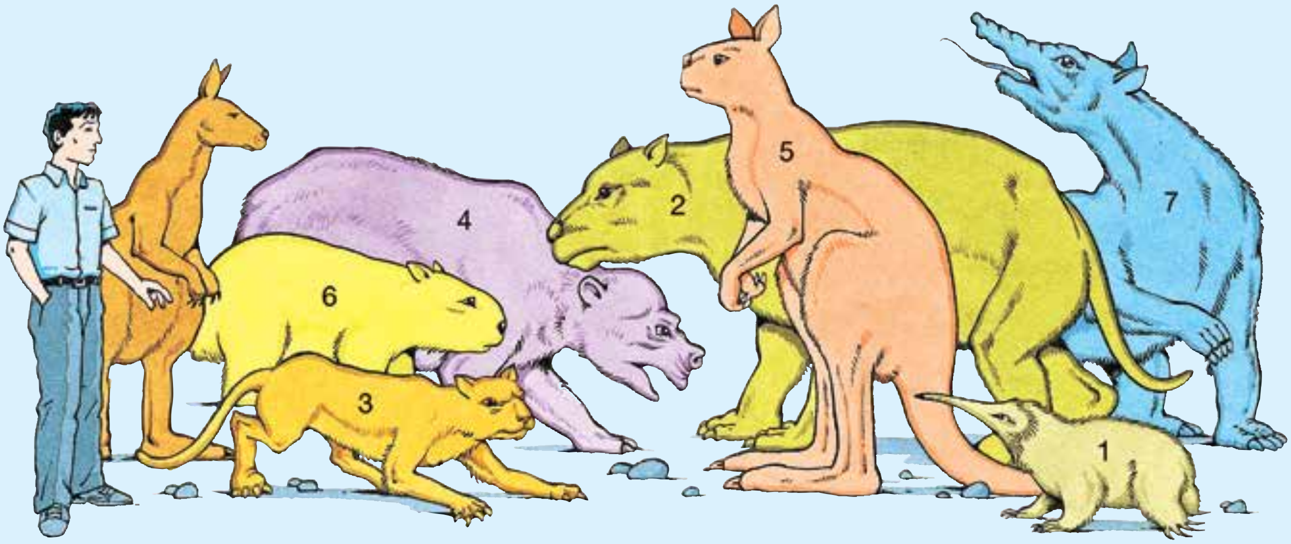



**FIGURE 10.26** Some of Australia's megafauna that lived during the Pleistocene epoch but became extinct from 55 000 to 10 000 years ago. At left are a person and a modern red kangaroo for scale. (Arbitrary colours have been used.)

**TABLE 10.3** Examples of megafauna, now extinct, that lived during the Pleistocene. Weights are estimates based on the known relationship between body length and mass in living animals. The number after a scientific name refers to the artist's representation of the megafauna in figure 10.26 above.

| GROUP | SCIENTIFIC NAME | DESCRIPTION |
|---|---|---|
| monotreme mammals | *Zaglossus hacketti* (1) | a sheep-sized echidna, probably the largest monotreme that ever lived |
| marsupial mammals | *Diprotodon optatum* (2) | a hippopotamus-sized herbivore, weighing over 1000 kg, with some reaching up to 2000 kg |
| | *Thylacoleo carnifex* (3) | a carnivorous marsupial 'lion' with powerful cutting teeth and a leopard-sized body of up to about 100 kg |
| | *Zygomaturus trilobus* (4) | a herbivore weighing up to 450 kg |
| | *Macropus titan* (5) | a giant flat-faced kangaroo about 2.5 m tall and weighing over 150 kg |
| | *Phascolonus gigas* (6) | a cow-sized wombat weighing up to 200 kg |
| | *Palorchestes azael* (7) | a marsupial 'tapir' that may have had a small trunk |
| | *Procoptodon goliah* | a giant kangaroo about 3 m tall and weighing up to 200 kg |
| birds | *Dromornis stirtoni* | the largest bird yet discovered, flightless, about 3 metres tall and weighing more than 300 kg |
| | *Genyornis newtoni* | a large bird, weighing 100–200 kg |
| reptiles | *Wonambi naracoortensis* | giant Australian python, a non-venomous snake up to 6 m long that killed its prey by constriction |
| | *Megalania prisca* | a giant goanna over 6 metres long and weighing up to 1000 kg |
| | *Meiolania* spp. | giant horned turtles, weighing between 50 and 200 kg |

# Fossils: evidence of past life

**ODD FACT**

The study of fossils is known as palaeontology and a person who specialises in that field is known as a palaeontologist.

Our window to life on Earth in the geological past is through fossils. Fossils provide evidence of the kinds of organisms that once lived on Earth and the fossil record illustrates how the life forms of this planet have changed over geological time.

## Types of fossils

Various kinds of fossils can be identified:

- Physical fossils are the remains of all or part of the structures of organisms, such as bones, teeth or leaves. Most commonly, the organic material of those structures is replaced by minerals (a mineralised fossil) or the organic matter has disappeared and the organism simply exists as an impression (an impression fossil).
- Trace fossils are preserved evidence of the activities of organisms, such as footprints, tooth marks, tracks, burrows and coprolites (fossilised dung).
- Biosignatures or biomarkers provide evidence from which past life may be inferred to exist.

### Physical fossils

Figure 10.27 shows some examples of physical fossils. The trilobite (figure 10.27a) is an example of an impression fossil. The original material of its exoskeleton has disappeared, but the impression has been preserved in the fine-grained mudstone. The ammonite fossil (figure 10.27b) is an example of a mineralised fossil. The original material of the ammonite shell has been replaced, molecule by molecule, by pyrite, an iron sulphide mineral.

**ODD FACT**

The Lark Quarry trackways were excavated over an 18-month period in 1976–77, revealing more than 3000 footprints.

### Trace fossils

Among the best-known trace fossils are sets of dinosaur footprints, called dinosaur trackways. Figure 10.28 shows part of the dinosaur trackway at Lark Quarry. The tracks tell the story of a Cretaceous day, 95 million years ago, when a mixed group of small dinosaurs were on the edge of a lake. On the approach of a large carnivorous dinosaur, the herd of small dinosaurs scattered, leaving their tracks in the mudflats beside the lake. This mud solidified, the lake level rose and the footprints were progressively covered by more and more sediment.

### Biosignatures

The terms 'biosignature' and 'biomarker' are both general terms for all possible traces of life and are used variously by palaeontologists, astrobiologists and geochemists. One suggested replacement term is 'bioindices'.

Fossils, including trace fossils, are definite evidence of past life. Another clue to past life comes from **biosignatures**. A biosignature is a physical or chemical sign preserved in minerals, rocks or sediments that *can be inferred* to be cellular or to have resulted from the metabolic activities of an organism. Biosignatures are important in the study of Archaean rocks when the only life forms were microbes that are unlikely to be fossilised.

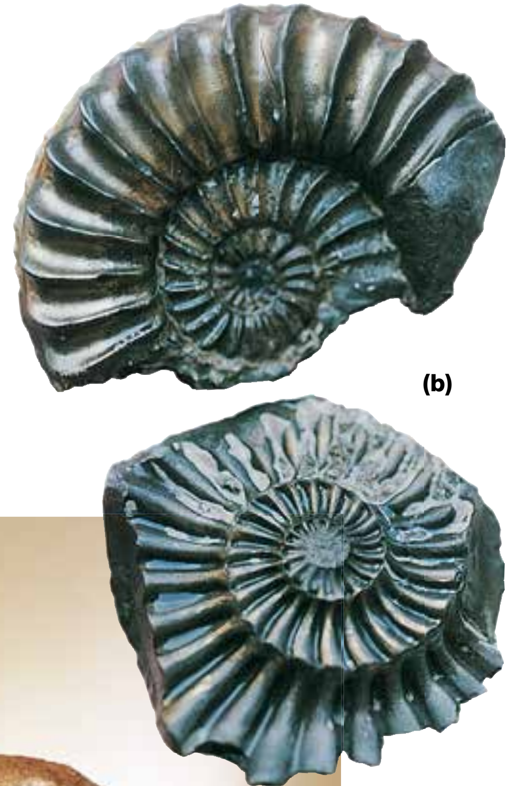
Examples of biosignatures include:

- corrosion pits in rocks caused by chemosynthetic microbes
- detection in rocks of isotopic ratios of bio-essential elements, such as carbon and sulphur, that are similar to the ratios produced by life processes in microbes
- stable molecules, such as steranes (compounds derived from plasma membranes), that can persist long after all other signs of the cells from which they came have disappeared.

Interpretation of a potential biosignature as a sign of past life must be made with caution as, in some cases, its origin might also be from a non-living (abiotic) source or it may be the result of later contamination. The older the microbial signatures, the easier to confuse them with those of an abiogenic origin and the more difficult to prove that they originated from the action of microbes.

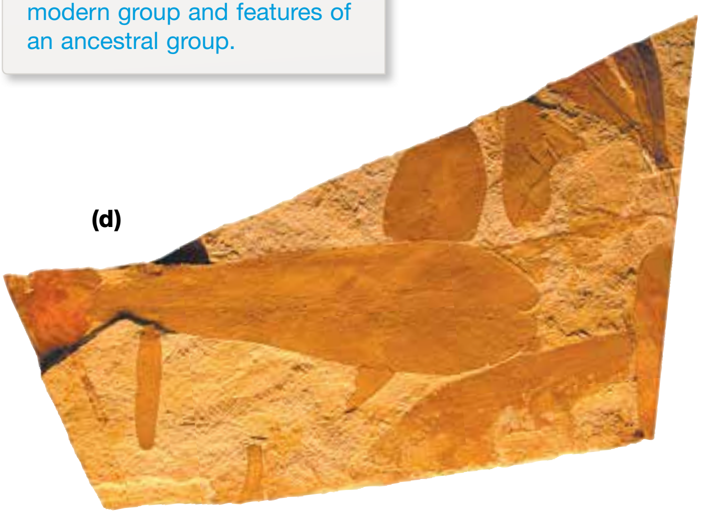
(a)

(b)



(c)

(d)



(e)

**FIGURE 10.27** Physical fossils. **(a)** Impression of the hard exoskeleton of a fossil trilobite, *Elrathia kingii*, that lived in the Cambrian period, about 550 Myr ago. **(b)** Fossilised ammonite shell (*Dactylioceras commune*). **(c)** Fossil of a large ammonite species (*Mesopuzosia* sp.). **(d)** Fossilised leaves from the extinct plant *Glossopteris*. (Images (a) through (d) courtesy of Judith Kinnear.) **(e)** Replica of the fossil organism *Archaeopteryx* (= 'ancient wing'). It shows bird-like features including feathers, but also shows reptile-like features including a long bony tail. This fossil is one example of a transition fossil that shows features of a modern group and features of an ancestral group.

**study on**

Unit 4 | AOS 1 | Topic 2 | Concept 2

**Fossils**
Summary screen and practice questions

**study on**

Unit 4 | AOS 1 | Topic 2 | Concept 2

**See more**
Fossils: evidence of past life

**FIGURE 10.28** A small part of the dinosaur trackway at Lark Quarry, near Winton in central Queensland. Some of the tracks in this figure were made by chicken-sized theropods with feet that had three toes of same length, with sharp claws. These footprints were on average 4.5 cm. (Image courtesy of Judith Kinnear.)

However, one unambiguous physical biosignature is the stromatolites from the Archaean eon. Fossil stromatolites are accepted as having a biological origin and signify the existence of microbial life *at least* 3.5 thousand million years ago. The microbes that formed these structures cannot be detected, but the laminated pattern of accumulated sedimentary grains in these ancient structures cannot be explained by any abiotic processes (refer back to figure 10.24a on page 464).

**ODD FACT**

Astrobiologists have a strong interest in identifying biosignatures that could be indicative of life on other planets in our solar system or beyond.

## Fossils: macro- and microscopic

Many fossils are **macroscopic** and can be seen and studied with an unaided eye or very low magnification. Some fossils, however, are microscopic. These **microfossils** require the use of either light or electron microscopy to visualise them and to identify the details of their shape and other morphological features.

The earliest life forms on Earth were prokaryotic microbes. As a result, evidence of these early life forms often involves (i) searching for rocks of a relevant age that might have been formed under conditions suitable for early life, and (ii) making thin slices of these rocks and searching for microfossils. (Refer to *Nature of Biology Book 1 Fifth Edition*, page 51 to read about Dr Dave Wacey, a palaeontologist who found the oldest fossil cells on Earth, in the Pilbara region of Western Australia.)

Another important group of microfossils is that of pollen and spores. (Pollen grains are the male sexual reproductive cells of flowering plants and conifers, while spores are the asexual reproductive bodies of mosses and ferns.) Because of their resistant outer coats, both pollen and spores are very durable and can be preserved in sedimentary rocks. Equally important is the fact that the outer coats of pollen and spores have distinctive patterns on their surfaces enabling identification at least to genus level and often to species level. These fossils are one tool that can assist in identifying the climate that prevailed in a region in the geological past.

### Bringing modern technology to the study of fossils

Today, high-resolution X-ray CT scanners are used to create detailed 3D images of fossils. These industrial CT scanners are similar to medical CT scanners, but they produce more intense X-rays that penetrate the fossil. The resulting images enable palaeontologists to view far more detail of a fossil, both external and internal, and from different angles.

Figure 10.29 shows the external and internal detail of a fossil as revealed by high-resolution X-ray CT technology. Note the four views of the digitally reconstructed cranium — frontal view of the cranium at right, and then, moving to the left, a more lateral external view, and then two internal views.

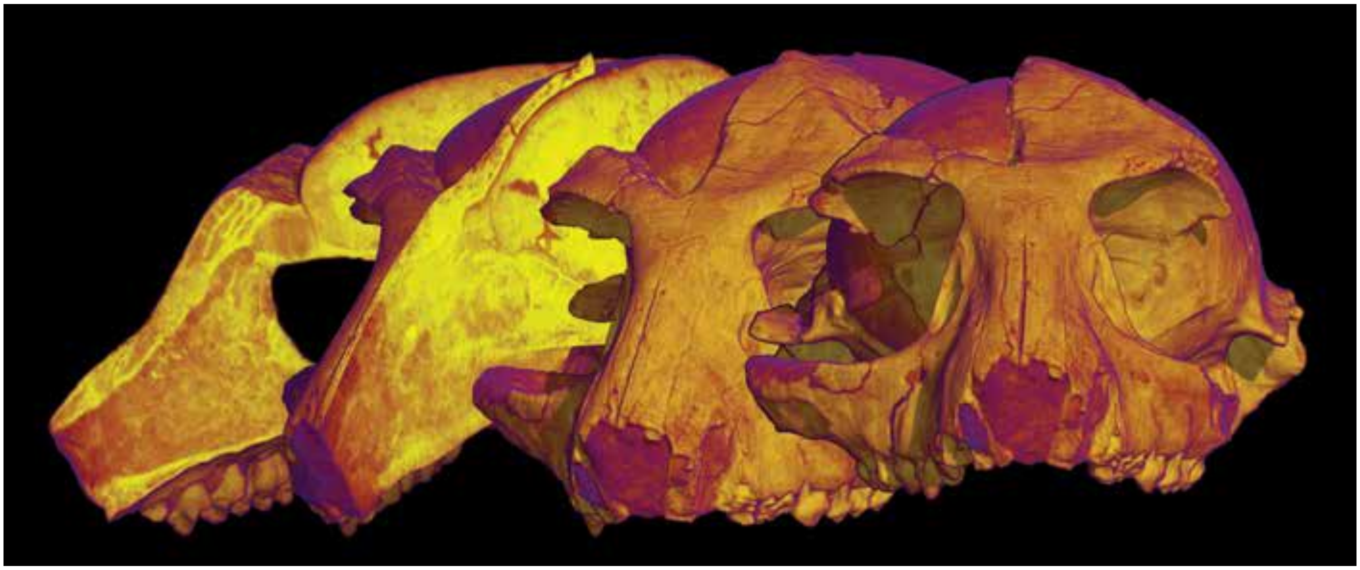


**FIGURE 10.29** False-coloured three-dimensional digital reconstruction of the skull of *Rooneyia viejaensis*, a primate from the Eocene of West Texas, derived from high-resolution X-ray computed tomography data. (Scanned at the University of Texas High-Resolution X-ray CT Facility, and courtesy of the Texas Memorial Museum.)

**eBook*plus***

**Weblink**
*Rooneyia*

## How fossils form

When organisms die, microbial action causes their decomposition and, after a period, no trace remains. Very rarely, the remains of organisms are preserved long after death. This process of preservation is known as **fossilisation**.

The chance of fossilisation of a dead organism is very low. As a result, the preservation of a significant proportion of an entire skeleton of a large vertebrate in one location is an unusual event. Predators and scavengers would be expected to consume tissues and, in so doing, disperse the remains. Great excitement accompanies the discovery of a near complete skeleton, such as the 1.6 million-year-old skeleton of the Turkana boy, now called Nariokotome Boy, (*Homo erectus*) in 1984 (see figure 12.56 on page 584), and the discovery of Penny the plesiosaur, in Queensland in 1989 (refer back to figure 10.22).

Conditions that enhance the preservation of tissues and increase the chance of fossilisation include:

- rapid burial in sediments after death; these sediments include:
  - silt deposits on a lake bed or on the flood plain of a river
  - sediments on the sea floor
  - ash from a volcanic eruption
  - windborne particles (loess) in a hot, dry desert.

• burial in conditions that are alkaline or oxygen-depleted or even anoxic (oxygen-free), as, for example, the mud at the bottom of a stagnant pool or a still body of water, a marsh, a peat bog or an asphalt seep
• occurrence of death in a very cold environment where the remains are preserved by freezing.

After a dead organism is buried, continued deposits of sediments bury it more and more deeply. Gradually, the weight of the overlying sediments compresses the original sediment layer so that it becomes rock. Mud deposits become mudstone or shale, sand deposits become sandstone, and volcanic ash becomes tuff.

When organisms are fossilised, most often only their hard tissues are preserved. In rare cases, however, the entire organism is preserved as a fossil. This can occur, for example, where insects become trapped in tree sap that becomes amber (see figure 10.30). Another situation is where organisms are buried and become frozen in the Arctic permafrost, such as has occurred with woolly mammoths. In these cases, DNA and other organic material can be extracted from these fossils.

**FIGURE 10.30** A winged termite fly preserved in amber. The insect last flew in the skies about 40 to 50 million years ago and rested on a tree and was then engulfed in tree sap that became fossilised as amber.

In other cases, organisms, or parts of organisms are fossilised in an altered form. This occurs when organisms are compressed under layers of sediment and their tissues are replaced by a thin carbon film showing their outlines (see figure 10.31a). In other cases, the organic material of the dead organism is replaced by minerals, as can be seen, for example, in petrified wood and in opalised snail and clam shells (see figure 10.31b).

Sometimes all or parts of dead organisms become covered by sediments that later form sandstone or mudstone. The organisms decay, leaving a cavity known as a **mould**. When the cavity within a mould is later filled by other material, such as fine-grained sediment, a three-dimensional model of the organism, known as a **cast**, is formed (see figure 10.32).


**FIGURE 10.31** **(a)** Carbonised fossil showing the outline of part of the leaf of a fern. (Image (a) courtesy of Judith Kinnear.) **(b)** Opalised clam and snails. The material of the shells has been replaced by opal.

**FIGURE 10.32** Cast and mould of a trilobite fossil. The image (at left) shows a trilobite cast. The cast formed when fine sediments filled the cavity within a mould. Note how the cast forms a 3D image, including the eyes at the sides of the head. The image (at right) shows part of the mould of the trilobite. The mould was formed when the trilobite tissue decayed, leaving a space, the walls of which carried the impressions of the trilobite body. (Image courtesy of Judith Kinnear.)

An unusual case of fossilisation occurs at the La Brea tar pits in Los Angeles, California. The tar pits are in fact asphalt seeps. This site is one of the world's most famous late Pleistocene fossil sites. Here, fossils are recovered, not from buried sediments, but from asphalt seeps. The sticky asphalt preserves the bones of any organism that is trapped in it (see figure 10.33).

More than 600 different species have been found in these asphalt seeps, with ages ranging from 50 000 to 10 000 years ago (late Pleistocene). Organisms found at La Brea include amphibians, reptiles, birds and mammals. Among the many predatory mammalian species found are dire wolves (*Canis dirus*), the most common species recovered, sabre-toothed cats (*Smilodon*

## ODD FACT

To date, fossil bones of more than 2000 individual sabre-toothed cats have been recovered from the La Brea tar pits, and more than 4000 dire wolves.

*fatalis*), the next most common species, and the Pleistocene coyote (*Canis latrans*). Other fossils recovered include ground sloths (*Paramylodon harlani*) and even mammoths (*Mammuthus columbi*).

In a typical ecosystem, the ratio of carnivores to herbivores is about 1 to 10. In the fossil assemblage recovered at the La Brea tar pits, the ratio of carnivores to herbivores is about 10 to 1. What factor(s) might have produced this atypical ratio? This ratio arises because if a herbivore became trapped in the asphalt seep, this attracted numbers of carnivores, such as a pack of dire wolves, many of which were themselves trapped.

**FIGURE 10.33** An excavator, Christina Lutz, at work for the Page Museum clearing bones recovered from the La Brea tar pits. (Image courtesy of La Brea Tar Pits and Page Museum, USA.)

## KEY IDEAS

- The biodiversity on Earth has changed over geologic time.
- Fossils provide evidence of the forms of past life on Earth.
- Evidence of past life may be categorised as structural, as trace fossils or as biosignatures.
- Structural fossils are often greatly altered, and may be impressions or may be mineralised.
- Fossilisation of a dead organism can only occur under special conditions.

## QUICK CHECK

**8** Identify the following statements as true or false:
  **a** Fossilisation is the most common fate of dead organisms.
  **b** Physical fossils may have little, if any, of the organic material of the original organism.
  **c** Hard structures of a vertebrate, such as the bones and teeth, are more likely to be fossilised than its soft tissues.
  **d** A terrestrial ecosystem in the Ordovician period would be expected to be similar to one in the Triassic period.

**9** What is the difference between:
  **a** a trace fossil and a physical fossil
  **b** a biosignature and a fossil
  **c** a cast and a mould
  **d** a fossil impression and a mineralised fossil?

**10** List two conditions that would favour fossilisation.

# Evidence of biological change over time

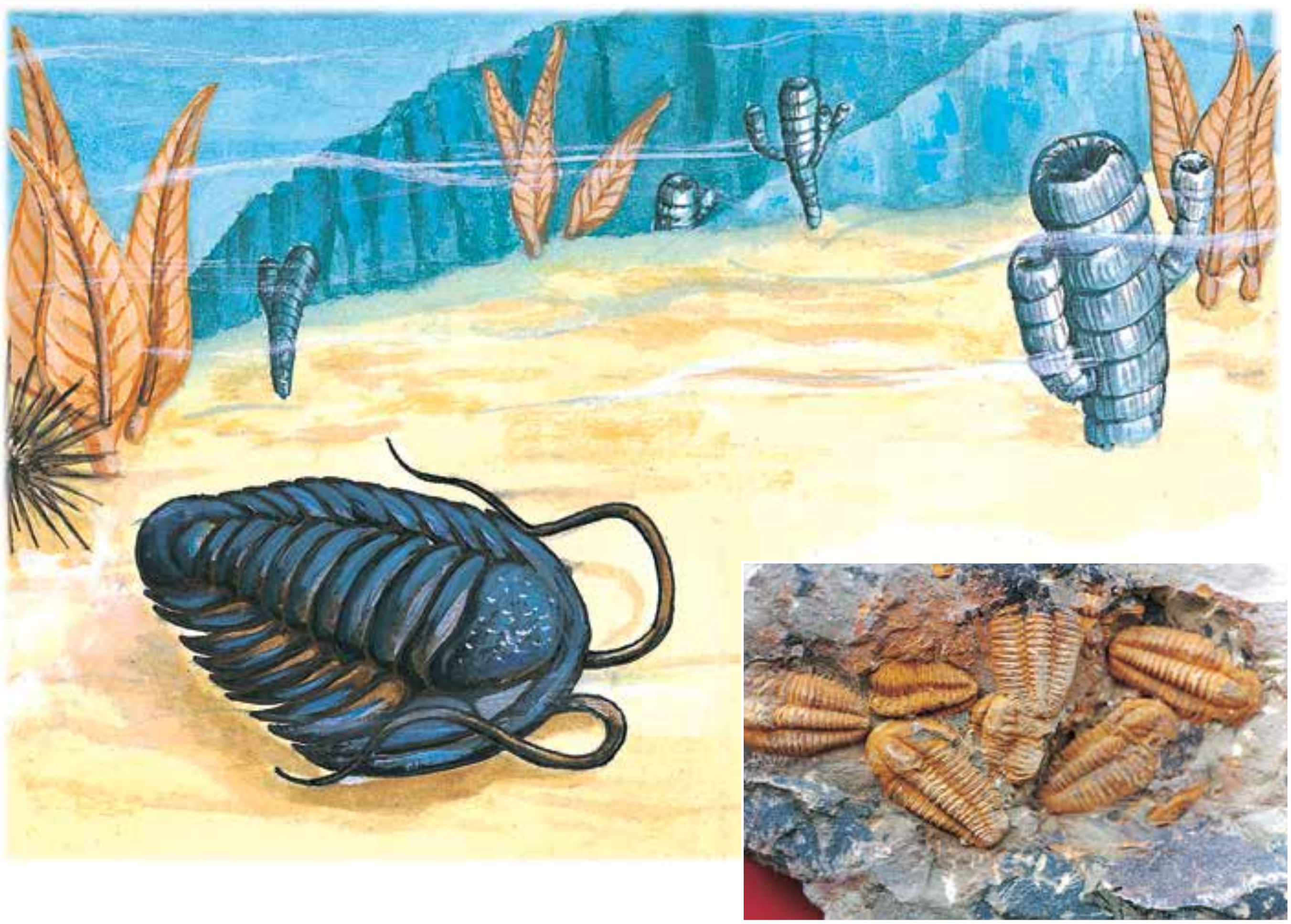
Let's examine some of the evidence of biological change over time:
- the fossil record
- structural morphology
- biogeography
- developmental biology.

## The fossil record

The Darwin–Wallace theory of evolution states that the various kinds of organisms living today evolved from ancestral kinds that lived in the geologic past. So, according to this theory, the rich diversity of the animals, plants, fungi, bacteria and archaea living on Earth today did not appear simultaneously but are the products of evolution over thousands of millions of years. The fossil record is one way in which this theory can be tested.

The fossil record reveals that, over time, changes have occurred in the types of organisms living on this planet. At one time, no terrestrial animals or plants existed. Conversely, many groups of plants and animals that were abundant in the past, such as trilobites (see figure 10.34), dinosaurs and giant club mosses, no longer exist. Other kinds of organisms living today, such as flowering plants and marsupial mammals, are absent from the fossil record of the distant geological past, and so are geological 'newcomers'.

**FIGURE 10.34** Fossils reveal evidence of past events. On a day about 450 million years ago, some trilobites were trapped in muddy sediment, perhaps in a pool that was drying out, where they died and were covered by further sediment and were eventually fossilised. They were revealed when the rock was fractured by a geologist's hammer 450 million years later.



Different kinds of organisms do not occur randomly in the fossil record but are found only in rocks of particular ages and appear in a consistent order; this is known as the law of fossil succession. For example:

- the first prokaryotes (microbes) appear in the fossil record long before the first eukaryotes
- the first single-celled organisms appear before the first multicellular organisms
- the first simple soft-bodied animals appear before the first complex animals with exoskeletons or with shells
- the first amphibians appear in the fossil record before the first reptiles
- the first ferns appear in the fossil record before the first flowering plants occur, and so on.

Sequences such as these can be seen in the fossil record worldwide. No exceptions have been found.

Figure 10.35 shows the approximate times of appearance of major groups of organisms.

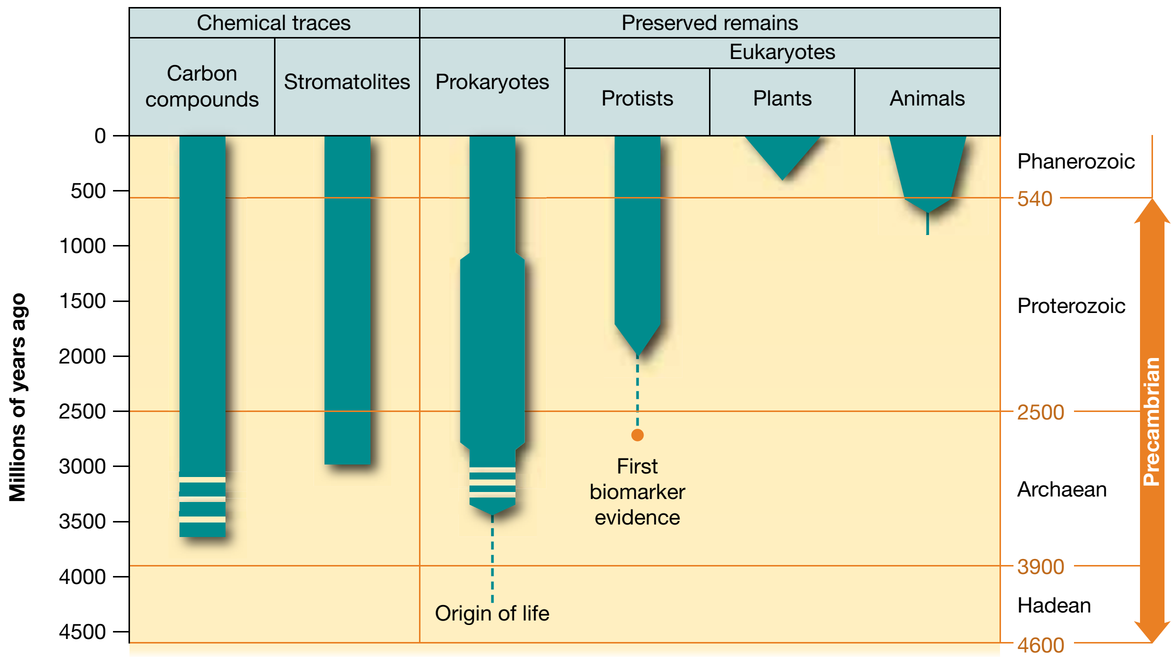



**FIGURE 10.35** Distribution of major groups over geological time.

Let's now look at the fossil record in terms of the major groups in more detail. Figure 10.36 shows a representation of the fossil record that identifies the time of first appearance in the fossil record of many different subgroups. This figure does not show extinctions.

The fossil record can also be examined for a particular group of organisms, such as the vertebrates. Figure 10.36 shows the fossil record for the vertebrates. Note that the first representatives of the vertebrates were the jawless fish, whose fossils first appeared in the Ordovician period.

If evolution has occurred, then it would be predicted that various life forms would appear in a predictable and consistent sequence in the fossil record, with ancestral species appearing in the fossil record before the species that evolved from them. Amphibians are believed to be ancestral to the reptiles and, consistent with this, the first amphibians appear in the fossil record before the first reptiles. Likewise, the first reptiles appear in the fossil record before the first birds. The fossil record provides evidence in support of the prediction that ancestral species will appear before the species that descend from them. In contrast, special creation of species would not necessarily be expected to produce species in such a sequence.

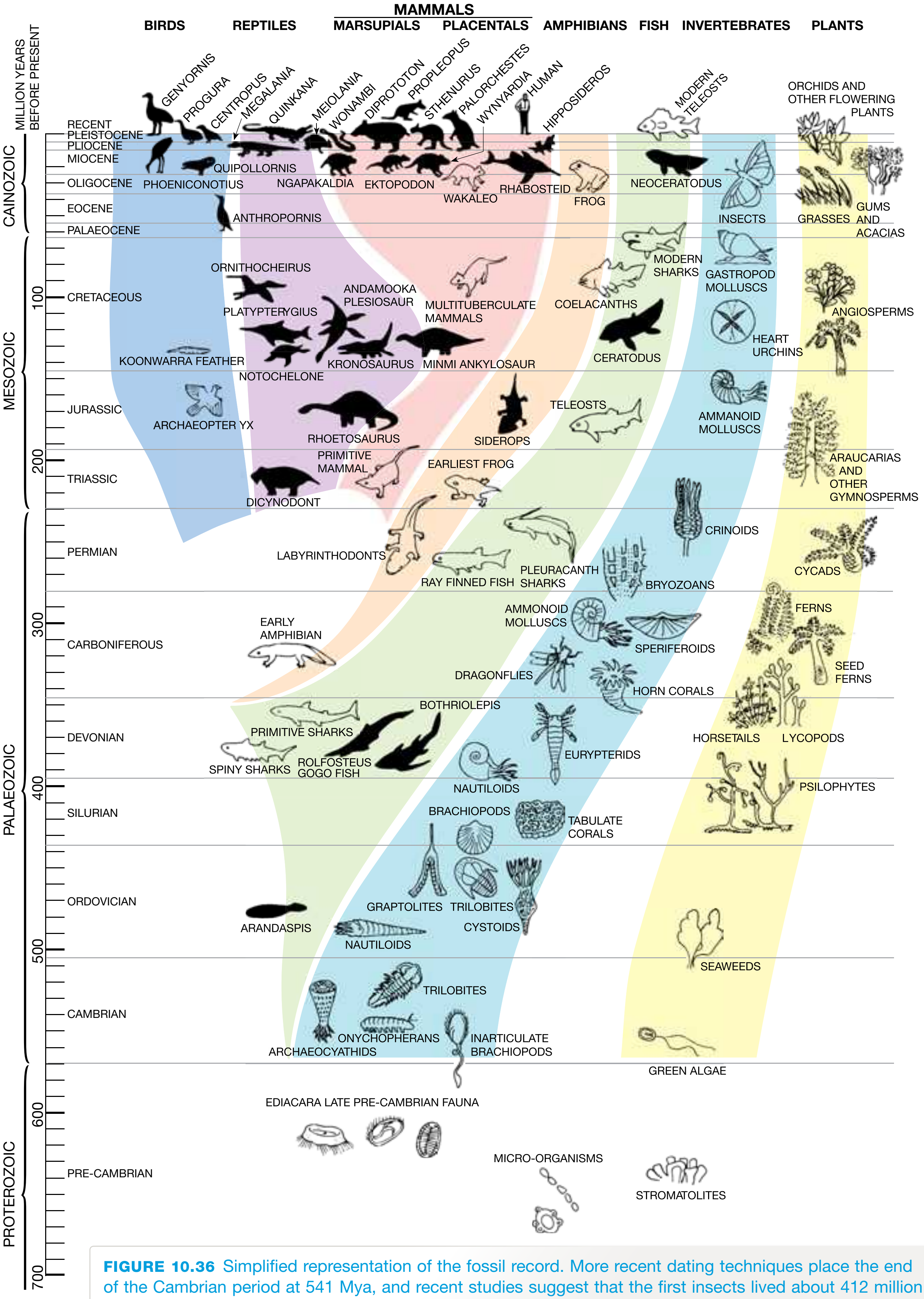
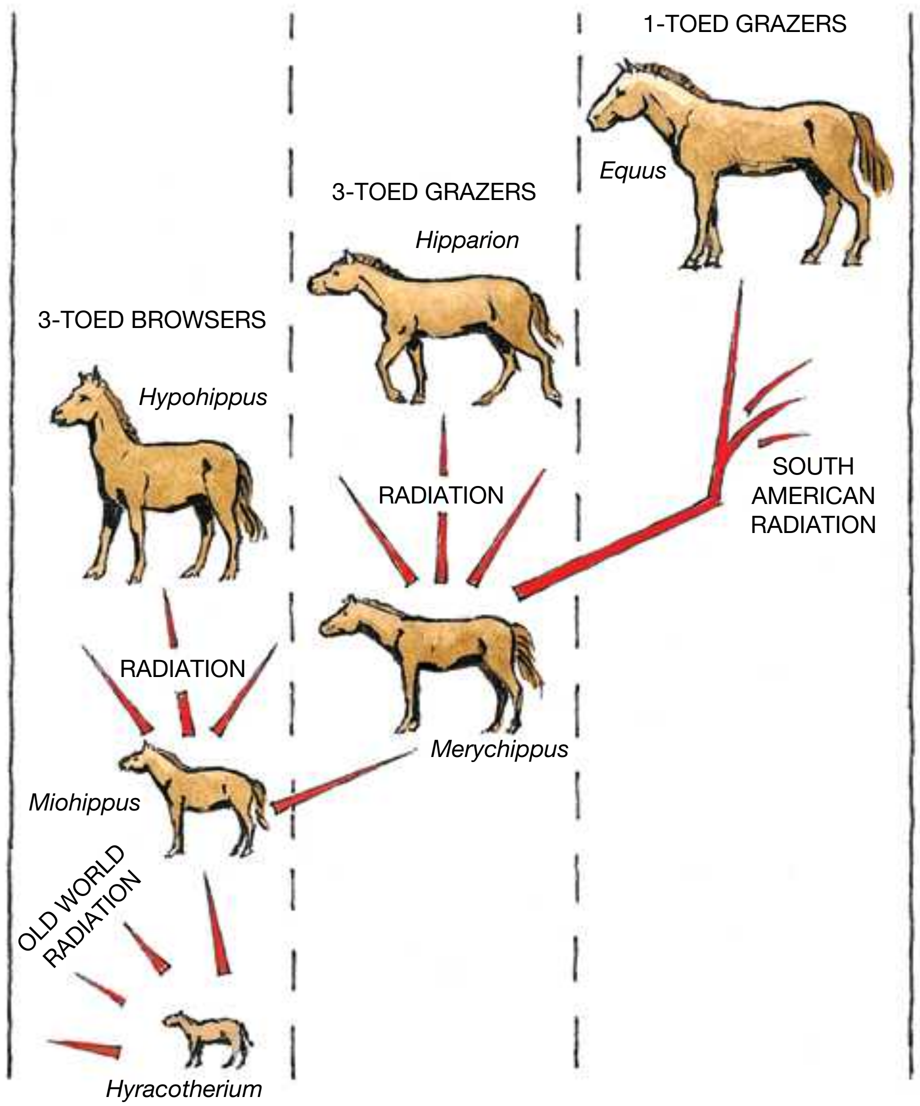



**FIGURE 10.36** Simplified representation of the fossil record. More recent dating techniques place the end of the Cambrian period at 541 Mya, and recent studies suggest that the first insects lived about 412 million years ago, during the Early Devonian Period. Overall, note the biological changes over time.

If evolution has occurred and if species can change over geologic time, then it would be predicted that the fossil record would reveal changes starting from an ancestral species that evolved into one or more new species, and these in turn evolved into yet different species. One example of evolutionary change can be seen in the evolution of the horse family (see figure 10.37). Members of the genus *Equus* are the only living representatives.




**ODD FACT**

*Hyracotherium* was originally known as *Eohippus* ('dawn horse').

**FIGURE 10.37** Evolution in the horse family involved many lines of descent and only some are shown. The various horse species are drawn to scale.

Figure 10.38 shows some of the structural differences in a few members of the horse family. These structures include feet, toe number, molar teeth and skull.

The first horse species (*Hyracotherium*) was about the size of a greyhound dog. It lived about 38–54 Myr ago in a forest environment, walked on four-toed feet and browsed soft leaves that it chewed with its small low-crowned molar teeth. In contrast, the modern horse (*Equus caballus*) is a much larger animal. It lives on open grassy plains and can move rapidly on its single-toed feet. It is a grazer that grinds tough grasses with its large high-crowned molar teeth that provide for greater wear during its lifetime. It is important to note that these changes are just samples of many different lines that evolved over geological time from isolated populations of various members of the horse family that lived under different environmental conditions. These changes do *not* represent a single line of evolution from *Hyracotherium* to *Equus*.

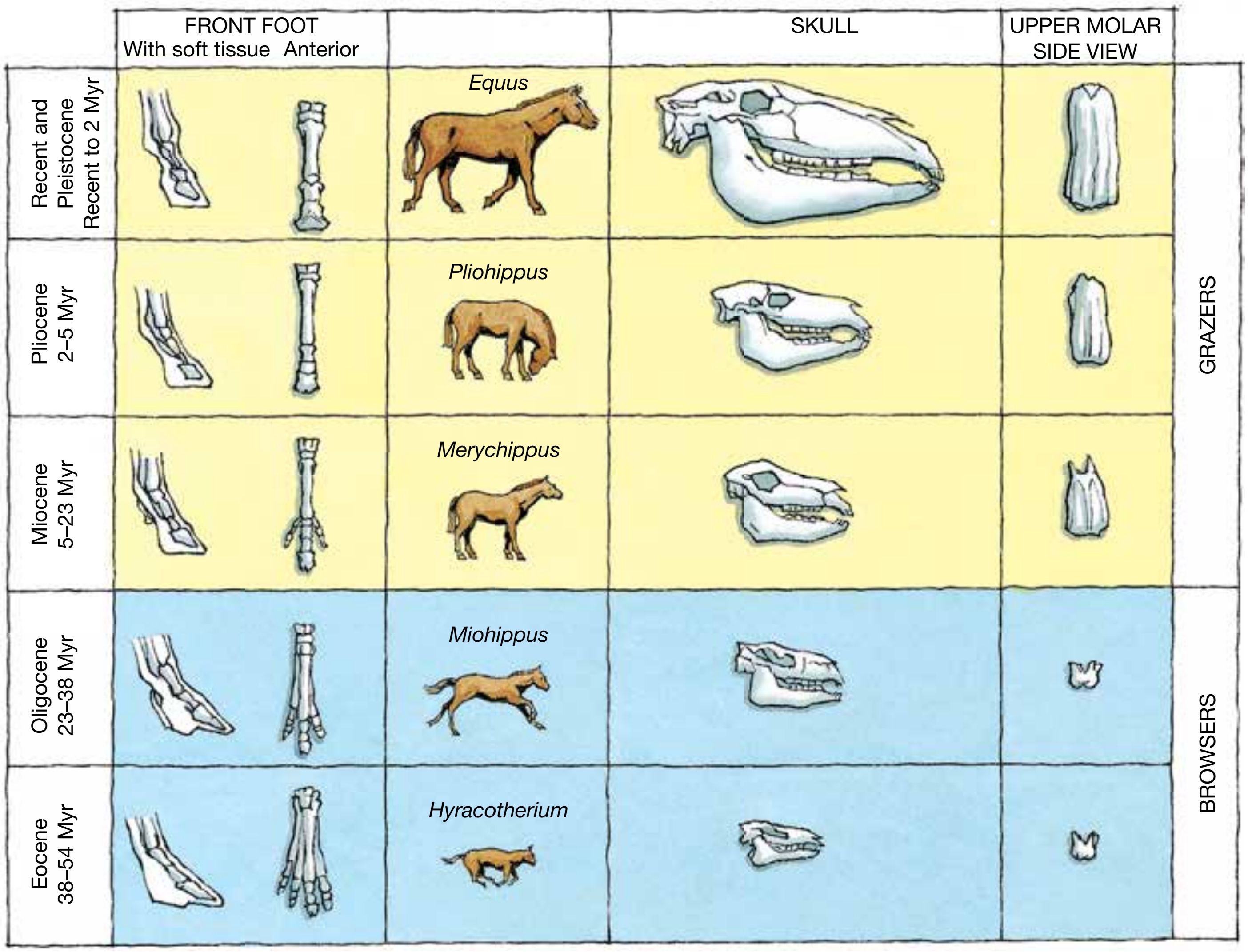



**FIGURE 10.38** Natural selection acted on some populations within the horse family over millions of years. It favoured features that equipped some populations for life on the grassy plains, including loss of some functional toes, which enabled speedier movement, and increased height of molar teeth, which gave a selective advantage for grazing tough grasses that wore away the tooth structure.

## Transitional fossils

If new species arise by evolution from ancestral species, it would be predicted that the fossil record should reveal some fossils that are intermediate between forms. Birds are believed to have evolved from a group of reptiles. If so, the fossil record might be expected to reveal organisms that show both the features of their reptilian ancestors and new features characteristic of birds.

*Archaeopteryx* = 'ancient wing'

**ODD FACT**

*Archaeopteryx* was about the size of a crow. Studies of their skeletons suggest that these early birds were not capable of the magnificent flight of modern birds.

The first unequivocal evidence of birds in the fossil record occurs in the late Jurassic period, about 150 Myr ago. A fossil skeleton of the earliest bird, *Archaeopteryx lithographica*, was found in a limestone quarry in Bavaria, Germany, in 1861. The fine-grained limestone preserved the faint impressions of feathers. In the absence of feather impressions, these organisms would have been classified as reptiles.

Like modern birds, *Archaeopteryx* showed the characteristic presence of feathers and a wishbone (furcula). However, it also showed some reptilian features now lost in modern birds; *Archaeopteryx* had teeth in its beak, claws on its wings, unfused (free) bones in its 'hand' and a long jointed bony tail (see figure 10.39).

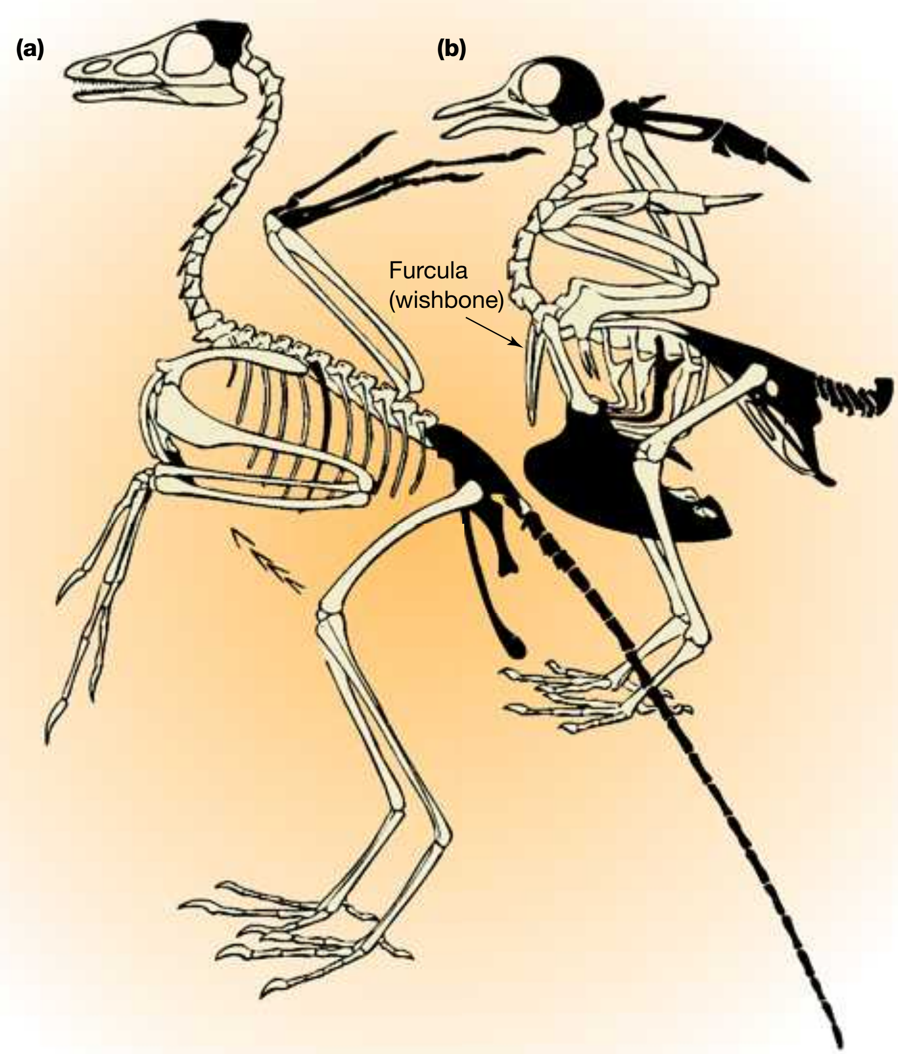



**FIGURE 10.39** Skeleton of **(a)** *Archaeopteryx* and **(b)** a modern flying bird. The black regions on the skeletons show distinctive reptilian features (at left) and bird features (at right). Biological change over time may be seen in the transition from ancient bird to modern bird.

study on

Unit 4
AOS 1
Topic 2
Concept 2

**Do more**
Transition fossil: evidence of evolution

There are many other examples of transition fossils or intermediate forms that relate an ancestral group with its descendants. These include:

- primitive amphibians that show a transitional stage between the simple pelvic (hip) girdle present in fish and the complex pelvic girdle in later more advanced amphibians
- fossil mammal-like reptiles that show a transitional stage between reptiles with simple conical teeth and mammals with teeth differentiated into incisors, canines, pre-molars and molars.

Fossils continue to be discovered, and the following box introduces the palaeontologist from the University of New England, NSW, whose fossil discovery featured on the front cover of the science journal *Nature* in December 2011.

## BIOLOGIST AT WORK

### Professor John Paterson — palaeontologist

Most people think that palaeontologists only dig up dinosaur bones. However, dinosaur experts are a minority in the field of palaeontology, with many specialising in organisms without a backbone. I'm one such palaeontologist, with my research focused on fossil arthropods — the group consisting of crustaceans, insects, spiders, centipedes, trilobites and kin. So, rather than spending hours on my belly armed with a paintbrush or arms covered in plaster, I typically have hammer and chisel in hand to split open rocks, parting the layers like the pages of a book. The goal of this arduous exercise is to reveal the remains of ancient organisms that haven't seen the light of day for millions of years and, on rare occasions, something truly spectacular.

I have been working with a team of palaeontologists on Kangaroo Island in South Australia on a deposit called the Emu Bay Shale. The fossils entombed within these 515-million-year-old rocks are exceptional, preserving a menagerie of marine organisms, including their soft parts, such as guts, muscle and eyes. One of our most exciting discoveries to date has been a pair of stalked compound eyes that belong to the world's first super predator, the metre-long *Anomalocaris*, with formidable grasping claws at the front of its head and a circular mouth lined with teeth-like serrations. But these are no ordinary eyes — each one is up to 3 centimetres in length and possesses over 16 000 lenses! This is impressive when you consider that a modern housefly has around 3200 lenses and these species are separated in time by a mere half a billion years. This find demonstrates that one of the earliest apex predators had very sharp vision.

Discovering the complex eyes of *Anomalocaris* also has important evolutionary implications. Firstly, it shows that a definitive arthropod trait — that is, compound eyes — exists in an animal that is on its way to becoming a true arthropod, from an evolutionary perspective. So *Anomalocaris* is considered to be a transitional form, often incorrectly referred to as a 'missing link'. Put another way, this find suggests that this particular type of visual organ appeared and was elaborated very early during arthropod evolution, originating before other characteristic anatomical structures of this group, such as a hardened exoskeleton and jointed walking legs, which are missing in *Anomalocaris*. It also tells us that large, sophisticated eyes evolved with astonishing rapidity, as there is virtually no evidence of eyes in rocks older than 520 million years of age.

**FIGURE 10.40** Dr John Paterson holding the grasping claw of *Anomalocaris*. Inset: reconstruction of *Anomalocaris*. (Images courtesy of Katrina Kenny and David Elkins.)

studyon

Unit 4
AOS 1
Topic 2
Concept 2

**See more**
Transition fossil

studyon

Unit 4
AOS 1
Topic 2
Concept 6

**Comparative anatomy**
Summary screen and practice questions

studyon

Unit 4
AOS 1
Topic 2
Concept 6

**Do more**
Comparative anatomy

## Structural morphology

The first mammals appeared about 200 Myr ago and all are believed to have evolved from a reptilian ancestor. The various kinds of mammals that are alive today share a common ancestry.

If mammals are related by evolution from a common ancestor, it would be expected that they would show similarities in their structural features, regardless of their way of life. This similarity of structure can be seen in the bones of forelimbs of various mammals (see figure 10.41).

**Each of these mammalian limbs has a similar number of bones arranged in the same basic pattern.** This similarity in basic structure exists, even though the limbs may serve different functions. The bones of the mammalian forelimbs are said to be homologous. So, the bones of a bat's wing are homologous to those of a whale's flipper, even though one structure is specialised for flying and the other is specialised for swimming.

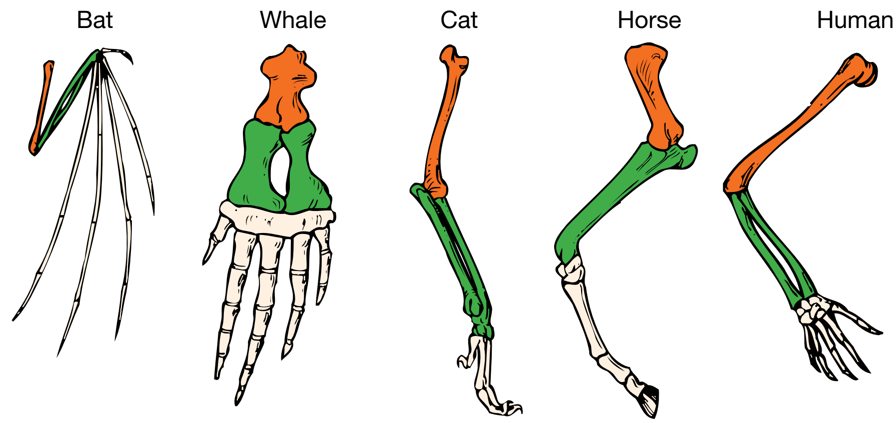



**FIGURE 10.41** The structures shown are homologous since they are all derived from a vertebrate forelimb. Do they have identical functions?

The various bones of the mammalian forelimb are equivalent and, by changing their sizes and proportions, one set of limb bones can be converted into another. Make one bone longer, another shorter, and so on . . . . . These morphological changes can be achieved through mutations that change the time or the place (when and where) genes are active. These mutations act, not on the coding sequence of a gene, but on their regulatory sequences. (This concept is discussed further in chapter 11, see page 505.)

Evolution commonly occurs by modification of pre-existing structures, not by the production of totally new structures. The homology seen in the pattern of bones in mammalian forelimbs illustrates how one basic pattern present in an ancestral species can, over time, be modified to produce a diversity of forelimbs adapted for flying, swimming, swinging and running.

The figure above relates just to mammalian forelimbs. In fact, if we look at the limbs of terrestrial vertebrates, we will see that their bone structures are homologous. In the case of evolution of the vertebrate limb, the pre-existing structures were the fins of lobe-finned fish. Fossils of these fish and also living lobe-finned fish have bones in their fins that are similar to the bones of modern terrestrial vertebrates.

Evidence of biological changes expressed in structural morphology may also be seen in the **vestigial organs** that are present in some species. If organisms are related by evolutionary descent, some species, because of changes in their way of life, will show the presence of functionless and reduced remnants of organs that were present and functional in their ancestor. A striking example of this change in structural morphology is seen in whales. Whales are mammals that evolved from a terrestrial mammalian ancestor, but they are now specialised for life in the seas, although they still breathe air. The front limbs of whales have become flattened flippers, and no hind limbs are apparent. However, their skeletons show the presence of a reduced pelvis and, in some cases, vestiges of the bones of the hind limbs. The presence of vestigial hind limbs in whales reflects the fact that whales evolved from four-limbed terrestrial mammals. In this case, the modification of the pre-existing pattern involves the reduction and/or elimination of some bones. The modification that results in producing vestigial bones does not involve new gene mutations. Rather, it can be achieved by changing the regulation of existing genes, as, for example, by delaying or stopping the activation of specific genes.

## Developmental biology

**studyon**

| | |
|---|---|
| Unit 4 | **Comparative embryology** |
| AOS 1 | Summary screen |
| Topic 2 | and practice |
| Concept 7 | questions |

Developmental biology is the study of the processes that result in the growth and development of multicellular organisms. These studies include comparing the development of different species in order (1) to identify their evolutionary relationships, and (2) to recognise how developmental processes have evolved. Over time, a developmental process that produces one structure in an ancestral species may be changed so that it produces new structures in descendant species.

Ancestral traits often appear and disappear at different stages of the embryological development of an organism. **Because they share a common ancestry, all vertebrate embryos display some common features at some point during their development.** Regardless of whether or not they are present in their adult structure, all vertebrates display the following features during at least some period of embryonic development:

- a tail, located posterior to the anus
- a cartilaginous notochord, located in the dorsal midline
- a hollow nerve cord, located dorsally
- pharyngeal arches.

Figure 10.42 shows embryos of three vertebrates with two of their common external features — pharyngeal arches and tails.

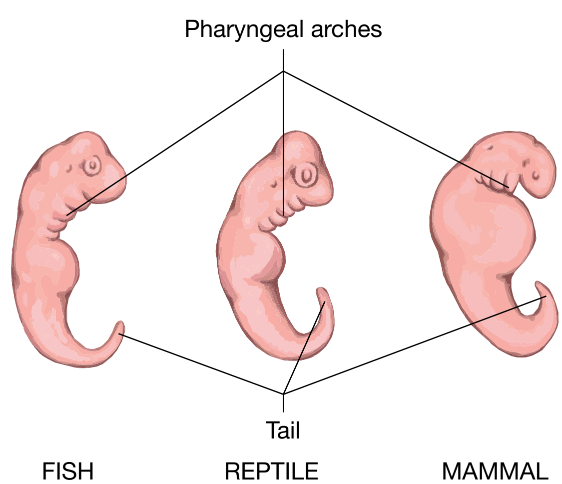



**FIGURE 10.42** Stylised diagram showing embryos of three types of vertebrate: fish, reptile and mammal (not to same scale). Note the presence of common structures, such as tail and pharyngeal arches.

**ODD FACT**

In human embryos, pharyngeal arches are visible as a series of five paired swellings from day 20 to day 35 of development. These arches are numbered 1, 2, 3, 4 and 6, because the fifth appears briefly and does not develop.

**Pharyngeal arches:** In aquatic vertebrates, such as jawless fish, sharks, rays and bony fish, the pharyngeal arches develop into gill arches that support the gill filaments, which take up dissolved oxygen from the water. A constant stream of water enters through the mouth of the fish, flows over the gill filaments and exits through the gill slits (see figure 10.43). So what happens to the pharyngeal arches in air-breathing vertebrates that do not have gills? In these terrestrial vertebrates, different structures are derived from the pharyngeal arches: in reptiles, some arches give rise to bones of the lower jaw, and, in mammals, pharyngeal arch derivatives are seen in structures of the face, lower jaw and ear. This is one example of how, during evolution, developmental processes that control ancestral features can be changed to produce new biological structures in their descendants.

**Tail:** Tails are structures present in most adult vertebrates, but are normally absent from human fetuses and newborn babies. However, human embryos at one stage have primitive tails. Likewise, the great apes — chimps, orang-utans and gorillas — also lack tails, but have them during embryonic development. We can reasonably explain these observations by assuming that all vertebrates share a common ancestor as part of their evolutionary history. The appearance of embryonic tails in tail-less vertebrates may simply be explained as a feature of their common ancestor that has been retained during embryonic development.

**Hollow dorsal nerve cord:** All vertebrates develop a hollow dorsal nerve cord during embryonic development and retain this structure throughout life.

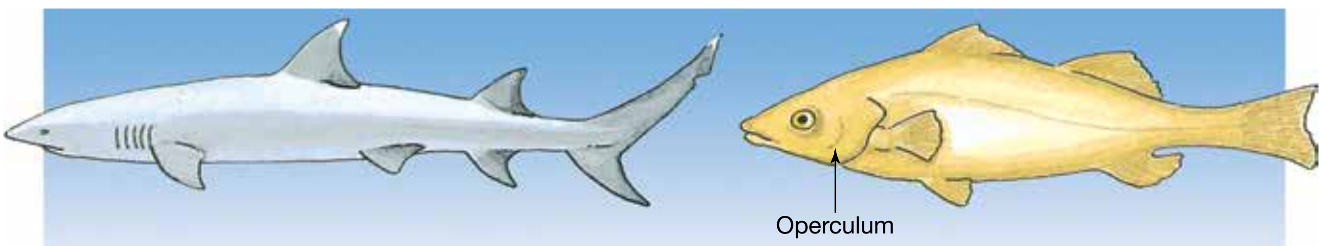



**FIGURE 10.43** Gill slits, which are the exit point of water from the gills, are visible in sharks and rays. In bony fish, the gill arches are concealed beneath a panel known as an operculum, which is derived from the second pharyngeal arch.


**FIGURE 10.44** Each isolated region has its distinctive native species of plant and animal. Where is this lost explorer?

**Notochord:** The notochord is a flexible rod made of cartilage. In jawless fish, such as lampreys, the notochord is retained as the major skeletal structure in adult life. However, in all other vertebrates, the notochord is only present for a short time, in embryonic development, and it then disappears and is replaced by the vertebral column (backbone). The notochord is a major skeletal element of the developing embryo that provides support, establishes the long axis of the embryo, and is the source of signalling molecules that shape surrounding tissues.

## Biogeography: biogeographic distributions

- Why are there no native cacti in Australia, yet Central and North America have many native cacti species?
- Why do the Hawaiian Islands have more than 20 species of honeycreeper birds that occur nowhere else in the world?
- Why does Australia have more than 100 different native marsupial species but Eurasia has none?
- Why do the Galapagos Islands have 13 species of finch that occur nowhere else in the world?

Unit 4 | AOS 1 | Topic 2 | Concept 5

**Biogeography**
Summary screen and practice questions

If each species was specially created, we might expect that the same ecological niche in different regions with the same environmental conditions might be occupied by the same species.

If, on the other hand, new species arise by biological change or evolution, we can predict that:

1. native species in different isolated regions will be distinctive, each group having evolved from different ancestral species
2. modern species native to a given region will be more similar to species that lived in that region in the geological past than to modern species living in a distant region with similar environmental conditions
3. the same ecological niche in different isolated regions will be occupied by different species (that are descended from different ancestral species that once lived in that region).

Prediction 1 above is supported by observation. For example, each discrete and isolated geographic region, such as an island continent or a cluster of oceanic islands, supports a distinctive group of endemic (native) species that are found nowhere else in the world. See a Sturt's desert pea (*Swainsonia formosa*) in a natural desert setting and you know you are in Australia, but see a saguaro cactus (*Carnegiea gigantea*) and you know that you are not in Australia but in the south-west United States (see figure 10.45).

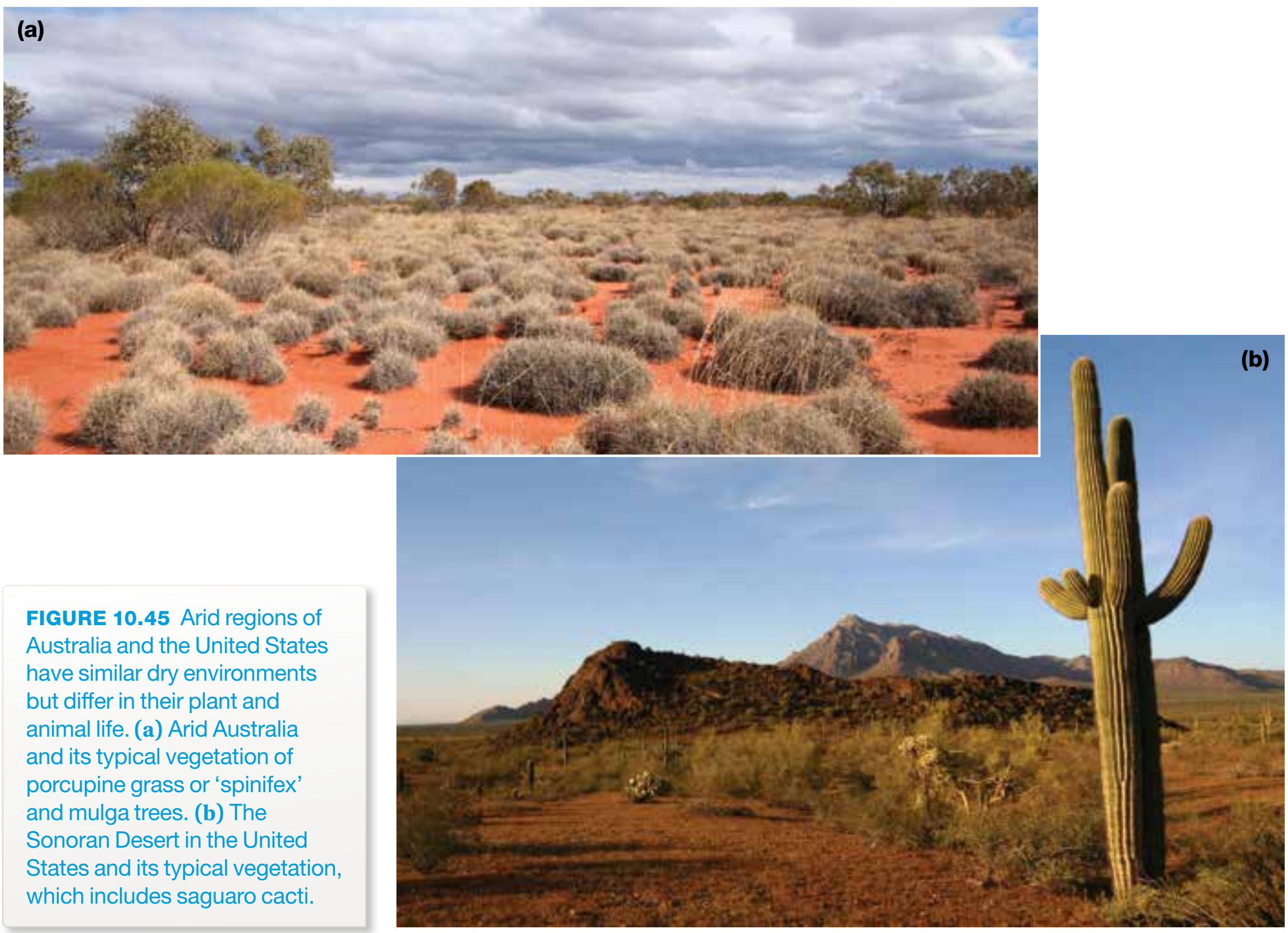



**FIGURE 10.45** Arid regions of Australia and the United States have similar dry environments but differ in their plant and animal life. **(a)** Arid Australia and its typical vegetation of porcupine grass or 'spinifex' and mulga trees. **(b)** The Sonoran Desert in the United States and its typical vegetation, which includes saguaro cacti.

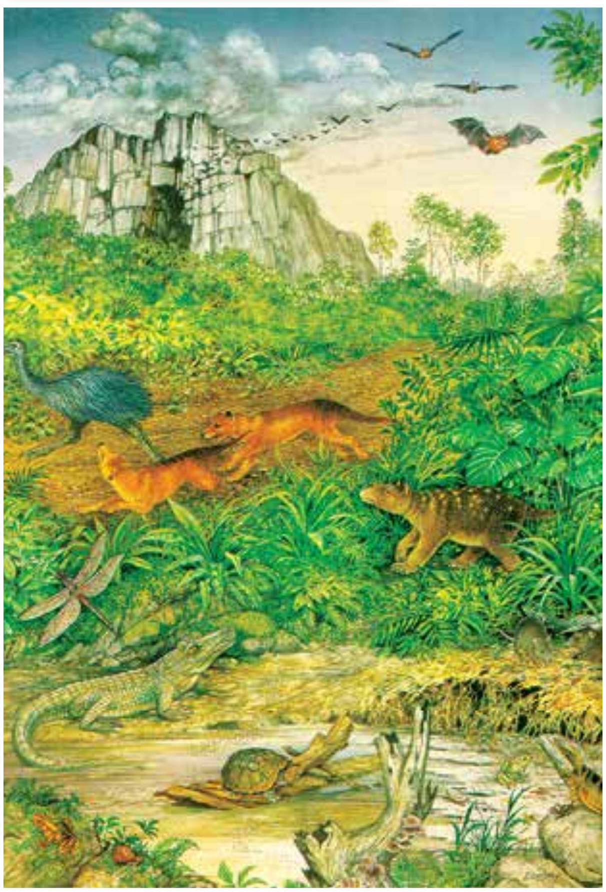
Both Australia and North America have large arid areas, and in each country the various plant and animal species found in these areas are distinctive and different. Table 10.4 and associated figure 10.45 show some of the species that are characteristic of each region.

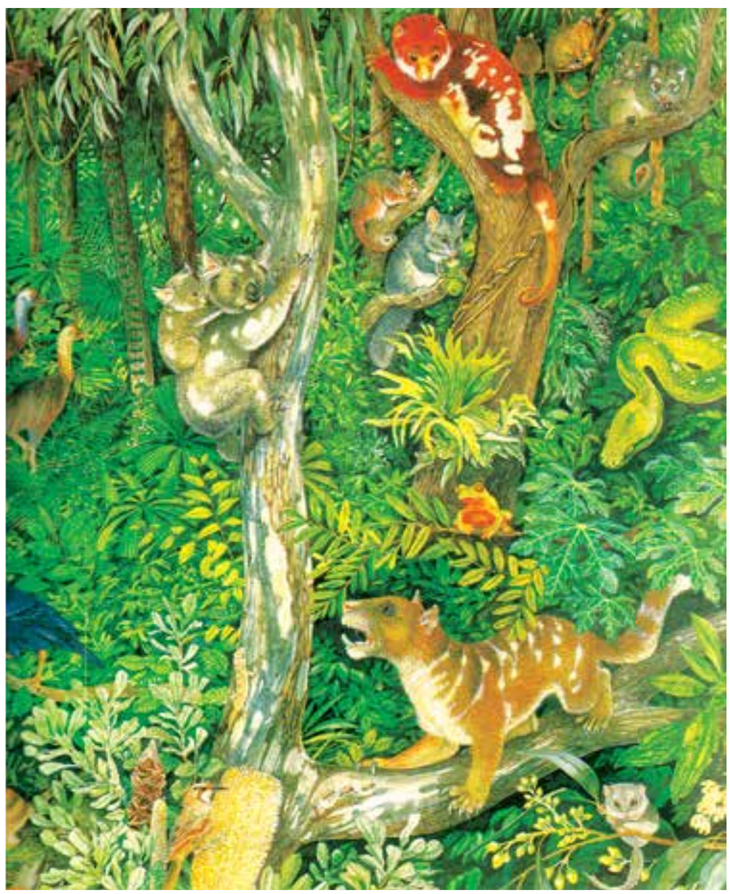
**TABLE 10.4** Species of two different geographic regions.

| Group | Region of arid Australia | Region of arid North America |
|---|---|---|
| Mammals | Marsupial:<br>bilby (*Macrotis lagotis*)<br>marsupial mole (*Notoryctes typhlops*) | Placental:<br>desert pocket mouse (*Perognathus penicillatus*)<br>black-tailed jack rabbit (*Lepus californicus*) |
| Plants | porcupine grass (*Triodia* spp.)<br>saltbushes (*Atriplex* spp.)<br>mulga trees (*Acacia aneura*) | saguaro cacti (*Carnegiea gigantea*)<br>creosote bushes (*Larrea tridentata*)<br>palo verde trees (*Cercidium microphyllum*) |

Prediction 2 (above) is supported by observation. For example, modern marsupial mammals of Australia are similar in structure to fossil marsupial species that are found in Australia; for example, at Riversleigh, fossils of a diversity of Australian species (reptiles, birds and mammals) have been found. There, different strata contain fossils from 25 million to just 40 000 years ago. Figure 10.46 shows a reconstruction, based on fossil evidence, of the Riversleigh ecosystem from 20 million years ago, when the area was a rainforest. Birds lived here, including the earliest known parrots, and *Emuarius*, which is thought to be ancestral to both modern emus and cassowaries. Reptiles included crocodiles, turtles, lizards and snakes. The dominant animals were the marsupials. Many kinds lived in the Riversleigh rainforests of 20 million years ago, including some that were ancestral to modern marsupials, such as kangaroos, possums, wombats, koalas and dasyurids. Other marsupials found in the ecosystem have no modern representatives; these extinct marsupial lines included so-called marsupial 'lions' and strange marsupials of the extinct genus *Yalkaparidon*. Monotreme mammals, including several ancient platypus species, lived in this ecosystem.

**FIGURE 10.46** A reconstruction of the Riversleigh ecosystem about 20 million years ago. Today, the region is harsh and dry, except during the monsoon season. This location is a World Heritage site.





*'Dasyurid' refers to members of family Dasyuridae, which includes carnivorous marsupials such as quolls, dunnarts and phascogales.*

Prediction 3 (above) is supported by observation. For example, the ant-eating niche on different continents is occupied by different species. In Australia, it is the echidna (*Tachyglossus aculeatus*); in South-East Asia, the pangolin (*Manis temminckii*); in South America, the giant anteater (*Myrmecophaga tridactyla*); and, in Africa, the aardvark (*Orycteropus afer*) (see figure 10.47).


**FIGURE 10.47** Different mammalian species in different continents feed on small colonial insects, such as termites or ants. **(a)** The Australian echidna is a marsupial mammal, while **(b)** the aardvark from Africa, **(c)** the pangolin from South-East Asia, and **(d)** the giant anteater from South America are placental mammals.

## KEY IDEAS

- **Evidence of biological change over time comes from the fossil record.**
- **The law of fossil succession states that fossils of plants and animals change over time and the order in which they appear remains the same.**
- **Evidence of biological change also comes from structural morphology, biogeographical distributions and developmental biology.**

## QUICK CHECK

**11** Identify the following statements as true or false:
  - **a** Bones of a whale flipper and those of a bat's wing are homologous.
  - **b** Vestigial organs are typically reduced and non-functional structures.
  - **c** All vertebrates, either as embryos or after birth, will display a tail.
  - **d** The order of appearance of different kinds of organisms in the fossil record is random.
  - **e** Species endemic to a region will be more similar to species that lived in that region in the past than to species that occupy the same niches in a geographically separated region.

**12** Identify an example of:
  - **a** a transitional fossil
  - **b** a mammal that occupies the ant/termite-eating niche
  - **c** a vestigial organ.

# Patterns of biological change

Over geological time, patterns of biological change have occurred; these include

- divergent evolution
- convergent evolution
- mass extinctions.

In this section, we will explore each of these patterns.

## Divergent evolution

**Divergent evolution** occurs when closely related species become more dissimilar over time, usually in response to different environmental conditions and different selection pressures. When divergent evolution occurs, structural similarities that are observed are true homologies.

Two species of hare found in North America are the snowshoe hare (*Lepus americanus*) and the black-tailed jack rabbit (*Lepus californicus*) (see figure 10.48). The snowshoe hare lives in the northern parts of North America that are snow-covered in winter. The black-tailed jack rabbit (it's really a hare) lives in desert areas. These two species are closely related and are descended from a common ancestor. Over time, in different habitats, they have diverged and eventually have become different species. The white-coated smaller-eared snowshoe hare is at a selective advantage in the cold north, while the large-eared jack rabbit has features that equip it for life in desert areas, such as its long ears that assist in losing body heat.


**FIGURE 10.48** Two related species in the genus *Lepus* are **(a)** the snowshoe hare, and **(b)** the black-tailed jack rabbit. Their structural differences are an example of divergent evolution of related species. How might these differences have originated?

In Australia, rosellas provide an example of divergent evolution (see figure 10.49). All of these rosella species share a common ancestor. The original populations in different regions over time have diverged in terms of their colouring so that they eventually became different species.

**Adaptive radiation** is a special case of divergent evolution (see figure 10.50). Adaptive radiation is the evolution of varieties of a species, each adapted for life in a different niche and each evolved over time from a single ancestral species, such as the more than 100 Australian marsupials.

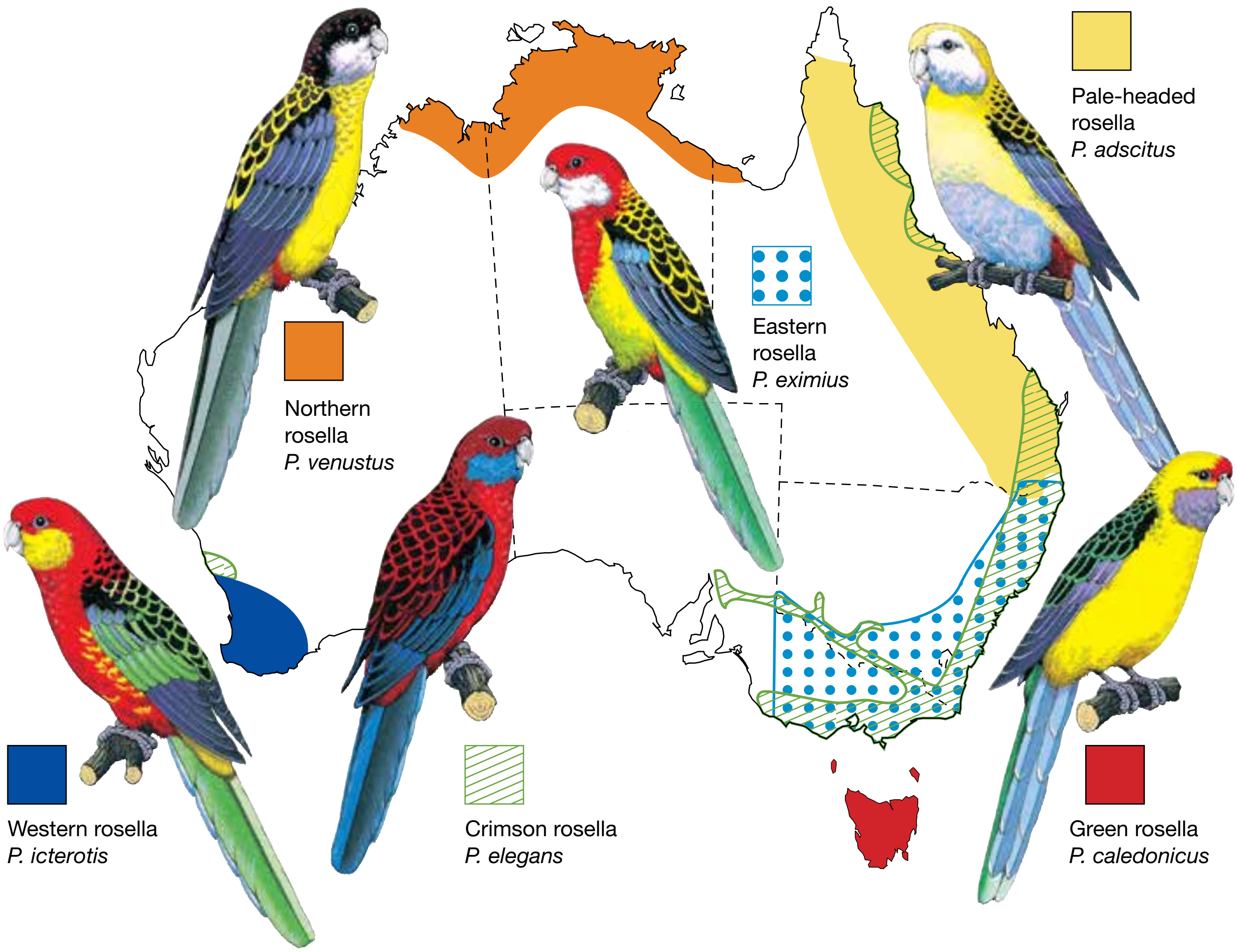



**FIGURE 10.49** Divergent evolution has produced several rosella species from an ancestral parrot. The map shows the distribution of the various rosella species (*Platycercus* spp.). In most cases, the distributions of the various species do not overlap. Where the species do overlap in their distribution, they do not interbreed.

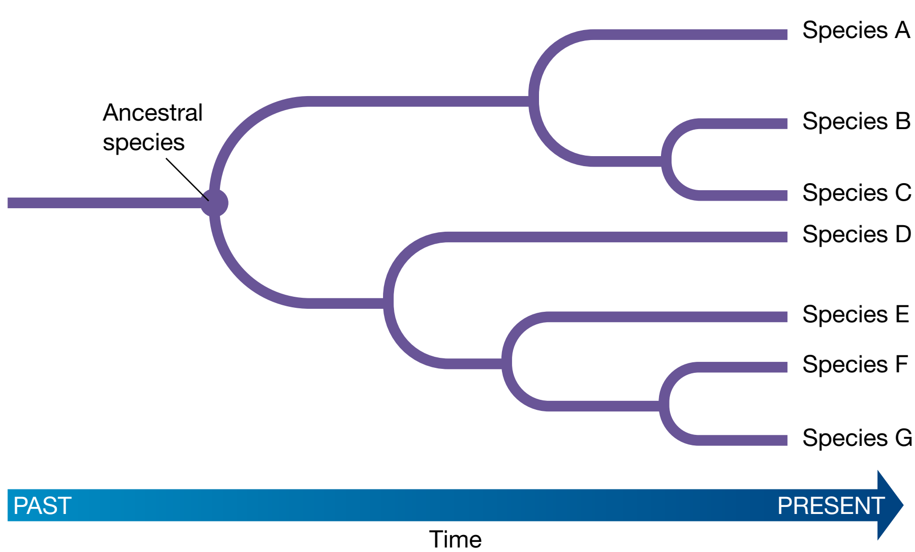



**FIGURE 10.50** Over geological time, one ancestral species can undergo evolution and give rise to several species. This is an example of divergent evolution.

Adaptive radiation commonly occurs when an ancestral species moves into an isolated region where a variety of habitats exists but where no competing species are present. Spectacular examples of adaptive radiation can be seen on oceanic islands or on island continents.

The role of master genes in producing beaks of different lengths and widths in the ground finches is discussed in chapter 11, page 534.

## Ground finches of the Galapagos Islands

The various finch species of the Galapagos Islands (also known as Darwin's finches) evolved from one ancestral species that spread to the various islands in that group. Natural selection acting on populations of this ancestral species over long periods of time produced many new species, each differing in terms of what it ate and how it obtained its food (see figure 10.51).

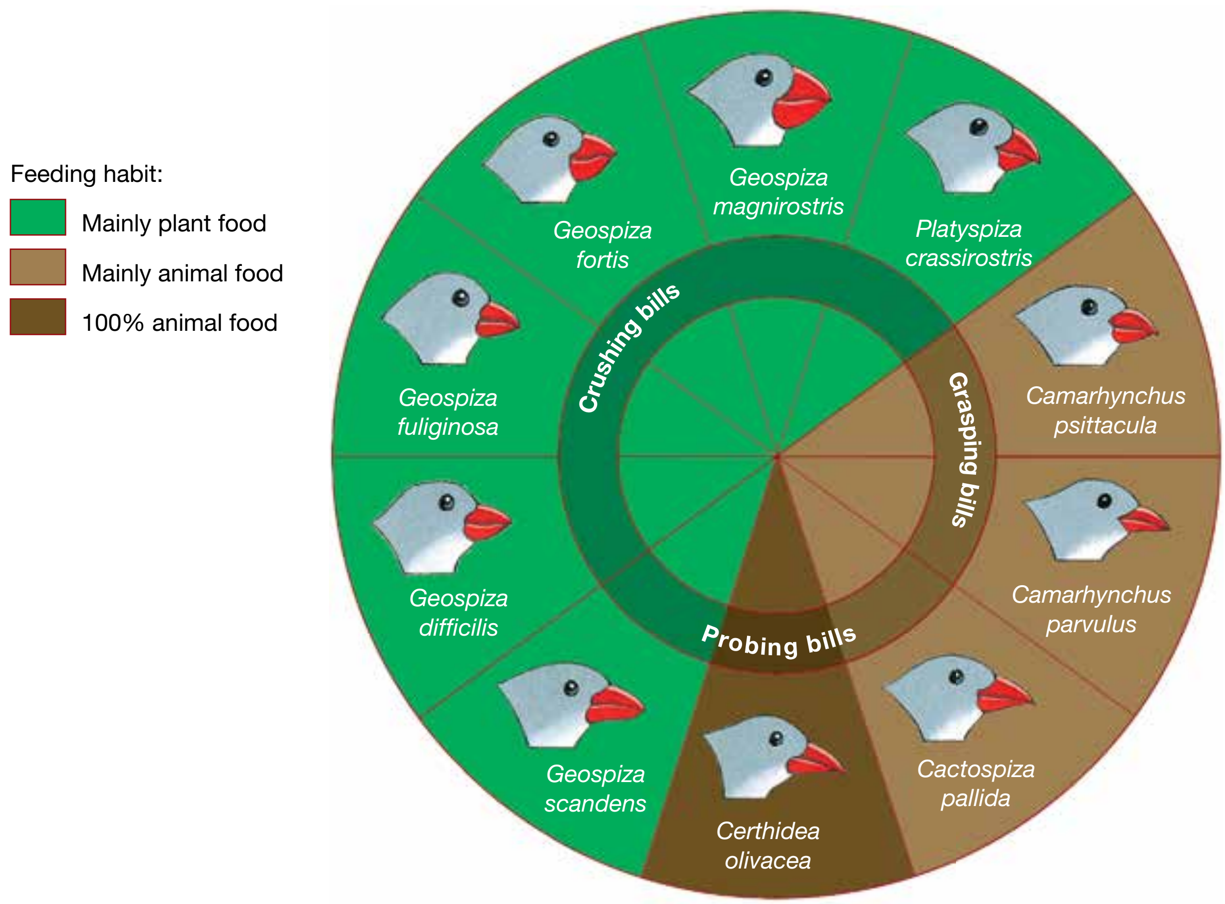



**FIGURE 10.51** One key difference between the Galapagos finches is in the shapes of their bills, which relate to their diets. The structure of a bill can equip the bird to crush or probe or grasp. Various finch species feed in different zones, so there are ground-feeding finches and tree finches. (Diagram adapted from Osbourne, R. and Benton, M., *Viking Atlas of Evolution*, Viking/Penguin Group, London, 1996, p. 19.)

## Marsupials of the island continent

From an original ancestral species, a large number of marsupial species have evolved, each showing features that equip it to live in a variety of environmental conditions and to occupy different niches (see figure 10.52).


**FIGURE 10.52** Some of the marsupial diversity that is unique to Australia. **(a)** The Tasmanian devil (*Sarcophilus harrisii*) is a nocturnal carnivorous marsupial. It feeds both as a scavenger (eater of dead animals) and a predator of small marsupials. **(b)** The koala (*Phascolarctus cinereus*) feeds by day exclusively on leaves of a few eucalypt species. **(c)** The kultarr (*Antechinomys laniger*) is a nocturnal hunter that feeds mainly on insects. **(d)** The feather-tailed glider (*Acrobates pygmaeus*) is an omnivore that feeds on small insects, nectar, pollen and sap; it is active by night and can glide for distances in excess of 15 metres. **(e)** The marsupial mole (*Notoryctes typhlops*) lives underground in sandy areas, is blind and has no external ears. Here, this marsupial mole is eating a gecko.

**studyon**

Unit 4 | AOS 1 | Topic 2 | Concept 8

**Divergent evolution**
Summary screen and practice questions

**studyon**

Unit 4 | AOS 1 | Topic 2 | Concept 8

**Do more**
Divergent evolution of vertebrates

The ancestors of kangaroos and wallabies were probably small tree-dwelling marsupials that ate fruits and leaves of various plants in the tropical rainforests. As these tropical rainforests shrank when the continent began to be more arid, various populations of these ancestral species moved from the forests into new habitats with new food sources. Over millions of years, natural selection acted on the different populations living in these new habitats and, over very many generations, a range of different species evolved. These new species had features that equipped them for life as grazers on grasslands and in open woodland, and as browsers in forests. The outcome is an example of adaptive radiation (divergent evolution) as seen today in the 35 living members of the family Macropodidae (see figure 10.53).

Evolutionary changes that occurred over millions of years produced features that enabled:

- hopping movement at significant speed, due to powerful hind limbs with long legs, hind feet and a fourth toe
- survival on abrasive and poor quality grasses, because of:
  - strong molar teeth with ridges
  - serial replacement of worn molar teeth
  - flattened pair of incisors on the lower jaw, the tips of these incisors fitting within the upper teeth
  - a large differentiated stomach in which microbial digestion of plant material occurs.


**FIGURE 10.53** **(a)** The tammar wallaby (*Macropus eugenii*) hides by day in dense shrubs in arid woodlands and feeds by night on native grasses. **(b)** The yellow-footed rock wallaby (*Petrogale xanthopus*) lives in rocky outcrops in arid scrublands; it sleeps by day between boulders and feeds by night on grasses. **(c)** The eastern grey kangaroo (*Macropus giganteus*) lives in forests, woodlands and shrublands, sleeping by day in a shady spot and feeding on grasses by night.

**FIGURE 10.54** Goodfellow's tree kangaroo (*Dendrolagus goodfellowi*). As their name implies, tree kangaroos are arboreal and feed on leaves.

Tree kangaroos (see figure 10.54) later reversed some of these changes when they evolved from ground-dwelling kangaroo species. The ground-dwelling species returned to the forests and became climbers that ate leaves of the rainforest trees. In the evolution of the tree kangaroos, their hind limbs became shortened and their hind feet became short and broad. Their sharper, more curved claws and textured soles of their hind feet enabled them to grasp and climb tree trunks.

### Snakes of the island continent

Adaptive radiation also occurred in other animal groups as the Australian island continent dried out. One striking example is seen in the elapid snakes (members of the family Elapidae), which are characterised by fixed front hollow fangs through which they inject their venom. Australia is home to 75 species of elapid snakes. The ancestral species of these modern elapid snakes reached Australia from Asia about five million years ago. Populations of these snakes moved into different habitats and evolved to produce the present diversity.

Today, elapid snakes are found in various habitats — tropical rainforests, cold mountain areas, swamps and deserts. Their ways of catching prey include pursuit over ground (as in the taipans), hunting in trees (as in Stephen's banded snake), ambushing with a lure (as in the death adders, see figure 10.55) or burrowing below ground (as in the burrowing snakes).

**FIGURE 10.55** The desert death adder ambushes its prey by using its thin tail tip as a lure to attract the prey. Note the positioning of the snake's thin tail above its head.

## Convergent evolution

Over geological time, natural selection may act on distantly related species to produce superficial similarities that are not due to a shared ancestry but reflect the fact that the species are adapted to a similar way of life, are exposed to similar selection pressures and typically occupy the same niche in their ecological communities. This situation is known as **convergent evolution** or **adaptive convergence**.

Similarities arising as a result of convergent evolution are common adaptations to similar selection pressures in a given habitat and have *not* been inherited from a common ancestor. Features that are similar in structure and function but that have arisen independently are said to be convergent (see figure 10.56).

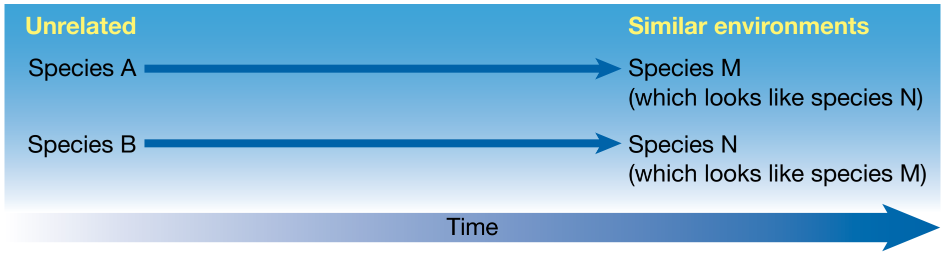



**FIGURE 10.56** Convergent evolution, or adaptive convergence.

Convergent evolution in action! Fish species in the Arctic and unrelated fish species in the Antarctic have independently evolved to life in frigid waters by producing different glycoproteins that act similarly as antifreeze agents. Different genes are involved in each group of fish.

In the plant kingdom, a spectacular example of convergent evolution is seen in some cactus species that are native to arid regions of the Americas and some euphorbia species that are native to arid regions of southern Africa (see figure 10.57). These plant groups are unrelated but they have independently evolved similar structures that equip them to survive in arid conditions. These similar features include succulent swollen stems (to store water), a leafless habit (to minimise water loss) and spines (to deter herbivores).



**FIGURE 10.57** Similar succulent stems and spines evolved independently in species of cacti and species of euphorbia. **(a)** Cacti belong to the family Cactaceae and are native to the Americas. These cacti (*Pachycerus pringelei*) are in a Californian desert. **(b)** Euphorbia belong to the family Euphorbiaceae and are native to southern Africa. This example is in the Namibian desert. Why is this an example of convergent evolution?

studyon

| Unit 4 | Convergent evolution |
|---|---|
| AOS 1 | Summary screen and practice questions |
| Topic 2 | |
| Concept 9 | |

The taxonomic rankings in order of decreasing inclusion are phylum (or division), then class, then order, then family, then genus, then species. Classification is discussed in *Nature of Biology Book 1 Fifth Edition* page 311.

## Mass extinctions

The fossil record provides evidence of changes in biodiversity over time, such as the appearance of new species that have evolved from ancestral species. In addition, the fossil record provides information about the loss or extinction of particular kinds of organisms from Earth's biodiversity. Extinctions may be global or local.

Global extinction occurs when all members of a particular taxonomic group are lost. For example, the death of the last marsupial lion (*Thylacoleo carnifex*) in the late Pleistocene epoch marked the complete loss of all members of the family Thylacoleonidae from this planet. Similarly, the death in 1936 of the last Tasmanian tiger (*Thylacinus cynocephalus*) marked the loss of the genus *Thylacinus* and the family Thylacinidae from this planet. Some extinctions are the complete loss of species, although other members of its genus still exist. So, the giant kangaroo, *Macropus titan,* became extinct, but other members of the genus are still living, such as the eastern grey kangaroo, *Macropus giganteus*, and the red kangaroo, *Macropus rufus*.

Local extinctions occur when members of a particular taxonomic group are lost entirely from a region, but are still in existence elsewhere. Many people are using their expertise to prevent local extinctions. You can read about the work of a zookeeper who is working to ensure that some of our Australian bird species do not become extinct (see page 495).

During Earth's geologic history, species became extinct and new species appeared. However, the fossil record identifies some periods of the Earth's geologic history during which extinctions occurred globally and on a massive scale. These are the so-called **mass extinctions**. **A mass extinction is a period in the geologic history of the Earth during which an abnormally high number of species die out over a relatively short time frame.** Five such periods of mass extinction have been identified from the fossil record as shown in figure 10.58.

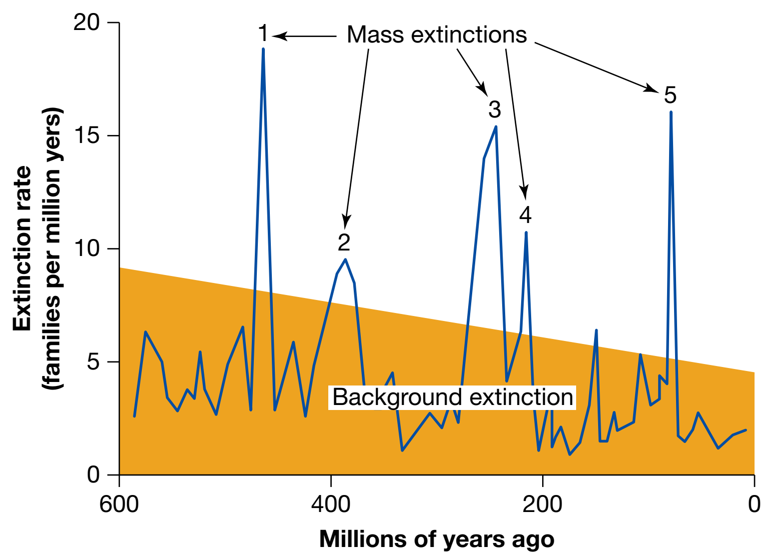



**FIGURE 10.58** Diagram showing the five major mass extinctions since the first appearance of animals and plants. These five extinctions are: 1. The late Ordovician mass extinction, 2. The Devonian mass extinction, 3. The Permian mass extinction, 4. The Triassic-Jurassic mass extinction, and 5. The Cretaceous-Tertiary mass extinction.

**studyon**

| | |
|---|---|
| Unit 4 | **Extinction: loss of species** |
| AOS 1 | Summary screen and practice questions |
| Topic 2 | |
| Concept 10 | |

The five periods of mass extinction are as follows:

- **The late Ordovician mass extinction:** occurred about 445 Mya and wiped out an estimated half of all marine species, reducing the numbers of trilobites, graptolites, and brachiopods; there were no terrestrial animal species at that time.
- **The late Devonian mass extinction:** occurred about 370 Mya, and an estimated three quarters of all species were eliminated.
- **The Permian mass extinction:** occurred about 250 Mya and was the greatest mass extinction in Earth's history, entailing the loss of an estimated 95 per cent of all marine species and about 70 per cent of terrestrial species. Species lost included all the trilobites, which had been the dominant marine invertebrates, some marine reptiles, some amphibians, many reef-building creatures, and many cephalopods.
- **The Triassic-Jurassic mass extinction:** occurred about 200 Mya.
- **The Cretaceous-Tertiary mass extinction:** also known as the **K/T extinction** and the **Cretaceous-Paleogene (K-Pg) extinction**, which occurred about 65 Mya. This extinction caused the loss of half of Earth's species and eliminated some animals groups completely, including dinosaurs, plesiosaurs and mososaurs (marine reptiles), pterosaurs (flying reptiles) and ammonites; other plants and animals groups, such as mammals, birds and insects, lost many, but not all of their species.

Each of these mass extinctions wiped out an enormous number of species, genera and families on Earth, greatly reducing Earth's biodiversity. However, these extinctions also created great evolutionary opportunities. Why? After the effects of the events that caused these extinctions had ceased, the surviving species were able to diversify and undergo adaptive radiation (see page 487). These survivors could expand into ecological niches that were left free with the loss of other species.

A classic example of diversification after a mass extinction may be seen in the adaptive radiation of the surviving mammals after the K/T mass extinction. The extinction of all dinosaurs, both large and small, and herbivorous and carnivorous, meant that many ecological niches in terrestrial habitats became vacant and competitor-free. From just a few small surviving species, the mammals underwent an intense period of adaptive radiation over the Cenozoic era, which produced the remarkable diversity of mammals that exists today on planet Earth, including the ocean-living whales, flying bats, gliding and hopping marsupials, grazing horses and the swinging primates.

**eBookplus**

**eLesson**
Zoos Victoria case study: the eastern barred bandicoot
**eles-2468**

## BIOLOGIST AT WORK

### Christine Tomlin — Zookeeper, Healesville Sanctuary

'Every day I fight wildlife extinction! I am a zookeeper at Zoos Victoria and work with critically endangered birds in recovery programs at Healesville Sanctuary.

The organisation is committed to saving 20 of Victoria's most endangered species, including the orange-bellied parrot (*Neophema chrysogaster*) and helmeted honeyeater (*Lichenostomus cassidix*) from extinction. The 20 species were determined based on their endangered status, and each species is unique and has a different recovery program. For the parrots and honeyeaters, we maintain a captive insurance population, which produces offspring that get released back into the wild. We hope these birds will then help to maintain or increase the wild population.

The main tasks that I undertake during the orange-bellied parrot breeding season is monitoring their nest boxes for eggs, chicks hatching and parental care. If parents are unable to care for their offspring, chicks may need to be fostered out to another family. We usually choose parents that have chicks of a similar age so they do not compete with one another. Eggs may also need to be artificially incubated if a female bird abandons her nest box.

Although the breeding season is interesting, the most rewarding part of my job is sending selected offspring back to north-west Tasmania. Here they are released into the wild and their progress is monitored carefully with the hope they will pass on their genetics to the next generation of wild orange-bellied parrots. Healesville Sanctuary is one of a number of institutions involved in the breeding program for orange-bellied parrots.

*(continued)*

I love birds so much that I studied ornithology at Charles Sturt University, New South Wales. One subject that I really enjoyed was speciation, which is when two populations of organisms change so that they become two new distinct species Learning about this biological process has made me realise how important biodiversity is and that every species on the planet is precious and deserves conserving.

Regularly, people ask me how did I get my zoo-keeping job? Although it's a highly competitive process, you can certainly improve your chances by completing an undergraduate course associated with the study of zoology or conservation science. I have travelled around the world working with many different organisations and different species, such as golden eagles (*Aquila chrysaetos*), merlin (*Falco columbarius*), jaguar (*Panthera onca*) and Baird's tapir (*Tapirus bairdii).* A highlight for me was a field study trip on horned larks (*Eremophila alpestris*), which inhabit grassland meadows in Washington, USA. I held the chicks while they were getting their legs banded for identification, certainly a memorable moment (see figure 10.59).

**FIGURE 10.59** Banding Horned Lark (*Eremophila alpestris*) chicks in Washington, USA. (Image courtesy of Christine Tomlin.)

Three million cubic kilometres of lava? Imagine an area 3000 km long and 1000 km wide covered in a one-kilometre-thick layer of lava.

## What caused mass extinctions?

The causes of these mass extinctions have been variously identified and often involve the action of several factors to produce the conditions that are hostile to life for most species. The late ***Ordovician mass extinction*** is believed to have been the result of an ice age that was accompanied by the formation of large ice sheets, falls in sea levels and climate change.

Other mass extinctions are associated with volcanic eruptions and the release of **flood basalts** that covered millions of square kilometres of the Earth's surface. One example of this is the ***Permian mass extinction***. This mass extinction appears to have been the result of massive volcanic eruptions. These were not small one-off eruptions but were eruptions on a large scale over long periods. In all, they produced about 3 million cubic kilometres of lava and other volcanic rocks, the evidence of which can be seen today in the **Siberian Traps** (see figure 10.60). When it flowed over the surface, the lava of the Siberian Traps released enormous volumes of gases — carbon dioxide, sulfur dioxide and hydrogen sulfide — into the atmosphere. This resulted in acid rain and impacted the global climate. It is estimated that, in just one year, the lava of the Siberian Traps released an estimated 1.46 thousand million tons of sulfur dioxide into the atmosphere.

### ODD FACT

Research published in 2014 has suggested that 'blooms' of methane-producing bacteria may have led to further increases in the carbon dioxide levels that, in turn, caused acidification of the ocean waters and even more global warming.

***The K-T mass extinction*** that wiped out the dinosaurs and many other life forms is associated with two cataclysmic events:

1. A major volcanic event: Ongoing eruptions produced flood basalts known as the **Deccan Traps** over a large area of India (see figure 10.61a). The flood basalt of the Deccan Traps covers about 500 000 square kilometres and has an estimated volume of 0.75 to 1.5 million cubic kilometres. In some areas, the basalt is three kilometres thick.
2. An impact event: The Earth was struck by an asteroid about 10 kilometres wide, which left the 180-kilometre-wide Chicxulub crater on an area around the Yucatan Peninsula in Mexico (see figure 10.61b). This impact event caused a shockwave that was felt worldwide, and tons of dust and debris from the impact blocked the sunlight, causing darkness and global cooling.

### ODD FACT

The asteroid impact that produced the Chicxulub crater is estimated to have released more than a billion times the energy of the atom bombs that destroyed Hiroshima and Nagasaki.



**FIGURE 10.60** **(a)** Map showing the extent of the flood basalts (lava and volcanic rocks) that form the Siberian Traps. **(b)** Photo showing part of the massive lava flow of the Siberian Traps that occurred about 250 million years ago and was part of the largest set of volcanic eruptions that are recorded in Earth's geologic history. These eruptions are believed to be the cause of the Permian mass extinction. This area is now a World Heritage site.

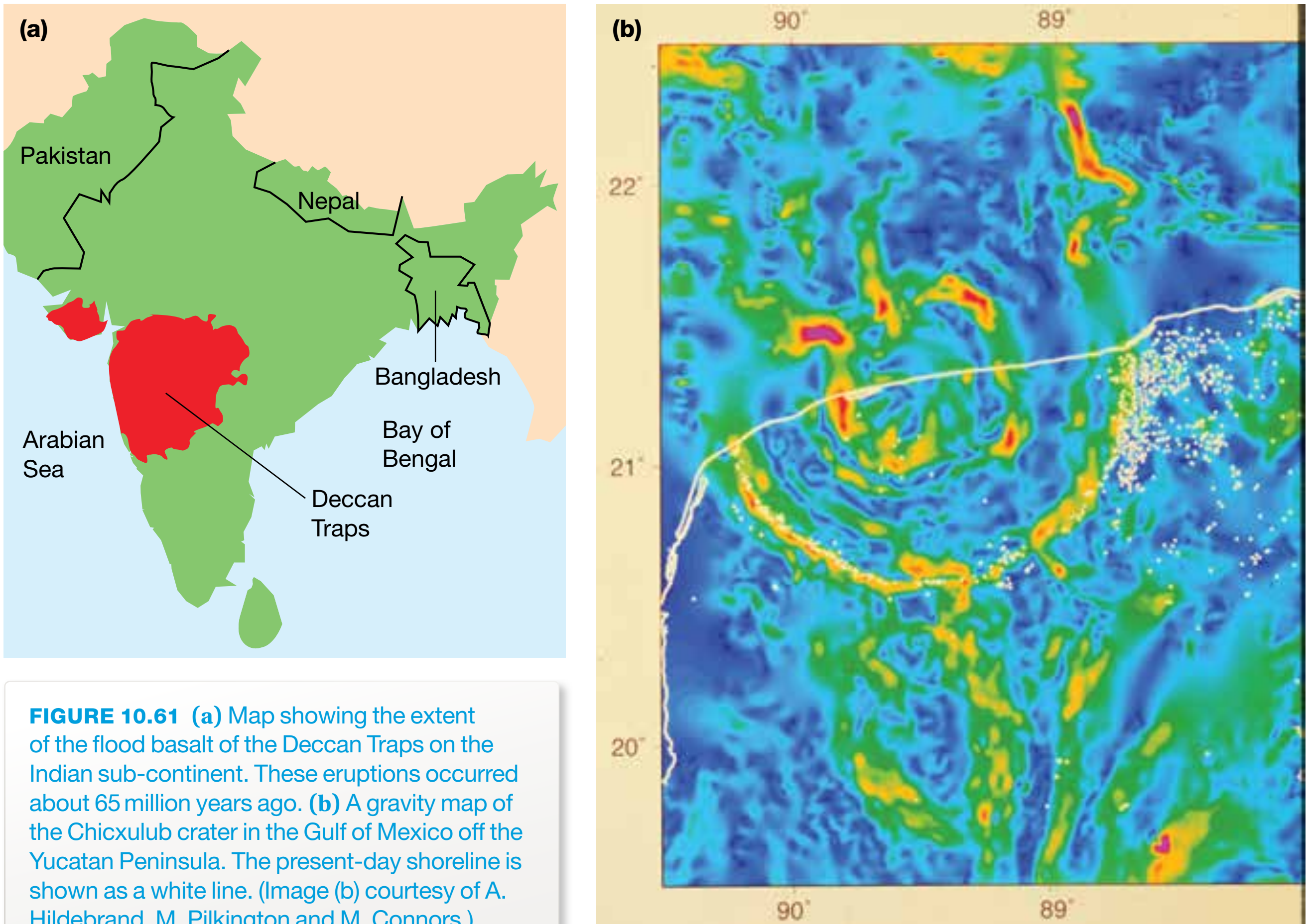



**FIGURE 10.61** **(a)** Map showing the extent of the flood basalt of the Deccan Traps on the Indian sub-continent. These eruptions occurred about 65 million years ago. **(b)** A gravity map of the Chicxulub crater in the Gulf of Mexico off the Yucatan Peninsula. The present-day shoreline is shown as a white line. (Image (b) courtesy of A. Hildebrand, M. Pilkington and M. Connors.)

## KEY IDEAS

- A number of different evolutionary patterns can be observed in nature.
- Divergent evolution results in the emergence of new species from a common ancestor, with each new species having features equipping them for survival in different environments.
- Adaptive radiation is a special case of divergent evolution that results in the emergence of many species adapted to different ways of life.
- Convergent evolution occurs when distantly related species display similarities because of their similar way of life, rather than common ancestry.
- Five periods of mass extinction have occurred in Earth's geological history.

## QUICK CHECK

**13** Identify the following statements as true or false:
  - **a** The mass extinction that resulted in the greatest loss of species occurred at the end of the Permian period.
  - **b** Cacti and euphorbia provide an example of divergent radiation.
  - **c** The diversity of elapid snakes in Australia provides an example of adaptive radiation.

**14** What is the difference between a mass extinction and a global extinction?

**15** List two structural changes that occurred during the evolution of modern kangaroos.

**16** Where would you expect to find the following:
  - **a** the Deccan Traps **b** the Chicxulub crater?

**17** When did the last ammonite live?

# BIOCHALLENGE

1 Examine figure 10.62, which shows rock strata in two widely separated regions of the world.
   a What conclusions, if any, can be made about the ages of stratum A4 and stratum B4? Explain.
   b What kinds of rock (igneous or sedimentary) have formed these strata? Give an example of a rock of this type.
   c Identify the youngest rock stratum in region A.
   d Identify the youngest rock stratum in region B.
   e What is the name of the method that you used to reach your answers to questions 3 and 4 above?
   f Do the terms 'oldest' and 'youngest' indicate relative or absolute age? Explain.
   g Is it reasonable to conclude that stratum A6 is the same age as stratum B8? Explain.

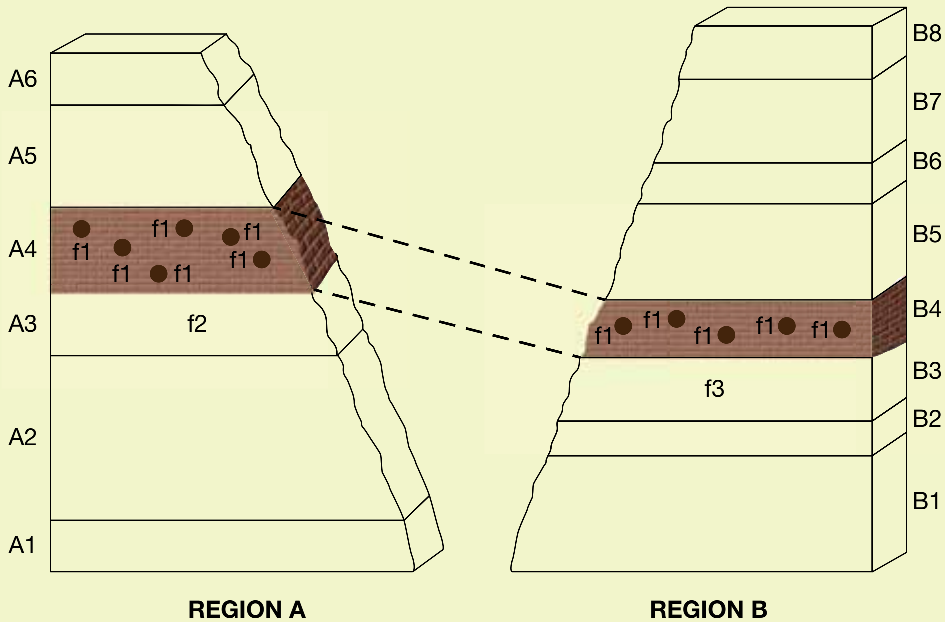



**FIGURE 10.62** Diagram showing two sets of rock strata (A and B) from widely separated regions in the world. An index fossil, f1, is identified in both strata A4 and B4. Fossil f2 occurs in region A only and fossil f3 occurs only in region B.

2 Palaeontologists in both regions examine these rocks and identify the same index fossil f1 in stratum A4 and in stratum B4. Fossil f1 is regarded as a 'good' index fossil.
   a List three properties that you would expect to apply to fossil f1.
   b Is fossil f1 likely to be microbial? Explain.
   c Is fossil f1 likely to a 10 000-year-old tooth from the La Brea tar pits? Explain.

3 Palaeontologists found a different fossil (f2) confined to stratum A3. Another fossil of a different species (f3) was found in stratum B3.
   a (Think carefully!) What conclusions, if any, can be made about the ages of fossils f2 and f3?. Explain.
   b Could fossil f3 also be present in other strata in region B? Explain.

4 The age of fossil f1 was independently identified in a separate study as being in the age range of 380 to 400 million years.
   a To what geologic period does fossil f1 belong?
   b Could this dating have been done using C-14 dating of organic material in the fossil?
   c Suggest, giving some detail, how this dating could in fact have been done?

5 Student J concluded that strata A5 and B5 must be from the Carboniferous period. Student K said 'Not necessarily!'
   a Do you agree with either student J or student K? Explain.

studyon

Unit 4 Changes in biodiversity over time
AOS 1
Topic 2

Practice questions

# Chapter review

## Key words

absolute ages
adaptive convergence
adaptive radiation
biosignatures
cast
convergent evolution
Cretaceous-Paleogene (K-Pg) extinction
Deccan Traps
divergent evolution
Ediacaran period
flood basalts
fossilisation
index fossils
K/T extinction
law of fossil succession
macroscopic
mass extinctions
microfossils
mould
radiometric
radiometric dating
radiometric dating technique
relative ages
Siberian Traps
special creation of species
stromatolites
transmutation of species
unconformity
vestigial organs

## Questions

**1** **Making connections** → Draw a concept map for 'fossils' incorporating key words from this chapter. You may add any other concepts that you wish.

**2** **Demonstrating knowledge** → Refer to the geologic time scale (figure 10.13 on page 455) and answer the following questions:
- **a** What is meant by the term 'pre-Cambrian'?
- **b** Identify the three periods of the Mesozoic era.
- **c** List the following periods in order from the earliest to the latest:
  Devonian, Cretaceous, Cambrian, Palaeogene, Silurian.

**3** **Interpreting pictorial information** → Refer back to figure 10.36 on page 476, which shows the times of first appearance in the fossil record of many different subgroups, and answer the following questions:

Which kind of organism appeared earlier in the fossil record:
- **a** trilobites or ammonites
- **b** birds or amphibians
- **c** cycads or ferns?

**4** **Evaluating alternatives** → The Hawaiian Islands formed at least five million years ago from volcanic mountains that rose from the ocean floor (see figure 10.63a).

Honeycreepers are endemic to the Hawaiian Islands. Today, 20 different species of honeycreepers are found on these islands. In addition, an estimated 45 honeycreeper species became extinct, mainly as a result of the introduction of alien species to the islands. The honeycreeper species have many structural similarities, but their bills differ in the size and the shape (see figure 10.63b). Some have finch-like beaks and eat seeds, insects and fruit; others have long thin beaks and brush-tipped tongues and feed on nectar.
- **a** Consider each of the statements below and comment, giving reasons, on the probability of each as a reasonable explanation for the diversity of honeycreeper species:
  - i special creation of each kind of honeycreeper
  - ii migration of many different kinds of honeycreepers to the islands from a distant mainland

(a)


(b)

**FIGURE 10.63** **(a)** Location of Hawaiian Islands. **(b)** Some of the honeycreeper species found on the Hawaiian Islands.

iii evolution of many species of honeycreeper from a few ancestral birds that reached the islands a long time ago.

**b** What name is given to the pattern of evolution outlined in (iii) above?

**c** A scientific hypothesis is one that allows predictions to be made and one that can either be supported or disproved by evidence. Which of the hypotheses above is/are scientific hypotheses?

**5** **Evaluating information** → Theophrastos of Eresos (368–284 BC) wrote about fossil fish that were found in rocks in hills at a great distance from the sea. He called these fossils 'dug-up fishes' and he explained their occurrence as being the result of fish eggs being washed by flood waters to high land where the eggs hatched in moist soil. Briefly outline an alternative explanation for these 'dug-up fishes' that would be given by scientists today.

**6** **Analysing information and drawing conclusions** → A group of students was discussing whether they could find the age of an isolated sample of igneous rock that they were given. Student P stated that she would use the principle of superposition. Student Q argued that searching for index fossils in this rock sample would be helpful. Student S proposed that carbon-14 dating would be useful.

Give your assessment, with reasons, of the validity of each student's proposal.

**7** **Applying knowledge and understanding** → From the various radiometric techniques for dating objects, name the one that corresponds to each of the following:

**a** can be used with a single mineral crystal from a rock

**b** produces argon-40 as a daughter product

**c** might be useful for the most ancient rocks on Earth

**d** can be used only for dating rocks with uranium-based minerals

**e** can be used for dating rocks that contain the mineral feldspar

**f** could be used for dating a wooden spear thrower

**g** might be useful to date blood on an ancient stone axe head.

**8** **Analysing information** → Refer to the account of the 1860 Oxford debate on Darwin's theory of evolution by natural selection (The Oxford debate of 1860) and answer the following questions:

**a** When was Darwin's *The Origin of Species* first published?

**b** Was Charles Darwin present at this Oxford debate?

**c** Suggest why the numbers attending the Oxford debate were high.

**d** Did Huxley reject absolutely the inference made by Wilberforce that he was descended from an ape?

**e** What was the significant point in Huxley's response to Wilberforce's question?

**f** What two points made by Hooker counted most against Wilberforce?

**g** What do you think is the meaning of Hooker's reference to 'his own conversion'?

**h** What were the 'old and new hypotheses' referred to by Hooker?

**9** **Developing logical explanations** → Suggest a reasonable explanation for the following observations:

**a** Igneous rocks that form from the solidification of molten magma or lava can be dated using radiometric techniques.

**b** Sedimentary rocks cannot be dated using radiometric techniques.

**c** Different people may interpret the same fossil in different ways.

**d** The end of some geologic periods are marked by the significant loss of many species.

**e** The fossil record shows no evidence of mass extinctions before about 443 million years ago.

**f** Australia has more than 70 species of elapid snakes.

**g** Sharks, porpoises and ichthyosaurs all have a similar streamlined body shape.

**h** Following a mass extinction, surviving species may undergo a period of adaptive radiation.

**10** **Discussing and sharing ideas** → Consider the following quotations taken from writings of Charles Darwin and identify what they appear to signal regarding Darwin's thinking about the origin of species (phrases in brackets have been added for clarity):

**a** 'It seems not a very improbable conjecture that the want of animals may be owing to none having been created since this country was raised from the sea.' (From manuscript of *Journal of Researches*, 1834.)

**b** 'This wonderful relationship in the same continent between the dead [extinct species] and the living, will, I do not doubt, hereafter throw more light on the appearance of organic beings on our Earth.' (From *Journal of Researches*, 1837.)

**c** 'The general conclusion at which I have slowly been driven from a directly opposite conviction, is that species are mutable, and that allied species are co-descendants from common stock.' (From Darwin's letter to L. Jenyns, October 1845.)

CHAPTER

# 11 How are species related?

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- recognise that species are related through evolutionary descent
- become aware that the degree of relatedness of species can be inferred using various techniques
- understand the use of DNA base sequences and amino acid sequences of proteins to estimate the relatedness of species and infer their evolutionary history
- become aware that other techniques, including DNA hybridisation and chromosome painting, can be used to infer relatedness between species
- gain knowledge of the concept of the 'molecular clock' and recognise its uses and limitations
- become familiar with the construction and use of phylogenetic trees and cladograms.

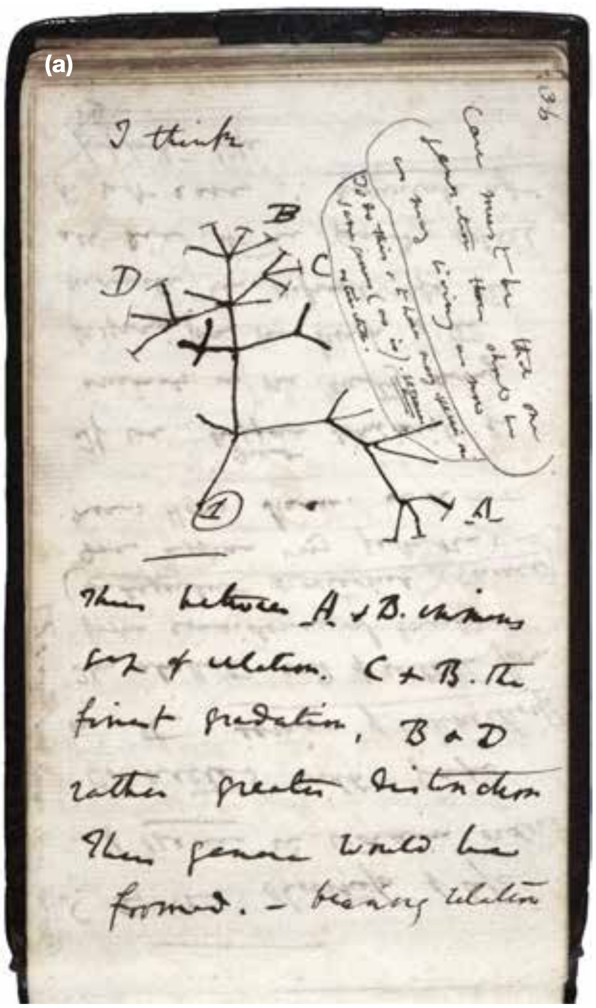
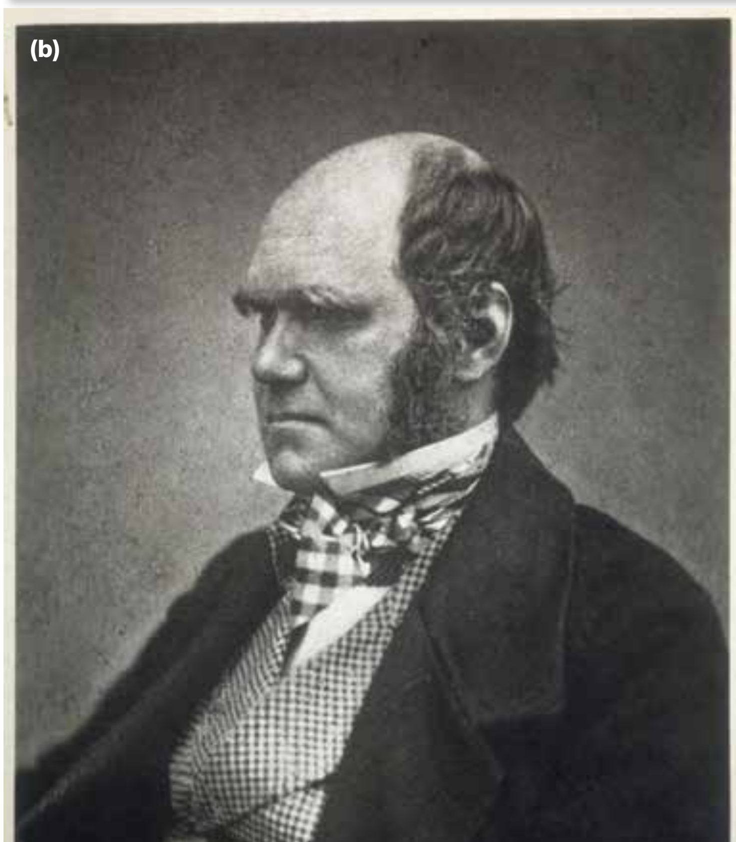



**FIGURE 11.1** **(a)** Copy of page 36 from Charles Darwin's Notebook B, written in 1837–38. On this page, Darwin drew a branching tree of life, showing an ancestral form, denoted by the numeral 1, and the groups descended from it. Four groups of living species, A, B, C and D, are shown (with a cross bar at the end of their lines). Other groups that became extinct are shown without a cross bar at the end of their lines. **(b)** Charles Darwin (1809–1882). In this chapter, we will examine how the relatedness between species can be identified and how the degree of relatedness can be shown using phylogenetic trees.



# Being related

In 2013, samples of mitochondrial DNA (mtDNA) from two sources were tested to see if they matched. One of the mtDNA samples came from Michael Ibsen, a London-based furniture maker. The second sample came from teeth recovered from a skeleton found in 2012 in Leicester under a carpark that in medieval times had been the site of the Greyfriars Church.

What was the purpose of seeing if a match existed between a London furniture maker and a medieval skeleton from a carpark in Leicester? If a match were to be found, this test result would be of historical significance.

Earlier evidence *strongly indicated* that the skeleton might be that of Richard III of England, who died in the Battle of Boswell on 22 August 1485. Richard's body was buried by Franciscan monks in their Greyfriars Church, which was later destroyed.

The evidence in support of this *possible* identification came from a detailed study of the skeleton that revealed:

(i) the bones were those of a male aged between his mid-twenties and mid-thirties — Richard was 32 years old when he was killed in the Battle of Bosworth

(ii) a skeletal abnormality, scoliosis, was detected, which causes a curvature of the spine — in life, Richard III had been described as being a 'hunchback'

(iii) the injuries on the skeleton matched those reported to have been suffered by Richard III in the Battle of Bosworth — some of the battle wounds on the skull were matched to those caused by particular medieval weapons

(iv) carbon-14 dating placed the age of the skeleton within the range of 1450–1540 AD.

Using a model of the skull from the skeleton found in the carpark, a facial reconstruction was prepared by a forensic expert (see figure 11.2).

**FIGURE 11.2** Facial reconstruction using a model of the skull that was part of a skeleton found in a Leicester carpark. This skeleton was shown to be that of Richard III of England.

An essential step was to find a relative of Richard III through his maternal line of descent and to test for a match between the skeletal mtDNA and that of such a relative. The test result would either support or disprove the conclusion that the skeleton was that of Richard III of England. Michael Ibsen, the furniture maker of London, was shown from family records to be a seventeenth-generation nephew of Richard III of England through an unbroken matrilineal line.

Recombination of chromosomal DNA is discussed in *Nature of Biology Book 1 Fifth edition*, page 451.

**Why use mitochondrial DNA?** Mitochondrial DNA samples were used for this matching because, unlike nuclear/chromosomal DNA, it can be recovered from ancient skeletal remains such as teeth and bones. Mitochondrial DNA passes from mother to children, generation after generation, unaltered by the recombination events that occur with chromosomal DNA in meiosis.

Mitochondrial DNA contains a non-coding region known as the D-loop. Within the D-loop are two regions of DNA that undergo mutational change at a much higher rate than the rest of the molecule. These two regions within the non-coding region are denoted as HVR1 and HVR2 (HVR = hypervariable region). Over the 200 000 years or so of the existence of modern humans (*Homo sapiens*), many mutations have occurred in these hypervariable regions. As a result, mtDNA sequences differ between populations and between individuals.

The particular mtDNA sequence of the D-loop of each person is called that person's haplotype. The various haplotypes worldwide fall into a number of large clusters known as **haplogroups**. Haplogroups are denoted by capital letters; for example, we can talk about haplogroup H, haplogroup T and haplogroup U. Figure 11.3 shows the major human haplogroups.

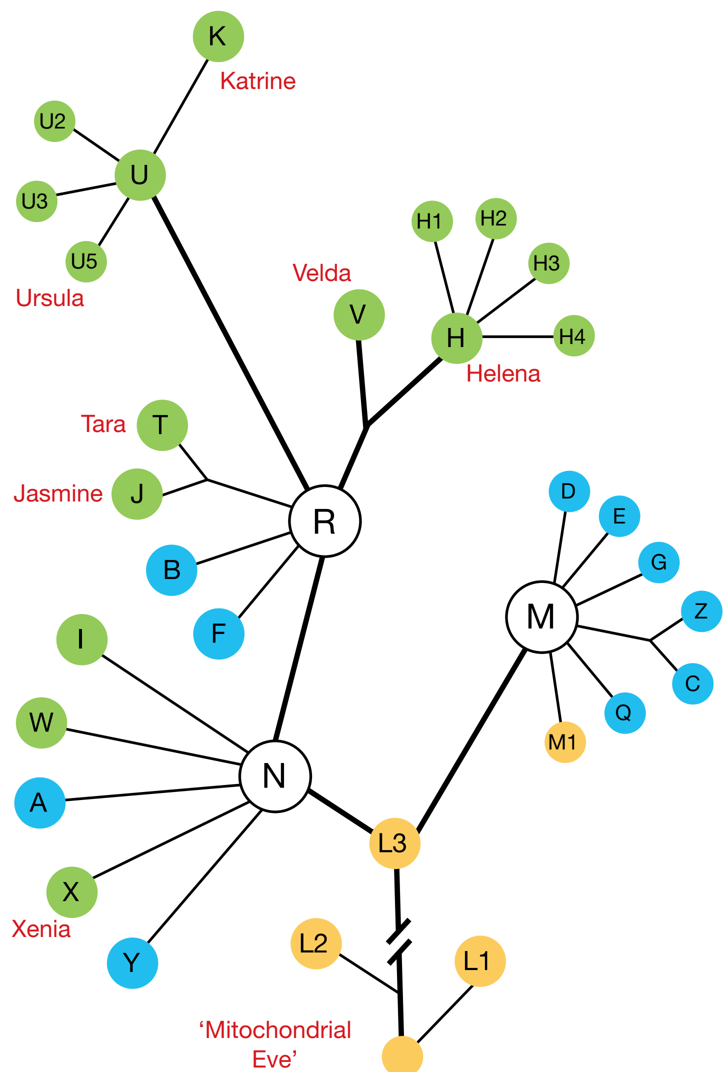



**FIGURE 11.3** Simplified version showing the evolution of and relationships between the major human haplogroups. The earliest human haplogroup is identified as that of 'Mitochondrial Eve'. In his book *The Seven Daughters of Eve*, Professor Bryan Sykes gave names to the founding females of some of the major haplogroups, and these names are shown. Note the three so-called superhaplogroups — M, N and R — from each of which other haplogroups have arisen by mutational events.

**What were the results of the test?** Michael Ibsen's mtDNA sample showed that he belonged to haplogroup J, which is not all that rare because about 17 per cent of the English population belong to haplogroup J. Michael's specific haplotype was identified as haplotype J1c2c. Only about 1.5 per cent of people with haplogroup J have this particular haplotype, so it is relatively rare. What about the mtDNA from the teeth in the carpark skeleton? On 3 February 2013, it was confirmed that the skeletal DNA was haplotype J1c2c. It was a match to Michael Ibsen's mtDNA! The probability of this happening by chance is about 1 in 500.

The following day, on 4 February 2013, the University of Leicester confirmed that the carpark skeleton was that of Richard III of England. The mtDNA match was the final piece of evidence that supported this conclusion. The remains were reburied in Leicester Cathedral in March 2015 in a coffin made by Michael Ibsen (see figure 11.4).

**FIGURE 11.4** The oak coffin made by Michael Ibsen in which the remains of Richard III were carried on a horse-drawn cart through the streets of Leicester en route to Leicester Cathedral, where Richard III was re-buried on 26 March 2015.

## mtDNA as a tool to identify relationships

We have seen that mitochondrial DNA is a valuable tool in identifying relationships between members of one species, such as occurred with Michael Ibsen and Richard III, especially when they are separated in time.

Studies of the sequence of all or part of the mtDNA genome, and mapping of restriction sites within that genome have been used for a variety of other purposes, including identifying relatedness of different species, in particular, where ancient DNA of extinct species is involved. Important examples of this were mtDNA sequencing of *Homo neanderthalensis* and *Homo denisova* and the assessment of their relatedness to modern humans, *Homo sapiens*.

The use of mtDNA is just one way in which relatedness between species can be identified. In this chapter, we will explore other ways that can be used to estimate the relatedness between different species, including:

- comparing proteins
- comparing DNA (nuclear as well as mitochondrial DNA) and RNA
  - base sequences
- comparing entire genomes.

studyon

Unit 4 | AOS 1 | Topic 3 | Concept 3

**DNA sequencing**
Summary screen and practice questions

Refer to chapter 12 for a discussion of genomic studies of modern humans, Neanderthals and Denisovans.

## What is relatedness?

The millions of different species of plants, animals and microorganisms that live on Earth today are related by descent from common ancestors. What does it mean to be related? How do we decide which species are the most closely related? How do we decide which species branched off from which?

In a biological sense, relatedness refers to how recently species split from a common ancestor. So, we may ask the question: *Is species A more closely related* *to species B or C or D?* The answer cannot come by comparing the similarity of habitat or the way or life (niche) or even similarity in appearance. (As we saw in chapter 10 (see page 492), convergent evolution can produce similarities in remotely related species.) The answer comes from re-phrasing the question: *Does species A share a more recent common ancestor with species B or with C or with D?* Whichever species, B, C or D, shares the more recent common ancestor with species A is the species that is most closely related to A.

Figure 11.5 shows a possible answer.

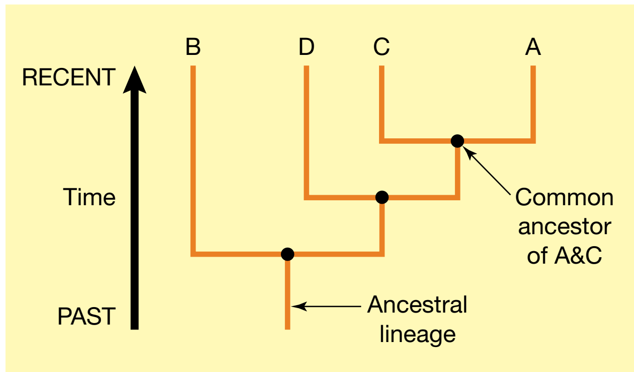



**FIGURE 11.5** Phylogenetic tree showing the evolutionary relationship (relatedness) between four species A, B, C and D. Common ancestors are shown as black dots. Can you identify the common ancestor of species A, C and D? Which pair of species shares the *most recent* common ancestor? They are the most closely related pair of species (and they are A and C).

If species are related by evolution, they should show similarities in their biochemical and genetic composition. Further, a greater degree of similarity should be evident in those species that are more closely related by descent and shared a common ancestor more recently than other species that are more distantly related.

Figure 11.6 shows two pairs of related species. Note that species L and M shared a common ancestor more recently than species J and K. This means that the gene pools of species J and K have been separated for a much longer period than has been the case for species L and M. As a result, more time has been available for changes in the DNA of species J and K to accumulate.

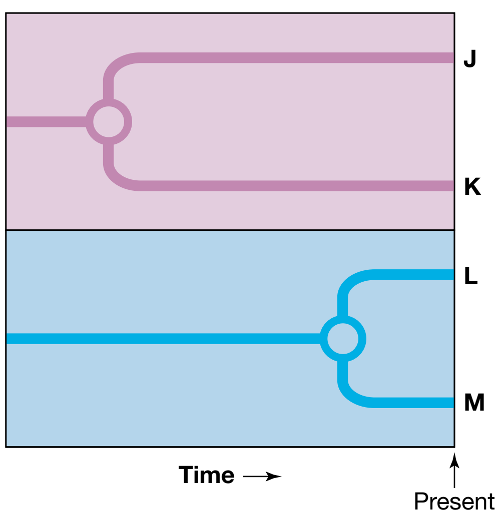



**FIGURE 11.6** Evolutionary history of four species. The most recent common ancestor of each pair of species is shown by an open circle. The horizontal scale denotes geologic time. Horizontal lines indicate the evolutionary path leading to the four modern species, not the time of existence of each modern species.

Table 11.1 shows the expected results of comparative biochemical and genetic studies for the J–K and the L–M pairs of species.

**TABLE 11.1** Comparisons between pairs of related species that differ in the time since they shared a common ancestor. (The species are those denoted in figure 11.6.)

| Molecular study | Expected result |
|---|---|
| Amino acid sequence studies of corresponding proteins | More differences in the amino acid sequences of homologous proteins from species J and K than from species L and M |
| DNA–DNA hybridisation studies | Less complementarity and lower melting temperature ($T_m$) for mixed DNA from species J and K<br>Higher complementarity and higher $T_m$ for mixed DNA from species L and M |
| Base sequence studies of nuclear DNA or mtDNA | More differences in the base sequences of comparable DNA from species J and K than from species L and M |
| Cytogenetic studies of chromosomes | Less similarity in banding pattern of chromosomes from J and K than from species L and M<br>More rearrangements and other structural changes between the chromosomes of species J and K than of species L and M |
| Chromosome 'painting' (spectral karyotypes) | Fewer matching segments identified between the chromosomes of species J and K than of species L and M<br>Fewer chromosomal rearrangements between species L and M than between J and K |

# Comparing proteins

The proteins of all organisms — whether they are jellyfish, tomatoes, lobsters, ferns or people — are composed of the same set of 20 amino acid building blocks. Likewise, the genetic code that carries the information for making proteins is essentially the same in all organisms. These observations are consistent with a common evolutionary ancestry for all living organisms.

Proteins from different species can be compared in terms of their amino acid sequences. **Species that are more closely related are expected to have fewer differences in the amino acid sequences of their corresponding proteins than species that are more distantly related**. Why? The longer the periods since two species diverged from their last common ancestor, the more time has been available for changes to occur in a protein present in both species. It is reasonable to conclude that, the more differences that are observed in the corresponding protein in two species, the further back their divergence in time.

Haemoglobin is composed of up to four chains: two identical chains (alpha chains) consisting of 141 amino acids, and two other identical chains (beta chains) consisting of 146 amino acids. The amino acid sequences of the haemoglobin chains have been identified for many mammals, birds, reptiles, amphibians, fish and invertebrates. Table 11.2 shows the number of differences in the amino acid sequence of the beta chain of haemoglobin between humans and several vertebrate species (see also figure 11.7). Based on these data, it is possible to estimate the relationships among the various species. Data in this table support the conclusion that gorillas are more closely related to humans than are Rhesus monkeys or mice. Degrees of evolutionary relationships identified through this means are in agreement with relationships inferred from fossil evidence and from structural comparisons.

**TABLE 11.2** Number of differences in the amino acid sequence of the beta chain of haemoglobin of various vertebrates compared to the human beta chain of haemoglobin. What pattern is apparent? Which group of species shows the closest relationship with humans?

| Vertebrate | No. of differences |
|---|---|
| chimp (*Pan* spp.) | 0 |
| gorilla (*Gorilla gorilla*) | 1 |
| gibbon (*Hylobates* spp.) | 2 |
| Rhesus monkey (*Macaca mullata*) | 8 |
| dog (*Canis familiaris*) | 15 |
| mouse (*Mus musculus*) | 27 |
| kangaroo (*Macropus* spp.) | 38 |
| chicken (*Gallus gallus*) | 45 |
| frog (*Rana* spp.) | 67 |
| lamprey (*Petromyzon marinus*) | 125 |
| mollusc (sea slug) | 127 |

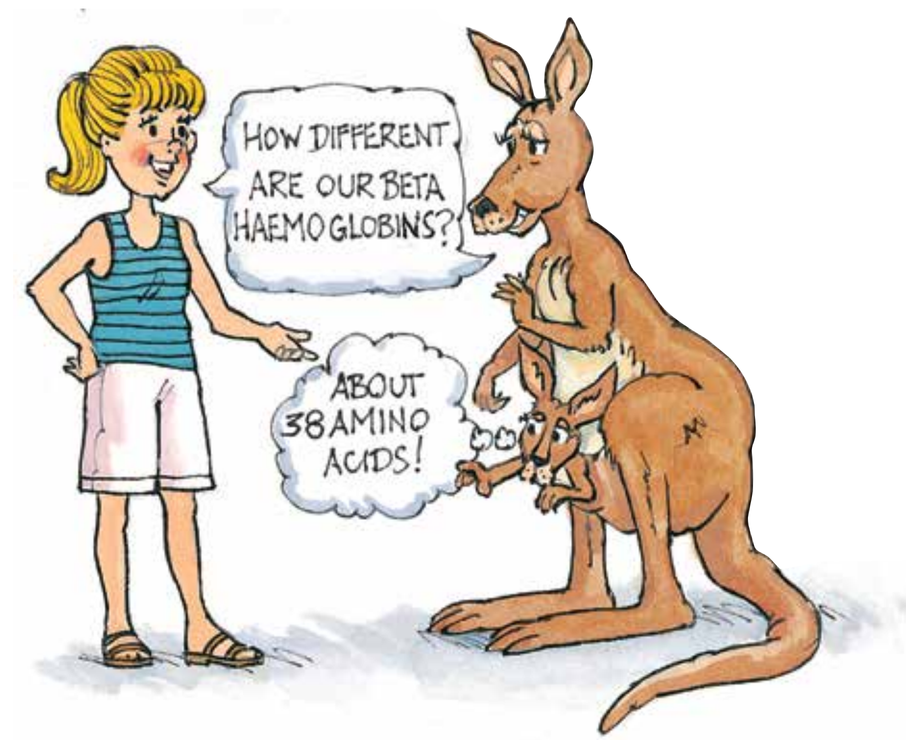



**FIGURE 11.7**

Likewise, if other proteins are examined, a similar picture of relationships between organisms can be inferred. An enzyme found in organisms from all of the five Kingdoms is known as *cytochrome c*. Human *cytochrome c* has been compared with that from other organisms (see table 11.3). Note that, compared with human *cytochrome c*, a single amino acid change is found in the corresponding enzyme from a Rhesus monkey, while 13 changes are found in that from a bird.

**TABLE 11.3** Comparison of the sequence of amino acid sub-units in *cytochrome c* from various species. *Cytochrome c* also occurs in yeast cells. If compared with human *cytochrome c*, what would you predict?

| Organism | Number of changes in amino acid sequence relative to the human *cytochrome c* |
|---|---|
| human | — |
| chimp | 0 |
| Rhesus monkey | 1 |
| whale | 7 |
| rabbit | 9 |
| chicken | 13 |
| fish (tuna) | 22 |
| wheat | 37 |

Species that show fewer differences in their amino acid sequences can be inferred to be more closely related than species showing greater differences. The data in table 11.3 above indicate that, of the species listed, chimps are most closely related to humans.

# Comparing DNA

studyon

Unit 4 | AOS 1 | Topic 3 | Concept 1

**Molecular studies–DNA**
Summary screen and practice questions

DNA sequences have been described as 'documents of evolutionary history'. Comparisons of DNA from different species may be made in several ways:

1. direct comparison of DNA base sequences
2. comparing whole genomes
3. DNA hybridisation
4. comparing karyotypes.

## Comparing DNA base sequences

studyon

Unit 4 | AOS 1 | Topic 3 | Concept 4

**Molecular studies–amino acid sequences**
Summary screen and practice questions

DNA molecules consist of a series of base pairs (bp) that form a base sequence (refer back to chapter 2, page 41).

If evolution has occurred, we can predict that species that are closely related by evolutionary descent will show more similarities in the base sequences of their common genes. Hence, direct comparisons of the DNA sequence of genes in different species can also be used to infer evolutionary relationships. For example, haemoglobin genes are present in all mammals. Sequences have been identified for the approximately 17 000 bases in this segment of DNA in human beings and other animals. The results show that these sequences are most similar between humans and chimpanzees.

Table 11.4 shows the DNA sequences from part of a haemoglobin gene.

**TABLE 11.4** DNA sequences from a segment of a haemoglobin gene from four mammalian species. Differences between the human DNA sequence and those of other species are shown by coloured letters. The dash (-) is used to keep the sequences aligned. Note that there are two differences between the human and the sequences of some other primates (orang-utan and monkey) but that there are more between the human and the rabbit DNA sequences. Why?

| Species | DNA sequence of part of a haemoglobin gene |
|---|---|
| human | TGACAAGAACA - GTTAGAG - TGTCCGAGGACCAACAGATGGGTACCTGGGTCCCAAGAAACTG |
| orang-utan | TCACGAGAACA - GTTAGAG - TGTCCGAGGACCAACAGATGGGTACCTGGGTCTCCAAGAAACTG |
| Rhesus monkey | TGACGAGAACA A GTTAGAG - TGTCCGAGGACCAACAGATGGGTACCTGGGTCTCCAAGAAACTG |
| rabbit | TGGTGATAACA A GACAGAG A TATCCGAGGACCAGCAGATAGGAACCTGGGTCTCTAAGAAGCTA |

If evolution has occurred, we can predict that many of the genes present in an ancestral species will also be present in the species that evolve from it. If a particular gene or DNA sequence is present in members of a related group, that DNA sequence or that gene is said to be conserved among those species. **Gene conservation** does not mean that the DNA sequences of the genes in different related species will be identical, since mutations will occur after their evolutionary lines diverged. **The DNA sequences of conserved genes will, however, be very similar**, such as the haemoglobin gene shown in table 11.4. It is reasonable to suggest that, when a DNA sequence (or a gene) is conserved, it has an important function.

Let's use a simple example to show how we can infer evolutionary relationships between species, based on differences in DNA sequences. Consider five modern species J to N. Table 11.5 shows the differences in a corresponding DNA segment in these species.

**TABLE 11.5** Base sequences of a common segment of DNA of five related species. Differences in DNA (base substitutions) are indicated by arrows.

| Species | DNA sequence |
|---|---|
| J | TCCAGCTCGT GCCTCGATGA AGACTAAGTG ATACCATAAA GACT |
| K | TCGAGCTCGT GCATCGATGA CGACTAAGTG ATACCATAAA GACT |
| L | TCCAGCTCGT GTATCGATGA CGACTAAGTG ATACCAAAAA GACT |
| M | TCCAGCTCGT GTATCGATGA CGACTAAGTG ATACCAAAAA GACT |
| N | TCCAGCTCGT GTATCGATGA CGACTACGTG ATACCATAAA GACT |
| | ↑ ↑ ↑ ↑ ↑ ↑ ↑ |
| | 1 2 3 4 5 6 7 |

Using the data from table 11.5, it is possible to construct a phylogenetic tree that indicates the relatedness of the different species (see figure 11.8).

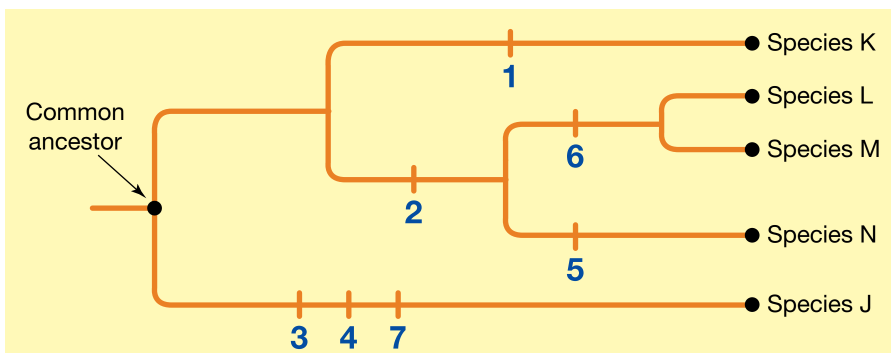



**FIGURE 11.8** Inferred evolutionary relationships of species J, K, L, M and N. The horizontal lines denote ancestry and do not denote the time of existence of the modern species. Mutations 3, 4 and 7 did not necessarily occur in that order.

**study on**

Unit 4
AOS 1
Topic 3
Concept 6

**Genome phylogeny**
Summary screen and practice questions

**study on**

Unit 4
AOS 1
Topic 3
Concept 6

**Do more**
Genome phylogeny

## Comparing whole genomes

It is now possible to compare the genomes of different organisms — a field of study known as **comparative genomics**. These comparisons can help to clarify the evolutionary history of species. Because the amounts of data are so large (for example, the human genome contains 3000 million base pairs), computer technology is necessary for these studies. Information gained from comparative genomics has applications in medicine and industry.

The genome of each species contains DNA sequences and distinctive features that have been conserved over millions of years of evolution. Because living species have evolved from common ancestors, the genomes of related species exhibit similarities. The more recent the divergence of two related species from a common ancestor, the greater the number of conserved DNA sequences and of their arrangement within the genome.

**By comparing the genomes of living species, it is possible to:**

- identify the degree of relationship between different species from the fraction of genes shared between their genomes
- make inferences about the phylogeny or evolutionary history of a species.

For example, human DNA was compared with DNA from 12 other vertebrate species (see figure 11.9). Note that DNA from the nonhuman species can be aligned or matched to the human DNA. Note also that the percentage of alignment varies, with the percentage being greatest between the most closely related species (as determined by other techniques). In this case, the greatest degree of similarity occurred in the human–chimp comparison.

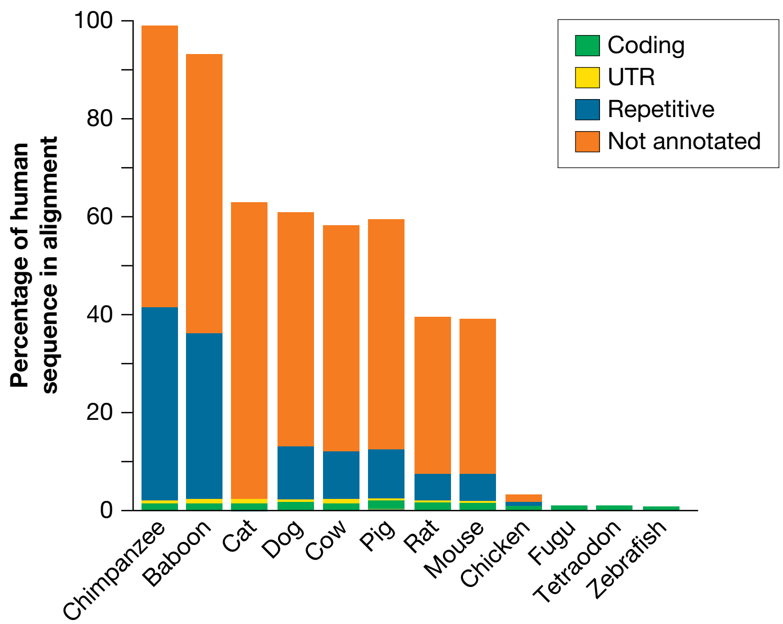



**FIGURE 11.9** Percentage of human DNA sequences that can be aligned in 12 vertebrate species. The different colours denote the DNA in different categories: coding DNA that is the known part of a structural gene, untranslated regions of DNA (UTR), repetitive DNA, and DNA whose function has not been identified. What alignment would be expected in a human–human comparison?

## Comparing DNA by hybridisation

study on
Unit 4 | AOS 1 | Topic 3 | Concept 2
**DNA hybridisation**
Summary screen and practice questions

All kinds of organisms encode their genetic instructions in DNA. DNA is normally double stranded but, if its temperature is raised, DNA dissociates into single strands. As the temperature falls, the single strands pair with their complementary strands to re-form double-stranded DNA.

This pairing tendency of DNA can be used to compare the similarity of the DNA of different species using the technique of **DNA–DNA hybridisation**:

1. The DNA is extracted from cells of the two species to be compared, purified and cut into fragments about 500 bp long.
2. The repeated sequences are removed, leaving the unique genetic information of the DNA gene sequences of the two species.
3. The DNA fragments of the two species are made single stranded, mixed and allowed to pair. Some pairing will occur between the single strands from the two species. This pairing is more complete and stronger when the DNA from the two species is more similar (see figure 11.10).
4. A measure of similarity is then obtained by heating the interspecific DNA double-stranded fragments and recording the temperature at which half of them become single stranded again. This temperature is known as the **melting temperature ($T_m$)**. When the degree of complementary pairing is lower, fewer hydrogen bonds must be broken and the $T_m$ is lower.
5. **A difference of one degree Celsius (1 °C) in melting temperature corresponds to a one per cent difference in DNA sequences.**

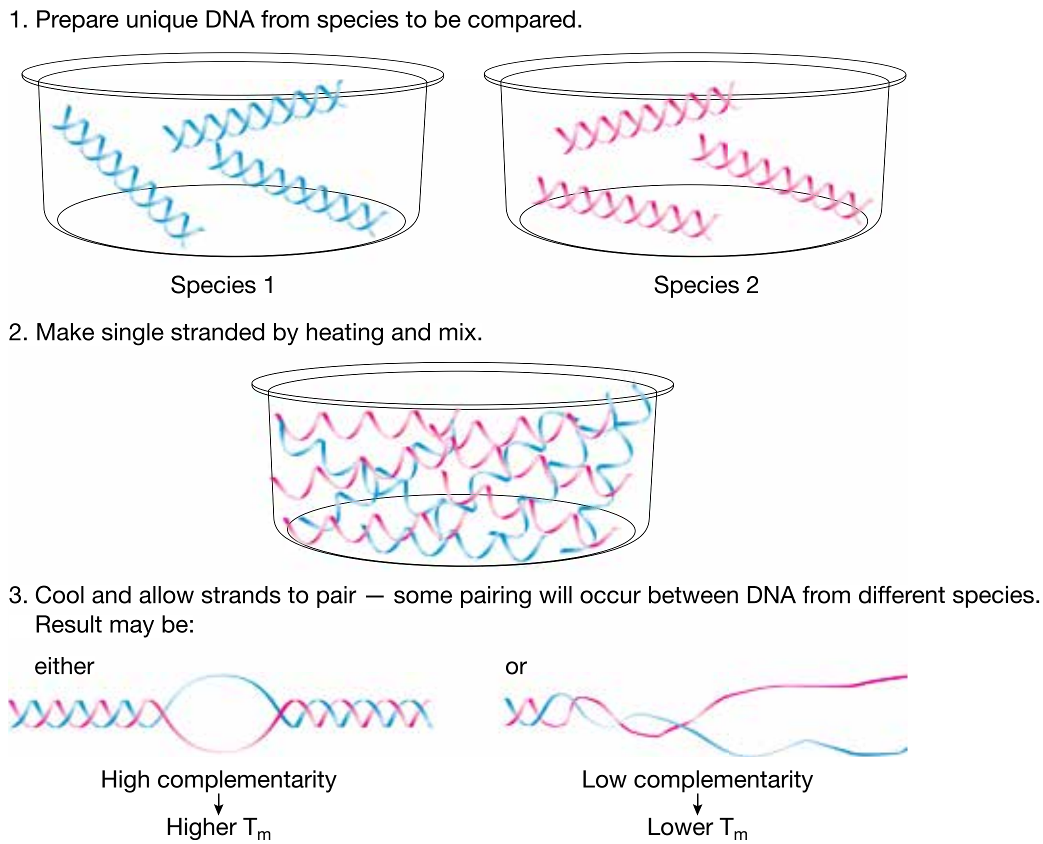



**FIGURE 11.10** DNA from two species may pair with a high degree of complementarity or with low complementarity. The strength of the pairing is greater when complementarity is higher because more hydrogen bonds form between the two DNA strands. $T_m$ = melting temperature.

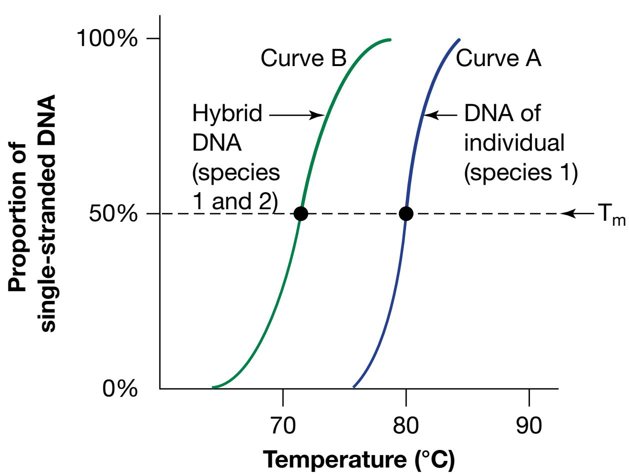



**FIGURE 11.11** Melting temperatures, $T_m$, for double-stranded DNA from an individual (curve A) compared to those for hybrid DNA comprising DNA from two different species (curve B)

The technique of DNA–DNA hybridisation provides a measure of the similarity of the genetic material of two species and gives an estimate of the genetic distance between them: that is, the time that has elapsed since the two species shared a common ancestor. The longer this period, the more time there has been for changes to occur in their DNA. Two species that diverged from a common ancestor a shorter time ago show a higher degree of pairing and have a higher $T_m$ than two species that diverged from a common ancestor a longer time ago.

To compare the DNA of two species using the DNA–DNA hybridisation technique, two melting temperatures are measured. The first $T_m$ is that of double-stranded DNA of an individual of one species (say, human). The second $T_m$ is double-stranded 'hybrid' DNA formed by combining single-stranded DNA from individuals from two different species (say, human–chimp). Compared with double-stranded DNA from the one organism, the strands of the hybrid DNA will not pair so well. As a result, fewer hydrogen bonds need to be broken to dissociate this hybrid into the single-stranded form (see figure 11.11).

**The smaller the reduction in $T_m$ for hybrid DNA from two species, the more closely related they are by evolution and the more recently they shared a common ancestor**. Table 11.6 shows the lowering of melting temperature for hybrid DNA from various living primates.

**TABLE 11.6** Reduction in melting temperature, $T_m$, when genomic DNA from different organisms, either the same species or different species, is mixed. The lowering is measured in relation to the $T_m$ for the DNA from one individual of a species.

| | Human | Chimp | Gorilla | Orang-utan | Gibbon |
|---|---|---|---|---|---|
| **Human** | 0.3 | | | | |
| **Chimp** | 1.7 | 0.3 | | | |
| **Gorilla** | 2.3 | 2.3 | 0.3 | | |
| **Orang-utan** | 3.6 | 3.6 | 3.5 | 0.3 | |
| **Gibbon** | 4.8 | 4.8 | 4.7 | 4.9 | 0.3 |

Data from DNA hybridisation studies can be used to infer the degree of evolutionary relationship between different species. For example, from the data in table 11.6, the evolutionary relationship between the primates can be inferred (see figure 11.12). By calibrating this tree against fossil evidence, it is possible to identify a time scale for this tree.

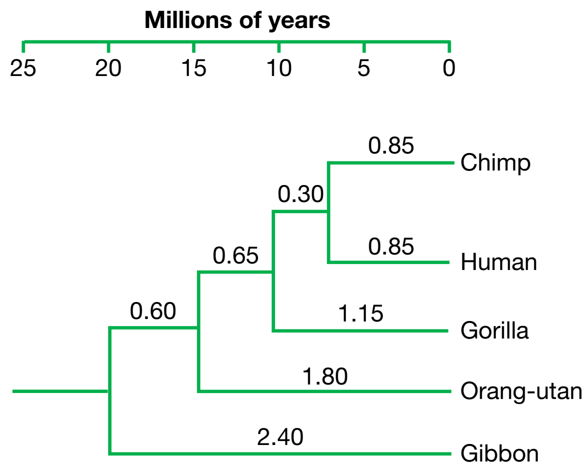



**FIGURE 11.12** Tree showing inferred evolutionary relationships between a group of primates based on DNA hybridisation evidence. For each pair of species, the difference in melting temperatures is split and shared between the two branches. Lengths of the lines are proportional to the time since the species last shared a common ancestor. The time scale is calibrated using fossil evidence that the orang-utan line diverged about 15 million years ago.

## Comparing chromosomes

The DNA of eukaryotic organisms is organised into chromosomes that are visible during cell division by mitosis and meiosis.

Chromosomes from different species can be compared in different ways and similarities identified by:

- examination of the **chromosome banding** patterns
- use of **chromosome painting**.

The chromosomes of various species can be identified by a three-letter code that consists of the first letter of the name of the genus plus the first two letters of the specific name. So, the human number-5 chromosome is denoted as HSA 5 and the common chimp number-7 chromosome is denoted as PTR 7. How would the number-11 chromosome of the pygmy chimp be denoted: PTR 11 or PPA 11?

**eBook*plus***

**Interactivity**
Mitosis
**int-3027**

Table 11.7 shows the diploid number for several ape species. Notice that the diploid number in humans is 46 while that in the other great apes (chimps, gorilla, orang-utan) is 48. How and when did this change occur? We can answer this by looking at the banding patterns of the chromosomes of these species.

**TABLE 11.7** Diploid number of chromosomes from various ape species.

| Species | Diploid chromosome number (2*n*) |
|---|---|
| human being, *Homo sapiens* | 46 |
| common chimp, *Pan troglodytes* | 48 |
| bonobo or pygmy chimp, *Pan paniscus* | 48 |
| gorilla, *Gorilla gorilla* | 48 |
| orang-utan, *Pongo pygmaeus* | 48 |

## Banding pattern of chromosomes

The banding pattern on the various human chromosomes can be matched to patterns on regions of chromosomes of other great ape species (see figure 11.13). This observation strongly suggests that the chromosomes of humans, chimps and gorillas share major **homologous** regions.

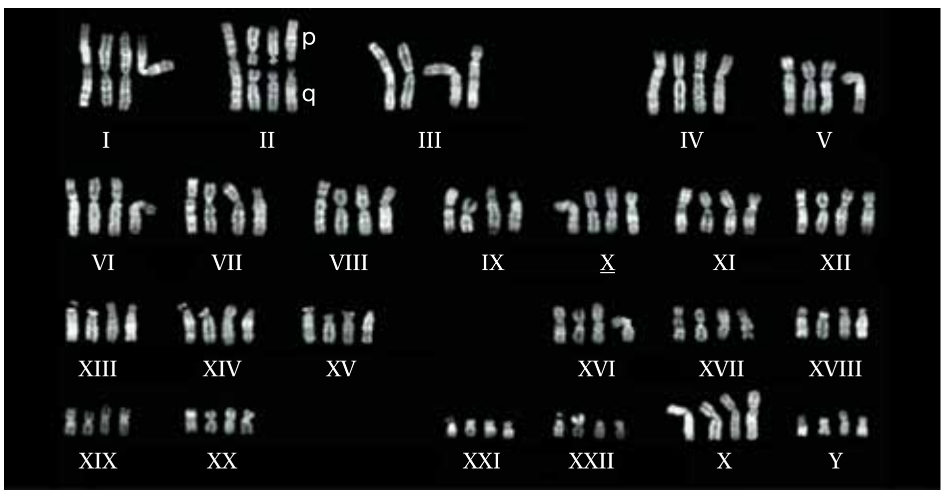



**FIGURE 11.13** Chromosome sets or karyotypes of four species of the great apes: humans, chimps, gorillas and orang-utans. In each case, the human chromosome is at the left, and alongside, in order, are the matching chromosomes of the chimp, the gorilla and the orang-utan. (Image courtesy of Dr Mariano Rocchi.)

**eBook*plus***

**Interactivity**
Mitosis and meiosis
**int-3028**

Figure 11.14 compares the banding patterns on four human chromosomes (numbers 2, 3, 4 and X) with the matching chromosomes of the chimp, gorilla and orang-utan. Note that the banding pattern of the human number-2 chromosome corresponds to two shorter chimp chromosomes, number-2A and number-2B, and to the same chromosomes in the case of the gorilla and the orang-utan. This observation explains the change from $2n = 48$ in the great apes to $2n = 46$ in humans. The common ancestor of apes and hominins had a diploid number of 48 chromosomes. We can speculate that, after the hominin line separated from this common ancestor, a chromosome **fusion** occurred between two nonhomologous chromosomes 2A and 2B so that two shorter chromosomes were replaced by one longer chromosome. As a result of this fusion, the diploid number of chromosomes was reduced from 48 to 46 in humans.

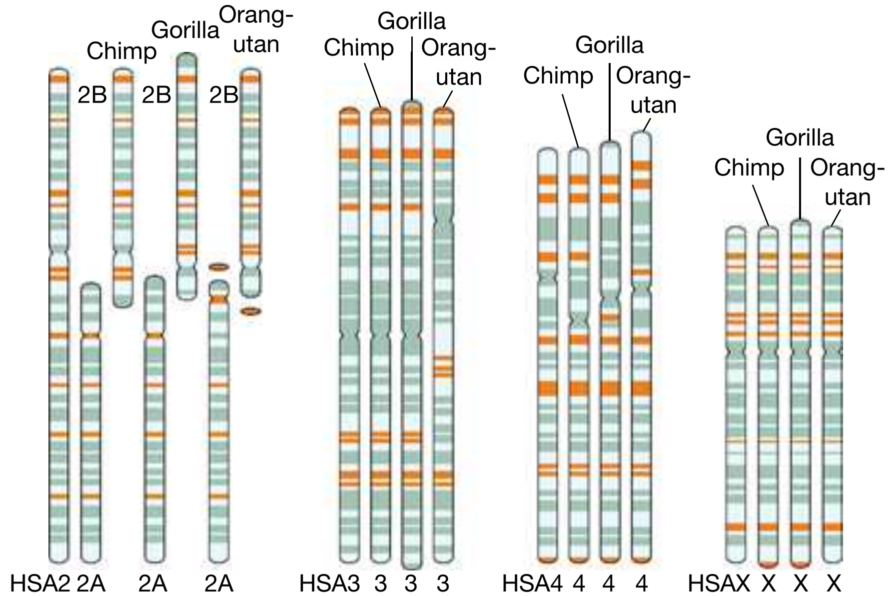



**FIGURE 11.14** Comparison of the banding patterns of some human chromosomes with those of the chimp, gorilla and orang-utan. For the banding pattern on the human number-2 chromosome, note that it matches that of two short chromosomes of the chimp, the gorilla and the orang-utan.

## Colour-coded chromosomes

A method of 'painting' chromosomes, developed in 1996 and known as multicolour spectral karyotyping, can reveal homologous chromosomes or regions of chromosomes in different species. For example, chromosome painting has shown that one group of genes on the human number-6 chromosome also exists on the number-5 chromosome of the chimpanzee, the B2 chromosome of the domestic cat, the number-7 chromosome of the pig and the number-23 chromosome of the cow. This group of genes controls aspects of the immune response and is present in all mammals. The painting method uses probes, linked to different numbers and combinations of fluorescent dyes, each of which binds to specific chromosomes. As a result, each different chromosome fluoresces with a unique wavelength signal. Figure 11.15 shows a human spectral karyotype.

It is now possible to purchase the probe 'paint' for the DNA of each specific chromosome. Each fluorescent 'paint' of a particular colour binds to one specific human chromosome.

**ODD FACT**

The orang-utan number-21 chromosome (PPY 21) corresponds in its banding pattern to the human number-21 chromosome (HSA 21). An orang-utan was found with three copies of its number-21 chromosome. This animal showed clinical signs corresponding to the human condition known as Down syndrome, which is due to trisomy-21.

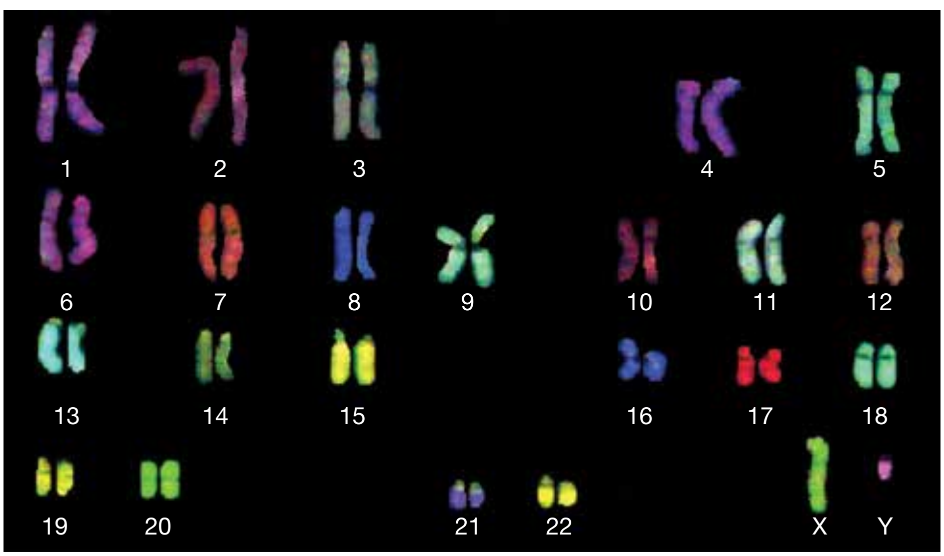



**FIGURE 11.15** A human spectral karyotype showing the distinctive colours of each chromosome. What colour is the paint for the number-3 chromosome? (Image courtesy of the National Human Genome Research Institute.)

**ODD FACT**

Chromosome painting has revealed that large regions of the chromosomes of sheep, goats and cattle carry matching genes.

As well as 'paints' for human chromosomes, paints can be prepared for other species. Figure 11.16 shows the chromosomes of the tammar wallaby (*Macropus eugenii*) treated with paints for three chromosomes from the common brush-tailed possum (*Trichosurus vulpecula*). Because these two marsupial species have both evolved from a common ancestor, it is not unexpected that segments matching the possum chromosomes can be identified on the wallaby chromosomes (see figure 11.16).

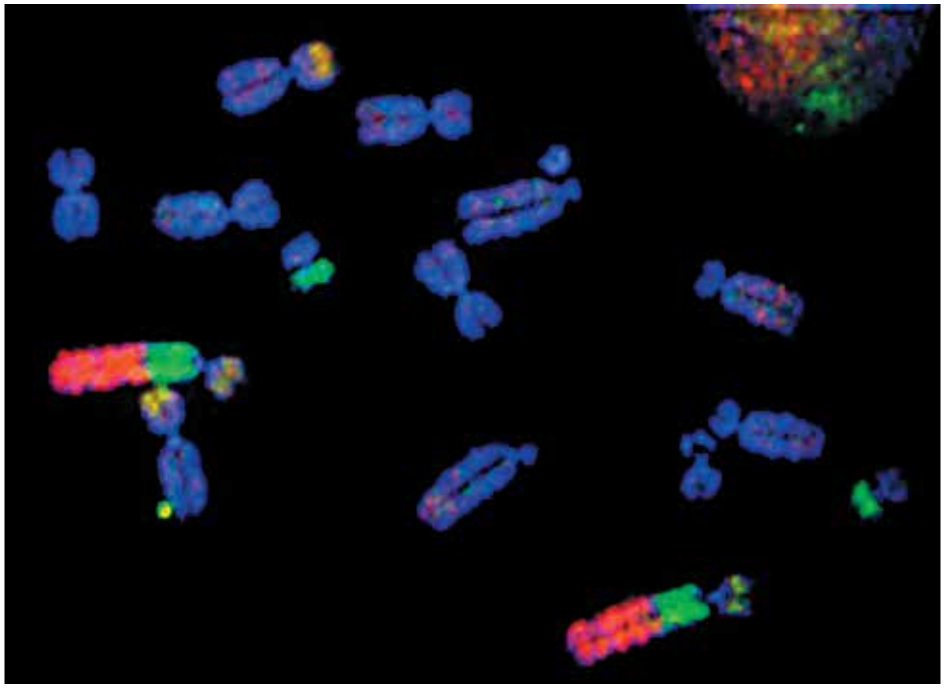


**FIGURE 11.16** Chromosomes of the tammar wallaby painted with chromosome paints for chromosomes of the common brush-tailed possum:
pink = chromosome 4
green = chromosome 7
brown = chromosome 9.
What conclusion can you draw about the wallaby chromosome that is pink and green? (Image courtesy of Dr Willem Rens.)

When human chromosome paints are applied to chimp chromosomes, almost every different chimp chromosome is painted with a specific colour. For example, the specific paint for the human number-1 chromosome (HSA 1) binds only to the chimp number-1 chromosome (PTR 1). Likewise, the specific paint for the human number-6 chromosome (HSA 6) binds only to the chimp number-6 chromosome (PTR 6). A similar situation occurs for the other chimp chromosomes. These results indicate that each chimp chromosome is homologous to one specific human chromosome. This finding confirms the very high degree of homology between single human and chimp chromosomes that was indicated by the banding patterns.

One exception is that the 'paint' for the human number-2 chromosome (HSA 2) binds to two different chimp chromosomes, PTR 2A and PTR 2B. This result confirms that the human number-2 chromosome arose during hominid evolution as a result of fusion of two smaller chromosomes (see figure 11.17).

In contrast, when gibbon chromosomes are 'painted' with human DNA sequences, different colours appear on many individual gibbon chromosomes. This indicates that they contain regions that are homologous to several different human chromosomes and indicates a greater evolutionary separation between gibbons and humans than that between the African great apes and humans.

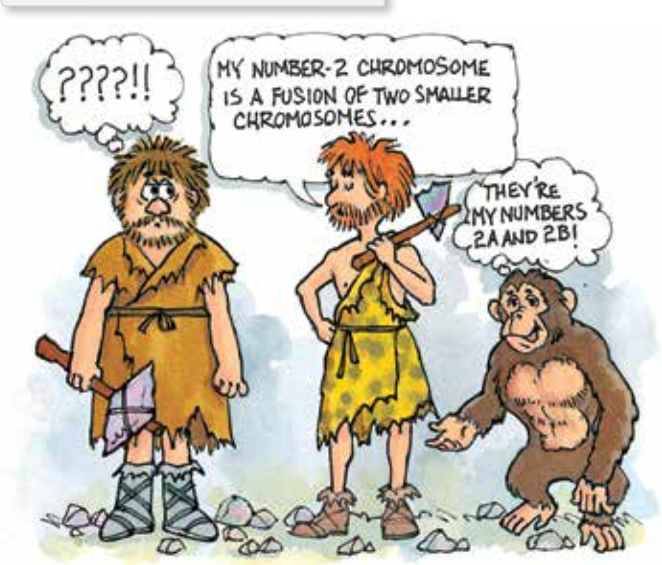



**FIGURE 11.17** During the evolution of the hominins, a fusion between two chromosomes reduced the diploid number ($2n$) from 48 to 46.

In summary, chromosome painting can show the degree of relatedness of species by:

- revealing homologous regions in the chromosomes of different species
- providing evidence about how particular karyotypes evolved.

In addition, a higher degree of colour matching between the chromosomes of two species indicates a shorter period since they diverged from a common ancestor as compared to two species that show a lower degree of matching.

## The 'molecular clock' concept

It was recognised that the number of differences in the proteins of two species might indicate the time that had elapsed since these species diverged from their most recent common ancestor. This is the concept of the **molecular clock**.

In 1966, two American biochemists, Vincent Sarich and Allan Wilson, looked at the amino acid sequences of corresponding proteins in related species. They noted that changes in the amino acid sequences appeared to occur at a steady rate. The longer that two species had existed separately, that is, the longer since they shared a common ancestor, the more differences in the amino acid sequences of their corresponding proteins. The molecular clock concept arose from this. It is the concept of estimating the time since two species diverged, based on differences in their biological make-up.

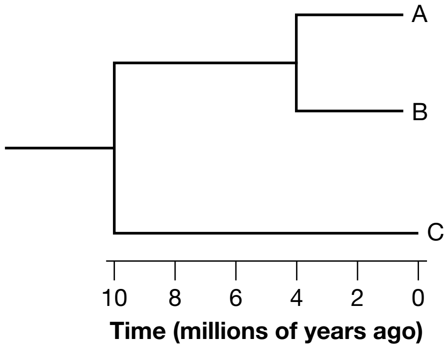



**FIGURE 11.18** Inferred time of divergence of species A, B, and C based on the molecular clock concept.

Assume that a specific protein is estimated to change at the rate of one amino acid per million years. This protein from species A, B and C is compared, and four differences are found between B and A and ten differences between B and C. From these data, we may infer that the divergence of the various species from common ancestors may have been as shown in figure 11.18.

Haemoglobin is an oxygen-carrying protein found in all vertebrate animals. When the alpha chains of the haemoglobin from a range of vertebrate species are compared, differences are observed in the amino acid sequences. When different pairs of vertebrates are compared, the number of differences is seen to vary, with more differences between less closely related species (see table 11.8).

**TABLE 11.8** Comparisons of the percentage differences in the amino acid sequence of the alpha chain of the haemoglobin protein of various vertebrates.

| | Chimp | Elephant | Platypus | Starling | Crocodile | Carp | Shark |
|---|---|---|---|---|---|---|---|
| **Human** | 3 | 18 | 28 | 29 | 33 | 48 | 50 |
| **Chimp** | | 19 | 28 | 30 | 33 | 47 | 52 |
| **Elephant** | | | 32 | 34 | 35 | 44 | 54 |
| **Platypus** | | | | 37 | 36 | 49 | 50 |
| **Starling** | | | | | 33 | 47 | 55 |
| **Crocodile** | | | | | | 49 | 55 |
| **Carp** | | | | | | | 62 |

Observations of these differences led to the idea that DNA mutates and, as a result, a particular protein changes over time at approximately the same rate in each evolutionary line. If this assumption is valid, it would be possible to use the molecular clock to identify evolutionary relationships and to estimate when various modern species last shared a common ancestor. Notice that the crocodile, a reptile, differs from the mammals by about 34 per cent. Notice also that the crocodile differs from the bird (starling) by a similar percentage. This indicates that the evolutionary line that gave rise to the modern crocodiles has been separated from the line that gave rise to the birds for about as long as it has been separated from the mammalian line. In contrast, notice that humans differ from chimps by a much smaller percentage (3%) than from the elephant (18%). From this, it follows that mammals are related and the human line has been separated from the elephant line for a longer period than from the chimp line.

Using these observed differences, the evolutionary relationship between these various species can be inferred, as shown in figure 11.19a. Looking at living organisms can assist us to infer evolutionary relationships, so the present helps us to interpret the past.

### Calibrating and testing the 'clock'

The molecular clock can be calibrated against the fossil record, provided adequate fossils exist. For example, in figure 11.19b, the time scale was inserted based on fossil evidence that the first sharks appeared in the fossil record in the lower part of the Devonian period about 350 Myr ago. It assumes that this is the time required to accumulate an average 53 per cent change in shark protein compared to the same protein in other groups.

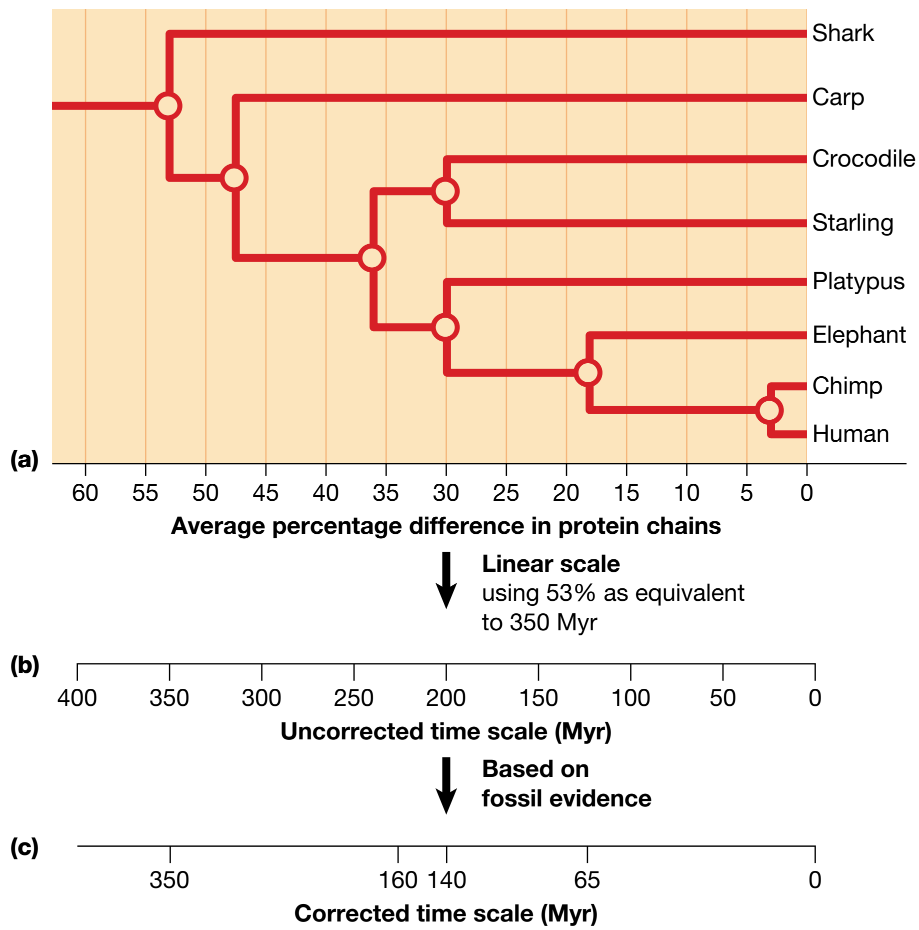



**FIGURE 11.19** **(a)** Evolutionary relationship based on percentage differences in the alpha chain of haemoglobin of various vertebrates. The horizontal lines do not identify the time span of existence of the modern species but identify the time span of the line leading to the modern species. **(b)** Uncorrected time scale based on two points, 0.0 and 53 per cent where the latter is equated to 350 Myr, the time of the first appearance of the shark line in the fossil record. **(c)** Time scale corrected for the fossil record appearance of all groups concerned.

In setting the time scale, it must be recognised that, when the percentage differences between the proteins of various species are large, these differences are underestimates. Why? Because, over long periods of time, some early changes can be reversed by later changes so that the evidence of the change is lost. A mathematical correction is made to take this into account (see figure 11.19c).

To test data from the molecular clock, three species are required:

- two that are closely related (such as two mammals, M1 and M2)
- a third species for reference (R) that diverged before species M1 and M2 diverged, such as a reptile.

The timing indicated by the molecular clock is valid if the difference between M1 and R is similar to the difference between M2 and R. In the case of the alpha haemoglobin protein, the molecular clock passes this test.

## Limitations of the molecular clock

A molecular clock based on percentage differences between corresponding genes or proteins does not keep perfect time and care must be taken in making inferences.

Research has shown that the following limitations can apply to the use of the molecular clock:

- The molecular clock 'ticks' at different rates for different proteins. The alpha haemoglobin protein appears to change at the rate of about one amino acid every five to six Myr, while a particular histone protein changes at the rate of only one amino acid every 5000 Myr. It is not valid to draw conclusions by combining the rates of change in different proteins.
- The rates of change of the same protein can differ in different groups. For example, the rate of change in a protein from plant groups has been shown to be slower than the rate for the same protein in animal groups.

However, the molecular clock remains a valuable technique provided it is used with appropriate care.

### KEY IDEAS

- Species are related through evolution, with the degree of relatedness differing according to when two species last shared a common ancestor.
- Differences in both DNA base sequences and the amino acid sequences of proteins can be used to assess degrees of relationship between species and infer their evolutionary history.
- Reductions in the melting temperatures of double-stranded DNA molecules hybridised from two species can provide a clue to the degree of relatedness of the two species.
- Chromosome painting can reveal homologous regions of chromosomes from related species.
- Chromosome painting is one tool that can be used to compare the genomes of different species.
- Within limits, the molecular clock is a technique for putting a time scale on speciation events.

### QUICK CHECK

1 Identify the following as true or false:
  - **a** The same protein would be expected to mutate at the same rate in different organisms.
  - **b** The smaller the reduction in the melting temperature of hybrid DNA from two species, the more closely related they are.
  - **c** More closely related species can be inferred to have shared a common ancestor more recently than less closely related species.
  - **d** A phylogenetic tree is a diagram that can be used to show the evolutionary relationships between species.
  - **e** The human number-2 chromosome (HSA 2) corresponds to PTR 2A and 2B.

2 Species F and G are related, and species D and E are also related.
  - **a** Identify a molecular technique that you could use to determine which pair is more closely related.
  - **b** Assuming that F and G are the more closely related pair of species, outline the results you would expect using your selected molecular technique.

3 Select the appropriate alternative, and in each case, briefly explain your choice:
  - **a** The longer the time since two species shared a common ancestor, the *<more/less>* similar are the amino acid sequences of a protein common to both species.
  - **b** The common ancestor of the great apes would be expected to have had a diploid chromosome number of *<46/48>*.

studyon

Unit 4 | AOS 1 | Topic 3 | Concept 7

**Phylogenetic trees**
Summary screen and practice questions

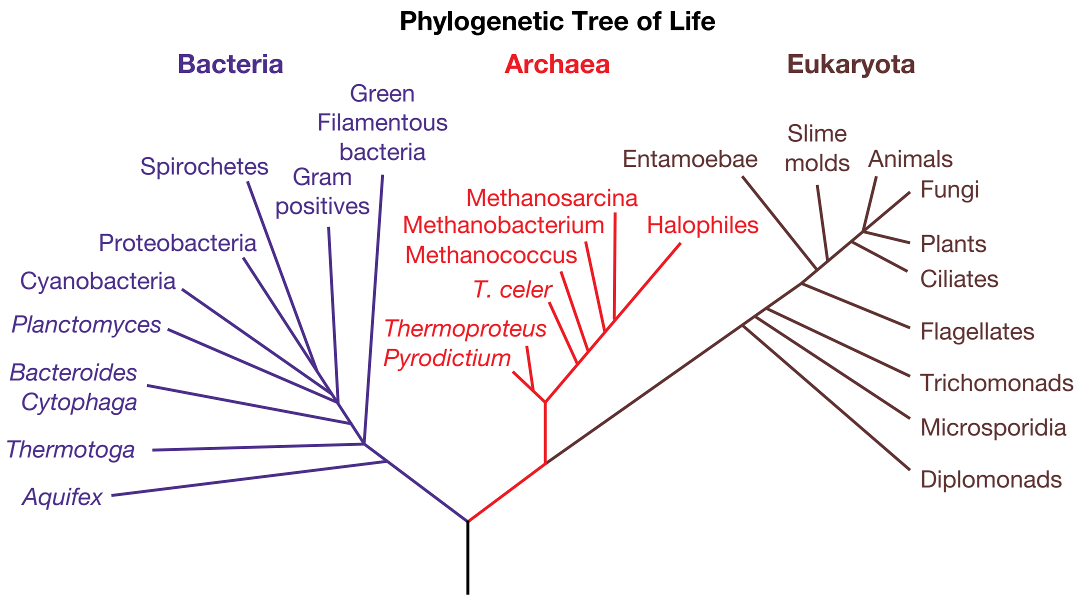
# Showing relatedness

Relationships between species (and larger taxonomic groups) may be shown visually by the use of diagrams termed:

- phylogenetic trees
- cladograms.

After realising that species are not fixed and unchanging, Darwin recognised that, over time, one species could change and that new species could emerge. In his 1837 notebook (refer back to figure 11.1), Darwin wrote '*I think*' and then made a sketch to show one form of life branching to give rise to multiple forms. This sketch shows the evolutionary relationship between four groups of species.

A diagram of this type is called a **phylogenetic tree**. Figure 11.20 shows the phylogenetic tree of life that captures all groups of living organisms. This tree is a product of Darwin's evolutionary theory that illustrates the inter-connectedness of all life forms through evolution.




Taxon (plural = taxa) is a taxonomic group of any rank, such as a species, family, or class.

**FIGURE 11.20** A tree of life is a diagram that shows the degree of evolutionary relatedness of all the groups of life forms on planet Earth. Diagrams like this are called phylogenetic trees.

studyon

Unit 4 | AOS 1 | Topic 3 | Concept 5

**Mitochondrial DNA**
Summary screen and practice questions

studyon

Unit 4 | AOS 1 | Topic 3 | Concept 5

**See more**
Mitochondrial DNA

## Phylogenetic trees

Phylogenetic trees are also called evolutionary trees. They are branching diagrams that show *inferred* evolutionary relationships or lines of evolutionary descent among biological groups or taxa (singular = taxon). These groups may be individual species or they may be larger groups. For example, Darwin's sketch in his 1837 notebook shows a phylogenetic tree for groups of species, and figure 11.20 above shows a tree for large groups that encompass all of Earth's life forms.

Phylogenetic trees illustrate evolutionary history as inferred from molecular data or other evidence. Molecular evidence includes amino acid sequences of proteins, RNA sequences, and DNA sequences. In the case of DNA sequences, these may be nuclear or mitochondrial DNA and may be coding or non-coding DNA. Phylogenetic trees are not fixed, but are subject to change as new research results are published.

Phylogenetic trees can be drawn using diagonal or horizontal or vertical lines (see figure 11.21). Each is correct.

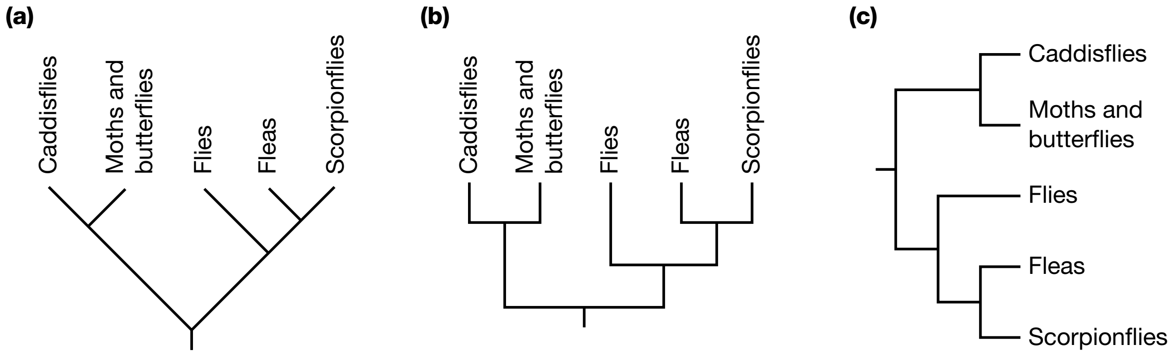



**FIGURE 11.21** Simple phylogenetic trees of living insect groups drawn **(a)** using diagonal branch lines, **(b)** using vertical branch lines, and **(c)** using horizontal branch lines. All versions are correct. Note that a time scale is not always included in every phylogenetic tree. However, time can always be assumed to run from the root of the tree (past) to the tips (present).

Figure 11.22, which follows, is a more detailed diagram identifying various features of a phylogenetic tree.

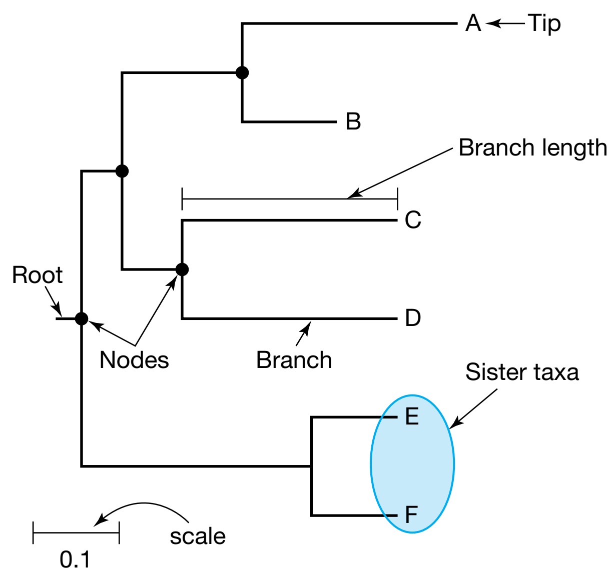



**FIGURE 11.22** Diagram showing the typical features of a phylogenetic tree.

Check out this figure and note each of the following features in this phylogenetic tree:

- The ***tips*** or terminals of a 'tree' are the descendant groups (**taxa**); the taxa are most commonly single species, but they can also be larger taxonomic groups, such as mammals or even larger, such as vertebrates. However, the tips can even be individual genes.
- A ***node*** denotes an ancestor of two (or more) descendants.
- A ***branch*** indicates a speciation event and shows the relationship between an ancestor and a descendant of that ancestor.
- The ***root*** is the common ancestor of all taxa shown in the tree.
- A ***sister taxa*** is, in fact, two groups, such as two species, with a common ancestor that is not shared with other taxa.

Note that the **branch length** conveys key information in a phylogenetic tree:

- In some trees, a time scale is included that may be a geological time scale in millions of years. In this case, the branch length denotes a period of time. If a time scale is not given, the longer a branch, the longer the time period it denotes.
- In other trees, a scale is included that indicates the amount of molecular change that has occurred between species (or other taxa). The amount of molecular change denoted by the branch length may be the number of differences in DNA base sequences between two species or the number of differences in the amino acid sequences of a protein.

  The scale may indicate either the actual number of changes or the percentage change (such as 0.1 or 10%). What is it in figure 11.22 above?
- The evolutionary distance between two species (or other taxa) is shown by the combined length of the two branches leading to each species.

The phylogenetic tree in figure 11.22 above is a very simple one. In reality, building a phylogenetic tree is an onerous task when multiple species are involved and when DNA sequences being compared may be up to 1000 nucleotides long.

Several websites are available where researchers can submit DNA sequences. The sequences are aligned and the 'best' or most robust phylogenetic tree is constructed based on the percentage identities of the DNA sequences submitted. Examples of these sites include *Phylogeny.fr* and *phyloGenerator*.

avian = related to birds

### Interpreting phylogenetic trees

Now, check out figure 11.23, which shows a phylogenetic tree of the taxonomic group called the vertebrates. You are included in that tree within the placental mammals.

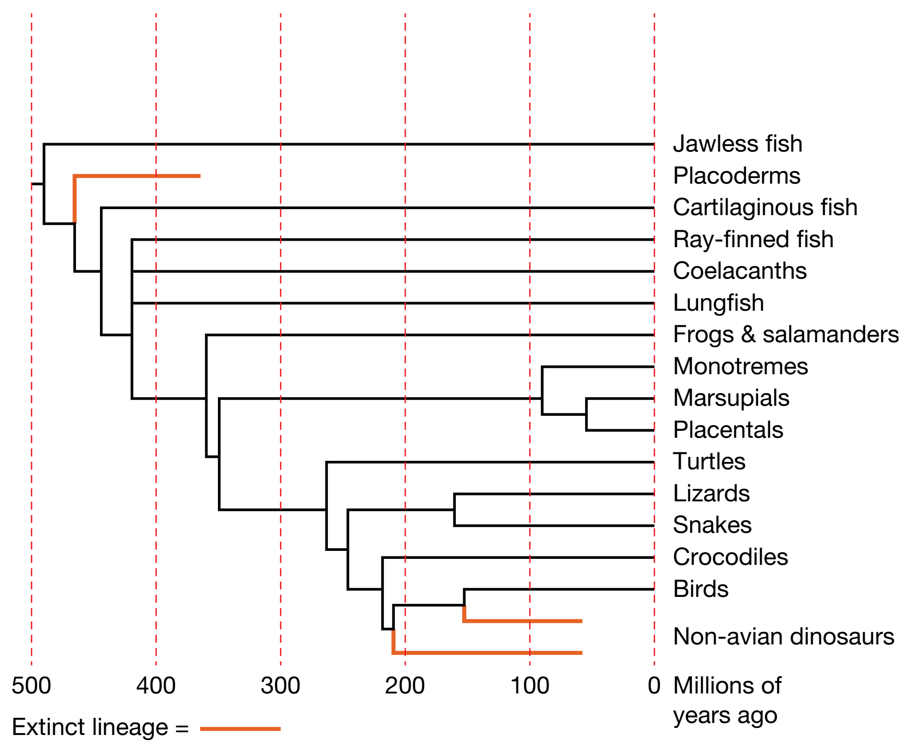
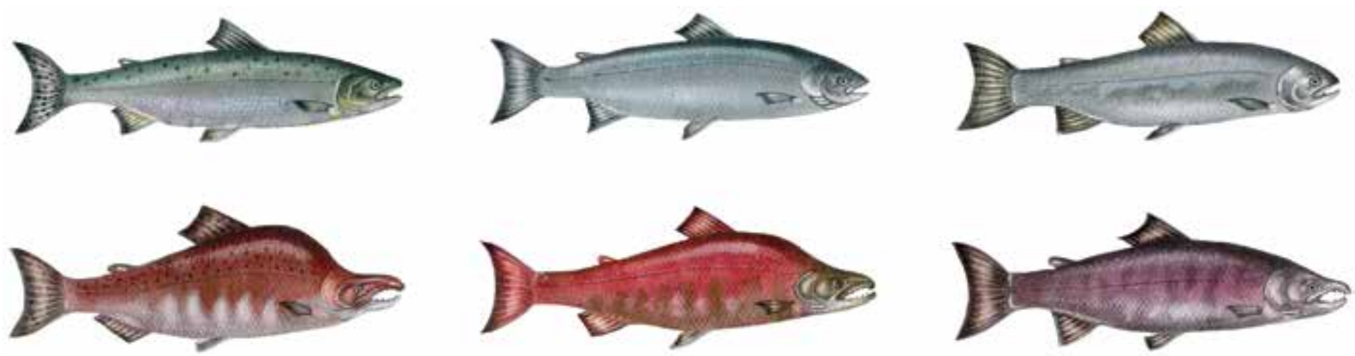



**FIGURE 11.23** Phylogenetic tree showing the relatedness of most vertebrate groups. Note the extinct lineage of the placoderms — the armoured fish. One line of dinosaurs known as the theropods diverged, giving rise to the first birds. Not all extinct groups are shown in this phylogenetic tree — among the absentees are the pterosaurs (flying reptiles) and the plesiosaurs (marine reptiles).

Let's look at a phylogenetic tree that shows the evolutionary relationship between species.

Pacific salmon and Pacific trout are members of the genus *Oncorhynchus*. Figure 11.24 shows mature males (lower row) and females (upper row) of Pacific salmon: from left, pink salmon (*O. gorbuscha*), sockeye salmon (*O. nerka*) and masu salmon (*O. masou*). Atlantic salmon are members of a different genus (*Salmo salva*).



*Oncorhynchus gorbuscha* *Oncorhynchus nerka* *Oncorhynchus masou*

**FIGURE 11.24** The genus *Oncorhynchus* contains 17 species, commonly called salmon and trout. This image shows three mature male and female species of Pacific salmon in the mating season. Note the hook on the upper jaw of the male fish, which appears in the mating season when salmon migrate from the ocean to their spawning grounds in upland rivers.

Figure 11.25 shows a phylogenetic tree of eight species in the genus *Oncorhynchus* that is the group of interest. The less closely related Atlantic salmon (*Salmo salva*) is included as an outgroup, that is, a taxon outside the particular group of interest. An outgroup branches directly from the base of the phylogenetic tree.

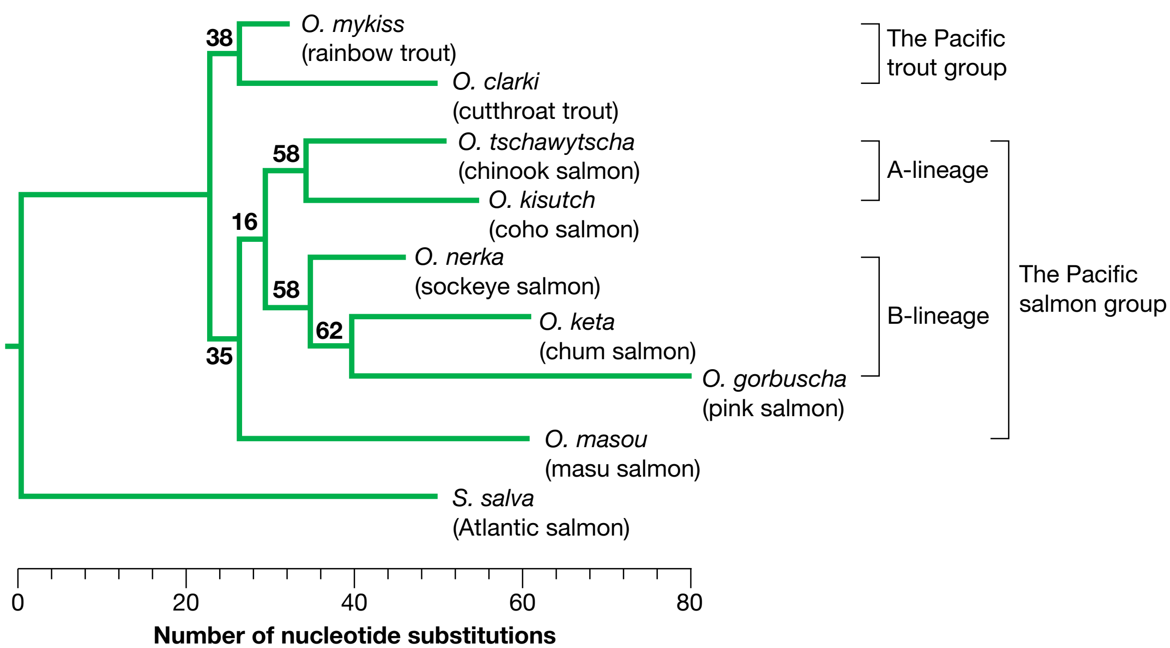



**FIGURE 11.25** Phylogenetic tree of eight species of salmon and trout of the genus *Oncorhynchus*, and the Atlantic salmon (*Salmo salva*) as the outgroup. The value given to each branch is the result of the bootstrap test (see marginal note on page 525). The scale bar shows the number of nucleotide substitutions. (*Source:* Kitano, T., Matsuoka, N. and Saitou, N. 1997, 'Phylogenetic relationship of the genus *Oncorhynchus* species inferred from nuclear and mitochondrial markers', *Genes & Genetic Systems*, 72, p. 30.)

**ODD FACT**

Estimates are that the *Salmo* genus (Atlantic salmon, brown trout and related) and the *Oncorhynchus* genus (Pacific salmon and Pacific trout) diverged by about the early Miocene (15 to 20 million years ago).

The bootstrap test checks the reliability of location of interior branches in a phylogenetic tree. In general, the higher the bootstrap value, the higher the probability that the branch position is robust and 'correct'.

Cladistics was introduced in *Nature of Biology Book 1 Fifth edition* on page 255.

**study on**

| Unit 4 | **Novel phenotypes** Summary screen and practice questions |
|---|---|
| AOS 1 | |
| Topic 3 | |
| Concept 8 | |

This tree was prepared using nucleotide sequence data from the D-loop of mitochondrial DNA. This example illustrates how phylogenetic trees are useful for organising information about biological diversity, for structuring classifications, and for providing insight to the evolutionary history of species. This phylogenetic tree illustrates the adaptive radiation of the genus *Oncorhynchus*, when an ancestral species reached the Pacific Ocean.

## Cladograms also show relationships

Another type of diagram that can be used to illustrate evolutionary relatedness is a **cladogram**.

Like a phylogenetic tree, a cladogram is a diagram that shows the relationships between groups of organisms. A cladogram is another way of organising groups of organisms based on similarities arising from their evolutionary relationships.

**In the cladistic method, organisms are organised into groups called clades on the basis of the presence of so-called derived characters or features that they share.**

- **A derived character is a novel feature that can be assumed to have evolved from an original character present in an ancestral organism.**
- **An original or ancestral character is the state of the character that was inferred to be present in the common ancestor of the clade.**

Characters can change from one state (original or ancestral) to another (derived or novel) through the process of evolution. In cladistics, the focus is *not* on the presence of all shared features. Cladistics only looks at the presence of shared **derived** features.

So, of starfish, jellyfish, and humans, which two are the most closely related? If you consider shared features, such as (i) habitat — aquatic or terrestrial, (ii) symmetry — radial or bilateral, and (iii) structure — invertebrate or vertebrate, you may well decide that starfish and jellyfish are the most closely related pair as they share all three of these features.

However, the features of radial symmetry, aquatic habitat and invertebrate structure are all traits that were present in the common ancestor of all animals. As such, these features are shared original or ancestral features, not shared derived features, and are not useful in a cladistics analysis.

In fact, the most closely related pair is the starfish–human pair! In inferring relationships using a cladistic analysis, shared derived features must be used. Figure 11.26 is a reminder that care must be taken in selecting the features to be used in a cladistics analysis.

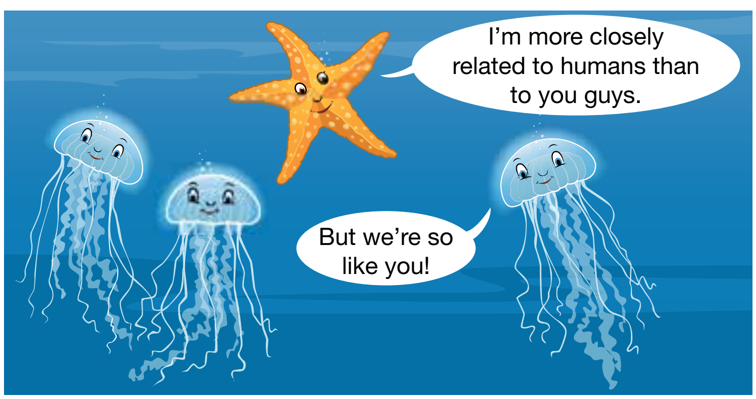



**FIGURE 11.26** Starfish (and other echinoderms) and chordates (including humans) share a derived novel feature that is not present in other animals. This derived feature relates to the timing of the formation of the mouth and the anus during very early embryonic development. In echinoderms and chordates only, the anus forms first and, later, the mouth.

Note that in a cladogram:

- all groups always appear at the endpoints, regardless of whether they are living or extinct. Contrast this with a phylogenetic tree, which shows extinct species at relevant positions on a time scale (refer back to figure 11.23 on page 523).
- the lengths of the branch lines in a cladogram do not convey information about evolutionary times or rates of evolution. Contrast this with a phylogenetic tree, which includes a time scale, perhaps in millions of years, or an axis showing rates of base changes or amino acid substitutions (refer back to figures 11.23 and 11.25 above).

In a cladogram, **a grouping that includes a common ancestor and all the descendants (living and extinct) of that ancestor form a clade** (see figure 11.27). Note the lizard–snake clade, which is shaded. Do crocodiles and plesiosaurs form a clade? There is a grouping known as clade Dinosauria. Can you identify it in this figure? What is interesting about this clade?

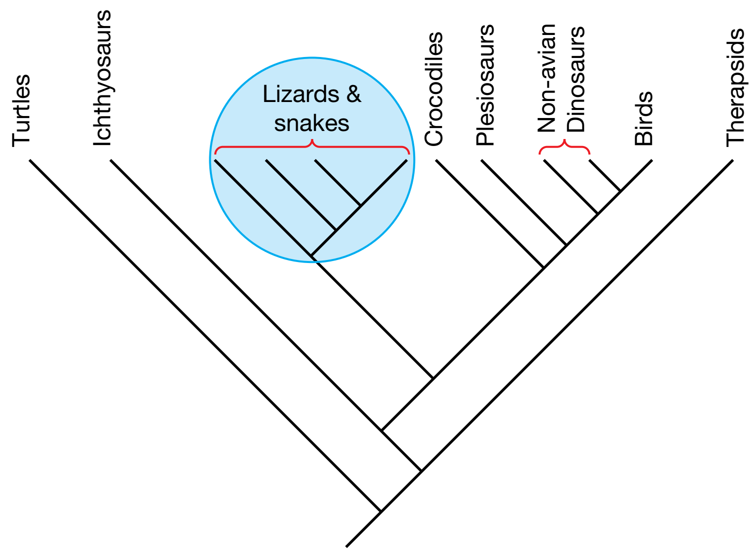



**FIGURE 11.27** A cladogram showing many reptilian groups. Birds are also included. (Why?) Note one of the clades (shown as a shaded area) within this cladogram. Can you identify one or more other clades? The therapsids were ancestral to the mammals.

Cladograms are constructed by finding the simplest arrangement of shared derived characters that are in a group of organisms. Let's look at an example of a cladistics analysis of vertebrates and the resulting cladogram that shows their evolutionary relationships.

We can start with a table that shows some relevant character states in a group of animals (see table 11.9).

**TABLE 11.9** Table showing character states for groups of vertebrate animals. In this table, a '0' means that the derived character is absent, and a '1' indicates that the derived character is present.

| character | sharks and kin | ray-finned fish | amphibians | primates | rabbits | crocodiles and kin | dinosaurs and birds |
|---|---|---|---|---|---|---|---|
| vertebral column | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| jaws | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| bony skeleton | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| four limbs | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| amniotic egg | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| hair | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| pre-orbital fenestra | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

A pre-orbital fenestra (also termed an antorbital fenestra) is a hole in the skull located in front of the eye socket. This state is present only in dinosaurs and birds (see figure 11.28).

**FIGURE 11.28** Diagram of dinosaur skull showing the pre-orbital (antorbital) fenestra.

The cladogram that is the best fit of the data in table 11.9 above is shown in figure 11.29 below.

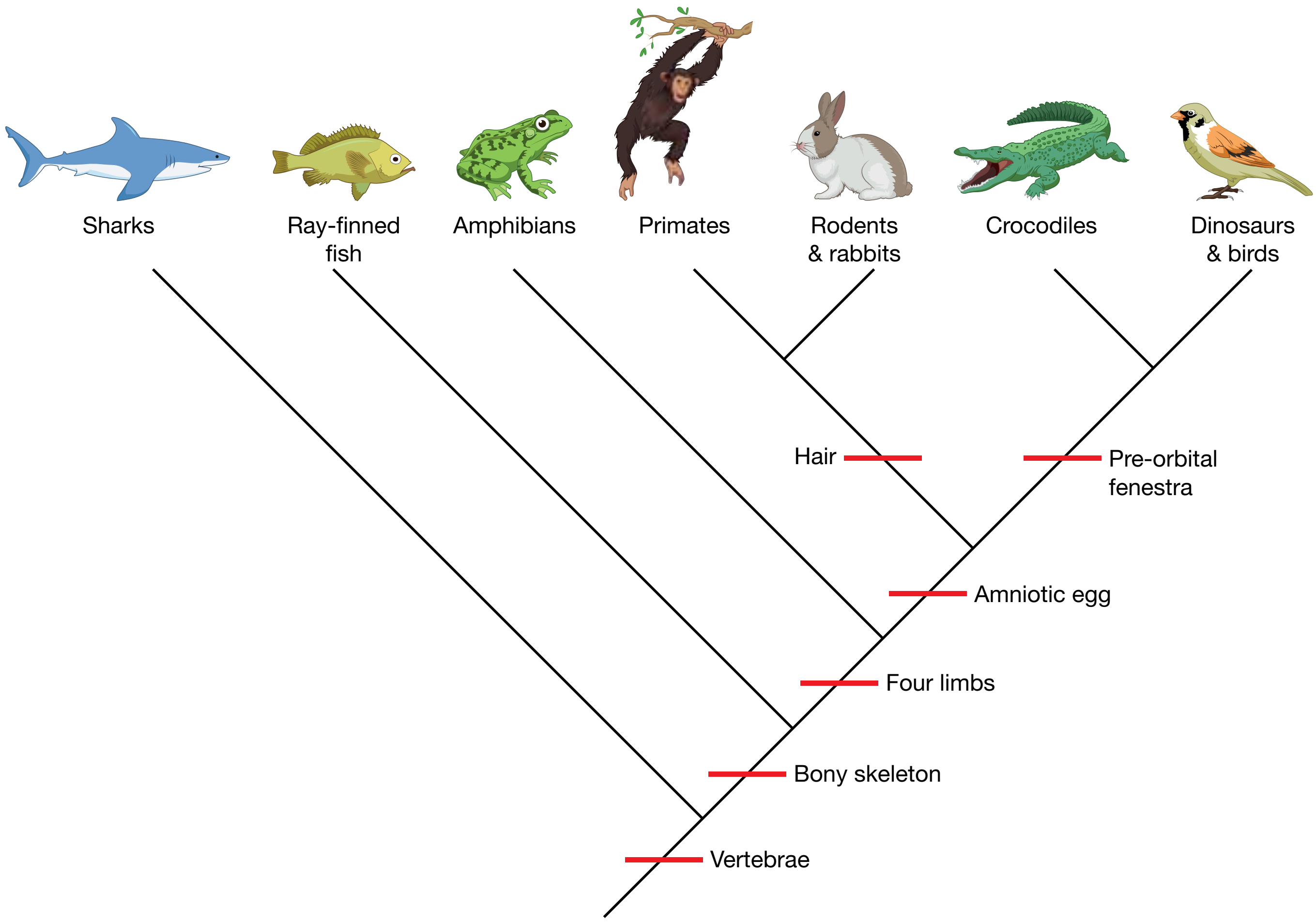



**FIGURE 11.29** Cladogram of vertebrate groups showing the 'best fit' order in which these various derived novel features appeared in descendants of the common ancestor of this group. The common ancestor can be inferred to have the original state of these features. An amniotic egg is one that has a tough outer shell.

The cladogram in figure 11.29 above is a simple example that shows just a small number of structural features. In a serious cladistics analysis, a large number of inherited characters are examined. While the character states may often relate to structural characters, other characters such as biochemical, genetic and inherited behavioural characters may also be included.

The cladogram shown in figure 11.29 above can also be shown using vertical branch lines (see figure 11.30).

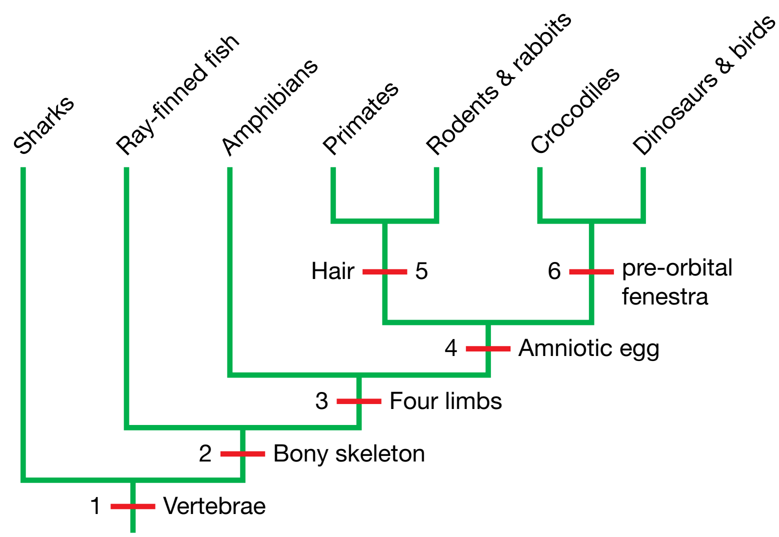



**FIGURE 11.30** Cladogram drawn with vertical lines showing the evolutionary relationships for a group of vertebrates.

Cladistic analysis is not confined to animal groups. Figure 11.31 is an example of a cladogram that shows inferred evolutionary relationships for plants.

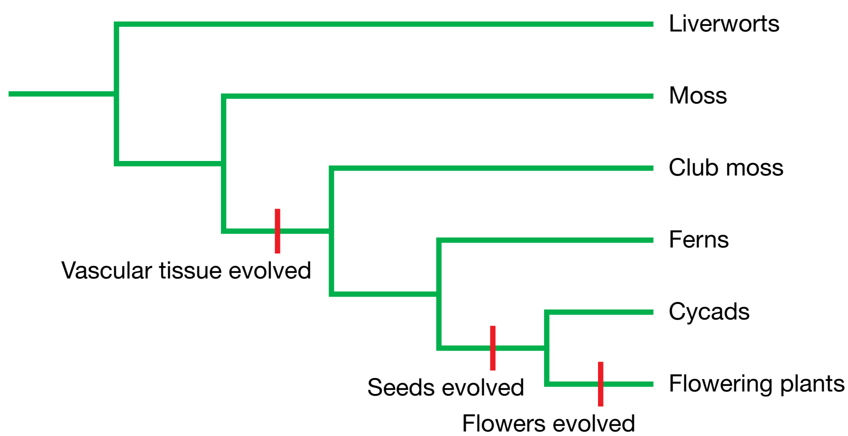



**FIGURE 11.31** Cladogram showing inferred evolutionary relationships of plant groups.

## Value of phylogenetic analyses

Phylogenetic studies have made valuable contributions to our understanding in many fields, including:

- identifying genetic variation within natural populations of a species and its subspecies
- reconstructing the evolutionary history of a species or a genus
- assisting our understanding of human prehistory, including early migrations of modern humans
- clarifying the relationship between modern humans, *Homo sapiens*, and the extinct species *Homo neanderthalensis* and *H. denisovans*
- identifying the origin of human viral diseases that have jumped from other species (zoonotic diseases).

## KEY IDEAS

- Relationships between organisms can be shown in different kinds of diagrams called phylogenetic trees and cladograms.
- Phylogenetic trees and cladograms enable inferences to be made about the evolutionary history of groups of organisms.
- Phylogenetic trees are typically based on molecular data, while cladograms use shared derived features (which can be molecular) that are inferred to have evolved from an original character present in an ancestral organism.
- Molecular data, including differences in DNA base sequences and amino acid sequences of proteins, are used as a basis for constructing phylogenetic trees.
- The greater the number of molecular differences between two taxa, the longer the period since they diverged.
- Inferences from phylogenetic trees and cladograms are subject to revision as new data become available.

## QUICK CHECK

**4** Identify the following as true or false:
 **a** A diagram showing evolutionary relationships that has an extinct organism at one of its tips would be a phylogenetic tree.
 **b** Species that are more closely related can be inferred to have shared a common ancestor more recently than less closely related species.
 **c** One kind of DNA that can provide sequences for constructing a phylogenetic tree is that of the D-loop of mitochondrial DNA (mtDNA).
 **d** A phylogenetic tree is a diagram that can show the evolutionary relationships between species.
 **e** In a cladistics analysis of different organisms, the organisms are separated into groups on the basis of shared ancestral features.

**5** Identify one relevant difference between the members of the following pairs:
 **a** a phylogenetic tree and a cladogram
 **b** a derived feature and an original feature (in cladistics)
 **c** a node and a branch (in a phylogenetic tree).

**6** Select the appropriate alternative, and in each case, briefly explain your choice:
 **a** The fewer the differences in the matching DNA sequences of two species that share a common ancestor, the <*longer/shorter*> the period since they diverged.
 **b** A group identified as a clade comprises a <*common ancestor/descendants of a common ancestor/both a common ancestor and its descendants*>.

# Cichlids of the African Lakes

The Great Rift Valley of East Africa is the location of the three largest lakes in Africa: Lake Malawi, Lake Tanganyika and Lake Victoria (see figure 11.32). (Lake Malawi is a World Heritage site.)

These three lakes are the habitats of a remarkable group of fish called **cichlids**. Lake Victoria is home to about 500 cichlid species, Lake Malawi is home to about 800 to 1000 other cichlid species, and Lake Tanganyika is home to about a further 250 different species of cichlid. Figure 11.33 shows a tiny sample of these cichlid species. The cichlids of each lake form what is termed a **species flock**. The term species flock is applied to a large group of closely related species living in a confined area, such as a lake or on an island. The cichlids in each of the African lakes have evolved independently so that each lake has a unique species flock.


**FIGURE 11.32** Lake Victoria, Lake Tanganyika and Lake Malawi are the habitats of the flocks of freshwater fish called cichlids (pronounced sick-lids).

The emergence of the species flock in each of the three great African lakes is an example of adaptive radiation. The adaptive radiation of the cichlids has been described by some researchers as being 'explosive'. The term explosive refers to the rapidity and the scale of evolutionary change that has generated so many species over a relatively short time frame. The speed with which this diversification has occurred is remarkable. The 500 cichlid species of Lake Victoria evolved within the past 10 000 to 15 000 years. No other animal group has evolved at anything like this rate. For example, the ground finches of the Galapagos Islands (refer back to chapter 10) have diversified with 14 species emerging in a period of several million years.

**ODD FACT**

Sadly, the introduction of Nile perch (*Lates niloticus*) in the 1950s and overfishing have resulted in a marked decline in many cichlid species, in particular in Lake Victoria.

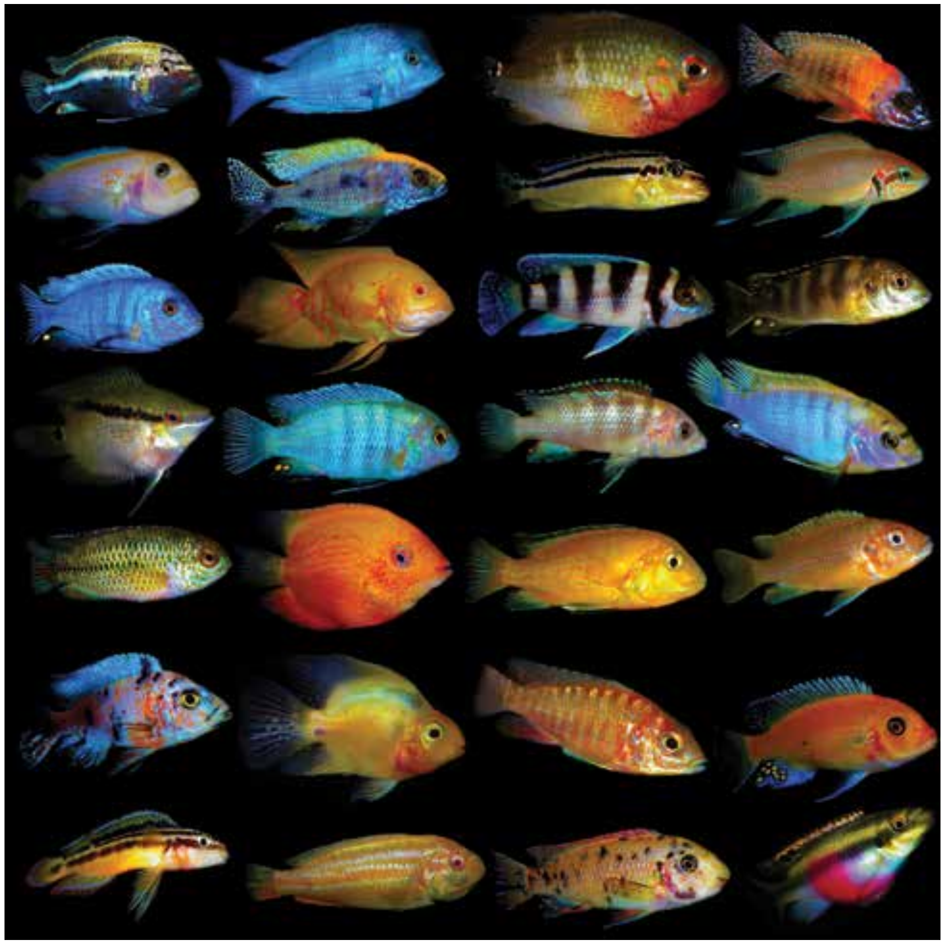


**FIGURE 11.33** A small sample of cichlid species. Note, in particular, the different shaped jaws and associated 'lips' that equip these fish to occupy different feeding niches. One set of cichlid lips (termed 'Angelina Jolie' lips by one researcher) act as a seal and enable the fish to suck prey from narrow crevices in rocks!

**ODD FACT**

The outer row of teeth in 'biting' cichlids comprises small, closely spaced teeth with multiple cusps. In cichlids that feed by suction, the outer row of teeth are large, with fewer cusps, and are widely spaced. All cichlids also have a second row of teeth in a smaller jaw located in their throats.

## Feeding habits of cichlids

The different cichlid species in each lake show great diversity in structure and behaviour, in particular in their feeding habits. Cichlids display a range of specialised feeding habits. Some graze on algae growing on rocks, some feed on filamentous algae growing on other algae, some prey on fish and dine on flesh — or scales that they tear from their prey, some feed on plankton that they suck from the water, some eat insects, and some crush snails. This diversity of feeding behaviours means that interspecific competition between the cichlids within a species flock in a lake is minimised. In addition, different cichlid species occupy different habitats in the lakes, including rocky areas, sandy shallows and deep water.

The specialised feeding behaviours of cichlids have been enabled by the development of many different phenotypes of jaw shape. These various jaw shapes enable the cichlids to bite or scrape or suck or crush. Are mutations in many structural genes required to produce this diverse range of jaw shapes? The answer is 'No!' Instead, what has been discovered is that changes in *when* and *where* and *how much* a few master genes are expressed can generate an enormous diversity of shapes and sizes.

## BMP4: a master gene

The jaw shape of cichlids is determined during their embryonic development by the level of expression of the **BMP4** gene. The BMP4 gene is expressed in a signalling molecule, the bone morphogenetic protein. A shift in either the time of expression of the BMP4 gene, the level of expression or where it is expressed can produce a diversity of jaw shapes.

eBook*plus*

**Weblink**
Cichlids from Lake Malawi

In cichlids, differential expression of BMP4 is correlated with dental patterning and with jaw development. When BMP4 is expressed at higher levels, as occurs in the biting cichlids, the jaw develops into a shorter, robust structure, with teeth that are small and closely spaced. In contrast, when BMP4 expression is reduced, as occurs in cichlids that feed by suction, the jaw is more elongate and the teeth are larger and spaced like a comb. Figure 11.34 shows the result of the differential expressions of BMP4 on the shape of the lower jaws of three members of the Cichlidae family that differ in their feeding habits. Note the differences in jaw shape and tooth number.


**FIGURE 11.34** Photos (not to same scale) of lower jaws (looking down from above) of three fish from the Cichlidae family. From left: **(a)** flesh-eating *Dimidiochromis compressiceps*, a large narrow-bodied fast-moving predatory cichlid with a long, narrow jaw with few teeth; **(b)** *Maylandia zebra*, a filter-feeding cichlid that feeds on free-floating plankton, diatoms and algae; **(c)** *Labeotropheus fuelleborni*, a 'biting' cichlid that uses its downward-oriented mouth with a short robust lower jaw with many teeth used to bite algae from rocks. (*Source:* Fraser, G.J., Hulsey, C.D., Bloomquist, R.F., Uyesugi, K., Manley, N.R. and Streelman, J.T. 2009, 'An ancient gene network is co-opted for teeth on old and new jaws', *PLoS Biology*, 7(2).)

## BMP4 in Darwin's finches

In chapter 10 (refer back to page 448), you were introduced to the ground finches of the Galapagos Islands, also known as Darwin's finches, as an example of adaptive radiation. Figure 11.35 shows a phylogenetic tree of Darwin's finches. Note that one group of six species of ground finch forms a very closely related group — the *Geospiza* group.

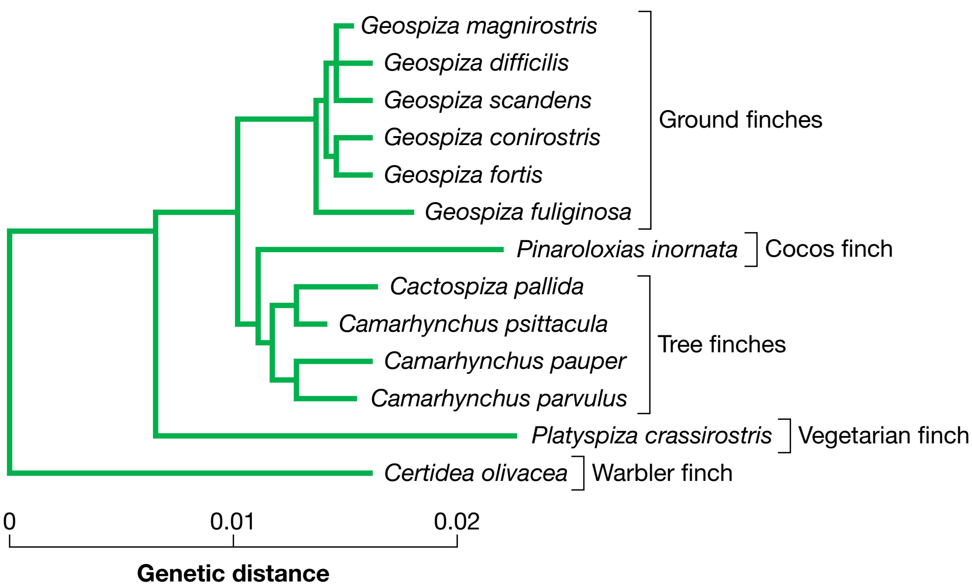



**FIGURE 11.35** The *Geospiza* ground finches species have diversified to exploit a range of foods, including seeds of various sizes, nectar of cactus flowers and cactus fruits, as summarised in Table 11.10. Figure 11.36 shows a sketch of four species observed by Darwin when he was on the Galapagos Islands.

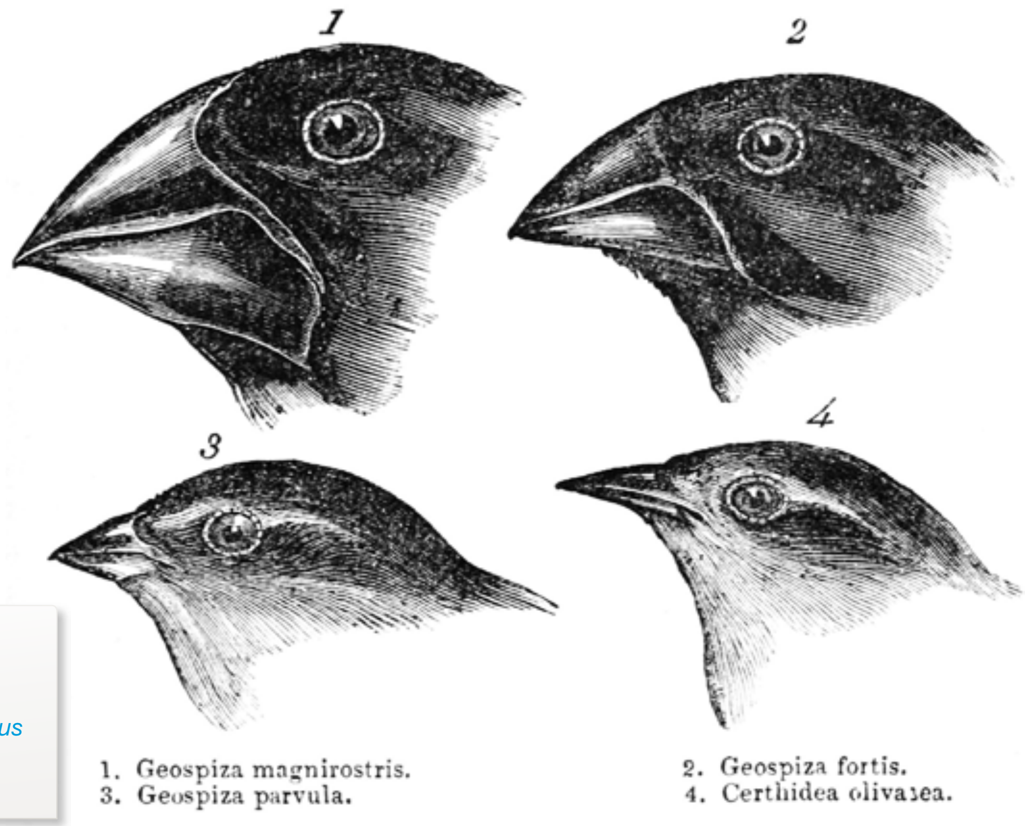



**FIGURE 11.36** Drawings of four species of Darwin's finches. *Geospiza parvula* is now known as *Camarhynchus parvulus*. Note the marked differences in beak sizes.

**TABLE 11.10** Six species of the *Geospiza* group of Darwin's finches, their food, and their beak shape.

| species | food source | beak shape |
|---|---|---|
| *Geospiza difficillis* | seeds and insects | small pointed beak |
| *G. fuliginosa* | small seeds | broad, deep beak |
| *G. fortis* | medium seeds | broad, deep beak |
| *G. magnirostris* | large seeds | broad, deep beak |
| *G. scandens* | cactus nectar and fruits | long, pointed beak |
| *G. conirostris* | cactus nectar and fruits | long, pointed beak |

The specialised beak shapes of the *Geospiza* finches are determined by differential patterns of expression of the BMP4 gene. The product of its expression is the bone morphogenetic protein, a signalling molecule involved in bone and cartilage development. Researchers found a correlation between beak shape (morphology) and the expression of the BMP4 gene in the upper beak primordia (early-forming tissue) of developing finch embryos.

In the early development of the seed-crushing species, low expression of BMP4 was detected in the mesenchyme of the upper beak, and this correlated with low beak width and depth. In contrast, high expression of BMP4 is correlated with beaks of greater widths and depths. (Mesenchyme is embryonic tissue that can develop into bone and cartilage.)

As would be expected, BMP4 gene expression is much greater in the developing embryos of the large ground finch, *G. magnirostris,* and its protein product is found in greater amounts as compared to finches with smaller, thinner beaks, such as *G. scandens.* Expression of this gene controls beak width and height.

Expression of another gene, the CaM gene that encodes the signalling molecule, calmodulin, has been found to control beak elongation. Calmodulin is a calcium-binding protein that regulates the action of other cell proteins involved in development.

In summary:

- The BMP4 gene controls beak width and depth and its expression results in the development of the broader, heavier beaks found in the seed-eating finches.
- The CaM gene controls beak length, and its expression produces the longer probing beaks of the cactus finches.

Figure 11.37 shows the interaction of the BMP4 gene and the CaM gene in producing different beak shapes.

Master genes, such as BMP4 and CaM, control aspects of embryonic development. Changes to their times of expression, their levels of expression and their sites of their expression can produce a diverse range of novel phenotypes.

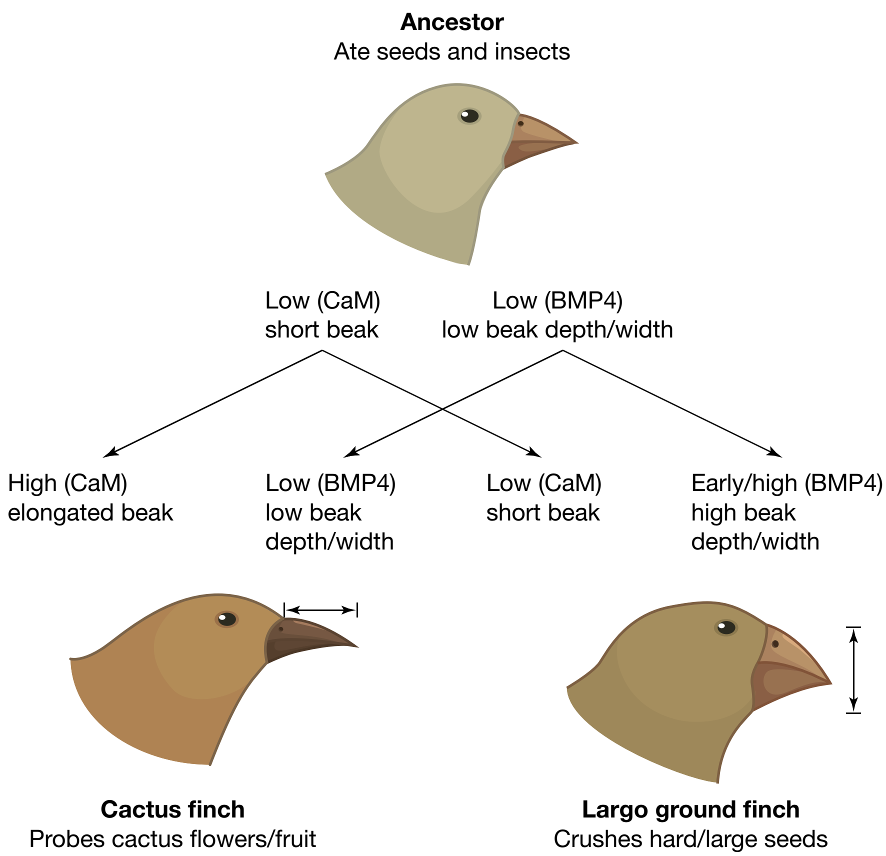



**FIGURE 11.37** Diagram showing how the differential expressions of two master genes, **BMP4** and **CaM**, can result in a variety of beak shapes. High levels of expression of the **BMP4** gene result in wider and higher beaks, and high levels of expression of the **CaM** gene produce longer beaks.

## KEY IDEAS

- Expression of BMP4, a master gene, during embryonic development of cichlid fish controls jaw shape in the developing cichlid fish embryos, and beak shape in finches.
- The product of the BMP4 gene is a signalling molecule, bone morphogenetic protein 4.
- Higher levels of BMP4 expression in the cichlid embryo result in a shorter and more robust jaw, and wider and deeper upper beaks in finch embryos.
- In finches, a second master gene, CaM, controls beak length.
- Slight variations in the timing, levels, and sites of expression of the BMP4 gene can generate a range of novel jaw phenotypes in cichlids and beak phenotypes in finches.

## QUICK CHECK

7 Identify the following as true or false:
  a Jaw formation in a cichlid feeding on snails would be associated with a high level of expression of the BMP4 gene.
  b In Darwin's finches, beak length is controlled by the BMP4 gene.
  c The embryonic development of an adult finch that feeds on nectar from flowers would be associated with a high level of expression of the CaM gene.
  d Embryonic beak development in *Geospiza magnirostris* would be associated with a high level of expression of BMP4.
  e Slight variations in the timing and the level of expression of the BMP4 gene could generate new phenotypes in cichlids.
8 Identify one relevant difference between the action of BMP4 and the action of CaM (in ground finches).

## BIOLOGY IN THE WORKPLACE

### Susie Moreton — Hospital Library Manager

**FIGURE 11.38** Susie Moreton. (Image courtesy of Susie Moreton.)

I am the Manager of the Epworth HealthCare Library in Victoria. My job is to make sure the doctors, nurses and other clinicians have the best and most current medical information in order to provide the best care to patients. Not only do I have to find the information, I must deliver it in the most applicable format; sometimes in print, often electronic and increasingly within the hospital's systems. Each day is interesting as I learn how technology is advancing health care in every way.

At school I really enjoyed biology, and majored in sciences as a teaching undergraduate. I soon qualified as a librarian and worked at several universities as a Reference Librarian, eventually supporting academics in senior liaison roles — this means understanding a particular discipline well, partnering with lecturers and senior students as they research and publish. Information literacy is central to this process, and I mentored skills in information discovery, access, analysis and synthesis. The great secret of librarianship is everyone values these abilities, so I collaborated with experts in education, business, literature, history, economics, psychology, fine arts, health, engineering, women's studies, etc. My favourite area was health sciences, and by drawing upon the scientific principles I learned at school, such as observation, investigation, validity, reliability, rigour and independence, I maintained credibility with the university staff and students. It was great to work with really clever people in an environment of enquiry.

A few years on I completed my Masters in Information Science, and commenced my role at the Epworth Library. The hospital is a very dynamic and supportive working environment, where initiative and intelligence are applied to help patients. Every day I learn more about therapies, conditions, procedures, interventions, instruments, technologies, and so many other aspects of healthcare. All the skills I developed in university roles are utilised at the hospital, especially as I help clinicians investigate emerging treatments and write up findings for publication. My daily activities vary from buying a specific book or journal, seeking the latest clinical trial report, teaching a clinician how to interrogate a database, to advice on writing for publication. Librarians' skills are highly valued, and the library is an active partner in the hospital's research, education and clinical activities. My job is very rewarding as I work with very clever and kind people, and every day I know I contribute to good patient care.

# BIOCHALLENGE

## Drawing a phylogenetic tree

In this Biochallenge, you will see how a simple phylogenetic tree can be built to show the evolutionary history of four living species. The data used will be the differences in the base sequences of the DNA that encodes *cytochrome b*, a mitochondrial protein, in each species (see table 11.11 below). In this exercise, the aim is to correctly order the various divergences that led to the emergence of the four species.

**TABLE 11.11** **Differences in DNA sequences that encode *cytochrome b* in four species. *Cytochrome b* is a protein found in the mitochondria of all eukaryotic species.**

| | Species A | Species B | Species C | Species D |
|---|---|---|---|---|
| Species A | — | 27 | 37 | 38 |
| Species B | 27 | — | 13 | 19 |
| Species C | 37 | 13 | — | 3 |
| Species D | 38 | 19 | 3 | — |

1 To build our phylogenetic tree, work the following example:

***Step 1:*** Identify the one species that shows the *most* differences with all the other species. The difference for each species is its column total. A quick check shows this to be species A with a *grand total* of 102 differences.

***Step 2:*** Now identify the species that shares a common ancestor with species A. This will be the species that has the *smallest* number of differences with species A. This turns out to be species B (with 27 differences between A and B).

Now we can start our phylogenetic tree by showing species A and B with a common ancestor as shown in figure 11.39a.

***Step 3:*** Now identify the species that shares a common ancestor with species B. This will be the species with which species B has the smallest number of differences. A quick check shows this to be species C (with 13 differences).

Now we add species C to our tree (see figure 11.39b).

***Step 4:*** This just leaves species D. Note that because it has the least number of differences with species C, we can assume that species D shares a common ancestor with that species.

We can now complete our simple phylogenetic tree (see figure 11.39c).

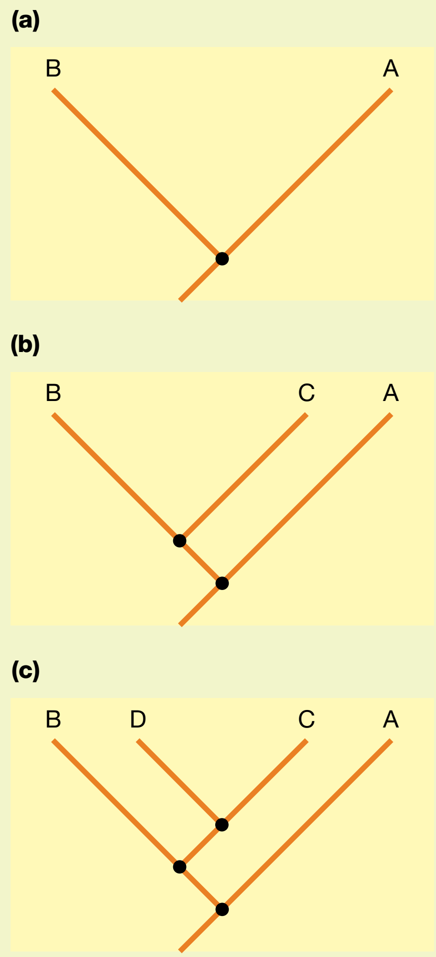
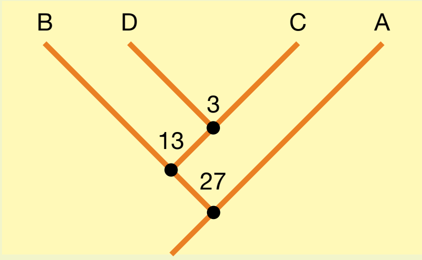



**FIGURE 11.39** **(a)**, **(b)** and **(c)**: diagrams showing the development of a phylogenetic tree. Common ancestors are denoted by black dots at the nodes.

Note that the branch lines are *not* drawn to scale. To draw this phylogenetic tree correctly to scale, the total length of the V-shaped branch connecting species A and B would be 27 units, the total length of the V-shaped branch connecting species B and C would be 13 units, and the total length of the V-shaped branch connecting species C and D would be a tiny 3 units. Instead, we can show these differences simply by inserting each value at the relevant node.

***Step 5:*** Complete the phylogenetic tree by inserting the number of differences between each species immediately above the node that denotes their common ancestor, as shown below:




**2** Now, construct a phylogenetic tree to show the evolutionary relationships of the five different species below, using the differences in the DNA sequences that encode *cytochrome c* (see table 11.12 below).

**TABLE 11.12** **Numbers of differences in the DNA sequences encoding *cytochrome c*.**

| | Horse | Donkey | Chicken | Penguin | Snake |
|---|---|---|---|---|---|
| **Horse** | 0 | 1 | 11 | 13 | 21 |
| **Donkey** | | 0 | 10 | 12 | 20 |
| **Chicken** | | | 0 | 3 | 18 |
| **Penguin** | | | | 0 | 17 |
| **Snake** | | | | | 0 |

# Chapter review

studyon

Unit 4 AOS 1 Topic 3 — Determining relatedness between species

Practice questions

## Key words

BMP4
branch
branch length
chromosome banding
chromosome painting
cichlids
clades
cladogram
comparative genomics
derived
derived character
DNA–DNA hybridisation
fusion
gene conservation
haplogroups
homologous
melting temperature ($T_m$)
molecular clock
node
original or ancestral character
phylogenetic tree
root
sister taxa
species flock
taxa
tips

## Questions

1 **Making connections** → Draw a concept map for 'evolutionary relatedness' incorporating relevant keywords from this chapter. You may add any other concepts that you wish.

2 **Developing logical explanations** → Suggest possible logical explanations for the following observations:
   - **a** The degree of hybridisation is greater between the DNA from two bird species than between the DNA from a bird and a frog species.
   - **b** When the haemoglobin of a species of mammal is compared with that of other mammals, some are more similar than others.
   - **c** The percentage difference in the amino acid sequence of one protein from a cat and a dog is far less than that between a lizard and either mammal.
   - **d** When a 'paint' specific to the number-7 human chromosome is added to chimp chromosomes and to gibbon chromosomes, only one chimp chromosome is painted but several gibbon chromosomes are painted.
   - **e** The diploid number of the great apes is $2n = 48$ but that of humans is $2n = 46$.

3 **Interpreting graphical data and communicating understanding** → Refer back to figure 11.23 on page 523, which shows a phylogenetic tree for the groups of animals called vertebrates. Answer the following questions based on the information in this tree:
   - **a** What were the earliest vertebrates to appear?
   - **b** Are representatives of this group still extant on Earth?
   - **c** Cartilaginous fish include sharks and rays, while ray-finned fish include the bony fish. Which group appeared earlier?

4 **Interpreting graphical data and communicating understanding** → Refer back to figure 11.27 on page 526, which shows a clade for the groups of animals popularly called reptiles. Answer the following questions based on the information in this cladogram:
   - **a** Give an example of another clade that includes at least two reptilian groups.
   - **b** Which, if any, of the groups shown in this cladogram are extinct?
   - **c** Identify the following statements as either true or false:
     - i Birds are more closely related to crocodiles than to mammals.
     - ii Turtles are a more ancient lineage than plesiosaurs.
     - iii The common ancestor of snakes and lizards emerged at the same time as the common ancestor of the dinosaurs.
   - **d** Consider the following statements and assess their validity:
     - i 'Birds are flying dinosaurs.'
     - ii Class Reptilia and Class Aves exist as taxa in traditional Linnean classification schemes. However, these classes do not exist in cladistics.

5 **Communicating understanding** → Identify a key difference between the members of the following pairs:
   - **a** a rooted and an unrooted phylogenetic tree
   - **b** immediate ancestor and remote ancestor
   - **c** a cladogram and a phylogenetic tree
   - **d** DNA hybridisation and DNA sequencing.

6 **Interpreting data and communicating understanding** → Go to figure 11.22 on page 522 and answer the following questions:
   - **a** What is the evolutionary distance between species A and B?
   - **b** In this example, what are the units of evolutionary distance?
   - **c** How is evolutionary distance measured in this phylogenetic tree?
   - **d** Complete the sentence from the choice below: Species pair E and F show (*a higher/a lower/ the same*) degree of relatedness compared to species pair A and B.
   - **e** Explain your choice in (d) above.

7 **Applying knowledge and understanding →** Arrange the following in order, from highest to lowest melting temperatures, if DNA from the following were tested using the technique of DNA–DNA hybridisation:
  - **a** a golden bandicoot (*Isoodon auratus*) and a southern brown bandicoot (*Isoodon obesulus*)
  - **b** a frog and a golden bandicoot
  - **c** a southern brown bandicoot and a kookaburra
  - **d** a golden bandicoot and a wallaby.

  Explain your decision.

8 **Interpreting data and applying understanding →** Examine the hypothetical evolutionary lineages shown in figure 11.40 below. The letters denote species represented by either organisms living at the present or extinct species represented by fossils.

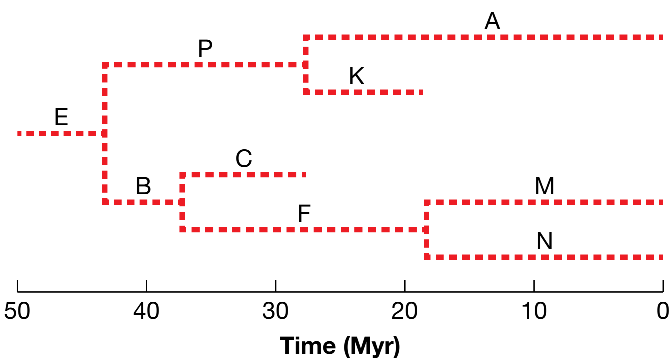



**FIGURE 11.40** An evolutionary diagram showing hypothetical species.

Examine this figure and answer the following questions:
  - **a** Is this diagram a phylogenetic tree or a cladogram? Explain.
  - **b** Approximately when did species M and N diverge from a common ancestor?
  - **c** Is it reasonable to conclude that species M evolved from species N? Explain.
  - **d** Which species is the common ancestor of M and N?
  - **e** Identify a remote ancestor of species M and N.
  - **f** Did species C evolve from species F? Explain.
  - **g** Did species B and species K live during the same period?
  - **h** What is meant by the term *common ancestor*?
  - **i** A species can be identified as an immediate ancestor or a remote ancestor of another species.

    In terms of species K, identify an immediate and a remote ancestral species.
  - **j** Is species C a remote ancestor of species M? Explain.
  - **k** Which species is the common ancestor of all the other species?
  - **l** Which species existed for the shortest period?

9 **Interpreting and re-presenting data →** The data shown in table 11.13 below are the percentage differences between the corresponding protein in four different species (A, B, C and D).

**TABLE 11.13**

| Species | B | C | D |
|---|---|---|---|
| A | 4% | 17% | 36% |
| B | | 16% | 32% |
| C | | | 35% |

  - **a** Use the data to draw a hypothetical evolutionary relationship similar to that shown in figure 11.39c, to show the evolutionary relationship of these four species.
  - **b** Assume that the four species are canary, cat, mouse and lion. Match these four species to the letters (A–D) giving reasons for your decision.

10 **Interpreting and re-presenting data →** Refer back to figure 11.39c on page 536.
  - **a** Would the reversal of the positions of C and D change the evolutionary story told by this phylogenetic tree? Explain.
  - **b** Redraw this tree to accommodate a fifth species, E, that shares a common ancestor with species C.
  - **c** Redraw the tree so that, instead, species E shares a common ancestor with species D.
  - **d** Which common ancestor appeared earlier in geological time: the common ancestor of species C and B or the common ancestor of species D and C?

11 **Demonstrating and communicating understanding →** Carefully consider the cladograms shown in figure 11.41 below and answer the following questions:
  - **a** Do these three cladograms tell the same story or not about the evolutionary relationships of mosses, ferns, pines and roses? Explain your choice.

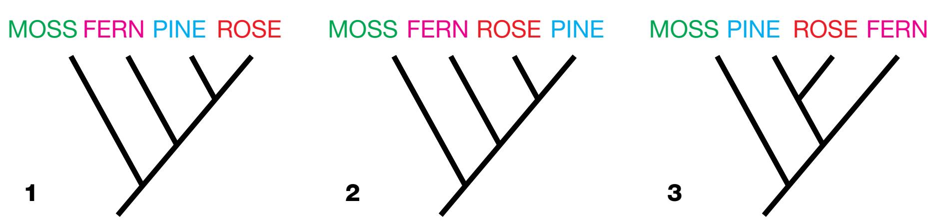



**FIGURE 11.41** Cladograms.

- **b** Redraw one of these trees using horizontal lines rather than diagonal lines.
- **c** Is it reasonable to suggest that a feature that was used to form these groups was the presence of chlorophyll? Explain.

**12** **Communicating understanding** ➜ A great diversity of jaw shape may be seen in the many different cichlid fishes of the great African Lakes — Victoria, Tanganyika and Malawi.

- **a** The evolution of this species flock is often termed 'explosive'. Suggest why this term is used.
- **b** Would you predict that identical cichlid species are present in the three lakes?
- **c** In a given lake, what factors act to minimise interspecific competition for food?
- **d** Student R stated that the great diversity of cichlid jaw shape is the result of mutations in one set of genes in cichlids with biting jaws, in other genes in cichlids with crushing jaws, and in yet other genes in cichlids with scraping jaws. Do you agree or disagree with student R? Give a reason for your decision.

CHAPTER

# 12 Human changes over time

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- identify characteristics that define the groups that comprise primates, hominoids and hominins
- develop awareness of major trends in hominin evolution, from the genus *Australopithecus* to the genus *Homo*
- give examples of the structural, behavioural and cognitive changes that have occurred in hominin evolution
- compare and contrast biological, cultural and technological evolution
- recognise that the human fossil record provides an explanatory model that is subject to revision as new fossils are discovered and as new interpretations are developed.

**FIGURE 12.1** All species are linked, either closely or remotely, depending on the time since they shared a common ancestor. Species are organised into groups, such as genera, classes and orders, according to their shared physical, physiological and genetic similarities. Here are two members of the order Primates.

# Our place in the living world

A visit to a zoo is a great experience. See the big cats, such as lions (*Panthera leo*) and cheetahs (*Acinonyx jubatus*), and inevitably you will see similarities with your pet cat at home. Seeing African painted dogs (*Lycaon pictus*) or a grey wolf (*Canis lupus*) will remind you of your pet dog. Stand and look at a great ape, such as a bonobo or pygmy chimp (*Pan paniscus*) (see figure 12.2). It may well look back at you with apparent equal interest. You might be tempted to say: '*What is it thinking?*'

**FIGURE 12.2** A bonobo (*Pan paniscus*), a member of one of the two species of chimpanzee. The other species is the common chimpanzee (*Pan troglodytes*).

**eBook*plus***

**Weblink**
Kyoto U chimp Ayumu and nine-number memory sequence

Refer to *Nature of Biology Book 1 Fifth Edition*, page 311ff. to review the principles of classification.

Indeed, studies have shown that chimps are capable of solving complex problems, can recognise symbols, and have extraordinary short-term memories. This remarkable short-term memory capability was demonstrated by a chimp called Ayumu at the Primate Research Institute of Kyoto University in Japan. Ayumu can remember the positions on a monitor screen of a sequence of nine numbers after being shown them for just one second.

Each different kind of organism is a species. The species that make up the living world are organised into groups that show relationships. How does the modern human species (*Homo sapiens*) fit into this biological classification scheme? In the following section, we will explore this.

# Classification of the human species

Let's look at the classification of modern humans. Biological classification is based on molecular data as well as structural and physiological data. Through classification, biologists:

- put order into the living world
- show the relationship between different species
- gain insight to the evolution of a species.

The animal world is organised into large groups called phyla (singular phylum). Members of one phylum share a common ancestor and show similarities in their body structures, both external and internal. About 36 different phyla are recognised, and modern humans are members of one of these phyla, namely the phylum Chordata. This phylum includes several classes, including the class Mammalia, which includes us — so, we are mammals.

## Humans are mammals

Figure 12.3 shows the classification of humans from phylum Chordata to order Primates.

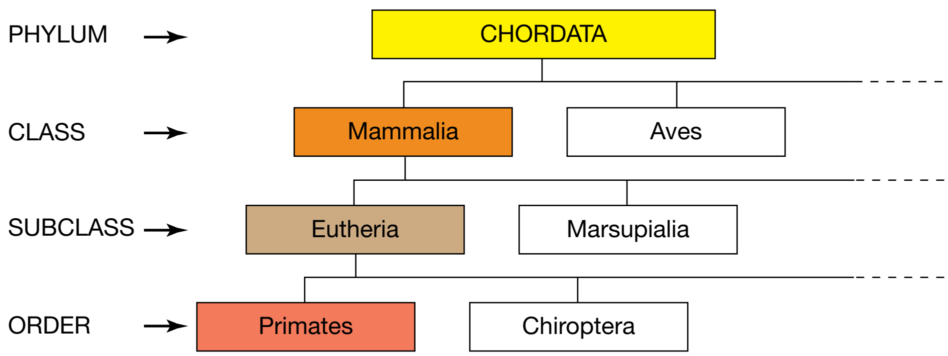



**FIGURE 12.3** Classification of humans from phylum Chordata to order Primates.

Humans, along with about 5400 other animal species, are members of class Mammalia within the phylum Chordata (see figure 12.3). Mammals are the familiar animals that include ourselves, common domestic pets such as cats and dogs, and farm animals such as sheep, cattle, pigs, horses and goats. Other mammals include species in the wild such as whales, bats, walruses and bears. How many different kinds of mammal can you identify in figure 12.4?

**FIGURE 12.4** How many different kinds of mammal are shown in this image?

## What is a mammal?

The distinguishing features of mammals are shown in figure 12.5. **Just one of these features is sufficient to identify a mammal**. As a result, a fossil jawbone or even a few teeth can identify whether an animal is a mammal.

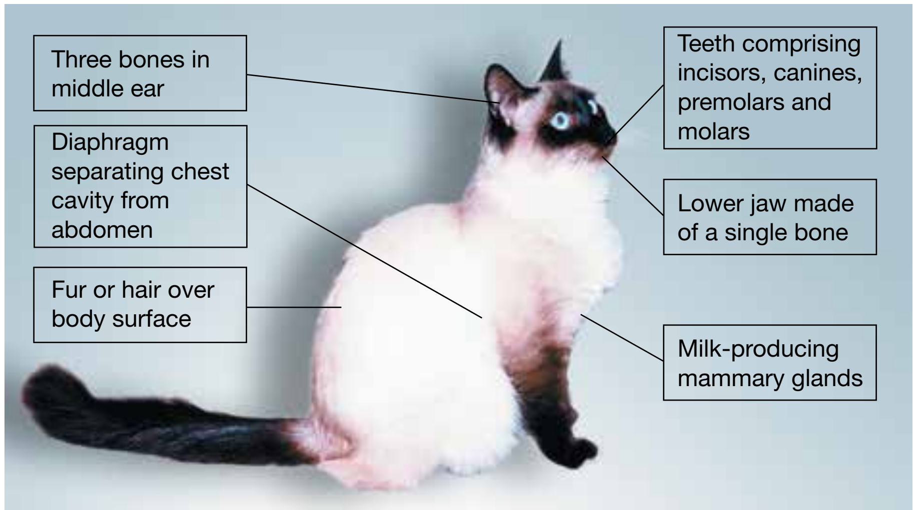



**FIGURE 12.5** Mammals show a variety of features. Observing just one of these features is sufficient to decide that an animal is a mammal. Some mammals, such as whales and dolphins, have lost their covering of hair. Can you suggest why? (Image courtesy of Judith Kinnear.)

The teeth of mammals are typically differentiated into incisors, canines, premolars and molars, which, between them, can cut, pierce, grip, shear and grind (see figure 12.6a). In contrast, the teeth of reptiles, such as crocodiles, are essentially simple cones — good for piercing and grasping (see figure 12.6b). In addition, the lower jawbone in reptiles, unlike that of mammals, is made of three bones.

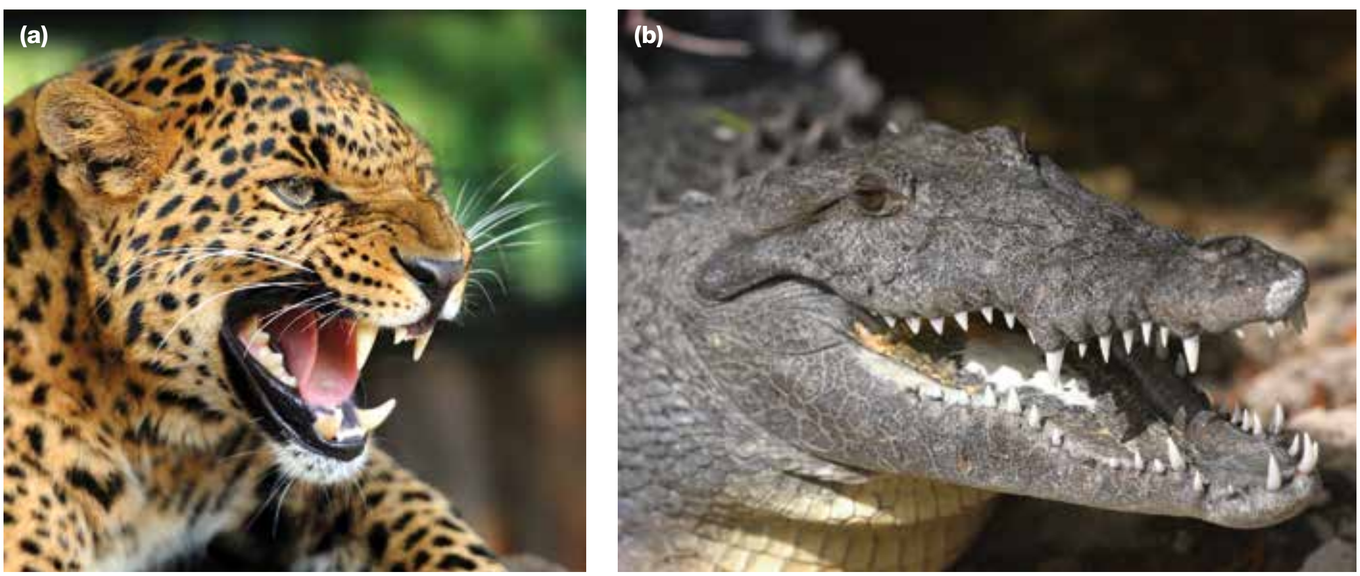



**FIGURE 12.6** **(a)** An example of a mammal, this leopard (*Panthera pardus*) shows its mammalian characteristics of differentiated teeth. Note the small front incisors, large canines and shearing molars (carnassial) teeth at the sides of the jaw. **(b)** An example of a non-mammal, this crocodile shows its undifferentiated conical teeth. What other physical features distinguish a member of class Mammalia from a member of class Reptilia?

The term 'primate', meaning 'first' or 'highest', was given to this group of animals by Carl Linnaeus in 1758. He was the first person to formally recognise primates as a distinct group.

## Humans are primates

Among the mammals, humans are primates or members of the order Primates.

About 400 different primate species live today on this planet (see figure 12.7). They range from the tiny mouse lemur (*Microcebus rufus*), which has an average mass of about 40 grams, to the western gorilla (*Gorilla gorilla*), with the adult male having a mass of about 175 kilograms. Most primates are tree dwellers and are found in tropical forest and woodland habitats. As well as living species, primates also include extinct species such as *Homo erectus* and various species of *Australopithecus*.

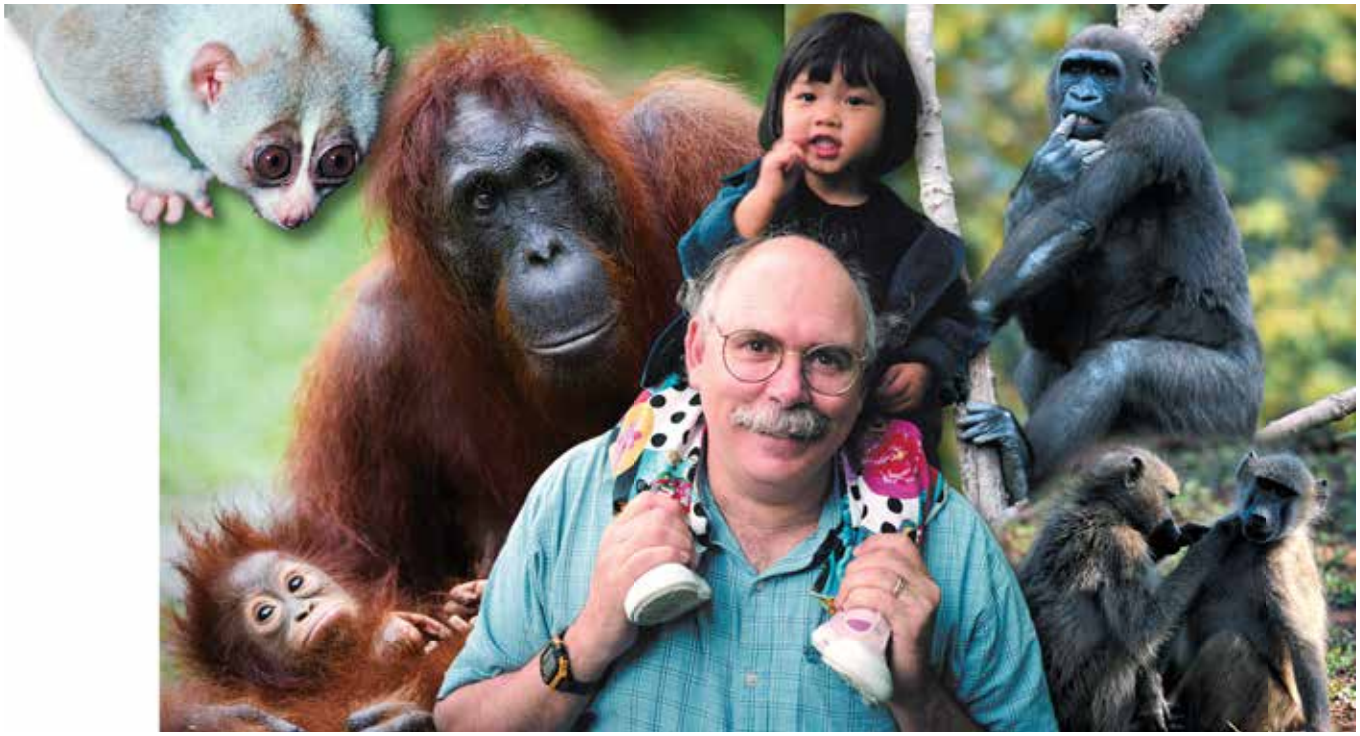


**FIGURE 12.7** Some members of the order Primates. One primate has successfully occupied a wide range of habitats in tropical, temperate and polar regions. Which primate is this?

Living primates include:

- prosimians — these include lemurs, lorises and tarsiers
- New World monkeys, from South and Central America — these include capuchin monkeys (*Cebus* spp.), howler monkeys (*Alouatta* spp.) and spider monkeys (*Ateles* spp.)
- Old World monkeys, from Africa and Asia — these include baboons (*Papio* spp.), mandrills (*Mandrillus sphinx*), and macaque monkeys (*Macaca* spp.)
- apes — these include gibbons (*Hylobates* spp.) of South-East Asia, orang-utans (*Pongo pygmaeus*) of Borneo and Sumatra, gorillas (*Gorilla* spp.), chimps (*Pan* spp.) of Africa, and humans (*Homo sapiens*).

**ODD FACT**

New World monkeys, from South and Central America, have nostrils that are widely separated and face sideways. Old World monkeys, from Africa and Asia, have nostrils that are closer together and face forward or downward. New World monkeys have prehensile (grasping) tails. Old World monkeys have tails, but they are not prehensile.

### What is a primate?

No single feature is diagnostic of a primate but, taken together, the following features identify a primate.

- Hands and feet of a primate typically have:
  - five digits that can grasp or curl around objects, and primate thumbs or big toes (sometimes both) are **opposable**; that is, they are able to be brought into contact with the other digits (see figures 12.8a and b)
  - flat nails on their digits, and, with the tips of the digits having many touch receptors, are able to gain information.

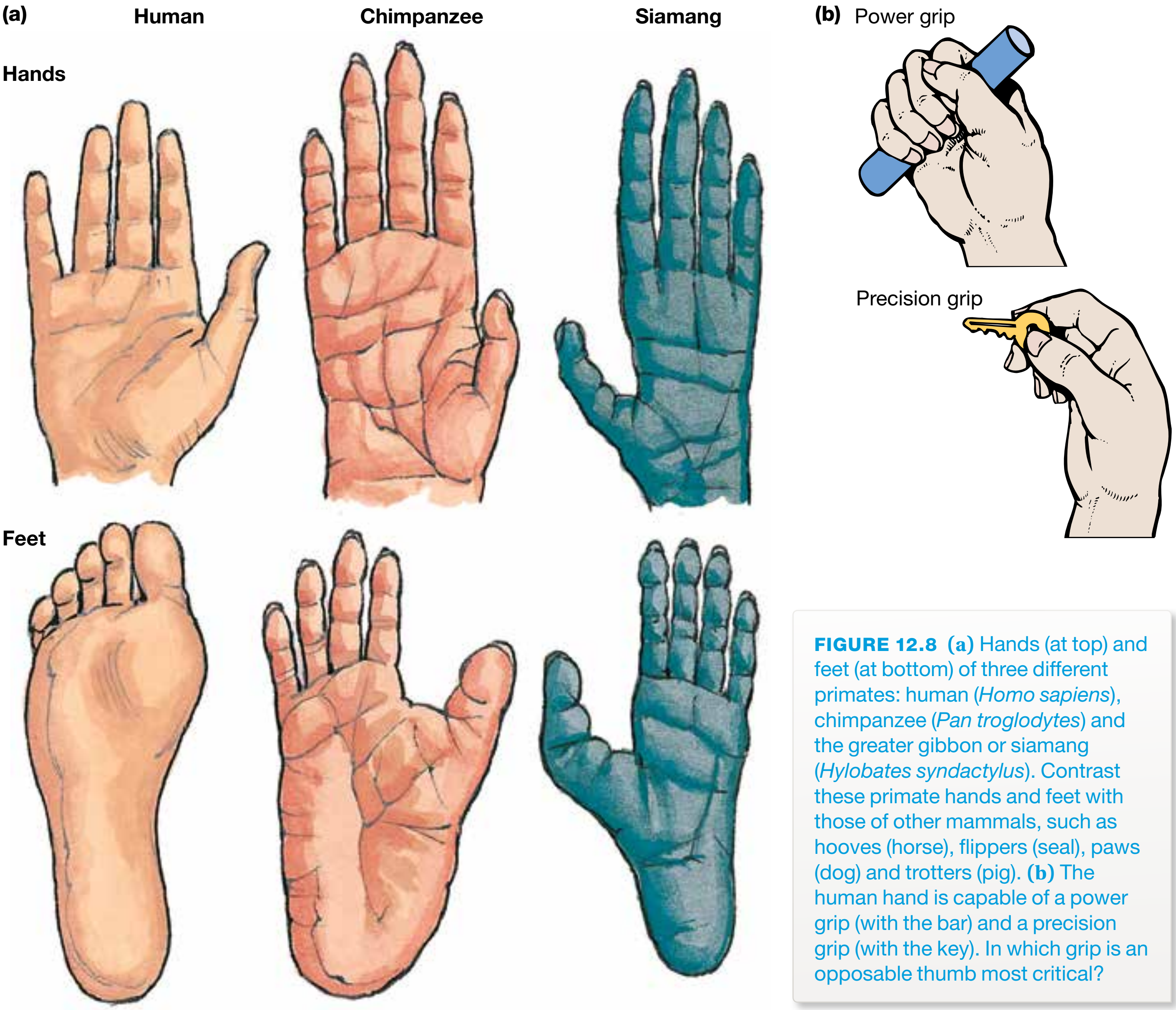



**FIGURE 12.8** **(a)** Hands (at top) and feet (at bottom) of three different primates: human (*Homo sapiens*), chimpanzee (*Pan troglodytes*) and the greater gibbon or siamang (*Hylobates syndactylus*). Contrast these primate hands and feet with those of other mammals, such as hooves (horse), flippers (seal), paws (dog) and trotters (pig). **(b)** The human hand is capable of a power grip (with the bar) and a precision grip (with the key). In which grip is an opposable thumb most critical?

**ODD FACT**

Of the primates, only humans have white sclera around the irises of their eyes. This feature enables people to tell when someone is looking at them.

- Primates have an emphasis on vision. They have:
  - large forward-facing eyes that give stereoscopic (3D) vision (see figure 12.9)
  - colour vision because they have three types of cone in the retinas of their eyes
  - a protective bone at the outer side of the eye socket.

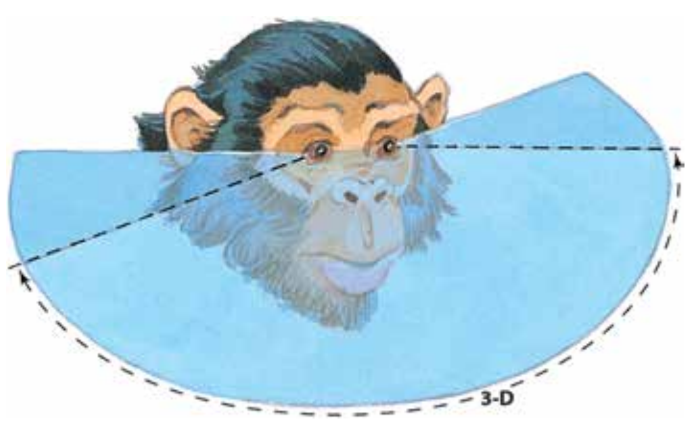



**FIGURE 12.9** Stereoscopic vision results from forward-facing eyes. The zone marked '3D' is the part of the visual field where depth of perception is possible.

studyon

| | |
|---|---|
| Unit 4 | **Primate characteristics** |
| AOS 1 | Summary screen and practice questions |
| Topic 4 | |
| Concept 1 | |

- Primate skeletons have flexibility adapted to life in the trees, such as the ability to rotate the arm in the shoulder socket so that the arm can extend behind the body.
- Compared to other mammals, primates have large brains relative to their body sizes — we primates are more brains than brawn! In addition, when a primate brain is compared to a brain of similar size from a nonprimate mammal, the primate brain is seen to have several times as many neurons as the nonprimate brain. Primate brains have a greater concentration of neurons per unit volume of brain tissue than is the case for nonprimates. Table 12.1 shows a comparison between the brains of several monkey species and brains of several rodent species. Note in particular the comparison between the capybara brain and that of the smaller brained capuchin monkey. Likewise, check the agouti brain with that of the similar-sized brain of the owl monkey.

**TABLE 12.1** Comparisons of the mass and the total number of neurons in brains from primates and from nonprimates, in this case rodents. (Data from Herculano-Houzel, S. *Frontiers in Human Neuroscience* 9 November 2009.)

| Species *Rodents* | Brain mass (gram) | Neurons (millions) | Species *Primates* | Brain mass (gram) | Neurons (millions) |
|---|---|---|---|---|---|
| * rat | 1.8 | 200 | • owl monkey | 15.7 | 1468 |
| * guinea pig | 3.8 | 240 | • squirrel monkey | 30.2 | 3246 |
| * agouti | 18.4 | 857 | • capuchin monkey | 53.2 | 3690 |
| * capybara | 76.0 | 1600 | • macaque monkey | 87.4 | 6376 |

- Primates are social mammals and typically live in groups that, depending on the species, may be as large as a troop of several hundred animals or as small as a pair (see figure 12.10).

**FIGURE 12.10** Primates, like these lemurs, are social mammals.

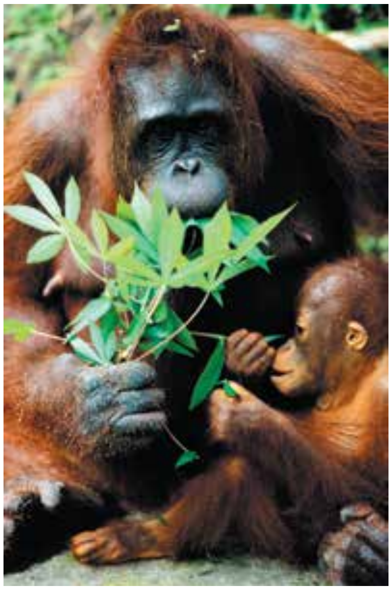


**FIGURE 12.11** All primate species, including the orang-utan, have a relatively long gestation period. Primates care for their young for an extended period after birth. Caring principally involves the mother but may also involve other female members of the social group.

- Primates have a relatively long **gestation period** compared to nonprimate mammals (see table 12.2). This longer gestation period allows for the more extensive fetal brain growth of primates. Primates typically produce a single young at each birth and provide parental care for an extended period (see figure 12.11), the longest period being in the case of human young.

**TABLE 12.2** Gestation periods for several mammals. Note that, for similar maternal body mass, a primate has a longer gestation period than a nonprimate mammal.

| Mammal | Average mass of adult female (kg) | Average gestation (days) |
|---|---|---|
| **Primates** | | |
| ring-tailed lemur (*Lemur cotta*) | 2.2 | 120 |
| owl monkey (*Aotus trivirgatus*) | 0.7 | 133 |
| white-cheeked gibbon (*Hylobates concolor*) | 7.5 | 210 |
| chimp (*Pan troglodytes*) | 40 | 237 |
| orang-utan (*Pongo pygmaeus*) | 36 | 251 |
| human (*Homo sapiens*) | 60 | 265 |
| **Nonprimates** | | |
| mouse (*Mus musculus*) | 0.3 | 21 |
| cat (*Felis catus*) | 5 | 60 |
| sheep (*Ovis aries*) | 70 | 150 |

## Humans are hominoids

Among the primates, humans are **hominoids** or members of the superfamily Hominoidea. Figure 12.12 shows the classification of members of the order Primates to the level of superfamily Hominoidea. **Humans share membership of the superfamily Hominoidea** with the lesser apes (gibbons, siamangs) and the other great apes (chimps, gorillas, orang-utans). So, all the apes are hominoids.

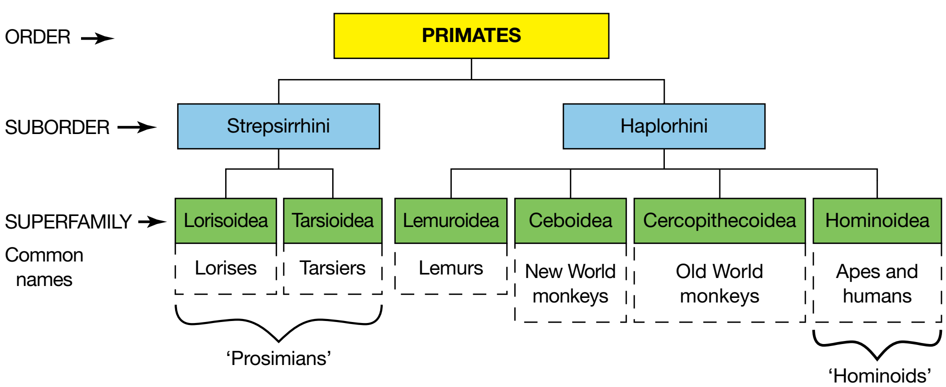



**FIGURE 12.12** Classification of primates to the level of superfamily. Note that some additional levels of classification are formed using prefixes such as 'sub' and 'super'. With which other animal groups do humans share membership of the superfamily Hominoidea?

In the Strepsirrhini group, the upper lip is firmly connected to the gums by a fold of tissue. As a result, the upper lips of lorises and tarsiers are far less mobile that those of monkeys and apes, making their faces less expressive than the Haplorhini, which have a free upper lip.

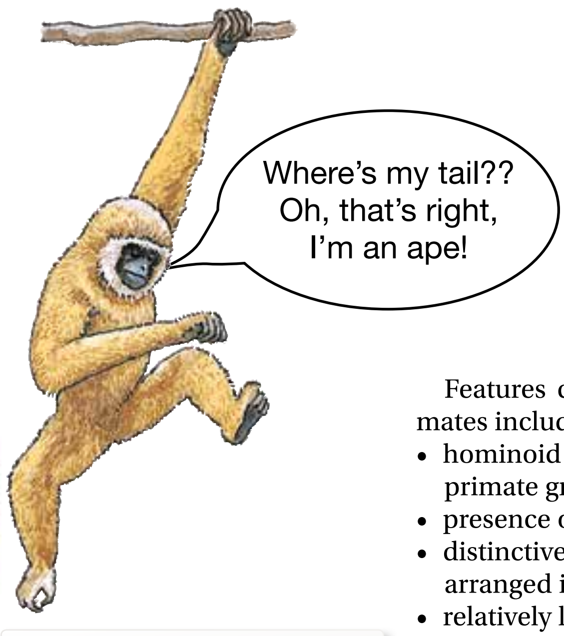



**FIGURE 12.13** Apes can be distinguished from monkeys by their wide chests and lack of a tail.

Note that the superfamily Hominoidea does *not* include monkeys. Can you tell a monkey from an ape? Easy! Monkeys have tails and narrow chests. In contrast, apes are tailless and have wide chests (see figure 12.13).

### What is a hominoid?

This taxonomic group, superfamily Hominoidea, includes all those primates that fit the biological definition of 'ape'. This means that the hominoids include the following living species: gibbons (14 species), orang-utans (2 species), gorillas (2 species), chimps (2 species) and humans.

Features common to hominoids that separate these apes from other primates include the following distinctive structural features:

- hominoid brains, which are larger and more complex than those of other primate groups
- presence of relatively broad palates and broad nasal regions
- distinctive molar teeth in the lower jaw with five cusps (raised bumps) arranged in a 'Y5' pattern (see figure 12.14)
- relatively long upper limbs
- ribcage that creates a wide chest
- axial skeleton that shows a reduced lumbar spine, an expanded sacrum and absence of a tail
- shoulder joints that permit the arms to be rotated around the shoulders, as in an overarm serve or bowling action.

**ODD FACT**

The action in a typical overarm tennis serve is possible only because the human shoulder joint permits the arm to be swung back behind the body. This is a relic of 'brachiation' in our primate ancestors — the ability to move by swinging from tree branches. All apes (and a few New World monkeys) can do this.

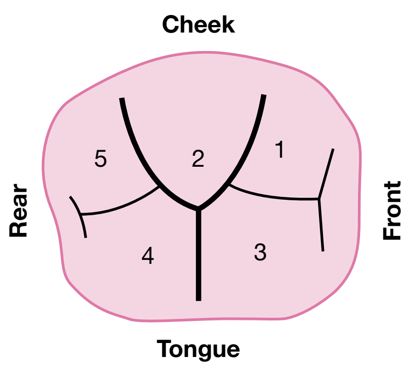



**FIGURE 12.14** Diagram showing an example of the Y5 pattern of five cusps (raised bumps) on the upper surface of a molar of a modern human. This molar pattern is also seen in apes.

**studyon**

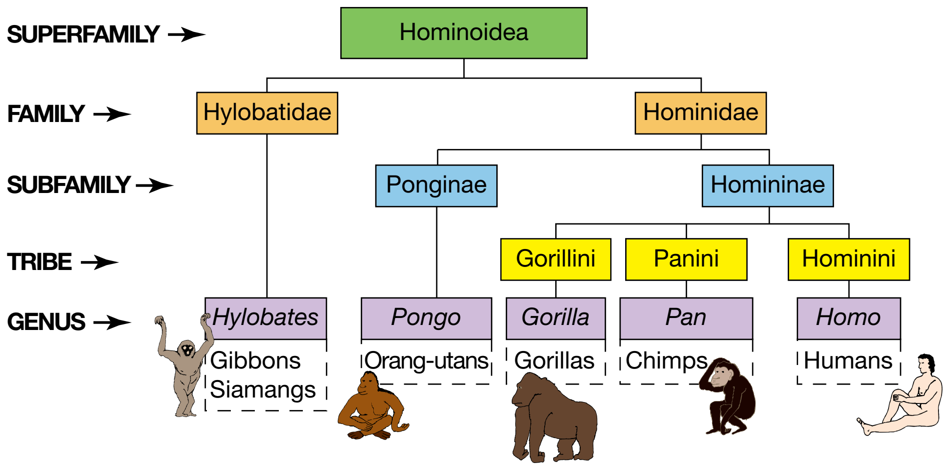
Unit 4 | AOS 1 | Topic 4 | Concept 2

**Hominoids and hominins**
Summary screen and practice questions

## Humans are hominins

Figure 12.15 shows the classification of hominoids or members of superfamily Hominoidea to the level of genus. Classifications are not fixed, but are subject to review and revision that can occur when new information comes to hand or when new interpretations provide a better explanation than the ones that they replace. Go to the box at the end of this section to see an example of how classifications can change. Older classifications placed modern humans on their own in the taxonomic group, family Hominidae, identifying us as the only living hominid species.

**Modern classification places humans in the taxonomic group known as tribe Hominini; that is, humans are hominins. It is only at this level of classification that humans are separated from all the other great apes**. Based on this classification, members of the genus *Homo* (humans) and their extinct erect-walking ancestors are termed *hominins*. Note that our closest living relatives are members of the tribe Panini, which includes the common chimp (*Pan troglodytes*) and the bonobo or pygmy chimp (*Pan paniscus*).




Members of tribe Hominini are called hominins, while members of the family Hominidae are called hominids.

**FIGURE 12.15** Classification of members of superfamily Hominoidea to the level of genus. Based on this classification, which group of great apes is most closely related to humans? Tribe Hominini includes modern humans as well as their extinct erect-walking ancestors; they are called hominins.

The currently accepted classification of modern humans starting at phylum Chordata is shown in figure 12.16.


**FIGURE 12.16** Modern classification of the human species. How would the classification of a chimp differ?

## What is a hominin?

A person walks across a room, sprints for a bus, stands in a queue. Each of these actions demonstrates that humans are hominins. **The term hominin refers to the modern human species and our extinct close relatives that could walk erect on their hind legs in a sustained fashion**. Such species have **bipedal** (meaning 'two-footed') locomotion, in contrast to other primates that walk on four limbs and are said to be quadrupedal (meaning 'four-footed').

**FIGURE 12.17** Gorillas use a so-called 'knuckle walk,' as do chimps.

The great apes, such as gorillas and chimps, typically walk with their knuckles on the ground (see figure 12.17).

## Footprints from the past

So how can we tell whether a fossil species walked erect?

Evidence of erect or bipedal walking can come from:

- the position of the hole (foramen magnum) in the base of the skull (see figure 12.25a on page 558)
- the arrangement of femur and tibia (see figure 12.31 on page 562)
- the shape of the pelvis (see figure 12.28 on page 560).

Other dramatic evidence for the erect walking of species within the genus *Australopithecus* comes from fossilised footprints. A trail of fossilised footprints extending for more than 20 metres was discovered in Laetoli in northern Tanzania in 1978. The co-discoverer of these footprints was Mary Leakey (see figure 12.18), a member of the Leakey family, three generations of whom have been active in the search for and discovery of hominin fossils.

The footprints were made one day about 3.6 Mya. One or possibly two adult hominins and a juvenile walked across a surface covered in damp volcanic ash (see figure 12.19). The sun dried the wet ash, preserving the footprints. These impressions were later covered by more volcanic ash and the material became compacted and hard. The feet that made these prints were like our feet, with the big toe parallel to the other toes. In contrast, footprints of African apes would show an opposable big toe (refer back to figure 12.8a on page 546).

**FIGURE 12.18** Mary Leakey, who worked for many years searching for hominin fossils. The Leakey family comprised Louis (1903–1972), his wife Mary (1913–1996), their sons Jonathan and Richard, their daughter-in-law Maeve, and their granddaughter Louise.

**FIGURE 12.19** The Laetoli footprints. This fossil evidence captured in rock supports the conclusion that hominins walked upright about 3.6 million years ago. These footprints were probably made by 'Lucy-type' hominins of the species *Australopithecus afarensis*, an erect-walking prehuman species.

## Classifications can change

Classifications are not fixed and unchanging. Classifications can change as additional information becomes available. Prior to 1980, for example, the accepted classification scheme identified humans as the only members of family Hominidae (see figure 12.20). This classification was based only on comparative anatomy and physiology and is no longer the commonly accepted view.

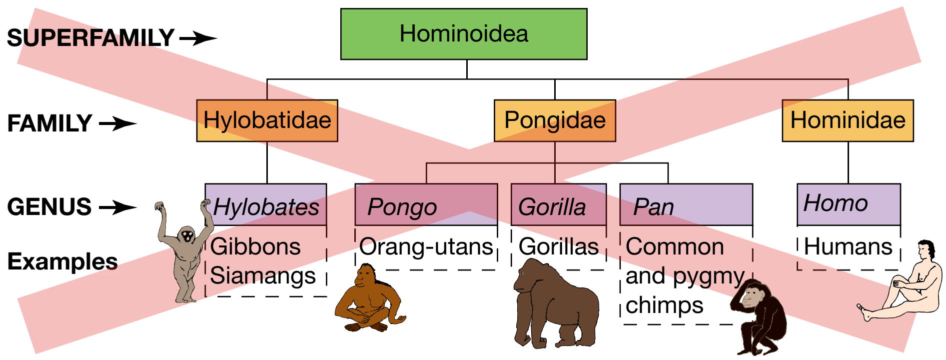



**FIGURE 12.20** A pre-1980 classification of the human species. This traditional classification placed humans in their own family (family Hominidae) and separated us from the other great apes.

According to this 'old' scheme, humans were the only living members of family Hominidae and were termed *hominids*. The other great apes were all members of the family Pongidae and they were termed *pongids*. However, molecular analysis, including genome sequencing, has identified a much closer relationship between humans and the other great apes so that they are now all placed in the family Hominidae (refer back to figure 12.15).

So, in the past, the term *hominid* referred to humans and their erect-walking ancestors, but *hominin* is now the preferred term that refers to this group (see figure 12.21). Are you a hominid or a hominin or both?

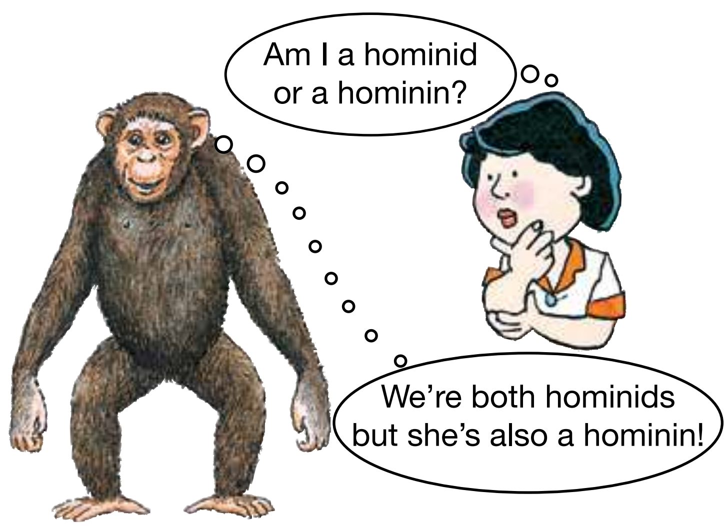



**FIGURE 12.21** Both the chimp and the person are members of the family Hominidae and so each is a hominid. Only the person is a member of the tribe Hominini and is the only living hominin.

## KEY IDEAS

- **Humans are members of the class Mammalia and share a distinctive set of features with all other mammals.**
- **Humans are members of the order Primates and share a distinctive set of features with all other primates.**
- **Humans and the other apes are members of the superfamily Hominoidea, and are called hominids.**
- **Humans are the only living members of the tribe Hominini, and are called hominins.**
- **Hominins, both living and extinct, are distinguished by the ability to walk upright.**
- **The term *hominin* includes modern humans and extinct species of the genus *Homo* and the genus *Australopithecus*.**
- **Classification schemes are not fixed but may change when new information becomes available or when new interpretations that provide better explanations are formulated.**

## QUICK CHECK

1. Identify the following as true or false:
   - **a** Chimpanzees have skeletal features that equip them for routine upright walking.
   - **b** The ability to walk upright is a distinctive feature of all hominins.
   - **c** All primates would be expected to show all the characteristic features of Class Mammalia.
   - **d** A monkey would be expected to have a greater brain mass than a dog with the same body mass.
   - **e** A mammal with stereoscopic vision must be a primate.
   - **f** Primate brains are characterised by a higher density of neurons than that in the brains of nonprimates.
2. What is the difference between:
   - **a** a hominin and a hominid
   - **b** bipedal and quadrupedal locomotion
   - **c** an Old World monkey and a New World monkey
   - **d** a monkey and an ape?
3. List three features that are characteristic of primates.
4. What is a power grip?

# Evolution of the primates

Humans are primates. What do we know about the evolution of the primates?

Today's primates include lemurs, tarsiers, New World monkeys, Old World monkeys and various apes, including humans. All the different kinds of primate living today gradually evolved over a period of at least 65 million years from ancestral primates.

At different times in the past, one line split or diverged to give rise to two different evolutionary lines, and so on. The form that is ancestral to two lines is said to be the **common ancestor** of the organisms that later evolved from these two lines. The times of divergence of various primate groups from a common ancestor have been inferred by comparing the groups in terms of their structure and shape (morphology) or in terms of the composition of particular proteins that they share (biochemical analysis), and in terms of similarities in their genome sequences and the organisation of their chromosomes. (Refer back to chapter 11, page 511.)

**ODD FACT**

An article in *Nature* (15 May 2013) reported the discovery of primate fossils that provides evidence for the split of the Old World monkey (OWM) line from the ape line approximately 25 Mya.

An early divergence produced one line that led to the primates represented today by lorises and tarsiers, which form a group known as prosimians (suborder Strepsirrhini), and the other line giving rise to the ancestral primates of suborder Haplorhini. Another divergence produced two lines, one leading to the New World monkeys and the second giving rise to the today's Old World monkeys and apes. The latter line subsequently diverged to produce the Old World monkey line (superfamily Cercopithecoidea) and the ape line (superfamily Hominoidea).

When the morphology and blood protein composition of the various apes were compared, it was concluded that the gibbon line diverged possibly 18 Mya, and the orang-utan (Asian ape) line possibly 14 Mya. Finally, the evolutionary line that led to humans diverged from the line that was ancestral to the African apes, first from the gorilla line (now represented by two species: *Gorilla gorilla*, the western gorilla, and *Gorilla berengi*, the eastern gorilla) and then from the chimp line (now represented by two species: *Pan paniscus*, the bonobo, and *Pan troglodytes*, the common chimp).

The phylogenetic tree of the major primate groups is well established, meaning that the order of the various divergences between these groups is accepted (see figure 12.22). However, universal agreement does not exist for the times of these divergences. Why? Genes (and their protein products) do not change at constant rates. This means that disagreements can arise, for example, when different scientists estimate times of divergence based on different molecular data. (Molecular data used to estimate divergence times include sequence changes in nuclear DNA or in mitochondrial DNA or in proteins.)

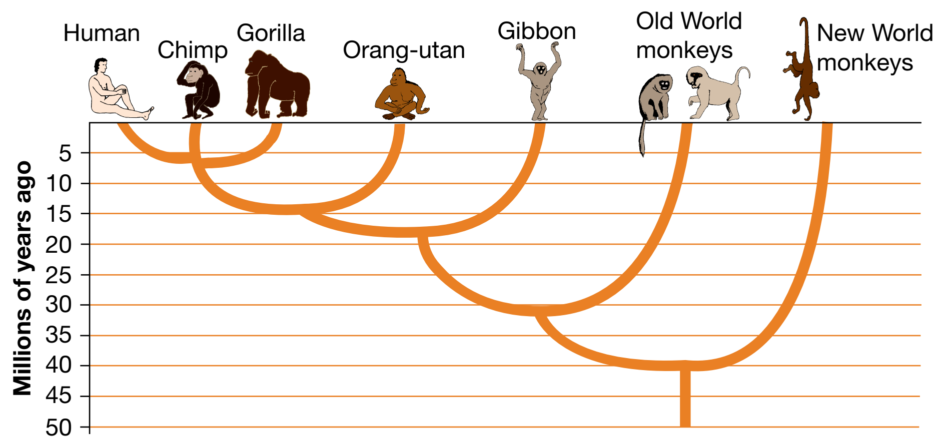



**FIGURE 12.22** One representation of the inferred evolution of primates. The pattern of divergence ('splits') of the various lines is universally agreed. Note that, based on different methods, the estimated times of divergence can differ (see marginal note for later estimates).

Estimates of divergences:

- human–chimp: 6 Mya
- human–gorilla: 7 Mya
- human–orang-utan: 13 Mya
- human–OWM: 23 Mya
- human–NWM: 33 Mya

Ref: Glazko, G.V. and Nei, M. 2003, *Mol Biol Evol* 20, pp. 424–34.

When scientists talk about the 'human' line diverging from the African ape line, this does *not* mean that humans evolved from chimpanzees or gorillas. It means that, at some time in the past, the 'human' line and the 'African ape' line shared a common ancestor. This common ancestor gave rise to two lines that evolved along different pathways, occupied different habitats and were subject to different selection pressures. The 'human' line evolved over millions of years and among its early representatives were prehuman hominins of the genus *Australopithecus*. Today, this evolutionary line is represented by one species, modern humans (*Homo sapiens*). The 'chimp' line, which evolved over the same period, is represented today by two species of chimp, the bonobo (*Pan paniscus*) and the common chimp (*Pan troglodytes*).


**FIGURE 12.23**

## When did the 'human' line diverge?

The 'human' line means an evolutionary line that finally led to humans via ancestral hominins, probably including ones of the genus *Australopithecus*, such as Lucy. When did this line diverge from the line that gave rise to the modern chimps?

The answer to this question can only be inferred. However, the time since divergence of the 'human' line from the African ape line can be estimated using the **molecular clock principle** formulated by Vincent Sarich and Allan Wilson in 1967. **The molecular clock principle assumes that changes in a given protein occur at a constant rate over time so that the greater the number of differences in the same protein from two different species, the longer the period since divergence of the ancestral lines of those species.**

Using the differences in the amino acid sequences of the same protein from humans and from chimps, Sarich and Wilson estimated that the human line diverged from the African ape line about 5 to 7 Mya; that is, the **last common ancestor (LCA)** of humans and African apes lived about 5 to 7 Mya. Comparisons of the DNA sequences of the living apes performed later also gave a similar result for the postulated time of divergence. Another view, taking into account the fact that the molecular clock does not run at a constant rate, suggests that this divergence might have occurred no earlier than 7 Mya, perhaps in the period 7 to 10 Mya.

The molecular clock concept was discussed in chapter 11, page 518.

## When did the first humans appear?

There is no universal agreement about the precise evolutionary history of our human species. However, there is agreement that small-brained hominins separated from the line that led to the gorillas and chimpanzees, possibly about 7 to 10 Mya, and developed the ability to walk erect. It is agreed that these hominins gradually became less ape-like and more human-like as generations of hominins spent more time at ground level and were subjected to various selection pressures, including climate change. Note that these first hominins were not humans; that is, they were not members of the genus *Homo*. Instead, they were prehuman hominins, one major group being members of the genus *Australopithecus* (see below).

Over millions of years following the divergence of the human line from the African ape line, the fossil record reveals many changes that occurred over time in these prehuman hominins, as follows:

- their brains enlarged
- their faces shortened and flattened
- their jaws and teeth became smaller
- their legs grew longer and became much stronger than their arms
- their feet became longer and developed arches
- their ability to fashion tools began.

Finally, some hominins reached a stage where they were identifiably human and could be classified as members of genus *Homo*. These first members of genus *Homo* were not modern humans like us, but they had features that distinguished them from prehuman hominins and qualified them as human. Many different *Homo* species evolved before the emergence of our species, *Homo sapiens*. We will explore the transition from prehuman hominin to human later in this chapter (see page 574).

**Fossil evidence announced in 2015 now places the time of appearance of the first members of genus *Homo* at about 2.8 Mya**. This fossil, a human jawbone, was found in Africa, in the Afar region of Ethiopia about 20 kilometres from where the Lucy fossil was found in 1974, and its discovery has pushed back the appearance of the first humans in the fossil record by almost half a million years. The previous oldest human fossils were dated at 2.3 million years old.

In the next section, you will meet some famous prehuman hominin fossils and identify some key locations where hominin fossils have been found.

**KEY IDEAS**

- Living primates evolved from an ancestral primate species over a period starting from at least 65 Mya and perhaps as long as 85 Mya.
- During primate evolution, lines that gave rise to different groups of modern primates diverged at various times.
- The molecular clock estimates the line that led to modern humans diverged from the line that led to modern chimps about 5–7 Mya.
- Estimates of times of divergence use molecular data such as changes in DNA sequences and in the amino acid sequences of proteins.
- Fossil evidence reveals a series of gradual physical changes in the early prehuman hominins making them less ape-like and more human-like.
- The first members of the genus *Homo* appear in the fossil record about 2.8 Mya.

**QUICK CHECK**

5 Identify the following as true or false:
  a The line that gave rise to Old World monkeys diverged after the appearance of the line that gave rise to orang-utans.
  b The oldest monkey fossil would be expected to be more recent than the oldest ape fossil.
  c Humans evolved from chimps.
  d The line that gave rise to modern gibbons diverged from the ape line before the orang-utan line.
  e The oldest fossils of *Homo sapiens* are dated at about 2.8 million years old.

6 Identify one line of evidence that may be used to estimate divergence times.

7 Identify three changes that can be seen in the prehuman hominin fossils over time.

# Hominin evolution: where did it begin?

With the acceptance of Charles Darwin's theory of evolution by the late nineteenth century, the scientific world awaited the discovery of the so-called 'missing link' that would be neither ape nor human but would be intermediate between both.

Based on physical and behavioural similarities that he observed between humans and modern African great apes, Darwin speculated that Africa was the cradle of humanity. Other scientists at that time thought that the Asian orang-utan was the closest relative of the human species and that Asia was where the human species had evolved from prehuman ancestors. So, where did hominin evolution begin?

Until the late nineteenth century, no hominin fossils had been discovered that might provide evidence to resolve this question and distinguish between the Africa and the Asia alternatives. The discovery of the fossil known as

Java Man (*Homo erectus*) in Trilin, Indonesia, in 1891 gave ammunition to the proponents of an Asian origin of the hominins.

However, it was the discovery of a fossil in a quarry in South Africa in 1924 that provided the first evidence that led to the conclusion that Africa was the continent where the earliest hominins emerged and where the line that led to human species diverged from prehuman hominins. This fossil, as well as later fossil discoveries in Africa, showed more primitive features than any hominin fossils found in Asia. These African fossils, in terms of both their physical features and their estimated ages, were undoubtedly prehuman hominins that lived far earlier than any Asian fossil candidate. Let's meet some of these significant fossils.

## The child from Taung

It is November 1924. An explosion rings out in a limestone quarry at Taung in Cape Province in South Africa. One rock fractured by the explosion reveals part of a skull and, remarkably, a petrified replica of a brain. Sixty years later, in 1984, this fossil skull, known as the Taung child, will be labelled as one of 'the twenty discoveries that changed our lives in the twentieth century'! Why?

The skull in the rock was taken to Johannesburg and given to Raymond Dart, a 32-year-old Australian who was Professor of Anatomy at the University of Witwatersrand (see figure 12.24). Because of his training in neuroanatomy, Dart's attention was drawn to the structure of the brain, which he immediately realised was different from that of living apes. He painstakingly extracted the remainder of the fossil from the rock and carefully examined the cranium (brain case) and the mandible (lower jaw).

**ODD FACT**

Raymond Dart (1893–1988) came from a cattle-raising family in rural Queensland. He is reputed to have neglected his farm duties one afternoon because he was dissecting a rooster. Dart graduated from the University of Queensland and then went to the University of Sydney where he completed a medical degree in 1917.

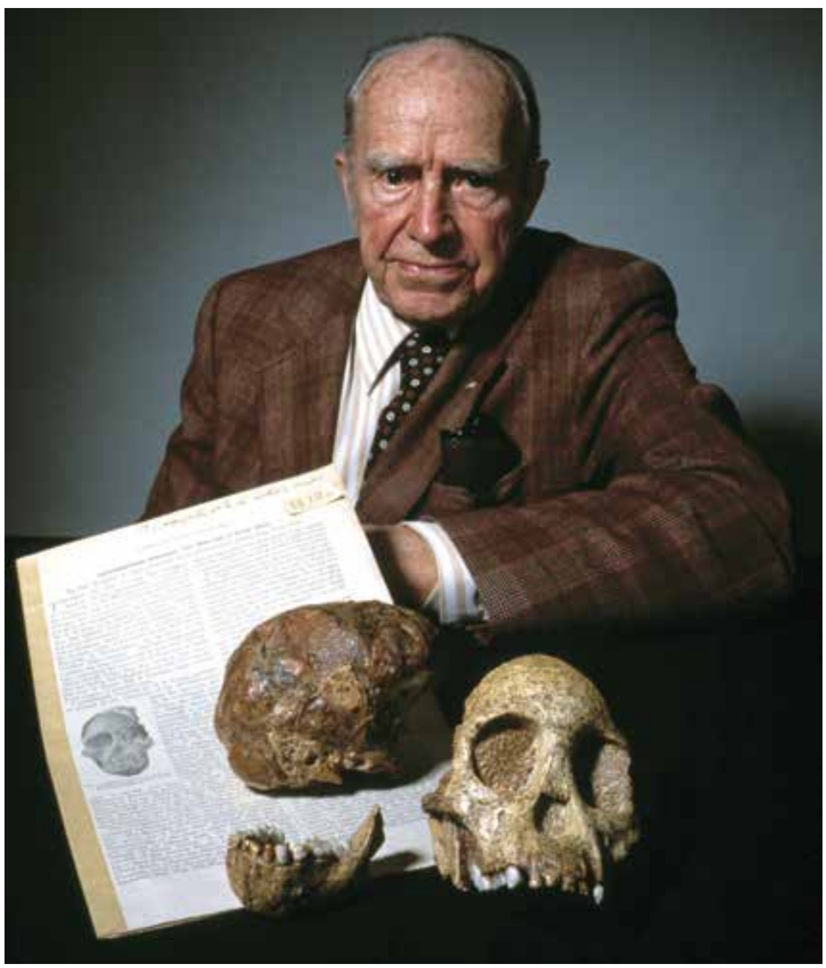


**FIGURE 12.24** Professor Raymond Dart identified and named the hominin species *Australopithecus africanus* based on his analysis in 1924 of the fossil skull of the Taung child. This photo shows Dart in 1978 with this skull. Dart was a palaeoanthropologist (from the Greek: *palaeo-* = ancient; *anthropos* = man; and *logia* = branch of knowledge).

Because the molar teeth were just erupting, Dart concluded that the fossil was that of a child and, because it possessed both ape-like features and human-like features, Dart concluded that the fossil skull was intermediate between living apes and humans (see table 12.3). The Taung child was classified by Dart as a new species and he gave it the scientific name *Australopithecus africanus* ( = 'southern ape from Africa').

**TABLE 12.3** Some comparisons between a human skull and those of the Taung child and a gorilla.

| Feature | Gorilla | Taung child | Human |
|---|---|---|---|
| brain size | small | larger | largest |
| size of canine teeth | largest | small | smallest |
| shape of dental arch | box-shaped | more rounded | parabolic |
| gap (diastema) in tooth rows | present | absent | absent |
| position of foramen magnum | towards back of skull | further forward | most forward |

From the evidence of the position of the foramen magnum (= 'big opening') on the undersurface of the cranium (see figure 12.25a ), Dart concluded that the Taung child walked upright. The foramen magnum is the hole at the bottom of the cranium where the spinal cord leaves the brain and enters the vertebral column (backbone). Its position indicates how the skull sat on the vertebral column. Where the foramen magnum is towards the front, the organism walks upright; where it is at the back of the skull, the organism walks stooped on all fours.

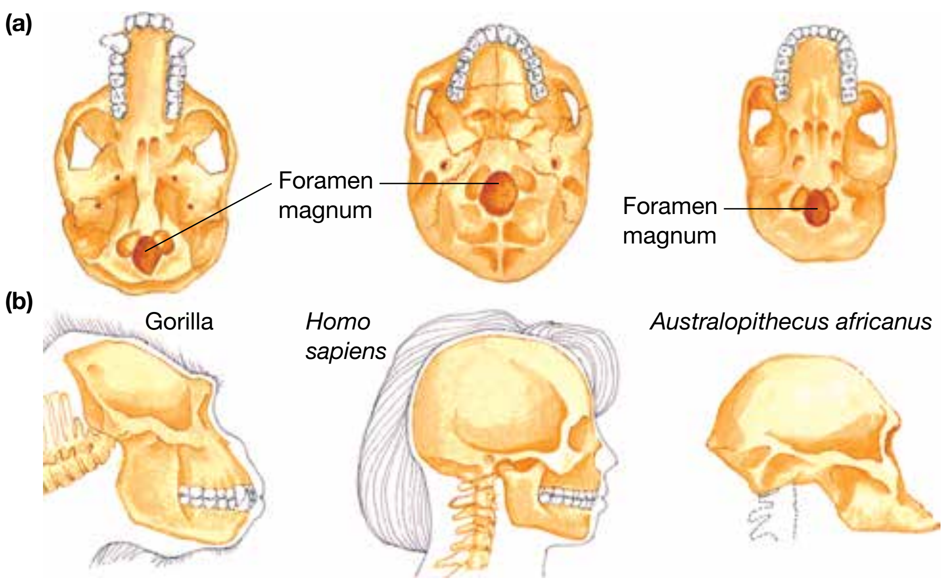



**FIGURE 12.25** **(a)** Basal view of cranium showing the relative position of the foramen magnum. Notice that the position is much further forward in humans, who walk erect, than in gorillas, who typically walk on all fours. In *A. africanus*, the position of the foramen magnum is closer to the front than in the gorilla and indicates that this species could walk upright. **(b)** Side view showing skull on vertebral column.

Also, refer to figure 12.25a above and look carefully at the upper jaw of each cranium, noting the shape formed by the teeth, the so-called dental arch. The shape of the dental arch in the gorilla is box-like and forms three sides of a rectangle. In contrast, the shape in the Taung child species (*Australopithecus africanus*) is more rounded, and the shape in a human is parabolic (as listed in table 12.3 above).

In February 1925, Dart's report on the skull was published in the journal *Nature*. His conclusion that the Taung child represented a new species and was a link between apes and humans was rejected by the leading British scientists

**FIGURE 12.26** Photograph of Robert Broom (seated at right) and a few of his fossil specimens and drawings. As well as his discoveries of *Australopithecus africanus* fossils, Broom also found fossils of another hominin, which he called *Paranthropus robustus*. This species is characterised by heavier facial bones and very large molar teeth.

of that time. As a result, the Taung child fossil was ignored until 1947. In that year, Le Gros Clark, a leading British scientist from Oxford University, visited Johannesburg where he examined many fossils, including the Taung child. He realised that this fossil was not an ape and that it had several human-like features. This fossil was again back in the spotlight as an accepted hominin on the human evolutionary line. Being on the human evolutionary line does not make a species a direct ancestor of the modern human species. This evolutionary line is not a straight line (like a pole), but is highly branched (like a bush), with many dead ends.

The Taung skull was the first hominin fossil to be discovered in Africa. It provided the earliest physical clue to possible ancestors of the human species and is now regarded as one of the twentieth-century discoveries 'that changed our lives'. Because quarrying destroyed the exact location where it was found, the age of the Taung fossil could not be determined precisely but it was estimated at 2.8 million years old.

The Taung child was *not* human like us. It was a hominin species within an extinct genus (*Australopithecus*), which, along with species in our genus (*Homo*), could walk erect.

Go to the box at the end of this section to read why the leading scientists of the time rejected the proposal that the Taung child fossil was a hominin.

## Then came Mrs Ples

Another hominin fossil hunter was Dr Robert Broom (1866–1951), a Scottish doctor (see figure 12.26). Broom was nearly 70 in 1936 when he and a co-worker unearthed a fossil skull at Sterkfontein, South Africa. Broom recognised that this skull was very similar to Dart's Taung child fossil. The skull was later identified as that of an adult of the same species as the Taung child, namely *Australopithecus africanus*.

A comparison of the skull of an adult *Australopithecus africanus* with those of a chimp and a modern human is shown in figure 12.27.

**study on**

| Unit 4 | **Early hominin fossils** |
|---|---|
| AOS 1 | Summary screen and practice questions |
| Topic 4 | |
| Concept 3 | |

**(a)** Chimp **(b)** *Australopithecus africanus* **(c)** Human

**FIGURE 12.27** Disarticulated skulls of **(a)** chimp (*Pan troglodytes*), **(b)** adult *Australopithecus africanus,* and **(c)** human (*Homo sapiens*) (not drawn to same scale).

In 1947, Broom, then in his eighties, and a co-worker found the first partial skeleton of an *A. africanus* fossil. This fossil is commonly known as Mrs Ples and has been dated at 2.5 to 2.1 Myr old. (Can you suggest why this fossil might have been identified as that of a female, rather than a male?)

The structure of its pelvis confirmed that this species was capable of bipedal locomotion (see figure 12.28). This conclusion had already been reached based on the position of the foramen magnum in the fossil skulls found earlier (refer back to figure 12.25a).

**(a)** **(b)** **(c)**

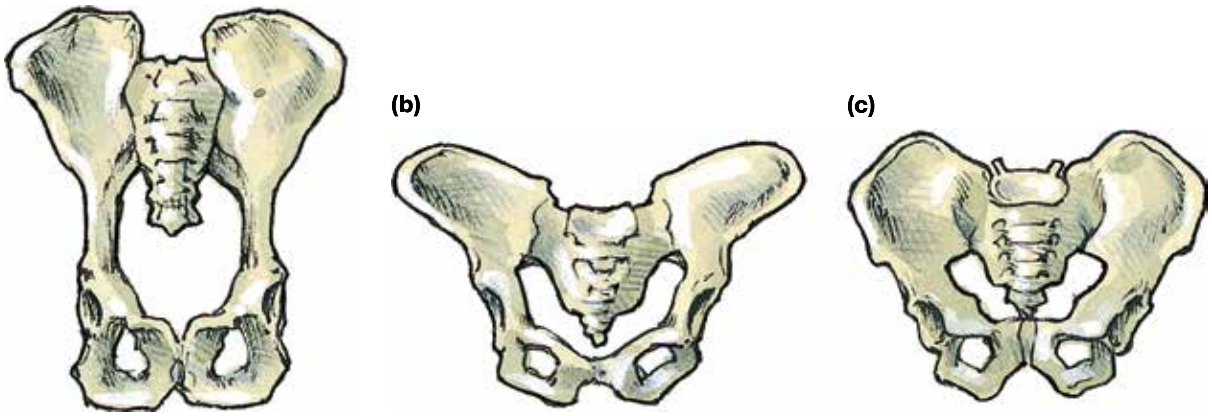


**FIGURE 12.28** Structure of the pelvis of **(a)** a chimp, **(b)** *Australopithecus africanus,* and **(c)** a modern human. In hominins, which can walk erect, the pelvis is bowl-shaped and the hip bone is short compared to that of the chimp. Note the similarities in **(b)** and **(c)**, which indicate that *Australopithecus africanus* was capable of bipedal locomotion.

The Taung child and the Mrs Ples fossils are hominin species classified as *Australopithecus africanus*. These fossils provided evidence that Africa was where the early hominin species appeared. Members of this species lived from about 3 million to 2 million years ago. What came before these hominins? It was the discovery of another hominin fossil, in 1974 in the Afar Triangle region of Ethiopia, that provided evidence of the existence of another member of this genus that lived even earlier. This hominin has the scientific name *Australopithecus afarensis*, and is popularly known as Lucy. Meet her in the following section.

**studyon**

| | |
|---|---|
| Unit 4 | **Hominin origins** |
| AOS 1 | Summary screen and practice questions |
| Topic 4 | |
| Concept 5 | |

## Lucy from the Afar Triangle

It is 1974. A group of scientists are on a field trip in the **Afar Triangle**, a remote area of Ethiopia about 160 kilometres north-east of Addis Ababa. Donald Johanson, leader of the expedition, drives to an ancient lake bed to look for fossils. By noon, the temperature is nearly 44 °C. In a gully he finds part of an elbow joint, then part of a skull, a femur (thigh bone), several vertebrae, parts of a pelvis (hipbones), some ribs, and other bones and bone fragments. The impact of this remarkable discovery gradually dawns on Johanson. He has found bones from the skeleton of a single individual.

Eventually, Johanson's team excavated several hundred pieces of bone that formed about 40 per cent of an entire skeleton (see figure 12.29). The skeleton was that of an ape-like creature who walked erect, who was a little more than one metre tall, and who lived in this part of Africa about 3 million years ago. Because part of the pelvis was present, it was possible to identify the bones as being from a female. The condition of her teeth suggested she was 25 to 30 years old when she died.

This fossil skeleton was given the popular name **Lucy**. Lucy was not a human but a hominin. **Lucy is classified as a member of the species *Australopithecus afarensis***. Lucy is accepted by many palaeoanthropologists as being a possible early ancestor of the human species and on the direct line from which the first human species gradually evolved.

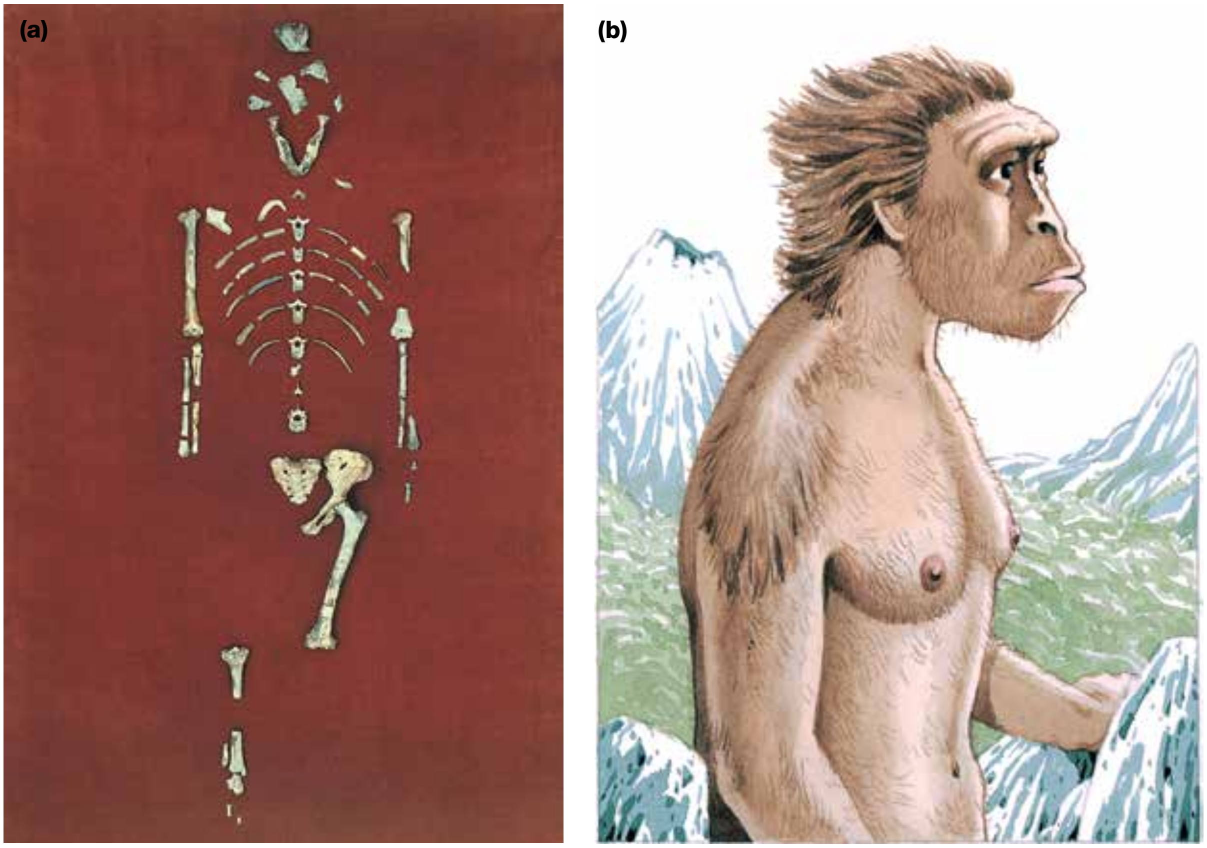



**FIGURE 12.29** **(a)** The remarkable hominin skeleton of Lucy, possibly a direct ancestor of the human species. (Image (a) courtesy of the Institute of Human Origins.) **(b)** A reconstruction of Lucy. This fossil skeleton obtained its popular name of 'Lucy' because, during the celebrations that occurred on the evening of her discovery, a tape recorder was playing the Beatles song 'Lucy in the sky with diamonds'.

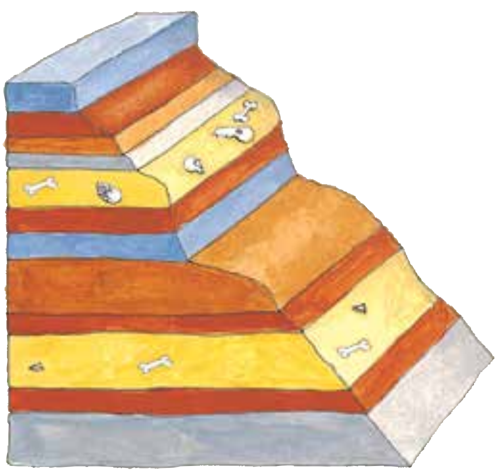


**FIGURE 12.30** Simplified version showing sediment interspersed by layers of volcanic debris known as tuff. Because the tuff layers result from volcanic activity, they can be accurately dated using radiometric dating. Bones found within a tuff layer are assumed to have the same age as the tuff layer.

When a fossil is found, the story that it tells is not apparent without careful examination and interpretation. We will now look at how the story of Lucy was revealed through the use of various techniques.

## When did Lucy live?

It is estimated that Lucy lived about 3 million or so years ago. The age of Lucy was determined indirectly by using the decay of radioactive isotopes to measure the ages of layers of volcanic material that lay above and below the sediment in which the fossil skeleton was located (see figure 12.30).

## Did Lucy walk erect?

**In primates that walk erect, such as humans, the femur (thighbone) is at an angle to the tibia (shinbone)** (see figure 12.31a). As a result, our knees and feet lie below the centre of mass of the body, and the main weight of the body falls on the outside of the knee joints. This arrangement, combined with a locking knee joint, enables us to stand erect with straight thighs and to walk erect because we can support our body weight on one leg as the other leg steps forward.

The angle made by the vertical with the femur forms the bicondylar angle (see figure 12.31b). In modern humans, the bicondylar angle is in the range of 8 to 11 degrees.

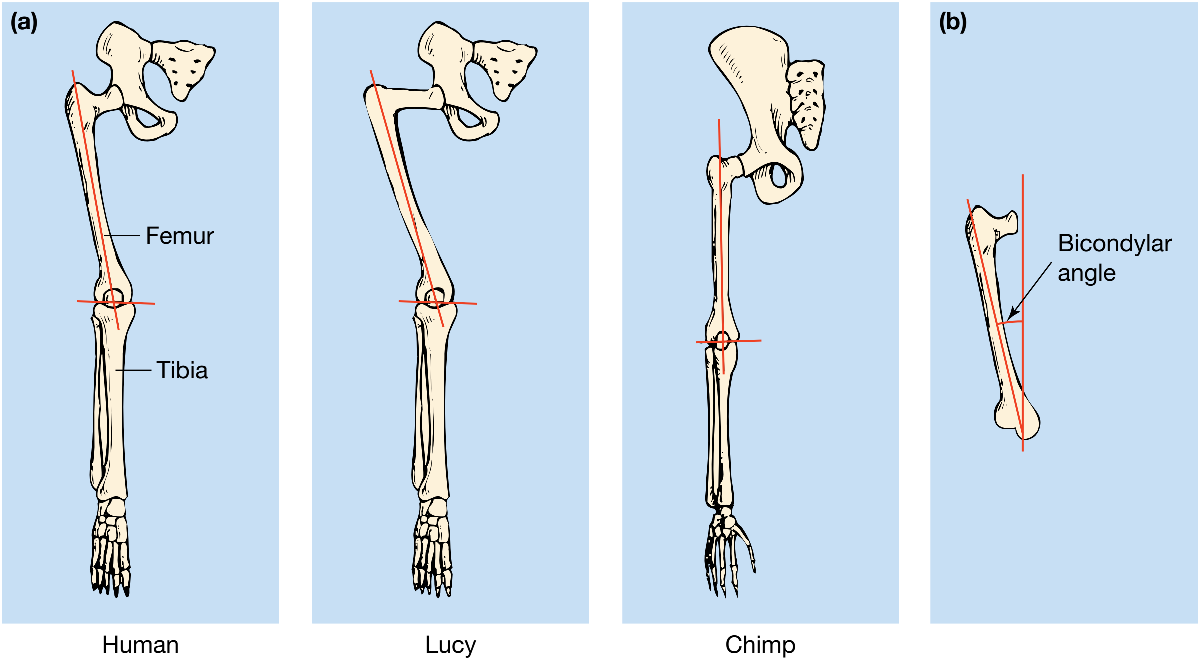



**FIGURE 12.31** **(a)** Arrangement of femur and tibia in (at left) a modern human, (at centre) in Lucy, and (at right) in a modern chimp. Note the outward slant of the femur relative to the tibia of the human and of Lucy, as compared with the straight-line arrangement of the chimp. **(b)** Viewing the bicondylar angle.

Chimpanzees (and the other great apes) cannot walk erect for a sustained period. Their femurs and tibias join in a straight line (see figure 12.31a, at right). The bicondylar angle in these prehuman great apes is about 1 degree. As a result, when erect, the knees and feet of the great apes lie outside the centre of mass of the body, and the main weight of the body falls on the inside of the knee joints. In attempting to walk erect, chimps lean forward with widely spaced legs and feet. Their natural mode of locomotion is a so-called **knuckle walk** in which the back feet and the knuckles touch the ground.

In Lucy's skeleton, the femur is angled out from the line of the tibia and provides evidence that she walked erect (see figure 12.31a, at centre). In Lucy's kind (*Australopithecus afarensis*), the bicondylar angle is larger, about 14 to 15 degrees. This angle is greater than in humans. This is the case because Lucy and her kind were much shorter than modern humans and a greater angle is required to bring the centre of mass within the outline of the feet.

Another fossil knee joint from about 3.5 Mya also shows that same angled arrangement of femur and tibia. So, **erect-walking primates have existed in Africa for at least 3.5 million years**. Evidence for erect walking also comes from the location of the so-called foramen magnum of the skull. Where would the foramen magnum have been located in Lucy?

Lucy and other *Australopithecus* species walked erect with their big toes oriented parallel to their other toes, as in the feet of modern humans. (Refer back to figure 12.8a on page 546 to see the contrasting position of the big toe in chimps.) While Lucy and her kind were capable of bipedal locomotion, it was not the long striding gait that is typical of modern humans. The skeletons of *Australopithecus* species show features that are also adaptations for tree climbing, such as long curved bones in the fingers and the toes, indicating that they still spent considerable time in trees.

Some scientists speculate that Lucy and her kind could move by swinging from branch to branch with their arms, in a mode of locomotion known as **brachiation**. Lucy's skeletal structure also allowed her to stand and walk erect when she descended to the ground, perhaps to move between clumps of trees or to seek food at ground level. In light of these findings, the conclusions now accepted are:

- the skeletal changes that allow erect walking evolved more than 3.5 million years ago, long *before* the first appearance of the genus *Homo* in the fossil record
- these skeletal changes were present in hominins that had a small brain size similar to nonhuman apes and much smaller than the human brain
- erect walking was present in hominins for several million years before the first stone tools appear in the fossil record.

These findings disprove an earlier hypothesis that human evolution was driven by selection pressure that gave an advantage to erect-walking hominins who could make greater use of tools and so hunt more effectively. The fossil evidence does not support this view.

## Discoveries of hominin fossils continue

Hominin fossils continue to be discovered and the pace of discovery has quickened in the last several decades. Each discovery of a fossil hominin may provide new information that can be used to test currently accepted hypotheses. Often this results in clarification and increased understanding of hominin evolution. In other cases, the discovery of a hominin fossil belonging to a new species may raise more questions about the pathway of human evolution. The dating of a newly discovered fossil of an existing species may extend the currently accepted age range of that species and lead to a new interpretation about human evolution; for example, in the preceding section, we saw that the discovery of a new fossil in 2015 changed the accepted date of the first appearance of humans in the fossil record. Some fossils are very fragmentary so that disagreement may exist as to whether they should be identified as hominins or as apes.

In the following sections, you will meet the prehuman hominins.

### REJECTING THE EVIDENCE OF THE TAUNG CHILD

Why was the Taung child originally rejected as being a link between apes and humans? This was due to the fact that the Taung child fossil did not fit the expectations and preconceptions of many influential British scientists of that time. These scientists expected that early hominins would:

1. have a large brain like a human and a jaw with large ape-like teeth
2. be found (that is, their fossils) in Asia, which was then believed to be the place where the first hominins evolved, because of the discovery of the Java Man fossil in 1897 in Indonesia.

Since the Taung child did not fit these expectations, it was ignored by most scientists for more than 20 years after its discovery. In particular, the Taung child had an estimated average brain volume in adults of 440 cubic centimetres compared with an average value of 1300 cubic centimetres for modern humans. As such, the Taung child did not meet the preconceived idea that early hominins would have large brains. In 1925, Sir Arthur Keith, a leading and highly influential British scientist, stated that the claim that the Taung child fossil represented a human ancestor was 'preposterous'. Keith's status and seniority meant that the Taung child fossil was regarded as an aberrant ape, and quietly ignored.

Because of this flawed preconception that early human ancestors must be large brained, one of the deliberate scientific frauds of the twentieth century was not immediately recognised as a fraud. This was the Piltdown forgery, a contrived hominin fossil, known as Piltdown Man, that was created using a modern human **cranium** and the lower jaw of an orang-utan that had been stained to look old**, and the molar teeth had been ground down to make them more human-like**. In addition, a key piece of the lower jaw where the jaw articulates with the cranium was missing. Had this been present, it would have shown that the lower jaw did not belong to the cranium.

The skull and jawbone of the Piltdown Man and various other objects including teeth and animal bones were 'discovered' during 1911 and 1912 in a gravel pit in southern England (see figure 12.32). Leading scientists of the day identified these bones as those of the missing link between apes and humans and named this new 'species' *Eoanthropus dawsonii* (Dawson's Dawn Man).

*(continued)*

Clearly, this fraud was designed to meet the preconceptions of the British palaeoanthropologists who celebrated its discovery as '*the most important ever made in England, and . . . any other yet made, either at home or abroad*'.

Doubts about the authenticity of the Piltdown Man began to circulate, but it was not until 1953, more than 40 years later, that, through the use of chemical tests to identify its age, the Piltdown Man was shown unequivocally to be a fraud, and indeed a clever and deliberate hoax.

Meanwhile, in 1947, Sir Arthur Keith (see above) finally conceded the rightful place of the Taung child fossil in hominin evolution, when he wrote: '. . . Professor Dart was right and I was wrong.'


**FIGURE 12.32** Photograph taken in the period 1912–1913 of the Piltdown site. Charles Dawson, the amateur scientist who made the 'discovery' of the cranium and the jaw is shown seated at left. At right is Sir Arthur Woodward, Keeper of Geology at the British Museum.

## KEY IDEAS

- The Taung child is a famous hominin fossil, species *Australopithecus africanus*, dated at about 2.8 million years old.
- Mrs Ples is a partial fossil skeleton of an adult *Australopithecus africanus*.
- Lucy is a famous fossil of an earlier hominin species, *Australopithecus afarensis*, dated at about 3 million years old.
- Certain features of a fossil skeleton allow conclusions to be drawn about whether the species was capable of upright (bipedal) locomotion.
- The Piltdown Man fossil is an example of deliberate scientific fraud.
- The ability to walk erect for sustained periods depends on the angled arrangement of the femur and the tibia.
- Hominin fossil discoveries continue to be made, including those of new hominin species.
- New hominin fossil discoveries may lead to new interpretations.

## QUICK CHECK

**8** Identify the following statements as true or false:
- **a** Robert Broom discovered the first *Australopithecus africanus* fossil.
- **b** The Taung child fossil was a member of the genus *Homo*.
- **c** The position of the foramen magnum can indicate whether a species walked upright or not.
- **d** In a chimp, the femur and the tibia join in a straight line.

**9** Identify a reason the Taung fossil skull was not immediately accepted as a link between humans and ancestral species.

**10** Give both the scientific name and the popular name of hominin fossils linked to the following famous palaeoanthropologists:
- **a** Raymond Dart
- **b** Donald Johanson
- **c** Robert Broom.

**11** A complete fossil skeleton of a hominid (family Hominidae) is available for examination.

Identify two regions that you would examine to assist you in deciding if the species represented by this fossil skeleton was capable of sustained bipedal locomotion.

# The beginning: the prehuman hominins

The prehuman hominin fossils are those hominins that evolved in the millions of years immediately following the divergence of the 'human' line from the 'African ape' line, before the emergence of the first *Homo* species.

The prehuman hominin fossils can conveniently be organised into four groups:
- the australopithecines: this term refers to any member of the genus *Australopithecus*
- the robusts: this term refers to those prehuman hominins belonging to the genus *Paranthropus*, characterised by heavily built skulls, heavy brow ridges and very large premolar and molar teeth
- the ardipthecines: this term refers to any member of the genus *Ardipithecus*
- the 'oldies': this term refers to the earliest of the prehuman hominins.

Table 12.4 shows the major fossil species of prehuman hominins. Look at the year of discovery and you will see that most of these fossil finds have occurred since the mid 1990s. Until the mid 1990s, the undisputed oldest hominin fossils that had been identified were those of Lucy's kind, *Australopithecus afarensis*, dated in the range of 3.0 to 3.7 million years ago (Mya). Since the mid 1990s, other prehuman hominin fossils have been discovered, including some that lived on the African continent millions of years before Lucy. In the first decade of the twenty-first century, more new hominin fossil discoveries were made, including some belonging to species dated to 4 to 7 million years ago.

**The oldies:**

The discovery of *Sahelanthropus tchadensis* appears to push the oldest identified hominin species back to *at least* 7 million years ago. However, varying interpretations of the *Sahelanthropus tchadensis* fossil have been made, so this species is not universally accepted as a hominin. One group of researchers who carried out a detailed analysis of this fossil concluded that *Sahelanthropus tchadensis* was not a hominin. In contrast, other researchers have identified this fossil species as possibly the earliest hominin fossil discovered to date.

**ODD FACT**

The year in which a fossil hominin is announced may be several years after the fossil is discovered. The interval between the two is the period during which the fossil is examined, compared to existing fossils, described and, if it is a new species, named.

**TABLE 12.4** Fossils representing species of prehuman hominins, the year of discovery or public announcement, and their estimated age ranges. The australopithecines are shown in pink, the ardipithecines in green, the robusts in blue, and the 'oldies' in orange.

| YEAR | Prehuman hominin species | Where discovered | Age range (Mya) |
|---|---|---|---|
| 2002 | *Sahelanthropus tchadensis* | Chad | 7.0–6.0 |
| 2001 | *Orririn tugensis* | Kenya | 6.2–5.8 |
| 2004 | *Ardipithecus kadabba* | Ethiopia | 5.8–5.2 |
| 2009 | *Ardipithecus ramidus* | Ethiopia | 4.4 |
| 1995 | *Australopithecus anamensis* | Kenya | 4.2–3.9 |
| 1974 | *Australopithecus afarensis* | Ethiopia | 3.85–2.95 |
| 1999 | *Australopithecus garhi* | Ethiopia | 3.5 |
| 2001 | *Kenyanthropus platyops* | Kenya | 3.5–3.2 |
| 1924 | *Australopithecus africanus* | South Africa | 3.3–2.1 |
| 2015 | *Australopithecus deyiremeda* | Ethiopia | 3.3 |
| 1985 | *Paranthropus aethiopicus* | Kenya | 2.7–2.3 |
| 1959 | *Paranthropus boisei* | Tanzania | 2.3–1.2 |
| 2010 | *Australopithecus sediba* | South Africa | 1.8–1.9 |
| 1938 | *Paranthropus robustus* | South Africa | 1.8–1.2 |

**The ardipithecines:**

Fossils belonging to a new hominin genus, *Ardipithecus*, were discovered in 2001 and 2004. *Ardipithecus ramidus* and *Ardipithecus kadabba* lived in East Africa in the period 5.8 to 4.4 million years ago, long before the appearance of the first australopithecines. It is not known whether *Ardipithecus* species were ancestral to the australopithecines or whether they were an evolutionary dead-end.

**The australopithecines**:

The most famous of the australopithecines are the Taung child (*Au. africanus*) and Lucy (*Au. afarensis*). Look back at figure 12.24, which shows the Taung child *Au. africanus* skull.

Since then, several more australopithecine species have been discovered, the most recent being *Australopithecus deyiremeda,* found in 2015 in the Afar region of Ethiopia, just 35 kilometres north of where the Lucy fossil was found in 1974. Fossils of Lucy's kind were also found at the same site, indicating that the two species overlapped, both in time and in space. The discoverer of *Australopithecus deyiremeda* has commented: '*They probably fought for resources. However, it is difficult to find out whether they hunted each other or interbred.*'

**FIGURE 12.33** Cranium of *Australopithecus sediba* found in South Africa in 2008. Note the sloping forehead and the projecting upper jaw.

The fossils of *Australopithecus sediba* (see figure 12.33) were discovered in South Africa in 2008, and the existence of this new species was announced in 2010. The interpretation of this fossil has been the subject of debate. Because it displays some human-like characteristics, this species has been identified as a possible direct ancestor of the first human species, but it also has some chimp-like features. Alternatively, it has been suggested that this species might be a later version of *Australopithecus africanus* and that differences between those species are simply due to changes that have occurred in the time interval that separates them. Refer to the box on the next page where you can read about a geologist whose expertise in dating rock strata using the uranium–lead procedure involved her in the dating of hominin fossils, including *Au. sediba*.

## SCIENTIST AT WORK

### Dr Robyn Pickering — determining timelines

**FIGURE 12.34** Dr Robyn Pickering in the clean laboratory in the School of Earth Sciences at the University of Melbourne, with a piece of two-million-year-old flowstone. (Image courtesy of Dr Robyn Pickering and Warrick Joe.)

'I am a geologist who works on understanding the timeline of human evolution. To be more specific, I do research into determining the ages of rocks associated with early human or hominin fossils. Knowing how old hominin fossils are is the key to placing them in our human family tree and figuring out who was related to whom.

I work in South Africa at the University of Cape Town, in the Department of Geological Sciences, where I am a lecturer. I use a technique called uranium–lead dating. What this actually means is that I extract the natural uranium and lead out of carbonate rocks and use the amount of the isotopes of these elements, together with the half-life of uranium, to work out how old the rocks are. I work mainly on cave sites that contain early human fossils and archaeological remains, so I collaborate with palaeontologists, palaeoanthropologists and archaeologists. Usually the fossils themselves are too old to date using carbon dating, so we have to date the rocks around them and work out how old they are based on the stratigraphy or layering of the rocks.

I grew up in Johannesburg, South Africa, and was always fascinated with rocks, animals and the great outdoors. Biology was my favourite subject at high school. After reading Richard Leakey's books about hunting for hominin fossils in East Africa and deciding that a career as an occupational therapist was not for me, I enrolled for a general science degree at the University of the Witwatersrand in Johannesburg. I studied archaeology and geology and enjoyed both but was drawn more to the geology. I began my research career working in an area known as the Cradle of Humankind. This series of caves is the richest source of early human fossils outside of East Africa and, while the fossils themselves were well known, little research had been done into the geology of the caves. I then went on to do a PhD at the University of Bern in Switzerland, where I was involved in developing the technique of uranium–lead dating and using this to date the Cradle sites. After this, I spent six very happy and productive years as a Postdoctoral Research Fellow at the University of Melbourne in the School of Earth Sciences working on dating cave sites from around the world.

I have been fortunate to be involved in a number of exciting fossil-related projects from around the world. One of the biggest discoveries of recent times was the partial skeletons of two individuals, a young male and an older female from the Malapa cave site in South Africa. These fossils were unlike any others and were classified as a new species, *Australopithecus sediba*. One of the big questions was how old these fossils were. I was able to date two flowstone layers above and below the fossils, which together with the palaeomagnetic signal trapped in the rocks around the bones, meant that we could calculate an age of $1.977 \pm 0.003$ million years for these fossils. The fossils of *Australopithecus sediba* are remarkably well preserved and complete and, at just under 2 million years old, they are the current best candidate as the ancestor to our own genus *Homo* and perhaps our oldest relatives. This discovery really shakes up our human family tree and no doubt there will be much ongoing debate.

My colleague Dr Helen Green and I recently dated the rocks surrounding a fossil monkey from the Dominican Republic in the Caribbean to be around 1 million years old. The fossils are found in amazing underwater caves — scuba divers work with the palaeontologists to recover the fossils. This age really changes our understanding of primate evolution in this region.

I love working as a scientist — it is a challenging, stimulating and exciting job. I am also a mum and, while it is not always easy juggling my busy days, I could not imagine doing anything else and I hope my children grow up to share my love and appreciation of the wonderful geological world around us. I am lucky that my work often takes me into the field, so I get a combination of working in a laboratory and visiting fossil sites around the world. There is no story greater than our own human story and I am honoured to contribute a small piece to the puzzle of our own origins.'

**FIGURE 12.35** Dr Maeve Leakey, who led a team that discovered the hominin fossil *Kenyanthropus platyops*. Its scientific name means 'flat-faced man from Kenya'.

**The robusts:**

The robusts are a group of prehuman hominins belonging to the extinct genus *Paranthropus*. These hominins are not on the direct line leading to the genus *Homo*, but are an evolutionary side branch that died out by 1.4 million years ago. The robusts are distinguished by their heavily built skulls with heavy brow ridges, large premolar and very large molar teeth, and a bony crest, known as a **sagittal crest**, on the top of the skull. This crest was the point of attachment for one of the muscles (the masseter muscle) that closes the jaw. (The box on page 575 gives more details of one of the robust species, *Paranthropus boisei*, and its diet.)

***Kenyanthropus platyops*:**

The one fossil (KNM-WT 40000) that lies outside these four informal groupings is *Kenyanthropus platyops*. This fossil was discovered in 1999 by a team led by Dr Maeve Leakey (see figure 12.35). The description of this fossil was published in 2001 and because of its differences from members of existing fossil groups, this fossil was assigned to a new genus, *Kenyanthropus*.

### Locations of major hominin fossil finds

Africa is indeed the cradle of humanity. The supporting evidence, based on fossil discoveries, is as follows:

- The oldest prehuman hominin fossils have been found in Africa.
- No fossils of prehuman hominins have been found outside Africa.
- Not one australopithecine fossil has been found outside Africa.
- The earliest fossils belonging to the genus *Homo* have been found in Africa.
- All hominin fossils found outside the African continent are species within the genus *Homo* and are younger than the earliest African *Homo* fossils.

Figure 12.36 shows the locations where many prehuman hominins (excluding the robusts) have been found. The majority of discoveries have occurred in the East African Rift Valley.


**FIGURE 12.36** Map showing the sites of discovery of prehuman hominin fossils in the period from 1995 to 2015. The East African Rift Valley encompasses the major sites of hominin fossil discoveries. This map does not include the fossils of the robust group of hominins.

Let's now look at the prehuman hominin family.

## The prehuman hominin family

The prehuman hominin fossil record covers a period from about 7 million years ago (*Sahelanthropus tchadensis*) to about 2 million years ago (*Australopithecus sediba*). The evolutionary relationship between the various species is not known for certain. By this time, the first humans had appeared in the fossil record and they shared the African continent with several prehuman hominins, including *Au. sediba* and two robusts, *Paranthropus boisei* and *P. robustus*.

Figure 12.37 is a diagram showing the distribution over time of the prehuman hominins. Many palaeoanthropologists suggest that *Au. afarensis* and *Au. africanus* are possible direct ancestors of the first human species. Was this the case? What is the relationship between *Kenyanthropus platyops* and the australopithecines? Further evidence is needed to answer these questions.

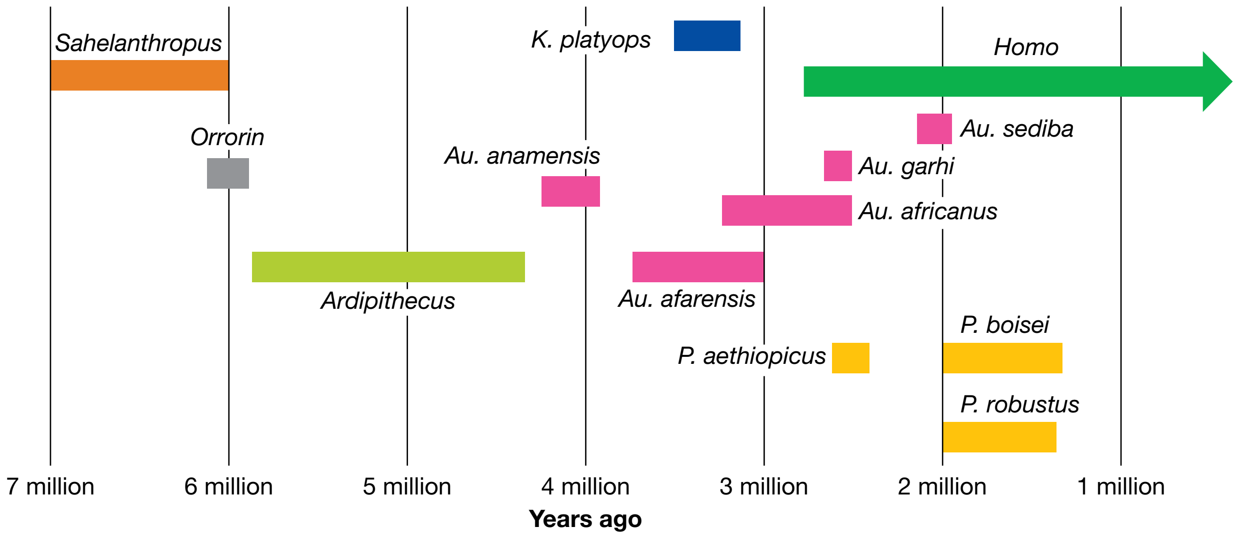



**FIGURE 12.37** Diagram showing the distribution over time of the major prehuman hominin fossil species that mainly preceded and coexisted with the first members of genus *Homo*. The label '*Ardipithecus*' includes two species: *A. ramidus* and *A. kadabba*.

**studyon**

Unit 4 | AOS 1 | Topic 4 | Concept 6

**Patterns in hominin evolution**
Summary screen and practice questions

## Erect walking in prehuman hominins

We know that hominins are characterised by the ability to walk erect in a sustained manner. Evidence of this ability is present in skeletal features such as the position of the foramen magnum on the base of the skull, the structure of the pelvis, and the angled arrangement of the femur relative to the tibia.

The prehuman hominins could walk erect, but they did not walk with the long easy strides of modern humans. Evidence from their fossil bones indicates that as well as spending time on the ground where they used bipedal locomotion, they also spent time in trees for several purposes, such as looking for fruit and sheltering there at night to avoid predators. The bipedal locomotion used when they were on the ground allowed them to move more efficiently across open grassland areas.

Possible advantages associated with bipedal locomotion (upright walking) as compared to quadrupedalism (knuckle walking) have been identified as follows:

- easier to view objects at a distance, such as watching for predators, locating sources of food or shelter, especially when moving across open grasslands
- reduced absorption of heat from the sun during the hottest part of the day because a smaller area of skin surface is exposed to direct sunlight, so it is easier to keep cool
- ability to walk greater distances when searching for food because the longer strides associated with bipedal locomotion require less energy than knuckle walking.

Bipedal locomotion requires strong hindlimbs ('legs') that can support the body weight in an upright position, stronger than those of a quadrupedal primate of comparable size. Stronger and longer legs increased the chance of

**FIGURE 12.38** Gibbons (*Hylobates* sp.) move by swinging from branch to branch, a form of locomotion known as brachiation. Note the arm length relative to the leg length.

survival and so would have given a selective advantage to individuals with those features. In contrast, for arboreal (tree-dwelling) primates, the selective advantage would have been for stronger and longer forelimbs ('arms').

A clue to the mode of locomotion of a primate may come from the ratio of its limb lengths. This ratio is obtained by dividing the length of the forelimb by the length of the hindlimb. This ratio allows us to make inferences about extinct primate species when only fossilised bones exist. A higher ratio, greater than 1.0, signals that the primate is **arboreal** (tree-dwelling) and is reliant on its forelimbs for climbing and swinging.

In contrast, primates that are capable of sustained erect walking have relatively shorter arms and longer legs. Table 12.5 shows the ratio of arm length to leg length for several species. The ratio of 1.32 for the gibbon indicates that its forelimbs are much longer than its hindlimbs — not surprising for a primate that moves by swinging from branch to branch (brachiation) (see figure 12.38). In contrast, for Lucy and her kind, the ratio of 0.88 indicates that the leg length is greater than that of the arms, an indication that these australopithecines could walk erect. For modern humans, who are more expert erect walkers, the ratio of arm length to leg length is even smaller, indicating a leg length that is relatively much longer than the arm length.

**TABLE 12.5** Table showing the ratio of forelimb or 'arm' length to hindlimb or 'leg' length for several primate species, including three hominins, *Australopithecus afarensis*, *Homo erectus* and modern humans. A ratio of 1.00 means that the 'arm' and the 'leg' lengths are equal. What does a value greater than 1.00 mean? What does a value less than 1.00 mean?

| Species | Ratio of arm (forelimb) length to leg (hindlimb) length |
|---|---|
| gibbon, *Hylobates* sp. | 1.32 |
| chimp, *Pan troglodytes* | 1.00 |
| Lucy, *Australopithecus afarensis* | 0.88 |
| *Homo erectus*, an early human | 0.70 |
| *Homo sapiens*, modern human | 0.68–0.70 |

It is seen from this table that the arm and leg lengths of adult chimps are equal, and that the arm lengths of the hominins are shorter than their leg lengths, indicating bipedal locomotion.

Because the amount of fossil material is very limited, evidence for bipedal locomotion in the earlier presumed hominins is less clear. The single fossil thighbone (femur) of *Orrorin tugenensis* suggests bipedal locomotion, as does the position of the foramen magnum in the 2005 reconstruction of the crushed fossil skull of *Sahelanthropus tchadensis*. If *Sahelanthropus* indeed used bipedal locomotion, at least some of the time, its discovery pushes the origin of hominins to 7 million years ago.

## How did early hominins live?

Fossil evidence allows us draw conclusions about the physical features of pre-human hominins. But how can we answer questions such as: What did they eat? How did they live? We cannot go back in time and observe them, but valid inferences about aspects of their way of life may be inferred from various lines of evidence such as:

- Relative body sizes of males and females provide clues to the social organisation of the species.
- Tooth enamel thickness, tooth wear, jaw features and carbon isotope studies provide clues to diet.
- Plant and animal fossils found in association with hominin fossils provide clues to the local habitats where hominins lived.

## Inferring social organisation

Where differences exist between the adult males and females of a species in average height, body mass or other features, that species shows **sexual dimorphism**. Studies of the social behaviour of living ape species reveal that apes with little or no sexual dimorphism, such as gibbons, have monogamous (one male – one female) pairings. In contrast, where a high degree of sexual dimorphism exists, as in gorillas and chimps, different social organisations are seen. These may comprise one male and a group of females, with the male controlling and defending the group, as in gorillas. Alternatively, as in chimps, it may comprise a group of related males who control and defend a group of females that may be spread out.

Because the degree of sexual dimorphism in living ape species relates in part to social organisation, these data can provide limited insight to possible social organisation in hominins. Table 12.6 shows the estimated body masses of a number of hominin species and several living ape species. Some degree of sexual dimorphism existed in these prehuman hominins and was greatest in *Au. afarensis*. (So, Lucy was a lot smaller than Luke!) Based on the data in table 12.6, scientists have speculated that the social organisation of early hominins may have been like that seen in modern chimps.

**TABLE 12.6** Estimated average body mass of hominin species of both sexes with comparative data from living ape species. The smaller the percentage value in column four, the greater the degree of sexual dimorphism in the species.

| | Adult body mass (kg) | | |
|---|---|---|---|
| Species | Male | Female | Female mass as percentage of male |
| *Au. afarensis* | 45 | 29 | 65 |
| *Au. africanus* | 41 | 30 | 73 |
| *P. robustus* | 40 | 32 | 80 |
| *P. boisei* | 49 | 34 | 69 |
| gorilla | 180 | 90 | 50 |
| chimp | 45 | 36 | 80 |

## Inferring diets

Based on evidence, such as the relative sizes of teeth, enamel thickness, jaw size and presumed jaw musculature, it is possible to make inferences about the diet of hominin species. For example, species in the *Ardipithicus* genus have small molars with thin enamel, suggesting that they ate fruit and soft leaves, while the more resistant teeth with thicker enamel of the australopithecines suggests a mixed diet including more fibrous plant materials.

Microscopic studies of the wear patterns on teeth surfaces can provide evidence about the diet of fossil hominins. This is done by comparing the microscopic texture on the surface of fossil molars with that on living primates with a known diet. For example, complex patterns with deep pits are seen in

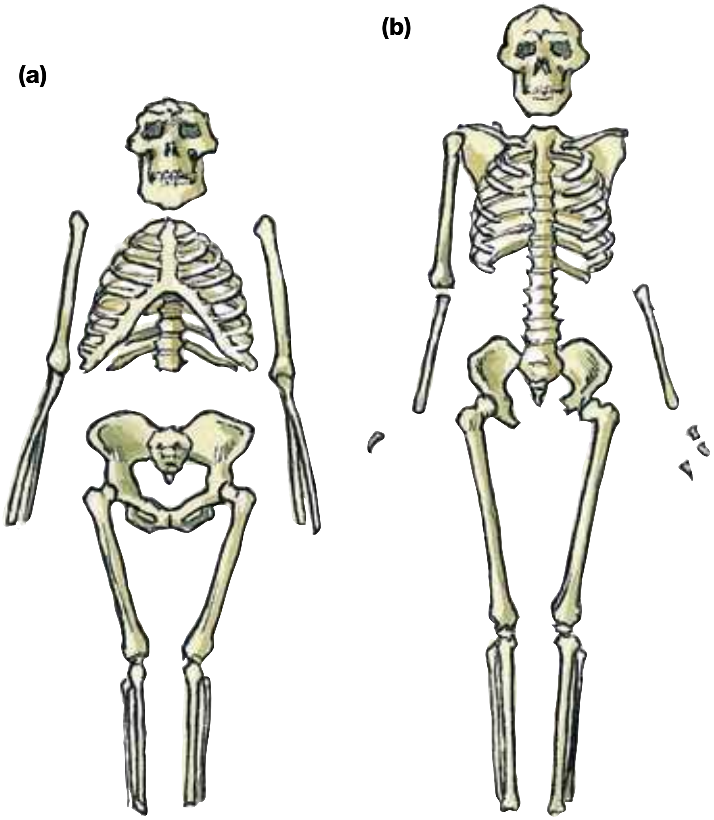



**FIGURE 12.39** **(a)** Skeleton of torso of *Australopithecus afarensis* ('Lucy'). Note how the wide rib cage gives a pot-bellied appearance. **(b)** Skeleton of torso of an early human species, *Homo erectus.*

primates that eat hard or brittle foods that require crushing, while simple patterns with long striations and scratches are seen in primates that eat soft or tough leaves and other plant material that require shearing. The simple tooth wear patterns of australopithecines indicate that they ate few, if any, hard and brittle foods.

While the exact nature of the diets of prehuman hominins continues to be researched, we can be sure that their diets were predominantly, or even exclusively, based on plant material. Much plant material has a lower nutrient value and takes longer to digest than meat. As a result, the digestive tracts of plant eaters are relatively much longer than those of carnivores (meat eaters) or omnivores (mixed meat/plant eaters). The box on page 575 describes how the analysis of carbon isotopes in the tooth enamel of a robust prehuman hominin gave a clue to its diet.

Because they were **herbivorous**, prehuman hominins had relatively large digestive tracts. To accommodate these large guts, australopithecines had wide waists and rib cages that were much wider at their lower ends than at the upper ends. These features gave australopithecines a pot-bellied appearance (see figure 12.39a). In contrast, early human species, such as *Homo erectus*, had a relatively shorter and smaller digestive tract, and the rib cage was more barrel shaped, so that they were relatively longer and narrower and had a defined waistline (see figure 12.39b), indicating a change in diet — meat! Because meat is energy rich and high in nutrients, this dietary change allowed for the diversion of energy and nutrients to the growth of other organs, particularly the brain.


**FIGURE 12.40** Map of Africa showing locations of discovery of various *Australopithecus* species. Where was Lucy discovered?

## Inferring habitats

Inferences can be made about the habitats where the various hominin species lived by examining the fossils of plants and animals found in association with australopithecine fossils (see figure 12.40 for locations of fossil discoveries).

For example, monkeys are forest dwellers and the presence of monkey fossils in an area indicates that the area was a forest or closed woodland in the period to which the fossils date. Likewise, the presence or absence of fossils of water-living animals, such as fish, crocodiles and hippopotamus, provide clues to habitat type. Other evidence that can be used to assist in inferring ancient habitats includes analyses of fossil pollen, which can identify the types of plants that grew in an area in the past.

Table 12.7 shows the inferred habitats of the areas in which various hominins lived. Fossils of *Ardipithecus* ('Ardi'), an early hominin, were found in association with other fossils including rats, monkeys and forest antelope, strongly indicating that *Ardipithecus ramidus* and her group lived in forest or closed woodland habitats. In contrast, many of the later hominin fossils, such as the australopiths, were associated with grassland savannah habitats (see figure 12.41). Because Ardi walked upright more than 4 million years ago, this finding has challenged the view that bipedal locomotion evolved when hominins lived in the open African grasslands and needed to cover distances at speed.

**TABLE 12.7** Inferred habitats in regions at the time when hominin species lived in the area

| Location | Habitat at time | Hominin species |
|---|---|---|
| Aramis, Ethiopia | Forest or closed woodland | *Ardipithecus ramidus* |
| Kanapoi, Kenya | Open habitat with nearby forest | *Australopithecus anamensis* |
| Allia Bay, Kenya | Forests near large river | |
| Hadar (Afar), Ethiopia | Riverside forests and grassland savannah | *Au. afarensis* |
| Laetoli, Tanzania | Grassland savannah | |

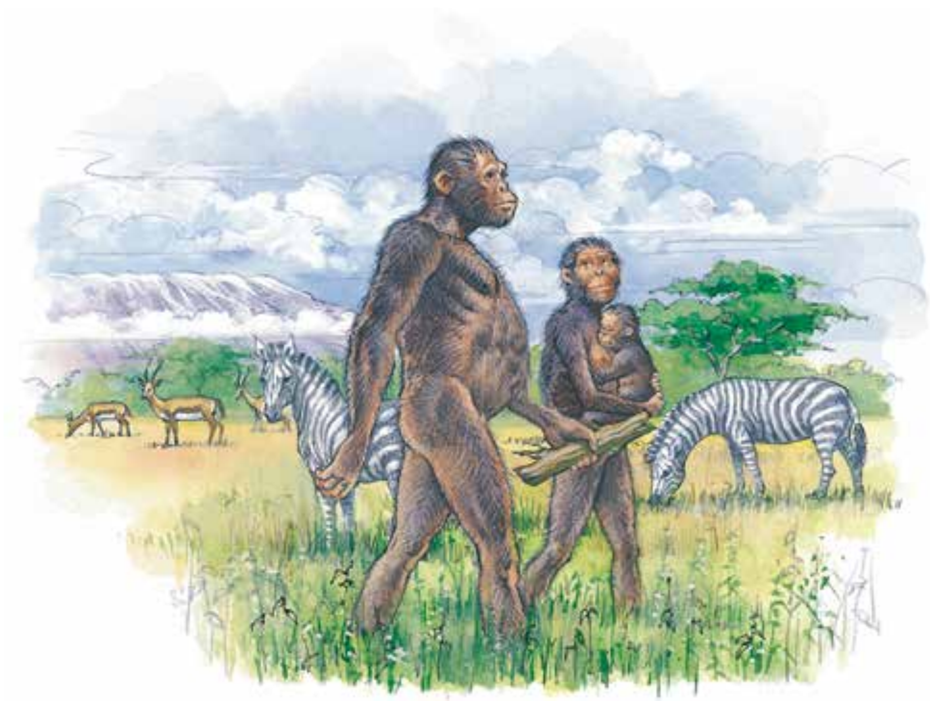


**FIGURE 12.41** A representation of prehuman hominins in a savannah grassland. This type of habitat was where *Australopithecus afarensis* is thought to have lived.

## Inferring culture

Because prehuman hominins could walk on two feet, their hands were free for other purposes. However, there is no definite evidence that any of the australopithecines made tools from natural objects in their environment. Did the australopithecines use tools? Evidence of tool use is limited if the tools are natural objects (such as sticks for digging or rocks for smashing bones) that are not modified into identifiable tools. However, australopithecines may have used objects as tools, perhaps carrying them or finding them in an area. This conclusion is based on the evidence of tool use of the modern apes, such as chimps using sticks to extract insects from crevices in logs.

Remarkably, evidence has been preserved of the use of an object as a tool by australopithecines. This evidence is in the form of smashed fossil antelope bones found in a region where *Australopithecus garhi* fossils have also been found and dated to the same age. This evidence allows us to picture an event that occurred about 2.5 Myr ago when one or more australopithecines used stones as tools to smash antelope bones, possibly to extract the bone marrow.

**ODD FACT**

Tools that have been crafted for use are always found in association with fossils of the genus *Homo*. It is generally concluded that toolmaking, as distinct from tool use, appeared with the earliest species of the genus *Homo*.

## How are the hominins related?

It is not possible to identify the exact evolutionary pathway that links pre-human hominin species to the first *Homo* species.

It is reasonable to speculate that the earliest prehuman hominins were ancestral to the australopithecines. From the australopithecines, one evolutionary line gave rise to the robust *Paranthropus* species, and a second line gave rise to the first members of the genus *Homo*, the first humans. The *Paranthropus* species were an evolutionary 'dead-end'. One simple hypothetical pathway is shown in figure 12.42.

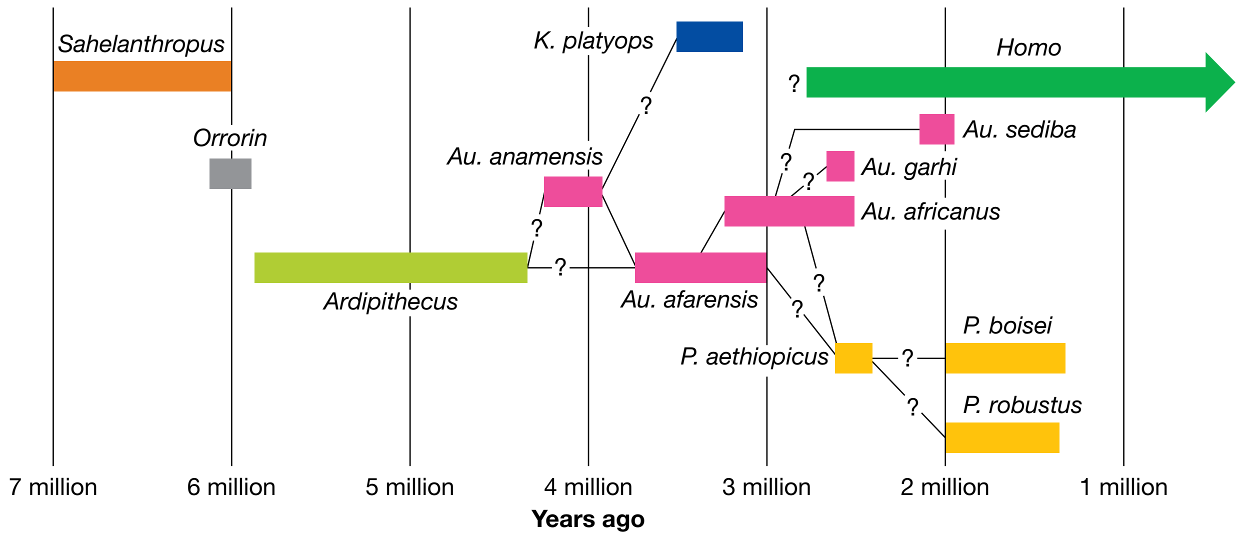
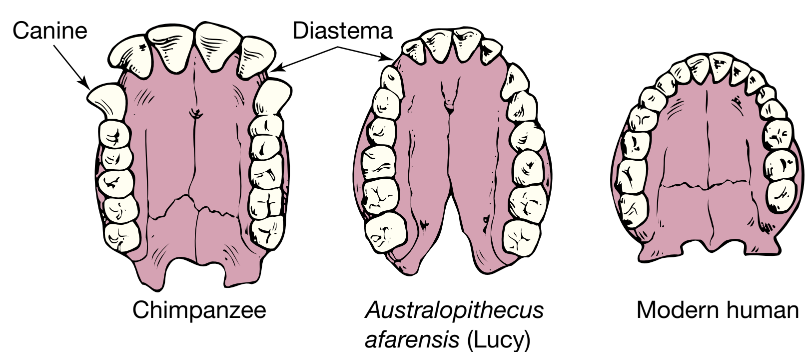



**FIGURE 12.42** Possible relationships are shown by '?' symbols between the various hominin species.

A transitional relationship between australopithecines and the first human species is supported by several lines of evidence including the upper jaw. Figure 12.43 shows a comparison of the upper jaws and teeth of a chimp, an australopithecine (*Au. afarensis*) and a modern human. Note the box-like shape of the chimp jaw compared to the parabolic shape of the human jaw and see that the rounded shape of the australopithecine jaw is intermediate between those of a chimp and a modern human. Australopithecines had smaller teeth than did chimps, making australopithecines more human-like in this regard. The earlier australopithecine, *Au. afarensis*, had a gap (diastema) in the dental arch, a chimp-like feature. However, a later australopithecine, *Au. africanus*, lacks this gap and its jaw is even more human-like.

**FIGURE 12.43** Diagram (not to scale) showing the upper jaw and teeth of a chimp, *Australopithecus afarensis* (Lucy) and a modern human. Note the shape of the dental arch in each species. The dental arch in *Australopithecus afarensis* is intermediate between the box-shaped arch of chimps and the parabolic arch of modern humans.



The *Australopithecus sediba* fossil found in 2010 is of particular interest. *Au. sediba* shares more derived features with *Homo* species than any other australopithecine; for example, the teeth of *Au. sediba* are reduced in size, as in *Homo* species, and the structure of its pelvis is similar to that in *Homo* species. The long legs of *Au. sediba* suggest that it was possibly the first hominin capable of running. However, *Au. sediba* retains some primitive chimp-like features. Its position within the human evolution story raises interesting questions such as: Could *Au. sediba* be a direct ancestor of the first *Homo* species? Might *Au. sediba* be reclassified as a member of the genus *Homo*?

## DID THE NUTCRACKER MAN EAT NUTS?

In Tanzania in 1959, a hominin fossil cranium was found and, later, a lower jaw. This fossil was given the scientific name *Paranthropus boisei* and is one of the robusts. What was unusual about the fossil was that its large jaw carried the biggest and flattest molar teeth yet seen in any hominin fossil (see figure 12.44). As a result, this fossil was given the popular name of the Nutcracker Man. The question is: *Did the Nutcracker Man eat nuts?*

**FIGURE 12.44** Casts of the upper jaws with molars of the Nutcracker Man, *Paranthropus boisei*, and a modern human. Note the thick tooth enamel and flared cheekbones (zygomatic arches) of *P. boisei*. (Image courtesy of Melissa Lutz Blouin (currently at University of Florida Health), when she was at University of Arkansas.)

Figure 12.45 shows the skull of *P. boisei* compared to that of *Australopithecus*. Note the flared cheekbones or **zygomatic arches** (arrowed) in *P. boisei* and the large spaces they enclose. These spaces accommodate the two muscles involved in closing the jaw, and the size of this space allows us to infer that *P. boisei* had powerful jaw muscles. The masseter muscle passes from the jaw and attaches to the zygomatic arch, and the temporal muscle passes from the jaw behind the zygomatic arch and through to the top of the skull where it attaches to the sagittal crest. *P. boisei* had an estimated biting force of about 770 kilograms (compared to 125 kilograms for a modern human jaw). Based on this evidence alone, it was assumed that the diet of *P. boisei* comprised hard, brittle objects, such as hard seeds, hard fruits and nuts.

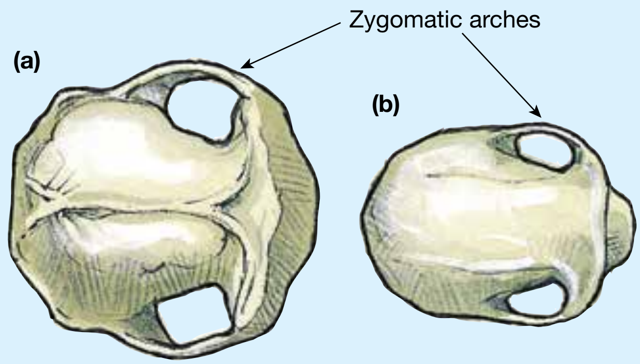



**FIGURE 12.45** Cranium of **(a)** *Paranthropus* sp. and **(b)** *Australopithecus* sp. viewed from above, showing the zygomatic arch or cheekbone (arrowed). Note how the arches in *Paranthropus* flare out widely. The space created is needed to accommodate the large muscles of the lower jaw — the larger the space, the bigger the chewing muscles. Note also the sagittal crest on the top of the *Paranthropus* cranium. What purpose did it serve?

In 2011, scientists published an analysis of the carbon isotopes present in 24 samples of tooth enamel from 22 *P. boisei* individuals. This analysis of carbon isotopes can identify whether a person's diet comes mainly from grasses and sedges that carry out photosynthesis using the so-called C4 pathway, or whether the diet comes mainly from the edible parts of trees, shrubs and herbs that use the C3 photosynthetic pathway.

The results of this analysis showed that the diet of the Nutcracker Man and other members of the *Paranthropus boisei* population comprised mainly grasses and sedges — not nuts. The grasses and sedges that were crushed between his large flat molar teeth constituted a diet unlike that of any other hominin.

## KEY IDEAS

- The prehuman hominins can be grouped as (1) earliest, including the ardipithicines, (2) the australopithecines, and (3) robusts.
- Inferences can be made about aspects of the way of life of prehuman hominins, including diet, habitat and social organisation.
- No evidence exists for toolmaking by australopithecines, but limited evidence exists for tool use.
- The evolutionary pathways among the prehuman hominins and their link to the first *Homo* species continue to be discussed and investigated by scientists.

## QUICK CHECK

**12** Identify the following statements as true or false:
  **a** The fossil record contains several species of *Australopithecus*.
  **b** The upper jaw of an australopithecine is identical to that of the chimp.
  **c** From an examination of the zygomatic arch on a fossil skull, it is possible to draw conclusions about the strength of the bite of the species.
  **d** The various australopithecine species lived in similar habitats.
**13** What evidence, if any, exists for tool use by the australopithecines?

# The genus *Homo*: human at last!

Africa is the site where the first members of the genus *Homo* appeared at least 2.8 million years ago. Many different species within the genus *Homo* evolved and existed before the first appearance of modern humans (*H. sapiens*) and some coexisted with our species. However, our species of *Homo sapiens* is now the only living member of this genus.

The early *Homo* species or humans were different from their australopithecine ancestors and showed the following physical changes:

- a marked increase in their brain size, with an accompanying increase in the size of the cranium, including the height and the width of the skull
- a reduction in the size of their teeth and jaw muscles
- a flattening and shortening of the face, making the face nearly vertical.

You can see the result of the flattening and shortening of the face in figure 12.46, which compares the skulls of a prehuman hominin (*Au. africanus*), an early human (*H. erectus*) and a modern human (*H. sapiens*). Note that as the human face became flatter and more vertical, the amount of the face that is visible from this perspective decreased. In prehuman hominins, the face tends to be concave and, because of their relatively much larger teeth and jaws, their faces project forward.


**FIGURE 12.46** Top view of skulls of **(a)** *Australopithecus africanus*, **(b)** *Homo erectus*, an early human species, and **(c)** *Homo sapiens*. Note that, as the face became flat and near vertical, as in *Homo sapiens*, the teeth can no longer be seen in this view.

Figure 12.47 shows side views of three skulls. The skull of modern human (*Homo sapiens*) is flanked on the left by the cranium of a robust prehuman hominin (*Paranthropus robustus*) and on the right by the skull of an australopithecine (*Australopithecus africanus*). Note the flat face of the human in contrast to the sloping faces of the prehuman hominins.

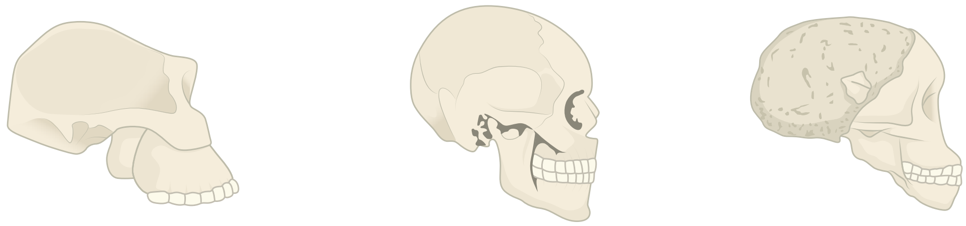



**FIGURE 12.47** Drawings (not to scale) of a side view of **(a)** cranium of *Paranthropus robustus,* **(b)** skull of a modern human, *Homo sapiens*, and **(c)** *Australopithecus africanus*, showing the contrast in the slope of the face. The evolution of the *Homo* species was accompanied by a flattening and shortening of the face.

## What does it mean to be human?

'Being human' is not easy to define. **No single feature defines humanity**. The evolution of the genus *Homo* is linked to structural changes such as enlargement of the brain. Over the period starting from 2.8 million years ago, after the appearance of the first human species, the evolution of the genus *Homo* continued and is most significantly associated with the development of complex cognitive skills and behavioural changes that include:

- making and using fire
- cooperating in group activities
- caring for aged and ill members of the species
- burying the dead
- the development of art, language, mathematics and music
- the use of symbols
- increasing use of technology that began with tools crudely fashioned from stone by *Homo habilis*, which has extended to the present-day use of computers, lasers and spacecraft by *Homo sapiens*.

## How many *Homo* species?

studyon

Unit 4 | AOS 1 | Topic 4 | Concept 8

**Difficulties in interpreting fossils**
Summary screen and practice questions

The genus *Homo* is represented today by a single living species, *Homo sapiens* (meaning 'wise man'), which includes people of all racial groups. Human populations show phenotypic variation (see figure 12.48). Although variation exists between populations, humans share most features in common and are members of a single species, capable of interbreeding to produce viable, fertile offspring.

For extinct organisms, decisions about defining species are made using evidence available either directly or by inference from fossil bones and teeth. For this reason, there is no universal agreement on the number of different *Homo* species within the genus *Homo* that have existed during its evolutionary history (over a period of about 2.5 Myr). Scientists who are 'splitters' tend to recognise more species, while those who are 'lumpers' tend to recognise fewer species. Recent analysis has led to the recognition of more species in the genus *Homo*. Disagreements between scientists are reminders that, even when they examine the same fossil material, scientists may differ in their interpretations and conclusions.

**FIGURE 12.48** Just a little of the phenotypic variation within *Homo sapiens.*

Table 12.8 shows the major fossil species of the genus *Homo*. The first entry in this table is a specific fossil identification and does not identify a species. This is a recently discovered jawbone that has been identified as belonging to the genus *Homo*.

**TABLE 12.8** Table showing the major *Homo* fossils, their year of discovery or public announcement, and the estimated age ranges for each species. Mya = millions of years ago. The fossil specimen LD 350-1 has not yet been assigned to a particular human species, so it is listed by its fossil identification number.

| YEAR | Species of genus *Homo* | Where discovered | Age range |
|---|---|---|---|
| 2015 | LD 350-1 | Ethiopia | 2.8 Mya |
| 1964 | *Homo habilis* | Olduvai Gorge, Tanzania | 2.4–1.4 Mya |
| 2016 | *Homo naledi* | South Africa | 1.98 Mya |
| 1986 | *Homo rudolfensis* | East Africa | ~1.9–1.8 Mya |
| 1891 | *Homo erectus* | Trinil, Indonesia | 1.89 Mya – 143 000 yr ago |
| 1908 | *Homo heidelbergensis* | Heidelberg, Germany | 700 000 – 200 000 yr ago |
| 1829 | *Homo neanderthalensis* | Neander Valley, Germany | 400 000 – 40 000 yr ago |
| 2003 | *Homo floresiensis* | Flores Island, Indonesia | 100 000 – 60 000 yr ago |
| 2010 | *Denisovans* | Siberia | 41 000 yr ago |

Figure 12.49 below shows the timelines of the major human species in relation to those of the prehuman hominin species. Can you identify two species that lived at the same time? Would they necessarily have interacted?

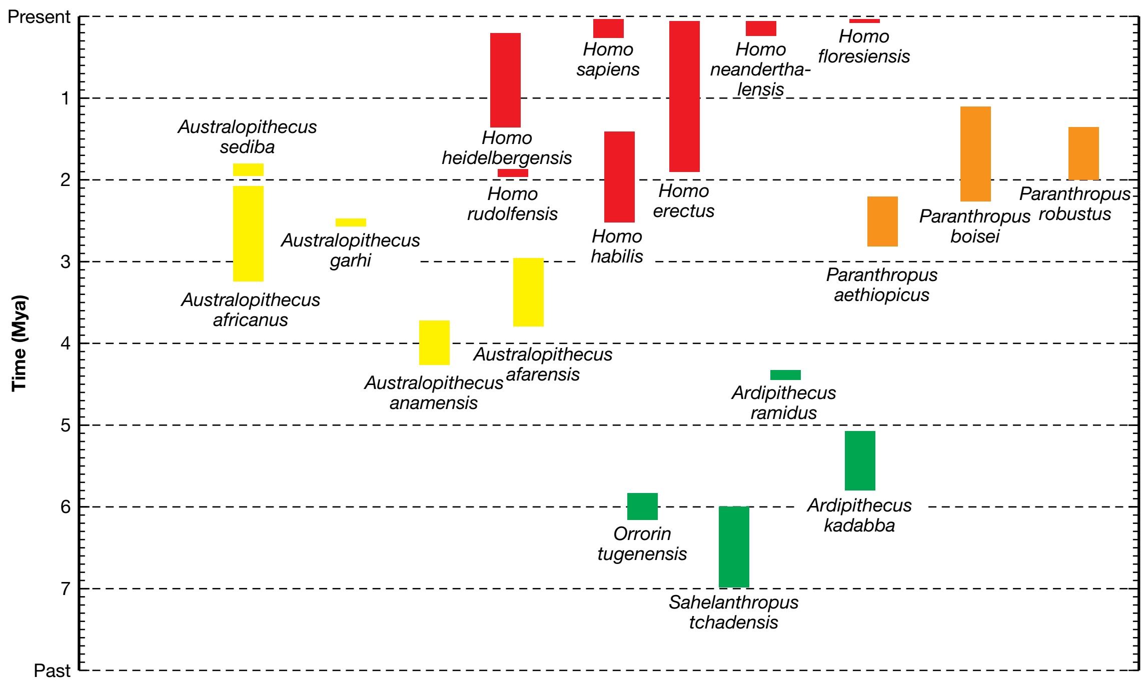



**FIGURE 12.49** Timelines of hominin species, both human and prehuman.

## Expansion of the human brain

Sir Arthur Keith (who originally discounted the Taung child fossil) believed that brain volume defined 'being human'. In his 1948 essay, 'Crossing the Rubicon twixt Ape and Man', Keith identified a brain volume of 750 mL as the critical lower limit for a fossil to be classified as a member of the genus *Homo*.

**Increased brain capacity, along with the greater cognitive capabilities it provides, distinguishes members of the genus *Homo* (humans) from the australopithecines.**

Figure 12.50 shows the brain volumes in various hominins, both prehuman and human. For about 2 million years (4.5 to 2.5 Mya), hominin brain volumes (based on measures of endocranial volume) remained at 400 to 500 millilitres, a little larger than that of the great apes.

However, after the appearance of *Homo habilis*, brain volume changed much more rapidly, more than doubling in the next 2 million years, but this was not accompanied by a doubling of the sizes of other body parts. This indicates that a strong selection pressure for increased cognitive processing was acting on populations of early humans, with improved problem-solving skills and the ability to plan, conferring a survival advantage. It should be noted that the power of an enlarged brain was only able to be used to advantage because of earlier evolutionary developments, such as the presence of an opposable thumb on the hands, giving remarkable manual dexterity, and the development of sustained bipedal locomotion, freeing the hands.

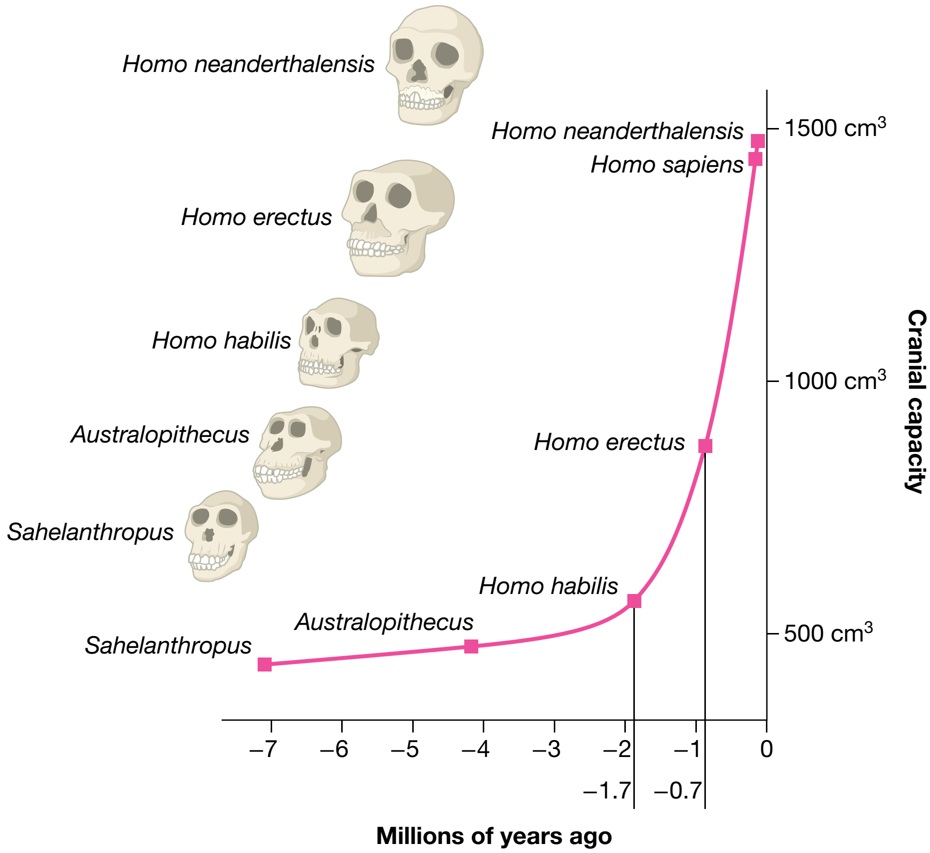



**FIGURE 12.50** Graph showing the rate of increase in size of the hominin brain over the 7 million years of hominin evolution. Note the rapid increase in the rate of growth starting just under 2 million years ago.

The disproportionate size of the brain relative to average body mass in members of genus *Homo* is seen even more clearly when average brain size is plotted against average body mass (see figure 12.51).

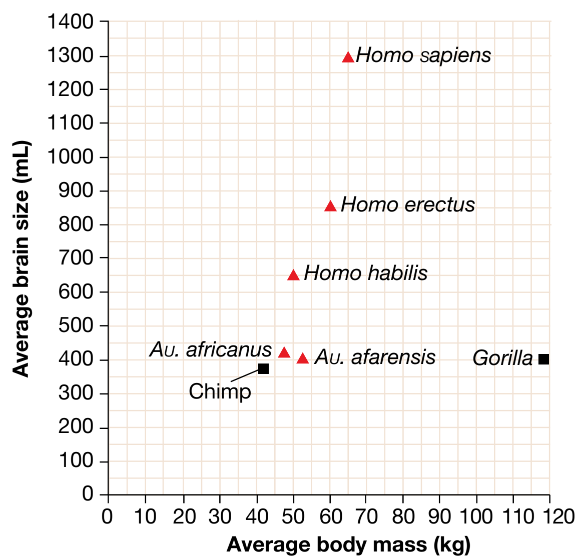



**FIGURE 12.51** Graph of average brain volume against average body mass. Note the disproportionate expansion of brain volume in various human species, as compared to prehuman great apes and the early hominins. Brain size doubled from *Homo habilis* to *Homo sapiens*. Did the average body mass also double?

The expansion of the brain in *Homo* species did not occur evenly over all its parts. Instead, the cerebral cortex (neocortex), the region of the brain concerned with cognitive abilities such as abstract reasoning, complex problem-solving skills, forward planning and social skills, expanded at a higher rate than most other brain regions, except for the cerebellum, which also expanded disproportionately. The cerebellum of the brain controls motor functions.

**ODD FACT**

The brains of apes, including humans, have a greatly expanded cerebellum relative to its size in monkeys, a difference that is equal to 16 billion extra neurons. (Source: *New Scientist* 11 October 2014.)

The brain is an energy-hungry organ. A large brain comes with a high-energy cost. In modern humans, the brain represents about 2 per cent of body mass but uses about 20 per cent of the total energy intake. In contrast, the figure for energy usage of ape brains is about 13 per cent, while that of other mammals is between 2 and 8 per cent.

The inclusion of meat in the human diet and the later use of fire for cooking produced the extra nutrients and energy for brain growth. What were the selective pressures on early human populations for this brain growth? Some researchers propose that a global shift in climate (cooling and drying) about 3 to 2.4 Myr ago reduced the area of lush forested areas of Africa and gave a selective advantage to individuals with greater cognitive skills. These individuals were better equipped to exploit the resources in the new habitats arising from the climate change. For example, hominins may have followed game herds and obtained meat by scavenging the kills of predators.

*Homo habilis* individuals are called habilines.

Let's now explore some of the species in the genus *Homo*, starting with the earliest species, *Homo habilis*.

## The first toolmakers: *Homo habilis*

In the early 1960s, the first human fossils were found by Louis Leakey in East Africa and were given the scientific name *Homo habilis*. The earliest of the known *H. habilis* fossils have been dated at about 2.4 Myr old. The evidence for being declared human was that these fossils had relatively greater brain size (approximately 640 millilitres) and smaller teeth than members of the genus *Australopithecus*.

Simple stone tools have been found at sites where *H. habilis* fossils have been found. This suggests that **habilines were probably the first hominins** who purposefully struck stone flakes off pebbles to create sharp-edged flakes (see figure 12.52) to use as tools. This procedure is termed **Oldowan** toolmaking.

**FIGURE 12.52** About 1.5 Myr ago, *Homo habilis* used a stone, known as the hammer stone, to strike another stone, known as the core, to knock off flakes of stone with sharp edges. These flakes were probably used for cutting the thick skin (hide) of dead animals and for scraping meat from bones.

These stone flakes were humanity's first 'knives' and they could be used in various ways: for example, as 'scrapers' to scrape flesh off bones or fat from skins, and as 'cutters' to cut through the tough outer hides of dead mammals to obtain the muscle tissue and to cut through tough tendons. Larger stones were used as tools to smash long bones. (Why do this?)

Discarded tools and broken animal bones associated with habiline fossil sites support the conclusion that the habilines were the first hominins to include meat as a main dietary item, which is more nutritious than tubers and roots. *H. habilis* may have obtained meat by collecting carrion (animals killed by other predators) on the African grasslands. In order to gather carrion, habilines may have worked in cooperative groups and used stones to drive predators, such as hyenas, away from their kill. Using their stone tools, the habilines could cut through tough hides, cut chunks of meat from a kill and carry food to a safer place for eating.

### Revision of the habilines

A great deal of variability existed in the fossils originally identified as *H. habilis* and, in 1992, a revision of the fossils identified as *H. habilis* took place. One important question to decide was: Are observed variations due to differences within a species, or are the variations a result of differences between species? The consensus view was that the fossils belonged to two different species, and so these habiline fossils were split into two groups:

- the smaller brained *Homo habilis* with a more prominent brow ridge
- the larger brained *Homo rudolfensis* with a less prominent brow ridge (see figure 12.53).

Fossils of these two species are restricted to East Africa and South Africa, consistent with the fact that the early evolution of the genus *Homo* occurred in Africa.

**This split means that, about 2 million years ago in East Africa, two different human species existed:** *H. habilis* and *H. rudolfensis.*

**ODD FACT**

Debate still continues about the classification of *H. habilis* and *H. rudolfensis.* In the case of *Homo rudolfensis*, the debate is whether it should retain its human classification or whether it should be reclassified, for example, as a member of the genus *Kenyanthropus*.

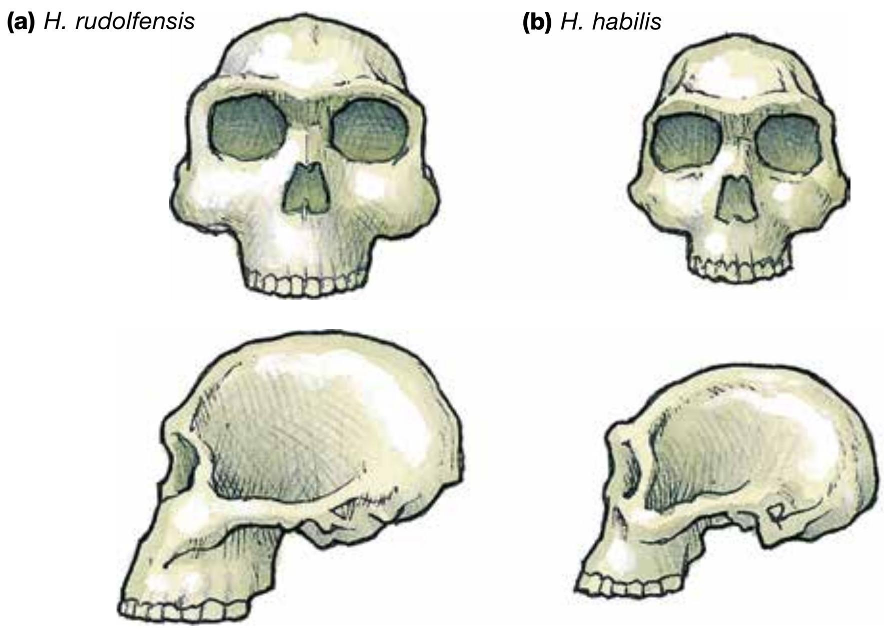



**FIGURE 12.53** Problems of the great variability in fossils originally classified as *Homo habilis* were resolved when this group was split into two species: **(a)** *H. rudolfensis* and **(b)** *H. habilis*. Which has the larger brain? The stronger brow ridges?

## The first human emigrants: *Homo erectus*

The earliest fossils of *H. erectus* are dated at about 1.8 Myr old and the latest up to about 143 000 years old. Figure 12.54 shows a *H. erectus* skull compared to that of a modern human (*H. sapiens*). Some of the key differences are noted. Of all the human fossil species, the *Homo erectus* species survived for about 1.7 million years. More importantly, *Homo erectus* is believed to have been the direct ancestor of later human species.

**ODD FACT**

Some scientists believe that *Homo erectus* should be split into two species, *H. erectus* and *H. ergaster*, with *H. ergaster* being the ancestral species in Africa, and *H. erectus* being the descendant species that spread out of Africa into Europe and Asia.

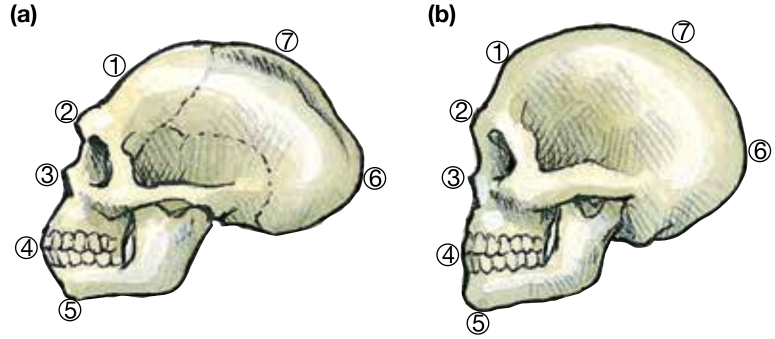



**FIGURE 12.54** Skulls of **(a)** *Homo erectus* and **(b)** a modern human, *H. sapiens*. Which species has a more vertical forehead? A larger cranium? More prominent brow ridges?

In this book, the term *Homo erectus* is used for both the African species and the species that spread from Africa.

Figure 12.55a shows some locations where fossils of *H. erectus* have been found. Note that the undisputed oldest specimen of *H. erectus* comes from Kenya in Africa. However, *H. erectus* fossils have been found in other continents and these fossils indicate that, early in its evolutionary history, this species migrated out of Africa to other continents (see figure 12.55a). **So, *Homo erectus* appears to be the first human emigrant from Africa to other parts of the world**.

In each geographic region where ***Homo erectus*** spread, populations developed local variations that can be seen in differences between the so-called 'Java Man' and the 'Peking Man' (see figures 12.55b and 12.55c).




**FIGURE 12.55** **(a)** Map showing some sites where *Homo erectus* fossils have been found. Approximate ages (Myr) of the fossils are given in brackets. **(b)** This cave in Zhoukoutien (near Beijing) in China was the site of the discovery in 1927 of the so-called Peking Man, later recognised as *H. erectus*. This site is now listed as a World Heritage site. **(c)** The Peking Man Museum at the Zhoukoutien World Heritage site features a representation of *H. erectus*. (Images (b) and (c) courtesy of Judith Kinnear.)

**FIGURE 12.56** Fossil skeleton of the Turkana Boy, now called Nariokotome Boy, a member of the species *Homo erectus.*

*Homo erectus* was the first human species with body proportions similar to modern humans. The most spectacular discovery of a fossil *Homo erectus* occurred in 1984 when a nearly complete skeleton of a boy (see figure 12.56) was found at Lake Turkana in Kenya. This skeleton was found in sediments dated at 1.6 Mya. The boy was estimated to have been about eight years old when he died, and his brain capacity was 880 mL. His height at that stage was about 162 centimetres — and as an adult he may have reached more than 183 centimetres.

## Making tools and controlling fire

The brain size of *H. erectus* was about 900 millilitres, considerably more than that of *H. habilis* (640 mL) and about 75 per cent that of modern humans.

*Homo erectus* was a smarter hominin and, as a result, **a number of new behaviours and skills** appeared for the first time in this species, such as:

- controlled use of fire
- crafting of more elaborate stone tools — the so-called Acheulean technology used by *Homo erectus* to make tools was far more sophisticated and skilful than the Oldowan toolmaking technology of *Homo habilis*
- evidence of systematic hunting.

Evidence supports the hypothesis that *Homo erectus* **with its large brain was the first human to use fire in a controlled way**. Evidence for the controlled use of fire comes from burnt bones and concentrated charcoal and ash deposits found in caves that were occupied by *Homo erectus*. These deposits can be distinguished from those of natural bushfires, which typically burn less intensely than a campfire and over a larger extent.

Fire may have been used by *Homo erectus* for a variety of purposes, such as:

- to provide warmth in cold climates
- for cooking meat and plant material so that it was more tender, more easily digested and hence more nutritious (see figure 12.57)
- for protection from predators
- for light at night.

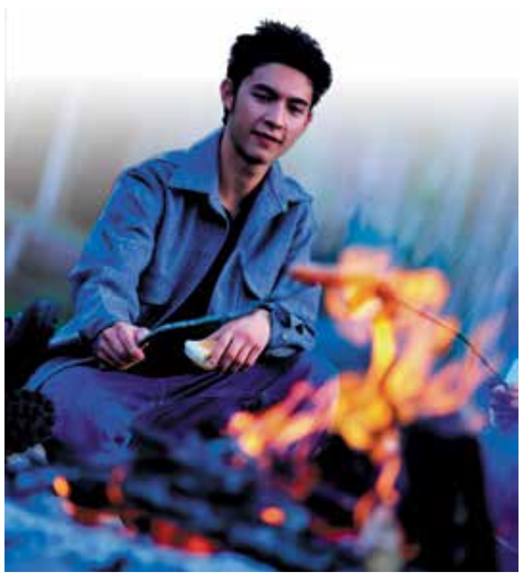


**FIGURE 12.57** The controlled use of fire goes back at least to the time of *Homo erectus*. Our use of fire for cooking food is an ancient human behaviour, so, when you are next at a barbeque, give a thought to the *Homo erectus* humans.

### ODD FACT

The best evidence of controlled use of fire, a behaviour seen only in humans, comes from cave deposits in China that have been dated at about 400 000 years ago.

*H. erectus* made tools according to a regular pattern. Beautifully crafted hand axes and other tools made of stone and bone are typical of the **Acheulean technology** of *H. erectus*. Figure 12.58 shows the process of making a hand axe using the so-called Acheulean technology of *H. erectus.* This process involves control and planning as a regular sequence of stone flakes is chipped from the two faces of a large pebble. The tools used to remove these flakes were hand-held 'hard' hammers made of stone and 'soft' hammers made of bone or antler. R. Lewin (in *Human Evolution, Fifth Edition*, Blackwell Publishing, 2005, p. 169) suggests that the Acheulean hand axe 'may have been the Swiss Army knife of the Lower Palaeolithic [early Stone Age]'. Research being carried out at Emory University in the United States shows that the skills needed to produce symmetrical hand axes with sharp, regular edges are acquired only after long periods of practice. This research has also shown that different areas of the brain are activated when making a stone tool using the Acheulean technique of *Homo erectus* as compared to those areas that are activated when making flakes using the Oldowan toolmaking technique of *Homo habilis*.

**ODD FACT**

The first evidence that *Homo erectus* produced not only tools but also made abstract patterns was recognised in 2014 from a deep zig-zag etched on the shell of a fresh water clam dated at about 500 000 years ago and found in Trinil, Indonesia. Until this Trinil etching was identified, the oldest human etching was a 75 000-year-old etching on stone found in South Africa.

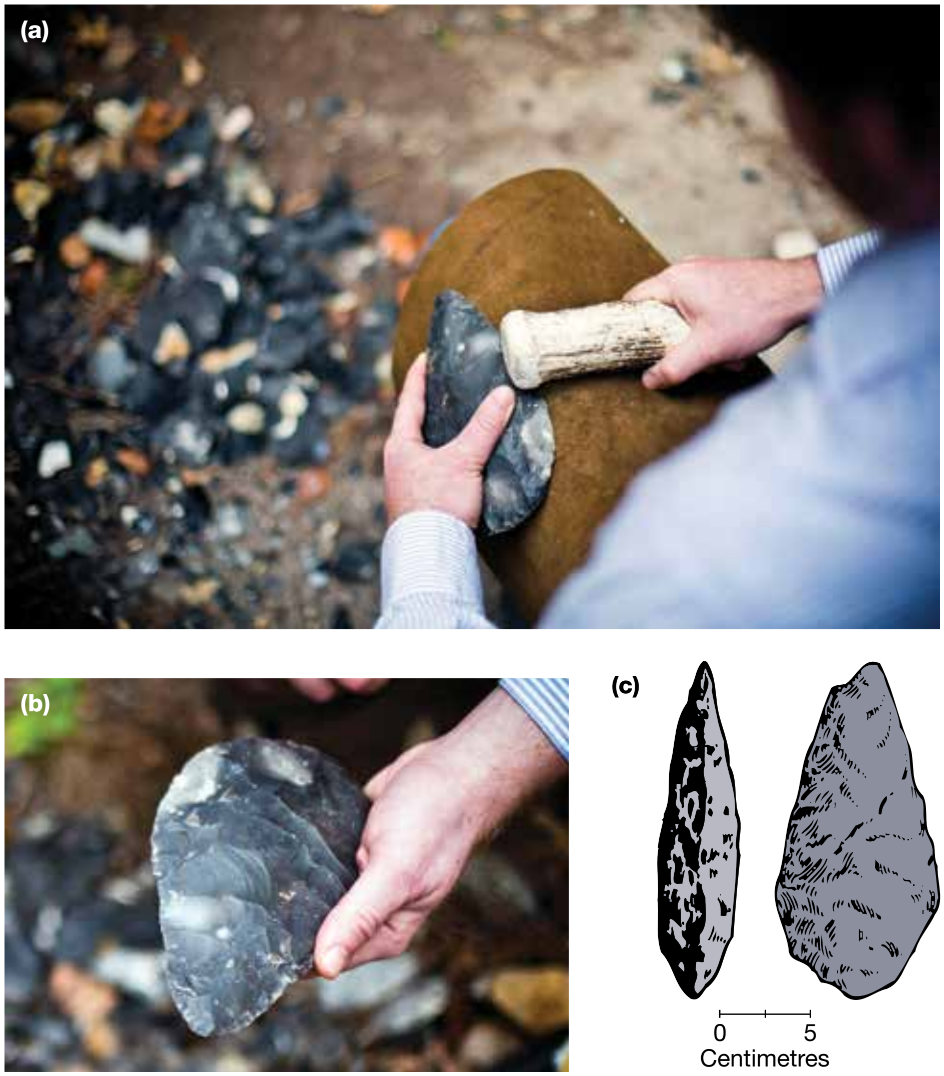



**FIGURE 12.58** **(a)** A researcher in the process of making a hand axe using Acheulean toolmaking technology. Note the piece of antler being used as a hammer to remove flakes from two faces of the core stone to shape it into a hand axe. **(b)** A modern hand axe made using the Acheulean toolmaking technique. **(c)** Edge-on view and side view of an Acheulean hand axe. Note the symmetry and the sharp edge of this tool. (Images **(a)** and **(b)** courtesy of Dr Dietrich Stout, Emory University, USA.)

Both the sophisticated toolmaking ability of *Homo erectus* humans and their controlled use of fire point to intelligence and planning. Relative to the earlier habilines, the increased brain capacity in *Homo erectus* gave this human species greater ability to plan and to act intelligently.

It appears that *Homo erectus* individuals were able to make sounds but could not articulate the varied and rich vocal patterns of true speech, which involves pronunciation, timing and intonation. Communication between early humans would probably have been a combination of facial expressions and vocalisations, but not spoken language. Supporting evidence for this claim came from the *Homo erectus* skeleton found in 1984. Each of the bony vertebrae in the spinal column (backbone) has a space through which the spinal cord passes. The spinal cord consists of nerve tissue including nerves that control muscles. In *Homo erectus*, the space in the vertebrae in the neck region is only half as big as in modern humans (see figure 12.59). This means that there was less nerve tissue in the spinal cord of *H. erectus*, and a detailed study concluded that this species lacked the fine muscle control of breathing that is necessary for speech.

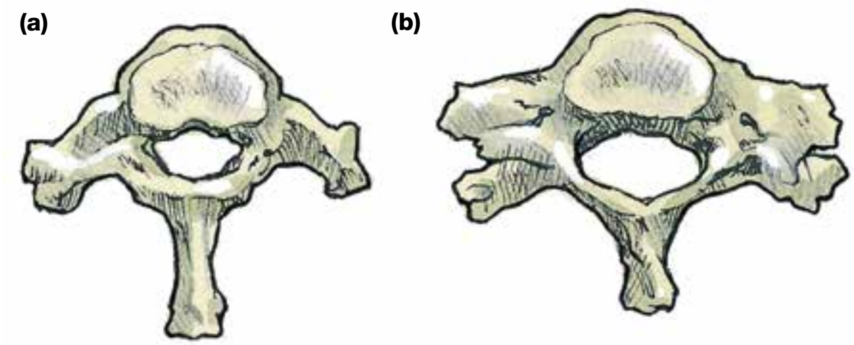



**FIGURE 12.59** Vertebrae from **(a)** *Homo erectus* and **(b)** *Homo sapiens*. Note that the space within the bone that accommodates the spinal cord is far larger in modern humans than in *H. erectus*. What conclusion was drawn from this difference?

## The Hobbit: *Homo floresiensis*

In October 2004, the discovery was announced of a remarkable hominin fossil found in a cave on Flores Island in Indonesia. It comprised the skull, teeth and jaw of an adult female and was given the scientific name *Homo floresiensis* (see figure 12.60). Later, other fossils of the same species were found in the same location. Based on recalculation of the age of these bones announced in 2016, these fossil bones have been redated from about 100 000 to 60 000 years ago. This redating indicates that this species became extinct much earlier than previously believed.

***Homo floresiensis* is a new hominin species**. Where does this species fit in the human family tree? One view is that this species evolved from *Homo erectus*. Further studies, possibly of ancient DNA extracted from fossil bones, will shed light on the relationships of *H. floresiensis*.

What was remarkable about this fossil was its size. The adult female *H. floresiensis* was just one metre tall and had a brain capacity of just 380 millilitres in a head 'about the size of a grapefruit' (compared to a brain capacity averaging 1400 mL in modern humans). *H. floresiensis* made tools and used fire.

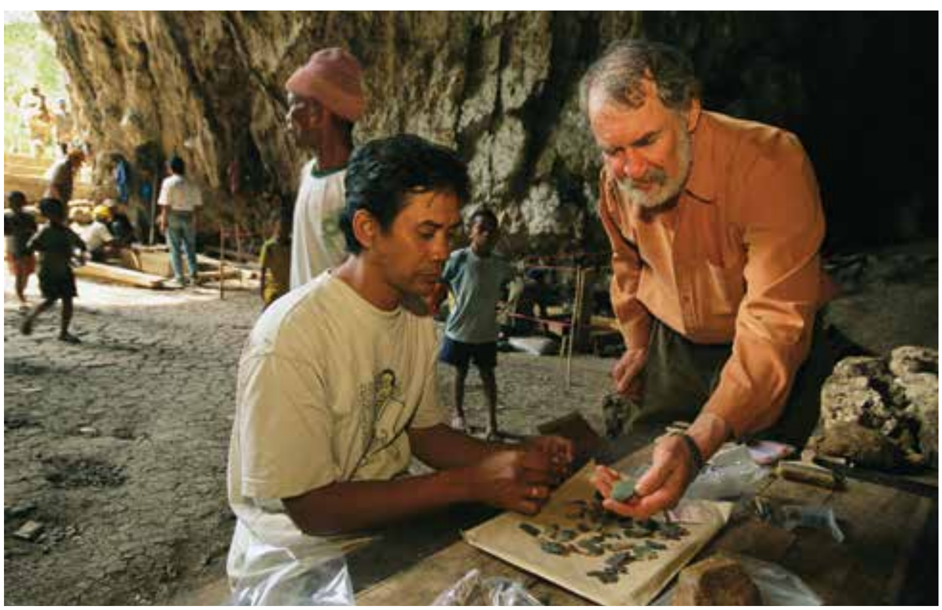


**FIGURE 12.60** Photo taken in 2007 showing Dr Mike Morwood (at right) with another scientist examining stone artefacts found in the Liang Bua cave on the Indonesian island of Flores. Dr Morwood was the leader of the Australian–Indonesian team that, in 2003, made the discovery in this cave of the hominin fossil *Homo floresiensis*, popularly called the Hobbit.

## *Homo heidelbergensis*

It is hard to identify where the fossil record of *Homo erectus* ends and that of *Homo sapiens* begins. This difficulty occurs because fossils exist that are intermediate between the two species. **The intermediate form is known as *Homo heidelbergensis*.** (see table 12.9).

**TABLE 12.9** Comparison of various *Homo* species. What single feature distinguishes *Homo heidelbergensis* from *Homo erectus*?
+ : feature present
– : feature absent

| Feature | *H. erectus* | *H. heidelbergensis* | *H. sapiens* |
|---|---|---|---|
| prominent brow ridges | + | + | – |
| rounded skull | – | – | + |
| big teeth | + | + | – |
| sloping forehead | + | + | – |
| large brain (≥1200 mL) | – | + | + |
| pointed chin | – | – | + |

*Source:* Based on information from the Australian Museum.

**studyon**

Unit 4
AOS 1
Topic 4
Concept 4

**Fossils of the genus *Homo***
Summary screen and practice questions

**studyon**

Unit 4
AOS 1
Topic 4
Concept 4

**Do more**
Hominin skulls

This human species can be differentiated from *Homo erectus* by its larger brain case (1100 to 1200 mL), which is about 93 per cent that of modern humans. *H. heidelbergensis* is generally accepted as the direct ancestor of the Neanderthals (*Homo neanderthalensis*).

The first fossil of *Homo heidelbergensis* was a heavily built lower jawbone with small human-like teeth that was found in a quarry at Mauer, a village near Heidelberg, Germany, in 1907. In 1908, this fossil was announced as a new human species, *Homo heidelbergensis*. However, it was only after the discovery of numbers of similar fossils in Europe, Africa and Asia that *H. heidelbergensis* became more widely accepted as a human species. The African fossils are the oldest, indicating an African origin for this species. Figure 12.61 shows a sample of the locations where *Homo heidelbergensis* fossils have been found.


**FIGURE 12.61** Map showing the location of some *Homo heidelbergensis* fossil finds. Where have the oldest of these fossils been found?

Fossilised skeletal remains found in Europe provide evidence of short stocky bodies, as would be expected for the colder climate, and large brow ridges. The cranial capacity that indicates brain size was typically in the range of 1200 to 1300 mL.

## *Homo neanderthalensis*

In 1856, parts of a fossilised human skeleton were found in a cave in the Neander Valley in Germany. These strongly built, heavy fossilised bones indicate that they came from a sturdy, stocky individual. This fossil was later called *Homo neanderthalensis*, after the site of its discovery. Later, similar fossils were found in many locations in Europe and the Middle East. As it turned out, Neanderthal fossils had also been found as early as 1829 but they were not recognised at that time as fossils of an extinct human species.

Neanderthal people had large brains with an average size of 1400 mL, were short and strongly built, had long, low and wide skulls, low sloping foreheads, heavy brow ridges, rounded chins, large eye sockets, projecting faces and a gap between the last molar and the back of the jaw (see figure 12.62). The Neanderthal inner ear bones are different from those of modern humans.

At least 440 000 years ago, the line that gave rise to the Neanderthals split from the line that was to give rise to modern humans. Some researchers place the time of this split much earlier, at 500 000 to 800 000 years ago, when the human line and the Neanderthal line both split off from a common ancestor, most probably *Homo heidelbergensis*. (In the rest of this chapter, the figure used for this divergence will be 'at least 440 000 years' ago.)

**ODD FACT**

A German anatomist, August Mayer, identified the Neanderthal fossil found in 1856 as the remains of a Cossack (Russian soldier) from the early 1800s wars who wandered around Germany and died in a cave. His thick-set bones were explained by Mayer as being the result of the Cossack spending most of his life on horseback, combined with the effects of old age and arthritis.

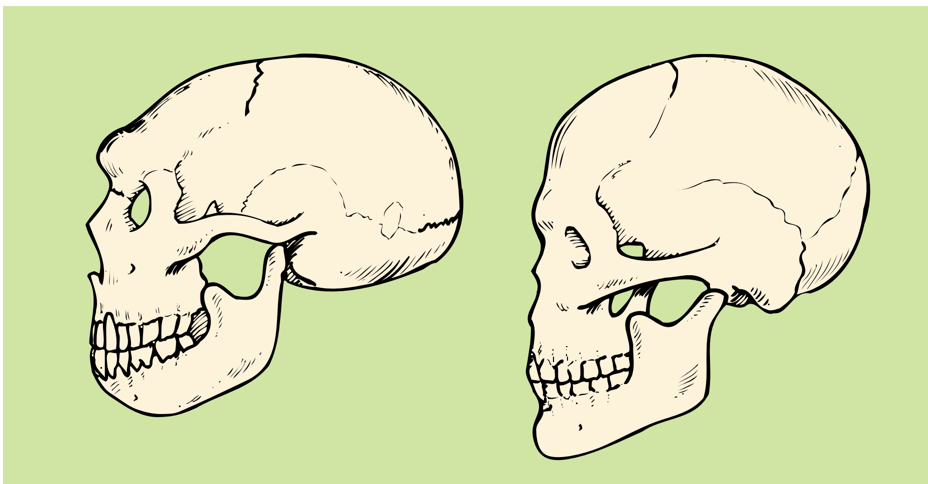


**FIGURE 12.62** Neanderthals (at left) can be distinguished from *Homo sapiens* (at right) by their lower, heavier skulls, the shape of their brain case (cranium), lower foreheads, more protruding jaws and prominent brow ridges. The average brain volume of Neanderthals (1400 mL) is greater than that of modern humans.

**ODD FACT**

Impacted wisdom (third molar) teeth occur in some people because the small jaws of *Homo sapiens* cannot always accommodate the full set of adult teeth. The larger jaw of *Homo neanderthalensis* meant that impacted wisdom teeth were not a problem for these people.

Neanderthal people crafted fine tools from flakes of stone struck from rocks, they produced ornaments and, evidence from a Neanderthal gravesite in Israel supports the proposal that they may have sometimes buried their dead. Neanderthals lived in caves but when caves were not available they made shelters of animal skins on wooden frames. Neanderthals also made clothing for their bodies and feet.

**The Neanderthals exist in the fossil record in Europe and the Middle East over a period from about 200 000 to about 30 000 years ago**. Sometime after this, Neanderthal fossils disappear from the fossil record. **Because Neanderthal**

**people overlapped in time with modern humans, they are *not* ancestors of modern humans, but are cousins. Neanderthals are believed to have evolved from *Homo erectus*.**

### Neanderthals are a distinct species

In 1997, a sample of mtDNA was isolated from a 40 000-year-old piece of arm bone of a Neanderthal fossil. The mtDNA was partly degraded but short DNA fragments were obtained. These many short fragments were amplified using the polymerase chain reaction (PCR). Overlapping regions were identified and, using these overlaps, a long sequence (379 bases) of Neanderthal mtDNA was obtained. The Neanderthal mtDNA sequence was compared with those of modern humans, and an average difference of 27 bases was noted. In contrast, the average difference between human mtDNA samples is eight bases. This finding strongly suggested that modern humans and Neanderthals are different species.

Later, in 1999, 2002 and 2004, longer segments of Neanderthal mtDNA were obtained from different Neanderthal fossils (see figure 12.63). All the Neanderthal sequences showed a high degree of similarity to one another, but were clearly distinct from the modern human samples.

**FIGURE 12.63** A scientist starting to extract DNA from a sample of Neanderthal bone.

In 2010, the first reconstruction of the entire Neanderthal nuclear genome obtained from the 38 000-year-old bones of several Neanderthals was achieved by Svante Pääbo's team at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. **Comparison of the Neanderthal genome with the genome of modern humans revealed many single base differences, sufficient to identify them as two different species. (The differences between Neanderthals and modern humans are much greater that the differences between any two modern humans.)**

Both the physical differences and the genetic differences support the conclusion that *Homo neanderthalensis* and *Homo sapiens* are different species and that the former diverged at least 400 000 years ago from the line that later gave rise to *H. sapiens*. The two species coexisted in several areas of the Middle East and Europe for tens of thousands of years before Neanderthals became extinct.

### Gene flow between Neanderthal and modern humans

The Neanderthals coexisted with modern humans in some regions. Did Neanderthals and our modern human ancestors interbreed? Fossil evidence to answer this question is not available, but genome sequencing has been crucial in providing an answer.

Comparisons of the nuclear genomes of modern humans with those of Neanderthals reveal that about **one to four per cent of the DNA of all non-African modern humans living today came from Neanderthals.** Only native Africans do not have any Neanderthal DNA.

**ODD FACT**

Evidence that modern humans and Neanderthals coexisted in the Middle East comes from the 2015 report of the discovery of a 55 000-year-old modern human skull in a cave in Israel. Neanderthals are known to have lived widely across this region.

This finding provides evidence that, after leaving Africa, modern humans interbred with Neanderthals. The major exodus of modern humans out of Africa occurred about 65 000 years ago. This interbreeding most likely occurred in the Middle East where both species coexisted for a long period and before modern humans spread into Europe and Asia from the Middle East. So, all modern humans, except Africans, today carry in their genomes evidence of gene flow from *Homo neanderthalensis* to *Homo sapiens.*

More recent evidence published in the science journal *Nature* in February 2016 indicates that some interbreeding between modern humans and Neanderthals occurred as early as 100 000 years ago. This evidence comes from the genome sequence of the skeletal remains of a female Neanderthal found in the Altaia mountains in Siberia and dated at 100 000 years ago. Her genomic DNA shows the presence of modern human DNA sequences, indicating gene flow from *Homo sapiens* to *Homo neanderthalensis*. Not only does this provide evidence of much earlier interbreeding between Neanderthals and modern humans, it also provides evidence that some groups of modern humans migrated out of Africa thousands of years before the main out-of-Africa migration that occurred about 65 000 years ago.

Figure 12.64 shows the two phases of interbreeding between Neanderthals and modern humans as described above.

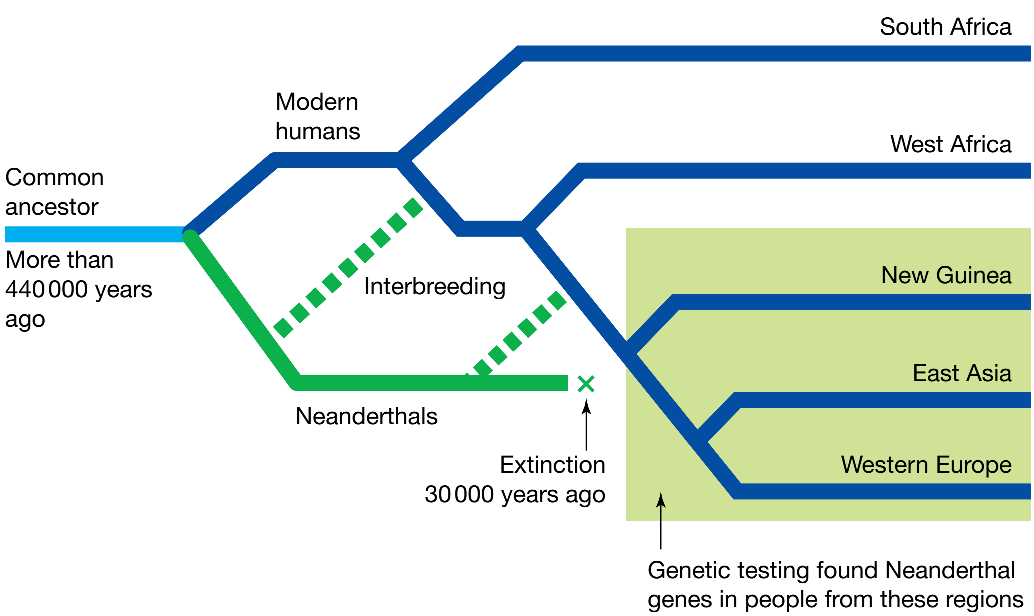



**FIGURE 12.64** Diagram showing the gene flow events or interbreeding that occurred between modern humans and Neanderthals (denoted by dotted lines connecting the human and the Neanderthal lines). Note the two time periods when this interbreeding occurred.

**ODD FACT**

The DNA extracted from the jawbone of a modern human male found in a cave in Romania was sequenced and found to include between six to nine per cent Neanderthal DNA. This high level of Neanderthal DNA indicates that this European male who lived 40 000 years ago had a Neanderthal ancestor between four and six generations back.

This small percentage of Neanderthal DNA that is present in today's non-African modern humans is not the same in each person. Different people today can have different segments of Neanderthal DNA. In all, about 30 per cent of the total Neanderthal genome is represented in the gene pool of non-African modern humans. Neanderthal genes identified in the genome of modern humans today include genes involved in features relating to skin and hair — thicker skin, paler skin, and thicker hair — and other genes related to the function of the innate immune system.

The Neanderthals disappeared from the fossil record about 30 000 years ago. The reason for the demise of the Neanderthals is not known. Several hypotheses are being explored, including competition from better equipped modern humans, and climatic instability.

. . . and just to complicate the human story, as we will see in the next section, another extinct group of humans, the Denisovans, have also left their mark on the gene pool of several modern human populations.

## The Denisovans

A tiny finger bone that was found in 2008 in the Denisova cave in south-western Siberia came from a human female who lived about 40 000 years ago. The DNA extracted from this bone yielded a partial DNA sequence and, unexpectedly, this DNA did not match that of modern humans or Neanderthals or any other known hominin. The DNA was identified as that of a new human group that has been called the **Denisovans**. So, another human group coexisted with modern humans and Neanderthals. This was the first time an extinct human group was recognised, not by fossil bones, but by comparative DNA analysis. Although Denisovan fossil material has so far been found only in Siberia, it is probable that these people spread over Asia. In the case of the Denisovan girl, her DNA included the gene variants (alleles) that would have made her brown-eyed, with brown hair and dark skin.

**ODD FACT**

The Denisovan cave is named after a Russian hermit named Denis, who lived in this cave in the 1700s.

Teeth recovered in the same cave were also identified as Denisovan in origin. From these various fossils, the German research team constructed an entire Denisovan genome sequence. Comparative genomic studies revealed that the modern humans of that period interbred with Denisovans. Traces of this interbreeding are seen today in the presence of Denisovan DNA in some human populations, such as present-day modern humans in the Oceania region and in mainland Asia. The highest percentages (3 to 5%) of Denisovan DNA are found today in Melanesian populations, such as those of Papua New Guinea. This interbreeding probably occurred when the ancestors of these present-day humans were migrating across southern regions of Asia.

Modern humans, Neanderthals and Denisovans all descended from a common ancestor that lived many hundreds of thousands of years ago. This common ancestor is thought to have been *Homo heidelbergensis*. The ancestors of modern humans then diverged to form a separate branch that has lasted to the present time and produced *Homo sapiens*, the modern human. The other branch was to give rise to the Neanderthals and the Denisovans. At a later time, the Neanderthals and the Denisovans split from each other to form two new evolutionary lines.

Figure 12.65 shows these divergences in the human evolutionary line and the interbreeding between the various human species that has been revealed by comparative genomic studies. These studies indicate that several gene flow (interbreeding) events occurred between Neanderthals, Denisovans and early modern humans. The possibility also exists of gene flow into Denisovans from an unknown archaic group.

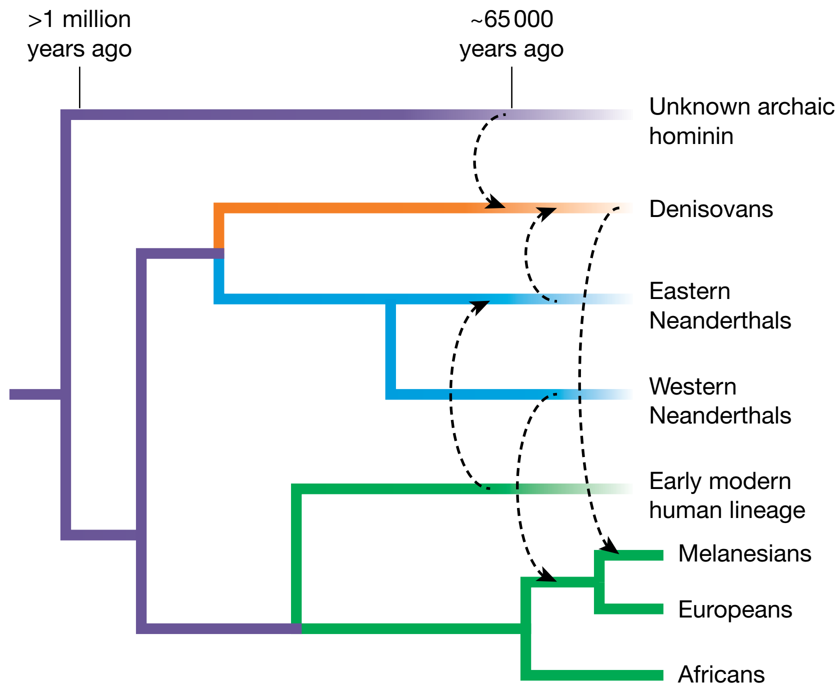



**FIGURE 12.65** Diagram showing gene flow or interbreeding (denoted by dotted arrows) between hominin species as revealed by comparative genomics. Neanderthals are shown as two geographically separated groups, but they are the one species. The fading of the lines for most groups denotes their extinction.

In various publications, the dates of the divergences shown in figure 12.65 above can differ. Why the differences?

The differences depend on the methods used to obtain an estimate of a particular date. This may be either fossil evidence or genomic evidence, and the DNA evidence may be from nuclear DNA or mitochondrial DNA. Based on comparisons of their genome sequences, the line leading to modern humans, and the Neanderthal–Denisovan line most likely diverged about 600 000 years ago in Africa. The split of the Neanderthals from the Denisovans is estimated to have occurred at least 440 000 years ago. Following this split, the Denisovans migrated east into Asia and some Neanderthals migrated west into Europe.

**studyon**

| | |
|---|---|
| Unit 4 | **Characteristics of humans** |
| AOS 1 | Summary screen and practice questions |
| Topic 4 | |
| Concept 7 | |

## Modern humans: *Homo sapiens*

We belong to the species *Homo sapiens*, sometimes termed 'modern humans'. Fossil evidence shows that modern humans evolved in Africa about 200 000 years ago. In the relatively short period of 200 000 years since modern humans emerged in Africa, as compared to the 7 million years since the human line split from the chimp line, the modern human species has spread across the globe. Modern humans now occupy almost all the world's terrestrial habitats. This geographic expansion has been supported by cultural developments and technological evolution as described on pages 600 to 606 below.

The first *Homo sapiens* people are sometimes called **Cro-Magnons** after a location in southern France where fossils of this type were found. A Cro-Magnon skull found in 1868 in a rock shelter in a limestone cliff was among the first fossils to be recognised as members of the modern human species, *Homo sapiens*.

Indistinguishable from today's human beings, Cro-Magnon people had:

- large brains with an average size of 1350 millilitres
- high, steep foreheads
- short, high, narrow skulls
- small or absent brow ridges
- small eye sockets
- pointed chins
- no gap between the third molar and the jawbone.

Cro-Magnons were very skilful toolmakers and produced a range of stone items that could be used to scrape, cut, chisel and pierce. They used their stone tools to craft other materials, such as bone and ivory, to produce a variety of instruments including fish hooks and needles. Evidence of the art of Cro-Magnons exists in caves in France and Spain where wall paintings depict animals that the Cro-Magnon people hunted (see figure 12.66).

**ODD FACT**

The world's greatest collection of prehistoric human art was discovered in 1994 in the Chauvet cave in southern France. This art is dated at 36 000 years ago and is spread over six chambers in the cave. Included among the 400 charcoal drawings are woolly rhinoceroses, lionesses in pursuit of aurochs (extinct cattle) and wild horses.

**FIGURE 12.66** Cro-magnon art. Cro-Magnons were the earliest modern humans in Europe. This image shows detail from the Black Frieze in the Pech Merle cave, Lot, France, from about 15 000 years ago. Note the mammoth (at left) and an auroch (at right) painted in black outline.

**Cro-Magnons buried their dead**. Burials dated at 30 000 years ago have been found where bodies were decorated with necklaces and surrounded by objects such as clay figurines and tools of bone and stone that were buried with them.

**Modern humans are capable of abstract thought, forward planning and complex speech, make use of symbols and conduct ceremonies** (see figure 12.67). We are the only living species that can contemplate the future, including our own deaths.

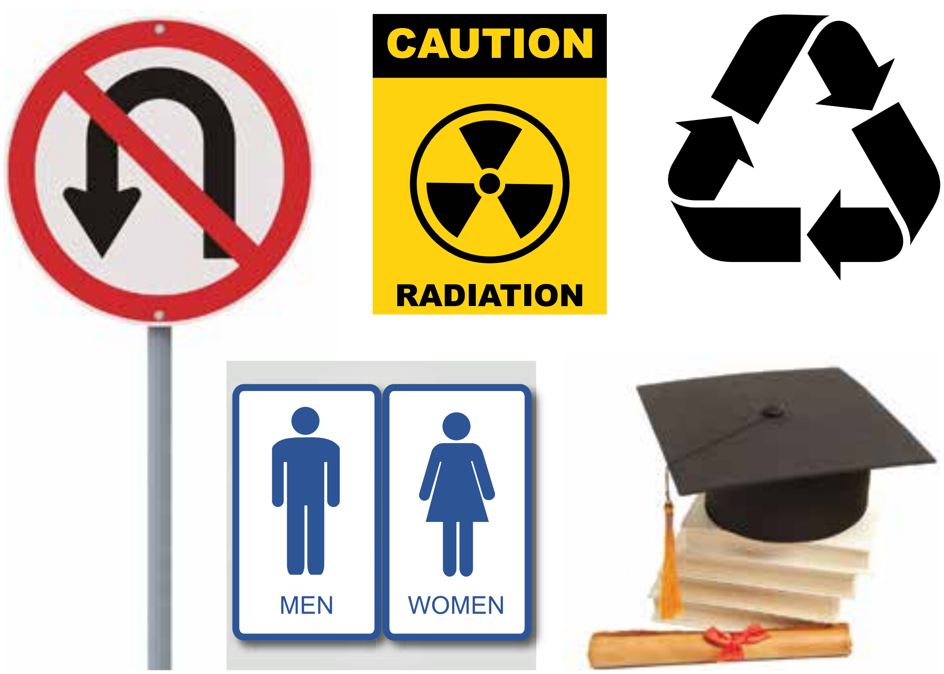



**FIGURE 12.67** Modern humans are characterised by the use of symbols and ceremonies.

# Views of human evolution

A nineteenth-century concept was that there was a 'missing link' between ape-like ancestors and modern humans, as shown in figure 12.68a. The human family tree was seen as a straight line from an erect-walking ape-like ancestor or 'missing link', through ancestral forms, each being more human-like and less ape-like, and finally producing modern humans, *Homo sapiens*.

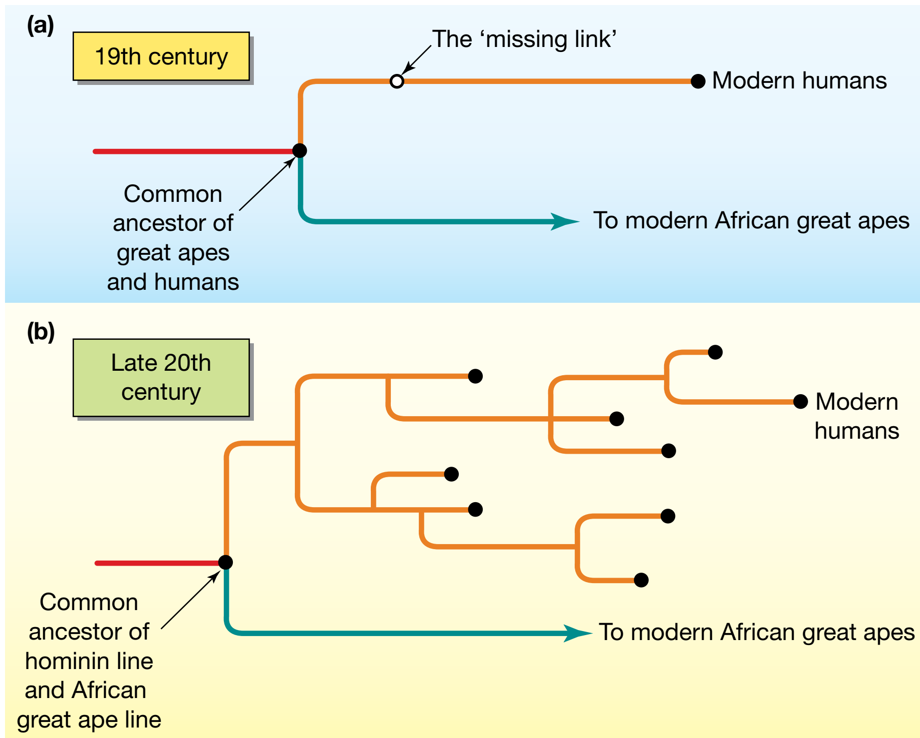



**FIGURE 12.68** **(a)** The nineteenth-century concept of a 'missing link' between the last common ancestor of modern African great apes and humans. **(b)** A late twentieth-century concept of an evolutionary 'bush' with many side branches. Note that this diagram illustrates a concept, *not* the precise number of side branches.

**studyon**

Unit 4 | AOS 1 | Topic 4 | Concept 9

**Current views on human evolution**
Summary screen and practice questions

Instead of a short straight evolutionary line, a spate of hominin fossil discoveries during the last third of the twentieth century produced **an evolutionary pathway leading to modern humans** that is a bush with many branches, mostly with dead ends (see figure 12.68b).

Figure 12.69 shows a view of hominin evolution that identifies the major branches on the evolutionary 'bush'.

Figure 12.69 does not to capture the most recent discoveries of new species, such as the Denisovans, which have made the hominin evolutionary bush even more tangled when interbreeding is brought into the equation.

New fossil discoveries, such as *Homo naledi*, first described in 2015, raise questions about their places in hominin evolution. Similar questions were asked when detailed descriptions of the fossil hominin *Australopithecus sediba* were first published in 2011, showing that it had some features similar to the australopithecines and others similar to early *Homo* species. As stated in the 9 September 2011 issue of the journal *Science*: 'Partial skeletons of 2-million-year-old hominin *Australopithecus sediba* leave researchers impressed by their completeness **but scratching their heads over the implications for our family tree**'.

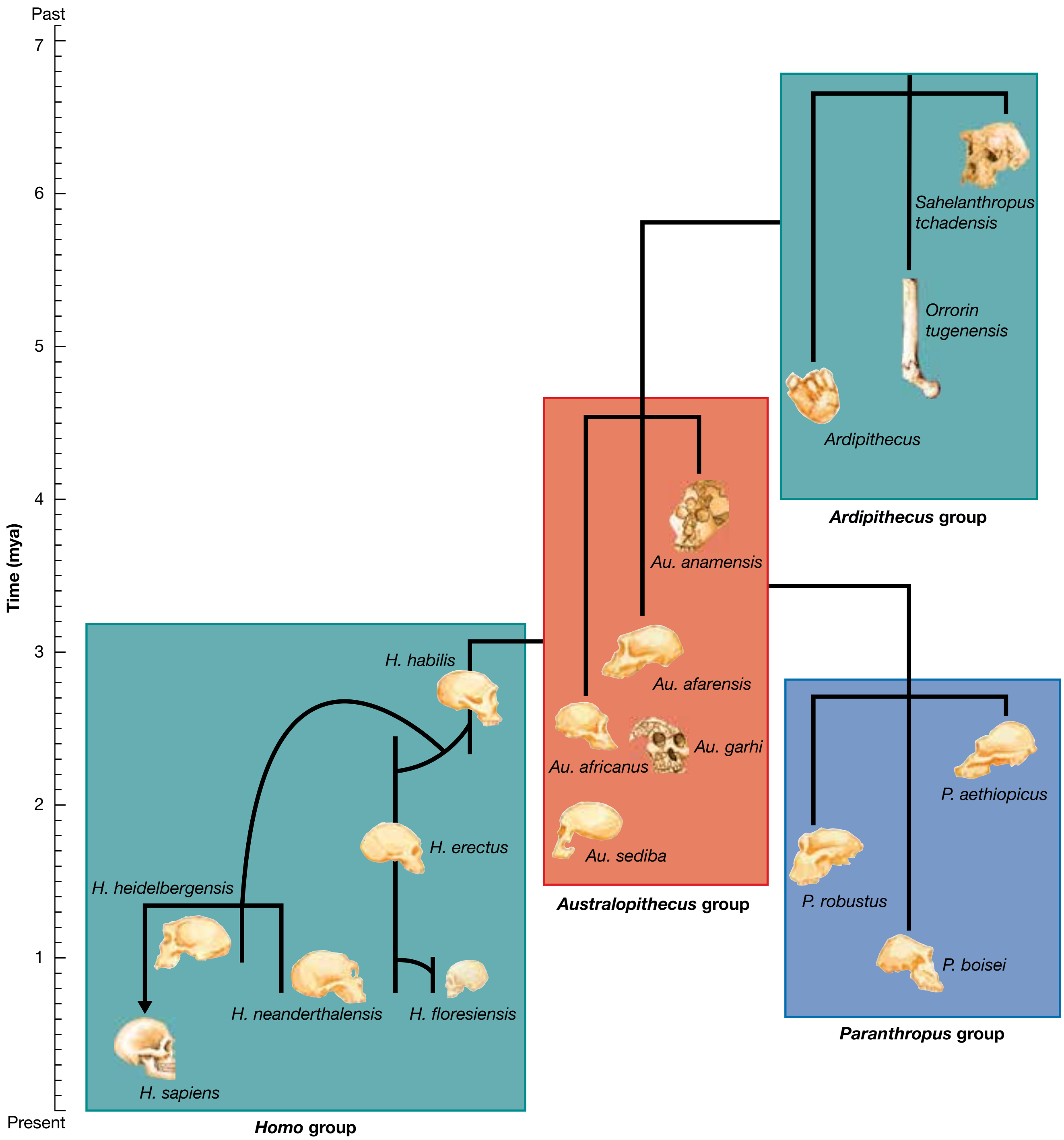



**FIGURE 12.69** Evolutionary links between major hominin groups. Precise ancestral species are not identified, apart from some inferred ancestral links between species within the genus *Homo*. Not every hominin fossil is included.

**ODD FACT**

In 2007, Meredith Small described the human family tree as '*now a tangled messy bush*'. What might she write a decade later?

Comparative genomics is providing new insights to hominin evolution. It was only through comparative genomics that the Denisovans, the new hominins in our family tree, were discovered. As at 2016, the Denisovans are represented by a tiny fingerbone and two teeth and have been described as 'a genome in search of a fossil.'

Just a few years ago, the existence of interbreeding between modern humans and Neanderthals was not known. Today, we know that modern humans interbred with Neanderthals and with Denisovans and that various groups of today's modern humans carry that evidence in their genomes, most vividly the Melanesian people of islands of the western Pacific.

## KEY IDEAS

- The earliest human species is generally accepted as *Homo habilis*.
- The genus *Homo* also includes *H. erectus, H. heidelbergensis, H. neanderthalensis* and *H. sapiens*.
- There is no agreement as to the number of human species; some scientists favour 'splitting' of some species; others favour 'lumping'.
- Evolution within the genus *Homo* involved physical trends, including an increase in brain size and cognitive abilities, a decrease in tooth size, and behavioural changes, including controlled use of fire and use of symbols and ceremonies.
- Modern humans are probably the first *Homo* species to develop speech in contrast to vocalisations.
- Modern humans emerged in Africa about 200 000 years ago.
- Comparative genomic evidence has identified the existence of a new human group called the Denisovans.
- Comparative genomics provides evidence that interbreeding occurred between modern humans and Neanderthals, and between modern humans and Denisovans.
- The hominin evolutionary tree is highly branched and is subject to change as new hominin fossils are discovered.
- Comparative genomics has become an important investigatory tool in hominin evolution.

## QUICK CHECK

**14** Identify the following as true or false:
- **a** *H. erectus* appeared earlier in the fossil record than *H. habilis*.
- **b** When the first human species appeared, all the australopithecines were extinct.
- **c** The term *Acheulean* refers to the stone tools made by *H. erectus*.
- **d** Interbreeding occurred between Neanderthal and the modern humans.
- **e** A 'splitter' would separate *H. erectus* into two species: *H. ergaster* and *H. erectus*.

**15** Name the following:
- **a** the first hominin species to make controlled use of fire
- **b** the first hominins to eat meat as a major component of their diet
- **c** the term used to describe the simple tools made by the habilines
- **d** the first human species to leave Africa
- **e** the first human group to be identified by comparative genomics.

**16** Identify two physical differences between *H. heidelbergensis* and *H. sapiens*.

# Changing lifestyles of *H. sapiens*

Throughout most of the 2.8 million years of its evolutionary history, members of the genus *Homo* obtained food by scavenging carrion, by hunting animals and by gathering plant materials. This lifestyle was nomadic, and communities moved during the year to follow the food supplies. **Human groups living a completely nomadic life as hunter–gatherers** probably existed in small groups, perhaps of 25 or so people.

**ODD FACT**

Neanderthal humans (*Homo neanderthalensis*) were efficient hunters. In a camp site in Europe dated at 55 000 years ago, animal bones from their hunts were found and were identified as reindeer (72%), woolly mammoth (14%), bison (5.4%), horse (4.6%), woolly rhinoceros (2%) and others (2%).

Success in hunting depended on mutual cooperation between small groups of hunters. A cooperative group would have been more successful in obtaining food than a solitary hunter–gatherer.

In a group:

- division of labour is possible
- more eyes can be directed to locating edible plants, watching for predators and tracking potential prey
- more hands can be directed to overcoming prey, cutting it up and carrying large quantities of meat back to the home base.

Success in cooperative hunting would have led to food sharing. Behaviours necessary for successful hunting imply the existence of a mental culture based on planning, ingenuity, foresight, cooperative behaviours and technological skill.

Like other *Homo* species that either lived before or coexisted with them, early modern humans were part of hunter–gatherer societies. They subsisted by cooperative hunting of large mammals, and, where they lived with access to the coast, they collected and cooked shellfish and caught fish. Vegetables were also part of the diet of early *Homo sapiens*.

Prehistoric modern humans made stone tools, such as flakes and scrapers. Later, prehistoric modern humans made a range of smaller specialised tools, using not only stone but also other materials, such as bone, ivory and wood. Many of the tools of these modern humans related to their hunting activities and included specialised tools such as harpoons, bows and arrows with stone tips, long thin stone flakes that were shaped into knives, and spears with stone tips. One invention that appeared about 17 000 to 15 000 years ago was the spear thrower or atlatl (see figure 12.70).

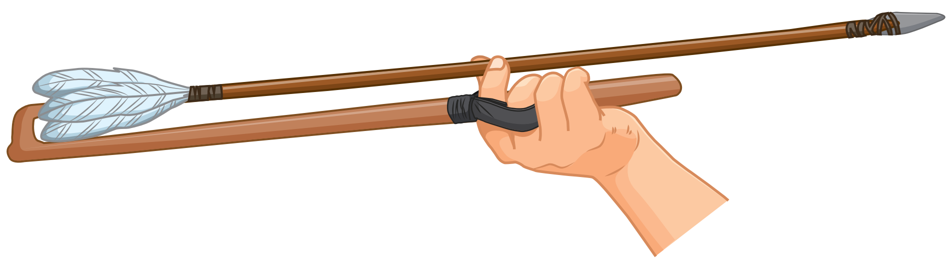


**FIGURE 12.70** Drawing showing a spear thrower or atlatl. Note the projectile, a spear, that fits into a cavity at the end of the spear thrower.

Spear throwers were rods made of wood or bone with a hook at one end to fit the end of a spear. Using a spear thrower effectively lengthened the arm of the thrower and enabled the spear to be thrown further and with greater force as compared to one thrown with no assistance. Their use enabled hunters to throw spears at the target from a greater and safer distance. By improving the leverage of the human arm, a spear thrower more than doubles the distance covered by a spear, and, depending on the thrower, increases the killing distance to between 18 and 27 metres.

Another way of life is that of permanent settlers who domesticated local wild animals and cultivated local wild plants as sources of food. Under these circumstances, humans can live in very much larger sedentary groups with a higher population density.

The transition of *Homo sapiens* from the hunter–gatherer lifestyle to the sedentary, food-producing agricultural lifestyle was not necessarily a sudden and complete transition. When and where were the first agricultural settlements established?

## Agriculture and the first villages

**About 10 000 to 12 000 years ago, modern humans first adopted agricultural practices**. The first evidence for the tending of domesticated animals and the growing of crops comes from areas of the Middle East (see the map in figure 12.71) known as the 'Fertile Crescent'.

**ODD FACT**

Pea seeds have been found in the remains of the ancient settlement of Jericho and in deposits from Turkey dated at more than 7000 years ago. Pea cultivation spread from there, across India, and reached China in the Tang Dynasty (AD 618–906).


**FIGURE 12.71** The 'Fertile Crescent' area of the Middle East. The first food crops to be cultivated in the Fertile Crescent were einkorn wheat, emmer wheat, barley, lentils, peas, chickpeas and bitter vetch (related to peas). Flax was also grown as a source of fibre. (Map provided by MAPgraphics Pty Ltd, Brisbane.)

The first plants to be cultivated were crops derived from the wild wheat, barley and pea plants that grew locally. The first animals to be domesticated and herded were the wild sheep and goats of the region. This change in lifestyle allowed permanent communities to develop with surplus food stores close at hand. **An agricultural lifestyle centring on permanent communities created greater opportunities for our species to engage in the creative pursuits**, such as arts, crafts, literature and music, that are an essential part of 'being human' (see figure 12.72).

**FIGURE 12.72** This Greek amphora attributed to the Leagros Group, made in 525–500 BC, is an example of the creativity of the human species.

Human groups in other regions also developed agricultural practices. These groups domesticated wild animals and plants from their regions, such as corn and cotton in the case of the people from central America (about 9000 years ago) and soya beans, yams and pigs in the case of the Chinese (about 7000 years ago).

**Cultivation of crops and farming of animals are more efficient** procedures for producing food than hunting and gathering. For a given area under cultivation over a given period, an agricultural society growing crops can produce about one hundred times more food for human consumption than the amount produced by hunting. As a result, the size and the density of populations living in permanent communities that relied on agriculture were much larger than those of nomadic groups that relied on hunting and gathering. The social structure in these larger permanent communities also became more complex with division of labour.

**The establishment of permanent communities marked the beginning of selective breeding**, when the early farmers selected:

- particular animals for breeding because they showed preferred features, such as high milk yields
- particular plants to provide the seeds for the next season's crop because they showed preferred features, such as ease of harvesting.

The establishment of permanent communities also led to the development of tools that could enhance food production, such as the plough to assist planting, the sickle for harvesting and the loom for producing wool.

**ODD FACT**

In a good season, one person in a permanent community could harvest nearly three kilograms of cultivated cereal per hour. At this rate, one year's supply for a family could be harvested in about three weeks.

**KEY IDEAS**

- Fossil discoveries support a multibranched human evolutionary line.
- Nomadic lifestyles involved small groups of cooperative hunter–gatherers.
- Prehistoric modern humans developed more complex tools using a variety of materials.
- About 10 000 to 12 000 years ago, evidence of agricultural practices first appeared in the Middle East.
- Sedentary lifestyles involved food production through domestication of wild animals and cultivation of wild plants.
- The sedentary lifestyle was associated with increased population size and a more complex social structure with division of labour.

**QUICK CHECK**

**17** Identify one difference between the nineteenth-century view of human evolution and that of the twenty-first century view.
**18** Why were spear throwers very useful inventions?
**19** Where is the 'Fertile Crescent'?
**20** Identify two differences between a food-gathering lifestyle and a food-producing lifestyle.
**21** Identify the following as true or false:
  **a** Nomadic societies have more complex social structures than sedentary societies.
  **b** Agricultural practices developed at the same time worldwide.
  **c** Selective breeding practices were used by nomadic hunter–gatherers.
  **d** The cultivation of wild rice began in the Fertile Crescent.

# Changes in human populations

A group of international students settles around a campfire. One girl from Norway has come for a year of study in Australia; she is fair-skinned and blue-eyed and has fine, blond straight hair. Another student is dark-skinned, brown-eyed with short frizzy hair and is from Nigeria. Among the others is a young man from Vietnam. All members of the group are warmly clothed in anticipation of the chill that will come after nightfall. In the background, someone uses a small axe to chop through a piece of wood that he then adds to the campfire. As the light fades, one member of the group puts down a book that she has been reading and relaxes by listening through her headphones to pop music played on her iPod. Another person tunes a portable radio to a frequency that picks up a broadcast of classical music, and yet another picks up a mobile phone to text a friend.

This scenario captures some different aspects of evolution or change that have affected human populations over time, namely:

- biological evolution
- cultural evolution, also termed 'cultural change'
- technological evolution.

Changes in these three areas over time resulted in the transformation of our hominin ancestors to modern human beings.

## Human biological evolution

**Biological evolution** refers to changes in human populations that have resulted from natural selection acting over generations on inherited phenotypes and producing differential survival and reproduction under a particular set of environmental conditions. **The result of biological evolution in human populations can be seen in certain inherited variations in the group around the campfire**, including those expressed in:

- physical traits, such as variation in skin colour and hair texture
- biochemical traits, such as variation in ABO and Rhesus blood types
- physiological traits, such as variation in tolerance to milk sugar (lactose).

It is important to note that, if the environmental conditions change, the selective advantage of one phenotype may be lost.

Let's move from the group around the campfire and look at human populations more generally. Variations can be seen between members of different indigenous populations — contrast the light skin and fair hair of Nordic populations with the dark skin and black hair of African populations. In addition, variation in certain traits can be seen within one population — contrast the differences in ABO blood type in members of one Caucasian population.

The inherited variations present in and between different modern human populations include: (i) many that were inherited from ancestral human species, (ii) some that have arisen from gene flow as a result of interbreeding with Neanderthal humans, and (iii) some that have arisen over the period of about 200 000 years that the modern human species has existed. **The variation or genetic diversity that exists today, within and between human populations, is the expression of biological evolution under particular environmental conditions**.

## Human cultural evolution

**Culture** refers to any form of learned behaviour, such as how to make a tool, learning a social rule, and learning to play a musical instrument. Learning can entail direct teaching or imitation or both. Cultural transmission has been enhanced through technologies, such as writing, printing and, now, computer systems.

When persistent and common changes occur in human populations as a result of transmission of learned or imitated behaviours, these changes are said to be the result of **cultural evolution**. **Cultural evolution or cultural change refers to changes in human societies over time where those changes are socially transmitted, not genetically transmitted**. These cultural changes include new ideas, knowledge, customs, symbols, values and skills. Their transmission from one generation to the next may involve teaching, imitation, conformity or any other factor that influences human behaviour. (Since the advent of mass media, advertising has become such a factor.) Most commonly, transmission of cultural data depends on the use of spoken, written or signed language.

The impact of cultural change can be seen in the group around the campfire (refer back to page 599). If we compare these students with their parents when they were teenagers, cultural changes will be seen, for example, in their style of clothing and their mobile phones and other electronic devices. The use of an axe by one student is part of a culture that reaches back to human prehistory. The modern axe may differ in its construction from a prehistoric stone hand axe, but its function is unchanged.

Music, art, sculpture and poetry are part of human culture. Musical instruments first appeared during the Upper Palaeolithic period (40 000 to 12 000 years ago). Figure 12.73 shows an ancient flute made during this period from a swan's wing bones. Art, such as cave paintings and carved figurines, also appeared during this time.

In contrast to the slow pace of biological evolution, however, cultural evolution is occurring at an ever-increasing rate. This is possible because cultural change is not transmitted genetically but can be transmitted by learning from others. The pace of cultural evolution is illustrated by the fact that, over a period of about 100 000 years, modern human beings have moved from stone hand axes to space probes, and from fire as a source of energy to nuclear power.

FIGURE 12.73 Photograph showing what was considered to be, until 2012, the world's oldest musical instrument: an ancient flute made from a swan's wing bones. The flute, dated at about 35 000 years old, was found in Hohle Fels Cavern in southwest Germany and is now held at the Landesmuseum in Stuttgart, Germany.

**ODD FACT**

Other ancient flutes, dated at 42 000 to 43 000 years old, have been found in southern Germany. Some are made of bird bone and others are made of ivory from mammoths.

Over the period of just ten thousand years since the first civilisations appeared, remarkable cultural change can be seen, including the developments of social institutions, transport, communication, food production and medical treatment. Some cultural changes may be trivial, such as changes in fashions (see figure 12.74), while other changes are significant, such as means of communication.

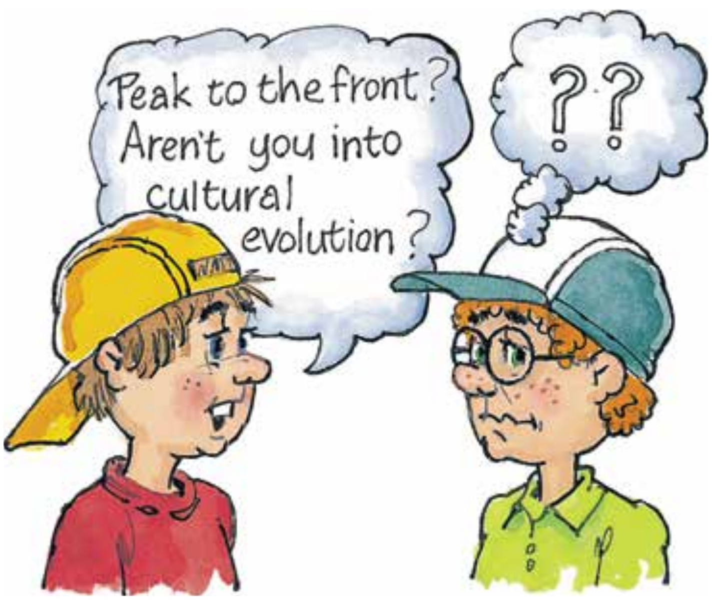



FIGURE 12.74 Cultural transmission of behaviours can occur rapidly and bring about changes in the behaviours of human populations.

## Biological and cultural evolution compared

Terms such as 'biological evolution' and 'cultural evolution' may suggest that the two processes are similar. In reality, there are major differences between the two, as shown in table 12.10. For these reasons, biologists such as the late Stephen Jay Gould suggest that the term 'cultural change' is preferable to the term 'cultural evolution'.

In contrast to biological evolution, **cultural evolution or change, often based on technological evolution, occurs very quickly in human populations**. For example, which of the following would occur more rapidly in a human population immigrating to a cold climate: (1) evolving a stocky body shape through the processes of biological evolution or (2) learning how to control fire and make fur clothing from the pelts of mammals through the processes of cultural evolution?

Cultural evolution has been a powerful agent of change in human populations. Cultural evolution can also occur in other primates. For example, a female macaque monkey from a troop of monkeys living on a Japanese island began to carry her food to a nearby stream and washed the sand off it prior to eating (see figure 12.75b). Eventually other monkeys copied this behaviour until it spread through the entire troop.


**FIGURE 12.75** Cultural evolution in two species. **(a)** Imitation is a powerful means by which cultural behaviours can be transmitted, as seen with this toddler and her mother. **(b)** The practice of washing food was transmitted rapidly through an entire macaque monkey population in Japan starting with the practice of just one monkey. Here we can see a Japanese macaque (*Macaca fuscata*) washing a sweet potato.

Table 12.10 summarises the differences between biological and cultural evolution.

**TABLE 12.10** Some differences between biological evolution and cultural change.

| Biological evolution | Cultural evolution/cultural change |
|---|---|
| Traits can be transmitted to a person only from parents. | Cultural traits can be transmitted to a person by many unrelated people. |
| Transmission can occur only from one generation to the next. | Transmission can be within or between generations and can occur between people widely separated in time and space. |
| It occurs at a slow pace, with many generations required to spread a trait widely in a population. | It occurs at a fast pace, as the spread can be by immediate learning and does not depend on inheritance. |
| Traits acquired during a person's lifetime cannot be transmitted through genetic inheritance. | Cultural traits acquired during a person's lifetime can be transmitted through processes such as teaching or imitation. |
| People cannot choose which genetic traits they will inherit. | People can choose to accept or reject some cultural traits. |
| Data transmitted are encoded in the genetic material. | Data transmitted are commonly in the form of written, spoken or signed language (see figure 12.76). |
| Biological evolution is unplanned and results from random processes. | Cultural change can be conscious and the result of deliberate processes. |

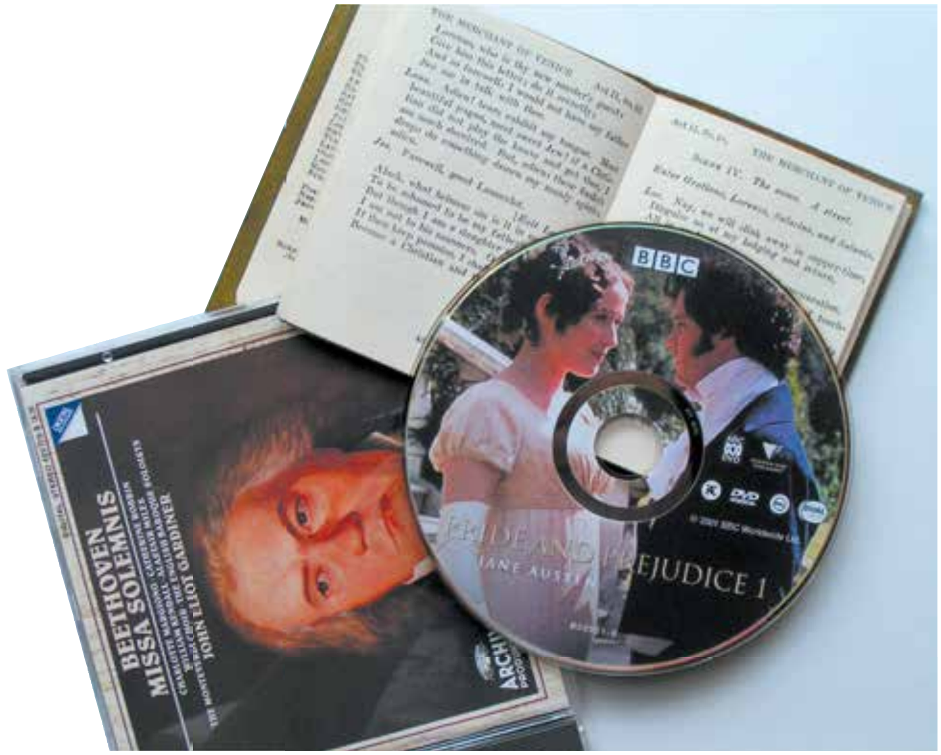



**FIGURE 12.76** Language in written or other forms is a powerful agent through which cultural data are transmitted. Language itself is a critically important human cultural trait. (Image courtesy of Judith Kinnear.)

**ODD FACT**

Tools (such as a pulley or a hammer) make human labour more efficient. Machines (such as a car or a plane) replace human physical effort but require a human operator. Automatons or self-controlled machines (such as a heart pacemaker or a digital clock) operate without any human control.

## Technological evolution

Technology is the application of science by human beings in order to gain better control of the material environment. Technology is expressed through objects, such as tools, machines and automatons (self-controlled machines), and through complex processes, such as medical procedures. **Technological evolution refers to changes in technology over time that give humans increased control over their environment**, such as the change from stone tools to metal tools, from steam power to electric power, from copper to fibre-optic cables for data transmission, and from typewriters to word processors.

Historical examples of technological evolution include:

- the invention of paper by the Chinese in about 105 AD
- the invention in 1456 of the western printing press, which enabled mass production of written material to replace handwritten documents produced in small numbers
- the invention in 1895 of radio, which enabled transmission of spoken language across geographic space
- the invention in the 1930s of television, which enabled images to be transmitted across space (see figure 12.77).

Printing, using ink-stained carved wooden blocks, was invented by the Chinese in the Sui Dynasty (581–618 AD).

In the second half of the twentieth century, technological evolution has led to many developments including fibre optics, lasers and space travel. The twenty-first century has remarkable technological achievements including the self-powered AbioCor artificial heart, which was first used to replace a human heart in an operation in 2001, and the landing of the European Space Agency (ESA) Philae probe on Comet 67P/Churyumov–Gerasimenko in November 2014.

**FIGURE 12.77** Television is a technology of mass communication that allows transmission of moving images and sound across geographic space. Television itself has also evolved from black-and-white to colour, and from analogue to digital.

## Technology extends capabilities

From the early foundations of human civilisation, humans have continued to develop and exploit new technologies that have changed human life.

Biological species are constrained by the limits of their anatomy and physiology — a cat cannot see further than the limits of its eyesight, a dog's bark cannot be heard beyond its immediate neighbourhood, the load carried by a donkey is limited by its skeletal structure and the speed of a horse is constrained by its anatomy. In contrast, humans have developed and used technology to override physical limitations and extend the capabilities of our senses.

Optical telescopes enable us to see stars at unimaginable distances from Earth, and microscopes enable us to see objects far below the resolution of our eyesight. Mobile phones enable us to have a conversation with a person in another country. Microphones linked to amplifiers allow our voices to be heard over greater distances or at higher volumes. Forklifts and cranes enable us to move loads that would crush a person. Cars and planes enable as to move at speeds beyond our own physical limitations.

**FIGURE 12.78** Technology enables us to extend our physical and sensory capabilities and has even enabled *Homo sapiens* to travel in space.

## Biology, culture and technology interact

Biological evolution, cultural change and technological evolution are not separate; they interact. Many of the cultural changes seen in human societies are based on technological evolution. For example, interpersonal two-way communication once depended on face-to-face conversations between two people in the same location. However, technology removed this limitation. First, fixed-line telephones and, later, mobile phones and videophones have enabled interpersonal two-way communication between people widely separated in distance. In the same way, mass communication once depended on a person speaking to a crowd. Technologies such as radio, satellite TV and the internet have changed mass communication and given it a global reach.

If we look at human prehistory, we can also see an example of a profound cultural change that was based on technological evolution and on biological evolution. This cultural change in early human species was toolmaking and it first appeared in the habilines (*Homo habilis*). Initially, toolmaking developed through a technology using stone tools to create new tools that were also made of stone: for example, using a hammer stone to make stone flakes to be used as cutting tools (refer back to figure 12.58). The significance of this cultural change is in the fact that tools, such as axes, enable the power of the body to be concentrated on a small area of an object so that it can be cut or otherwise manipulated, and tools, such as levers, provide a mechanical advantage.

The development of a toolmaking culture by early human species depended on both technological evolution and biological evolution. Why? Toolmaking was possible only because, much earlier, biological evolution in hominins had resulted in two important developments:

1. bipedal (two-footed) locomotion that freed the hands for other purposes
2. increased brain size that gave the capacity for greater problem-solving skills.

So, **biological evolution provided the physical and cognitive framework in which cultural change based on technology could take place in human species**.

## DEVELOPMENTS IN TOOLMAKING

Lucy and the other australopithecines probably used sticks and other natural objects as tools rather like the way chimpanzees do today. In contrast, members of the genus *Homo* not only used tools but also made them, using tools to make other tools. *Homo habilis* broke stone pebbles to produce crude tools for scraping and cutting. *H. erectus* fashioned regular tools from stone. *Homo sapiens* continued the use of stone and produced even more elegant tools. This has been a common theme throughout human evolution.

So, toolmaking using tools is an essentially human trait that distinguished us from other animals (see figure 12.79).

Toolmaking is a part of human culture that is preserved in the fossil record, beginning with the simple stone tools of Oldowan culture. Technology is the use of tools. The early tools, both Oldowan and Acheulean, used by hunter–gatherers were stone tools. Table 12.11 shows the major periods of technological development in human prehistory and history.

In the Middle Palaeolithic period (200 000 to 40 000 years ago), toolmaking became more sophisticated, and stone points were combined with sticks to form lances and spears. This technological development may have marked a change from gathering carrion to hunting large animals. Stone scrapers were probably used to prepare hides to make clothing, shoes, rugs and containers.

**FIGURE 12.79** Toolmaking is an essential part of 'being human'.

*(continued)*

**TABLE 12.11** Major periods of technological development in human prehistory and history. The time ranges shown apply to Asia and Europe. Different time ranges apply to Africa. The Palaeolithic periods together are also known as the Stone Age.

| Prehistoric period | Approx. time range (years before present) | Technology |
|---|---|---|
| Lower Palaeolithic | 2 million–200 000 | Oldowan technology: stone choppers and flakes (refer back to figure 12.52) Acheulean technology: bi-face hand stone axes (refer back to figure 12.58c), which first appeared about 1.5 million years ago |
| Middle Palaeolithic | 200 000–40 000 | Mousterian technology: prepared stone cores used as raw materials for the manufacture of smaller tools, including scrapers for dressing animal hides and points for spears |
| Upper Palaeolithic | 40 000–12 000 | Stone blade tools consisting of thin narrow flakes with long parallel sides. Other tools included needles, fish hooks, harpoons and spear throwers |
| Mesolithic | 12 000–10 000 | Small pieces of flint worked into various shapes (microliths) that were cemented with resin to bone, wood or antlers to form tools such as arrows, spears and barbed rods; other tools included axes and adzes |
| Neolithic | 10 000 | Domestication of plants and animals and the rise of agricultural communities |
| Bronze and Iron Ages | 5000 | Beginning of technology based on metals: copper, then bronze and later iron |
| Industrialisation | 150 | Beginning of the industrial age |

In the Upper Palaeolithic period (40 000 to 12 000 years ago), a greater variety of stone tools appeared and also tools made of other materials, such as ivory, bone and antlers, which included multibarbed harpoon heads, spear throwers, engraving tools and bone points. Other objects from this period include the first bows and arrows, ornaments including ivory beads and necklaces made of teeth, and household items, such as bone needles and thread made of sinew or plant fibre.

In the prehistoric development of toolmaking, members of the genus *Homo* moved from using just stone as a material for their tools to using a range of different materials. With the establishment of agriculture, new tools appeared, such as grinding stones for crushing grains, and new technologies appeared, including the production of pottery, weaving and basketry.

About 5000 years ago, people discovered that copper-containing rocks could be heated and melted to release the metal. Goods began to be manufactured, first from copper and later from bronze. One advantage of metal is that damaged or old items can be melted down and new items recast. Metals are much easier to fashion than stone!

About 2500 years ago, iron-based technology was introduced and iron tools were forged. By about the same date, the use of money was firmly established as a basis for trade and business. This replaced the earlier system of barter in which goods were exchanged for other goods.

Today's technology includes electronics and lasers (see figure 12.80) and uses materials such as ceramics and a wide range of rare-earth elements.

**FIGURE 12.80** Laser beams have many technological applications, including their use in creating holograms in which a laser beam etches an image onto a light-sensitive plate.

## KEY IDEAS

- **Factors influencing human populations over time are biological evolution, cultural change and technological evolution.**
- **Biological evolution, cultural change and technological evolution interact.**
- **Cultural change in many cases has been based on technological evolution.**
- **Cultural change refers to changes in human societies transmitted by social agents, such as imitation and teaching.**
- **Technological evolution refers to changes over time in technology that give humans increased control over their environment.**
- **Biological evolution and cultural evolution differ in many ways.**
- **Technology extends human capabilities.**

## QUICK CHECK

**22** Identify two ways in which biological evolution differs from cultural evolution.
**23** Give one example of cultural evolution or cultural change in communication.
**24** Give one example of how technology has extended human capabilities.
**25** Give one example of an interaction between biological, cultural and technological evolution.

# BIOCHALLENGE

1 *Kenyanthropus platyops*, described in 2001, is a prehuman hominin fossil that appears different from existing fossils (see figure 12.81). Is it a new species within a new genus?

Consider the following statements made about this fossil:

i 'There is always, however, the possibility that it [the difference] is an example of extreme variation within an already established taxa.'

'Despite incredible advances in computing technologies, the bias of post-mortem taphonomic [after-death] processes causing deformation of fossils is very real . . . . . It [the fossil] presents features such as a non-prognathic [flattened] face and relatively small molars, which led its discoverers to conclude it belongs to a new taxon. On the other hand, some of these traits could be due to distortion or variation within species.' (Source: 'Kenyanthropus Platyops: Description, Classification, and Controversy', *CJA Anthrojournal*, 19 Feb 2012.)

ii ' . . . the acceptance of this new species is unresolved. Many experts argue that this skull was extremely distorted and has been badly reconstructed. They claim it is a Kenyan variant of *Australopithecus afarensis*.'

'More material needs to be recovered, especially a skull with no apparent distortion, in order to resolve this debate.' (Source: *Kenyanthropus platyops*', Australian Museum.)

**FIGURE 12.81** Skull of *Kenyanthropus platyops.*

Answer the following questions:

a What are the key issues in the debate about the *Kenyanthropus platyops* fossil?

b How would the discovery of more fossil material assist in resolving this debate?

2 A collection of hominin bones was found in Sima de los Huesos (Pit of Bones), a series of caves in the Atapuerca Mountains in Spain (see figure 12.82). These bones have been dated to about 430 000 years ago. Over a period of many years, thousands of bones have been recovered that correspond to the skeletal remains of 28 individuals. Where do these fossils fit in the human family tree?


**FIGURE 12.82** Map showing the location of the Atapuerca Mountains in northern Spain. This is the site of the Sima de los Huesos (Pit of Bones) where many hominin fossils were found.

The 24 March 2016 issue of the science journal *Nature* reported that nuclear DNA sequences recovered from two Sima specimens show that they are related to the Neanderthals rather than to Denisovans.

For each of the questions below, select the most appropriate alternative and give a reason for your choice:

a From the data given above, it is reasonable to conclude that the Neanderthal line split from the Denisovan line:
   i sometime before 430 000 years ago
   ii sometime after 430 000 years ago
   iii these data provide no evidence about the time of the split.

b While they have not yet been assigned to a species, it is reasonable to suggest that the Sima hominins:
   i are early *Homo sapiens*
   ii are very early Neanderthals
   iii are a newly discovered australopithecine species
   iv are Denisovans that have evolved into a new species.

# Chapter review

studyon
Unit 4
AOS 1
Topic 4
Human change over time
Practice questions

## Key words

acheulean technology, Afar Triangle, arboreal, biological evolution, bipedal, brachiation, common ancestor, cranium, Cro-Magnons, cultural change, cultural evolution, culture, Denisovans, gestation period, herbivorous, hominoids, *Homo erectus*, knuckle walk, last common ancestor (LCA), Lucy, molecular clock principle, Oldowan, opposable, sagittal crest, sexual dimorphism, technological evolution, zygomatic arches

## Questions

1 **Making connections** → Construct a concept map for 'hominin evolution' using any of the key words above. Add other terms as needed.

2 **Communicating your understanding** → Outline the trends in hominin evolution in terms of:
- a brain size
- b tooth size
- c shape and angle of forehead
- d size of brow ridges.

3 **Analysing information and applying your understanding** → Examine the fossil hominin skulls in figure 12.83 and place them in order from most ancient to most recent. Explain your choice.

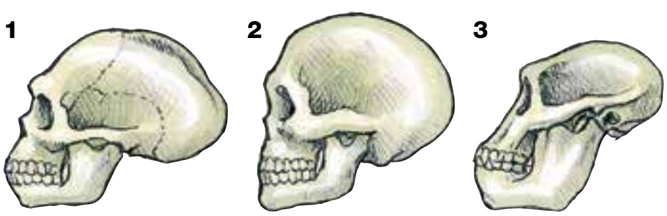



FIGURE 12.83

4 **Applying your understanding** → A skull is found at a dig. One person says that it is the skull of a chimpanzee while a second person says that it is a hominin skull. Identify two key features that you might use to distinguish a chimp skull from that of a hominin.

5 **Analysing information and drawing conclusions** → Suggest possible explanations for the following observations:
- a Fossils classified in the genus *Australopithecus* and other early hominin fossils have been found in Africa only and nowhere else in the world.
- b Crafted stone tools have been found in locations in association with *Homo erectus* fossils but not in locations in association with australopithecine fossils.
- c Prehistoric cave drawings by *Homo sapiens* in certain regions show elephants and mammoths but not dinosaurs.
- d The African great apes do not walk erect but typically use a knuckle walk.
- e The discovery of a new hominin fossil may be the source of many new questions about the human family tree.
- f In a nomadic society, all members of a human group are concerned with the hunting or gathering of food, but this is not the case in a sedentary society.

6 **Making connections between concepts** → Of the groups of five items listed below, four are related, and one does not belong. Identify the 'odd one out' and explain how the other items are related.
- a chimpanzee, *P. robustus*, Handy man, Cro-Magnon, *Au. anamensis*
- b Handy man, Piltdown Man, Cro-Magnon, Neanderthal, Taung child
- c *Au. garhi*, *H. habilis*, Lucy, *Au. africanus*, *Au. sediba*
- d Oldowan, Mousterian, Ardipithecan, Acheulean
- e gorilla, lemur, gibbon, macaque, Rhesus monkey.

7 **Analysing information and solving problems** → Look at figure 12.84, which shows layers of sediments and some fossil skulls.
- a Which fossil is the older, M or P? Explain.
- b If fossil R were identified as *Au. afarensis*, could fossil Q be *Homo habilis*? Explain.
- c Could fossils J and K be *H. sapiens* and *P. robustus* respectively? Explain.

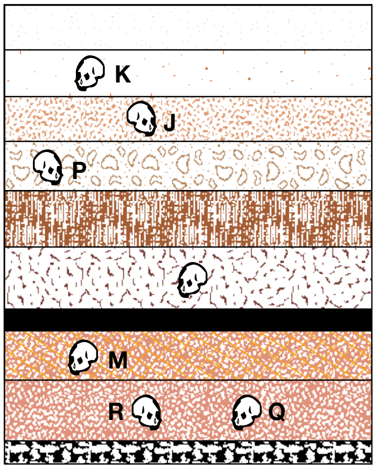



FIGURE 12.84

**8** **Applying your understanding →** Identify the following.
- **a** a new hominin genus based on a fossil discovery in the late 1990s
- **b** a famous family of hominin fossil hunters
- **c** two kinds of prehuman hominin that coexisted
- **d** three kinds of human that coexisted
- **e** the first humans to make controlled use of fire
- **f** a member of tribe Panini.

**9** **Demonstrating knowledge and understanding →** Examine figure 12.85.

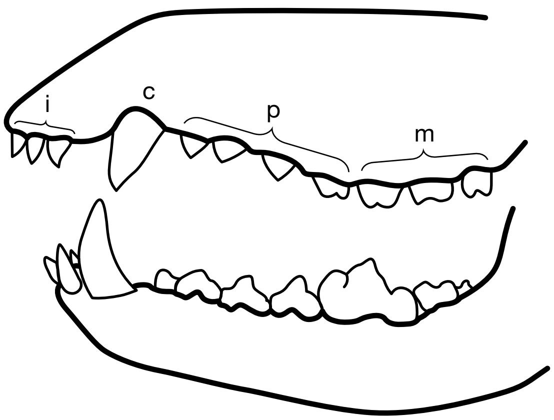



**FIGURE 12.85** Line drawing showing parts of the upper and lower jaws of an animal. i = incisor, c = canine, p = premolar, m = molar.

- **a** On the basis of the information available, which of the following conclusions is reasonable:
  - i this jaw is from a reptile
  - ii this jaw is from a mammal
  - iii it is not possible to draw a conclusion.
- **b** Identify the reason for your decision.

**10** **Communicating knowledge and understanding →** Identify the following statements as true or false. Where false, rewrite the statement so that it would be judged to be true.
- **a** A mammal with binocular vision must be a member of the order Primates.
- **b** A member of phylum Chordata that has a diaphragm must be a mammal.
- **c** The human fossil record is now complete.
- **d** Aspects of the diets of extinct hominins can be inferred from an analysis of the C13/C12 isotope ratio of their teeth.

**11** **Applying your understanding →** Examine figure 12.86 carefully. Identify the evidence visible in the image that would enable you to classify this animal into:
- **a** class
- **b** order.

FIGURE 12.86

**12** **Discussion questions →** Consider the following statements and assess their validity or otherwise, giving a reason for your choice:
- **a** Person A says: 'I'm shocked. The human family tree shows that I evolved from a chimp.'
- **b** Examine figure 12.87, which shows the outline of a modern human on the right. Person B says: 'The image on the left looks as if it is a chimp.'

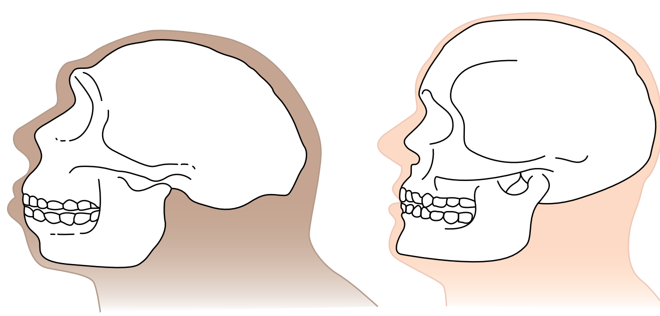


**FIGURE 12.87** Outlines (drawn to scale) of the profile and the skull of two primates. At right is a modern human (*Homo sapiens*).

CHAPTER

# 13 DNA manipulation techniques

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- recognise that DNA can be manipulated in various ways, such as by cutting, joining, and copying
- identify various techniques used to manipulate DNA
- distinguish between endonucleases and exonucleases
- become familiar with the role of the polymerase chain reaction in producing copies of DNA
- develop awareness of the use of plasmids as vectors to transform bacterial cells.

**FIGURE 13.1** Dr Jennifer Doudna, a Howard Hughes Medical Institute Investigator and Professor of Molecular and Cell Biology and Chemistry at the University of California, Berkeley, and co-discoverer of the revolutionary CRISPR-Cas9 gene-editing technology. In this chapter, we will explore some of the tools for manipulating DNA. (Image courtesy of Dr Jennifer Doudna.)

# A recipe for gene editing

The first attempts at genetic manipulation in the 1970s involved adding DNA to the plant and animal genomes.

Early techniques used to add DNA to target cells included:

- the use of so-called 'gene guns' in which gold atoms coated with DNA were blasted into cells, a technique used mainly to insert genes for pesticide or herbicide resistance into plant cells
- the use of modified viruses to act as vectors to carry a functional gene into cells with a defective gene. Viruses used as vectors include retroviruses and adenoviruses.

These techniques were clearly very limited because there was no control over where any added DNA would be inserted into the genome of the target cells. Instead, DNA was simply added randomly into the genome. What would be expected to happen if the added DNA was by chance inserted into a gene that was essential for cell survival, causing the gene to become non-functional? What would be expected to happen if the added DNA was by chance inserted into a cancer-suppressing gene so that this gene was disabled?

CRISPR stands for clustered regularly interspaced short palindromic repeats. These are short repeated segments of DNA, with each repeated segment being separated by a length of spacer DNA.

Another major limitation was that, although these early techniques could carry a functional replacement copy of a gene into cells with a defective gene, these techniques were not able to mend a defective gene by editing or disabling it. Editing is a process of correction, such as occurs when an editor of a manuscript corrects a misspelt word, or adds or replaces a word or phrase, or deletes a sentence.

What was missing was an efficient and reliable technique for making precise and targeted changes to the genome of living cells. Such a process is called **gene editing** or **genome editing** or **DNA editing**. Gene editing refers to changing a genomic DNA sequence in some way. This might involve making a single base change to a gene, such as a base substitution, addition or deletion. Gene editing on a larger scale might involve disabling, replacing or adding a gene, or changing an upstream DNA sequence that regulates a gene. However, in 2012, a 'game-changing' discovery was made.

**FIGURE 13.2** Dr Emmanuelle Charpentier, Director of the new Max Planck Institute of Infection Biology, co-inventor, with Dr Jennifer Doudna, of the CRISPR gene-editing technology. (Image courtesy of Hallbauer & Fioretti.)

In 2012, an article in the journal *Science* (Jinek, M. *et al. Science* **337**, 816–821) described a gene-editing technique that is called **CRISPR-Cas9** (often just shortened to CRISPR and pronounced 'crisper'). Interestingly, CRISPR is part of the adaptive immune system of bacteria that chops any DNA of invading viruses, and has been operating in the microbial world for many thousands of millions of years.

**What was most significant was the recognition that CRISPR-Cas9 technology could be readily adapted to provide an inexpensive and easy-to-use means of genome editing for use in an endless range of genes from any organism**. The co-inventors of this exciting gene-editing technique were two women scientists, Dr Jennifer Doudna (refer back to figure 13.1) and Dr Emmanuelle Charpentier (see figure 13.2).

This tool, borrowed and adapted from the bacterial adaptive immune system, is now being put to use in editing faulty genes and in silencing genes in plants and animals. Many applications in agriculture are being explored (see chapter 15, page 706). Research is already underway on the use of the CRISPR technology as a safe and reliable means of editing the defective alleles responsible for human diseases, such as cystic fibrosis and sickle-cell anaemia.

## How does CRISPR-Cas9 work?

The guide RNA is typically denoted by sgRNA, where sg = synthetic guide.

**ODD FACT**

Cas9 is an *endo*nuclease (*endo* = within), which can cut double-stranded DNA from within the DNA. This is in contrast to *exo*nucleases (*exo* = outside), which can only remove nucleotides from the ends of DNA molecules.

PAM = protospacer adjacent motif. PAM is a three-base sequence (NGG) in one strand of the DNA (N stands for any base). The PAM sequence is just downstream from where the Cas9 nuclease makes one of its two cuts in the DNA.

The CRISPR gene-editing technique simply requires molecular 'scissors' to cut the target DNA and a 'guide' to direct the 'scissors' to the site where the cuts in the DNA will be made.

These two tools are as follows:

1. The 'scissors' are **Cas9 nuclease**, a bacterial enzyme that:
   (i) can unwind double-stranded DNA
   (ii) can cut both strands of double-stranded DNA at a precise location.
2. The 'guide' is a segment of an artificially synthesised single strand of RNA. This **guide RNA** is designed to include a 20-base sequence that is complementary to part of the target DNA.

*What gives the CRISPR-Cas9 technology its potential remarkable range of use?*
Firstly, guide RNA molecules can be designed with a 20-base sequence that can home in on any segment of genomic DNA from any organism. Secondly, the Cas9 enzyme can cut any double-stranded DNA. This is a reminder of the universality of DNA as the material of inheritance.

*How does the CRISPR technology work?*
The cell containing the DNA to be edited is transfected with the Cas9 nuclease enzyme and the guide RNA. The guide RNA hybridises with its complementary DNA sequence and this identifies the position where the Cas9 nuclease enzyme will cut the DNA. The cut is made by the Cas9 nuclease at a location just upstream from a specific three-base sequence called PAM.

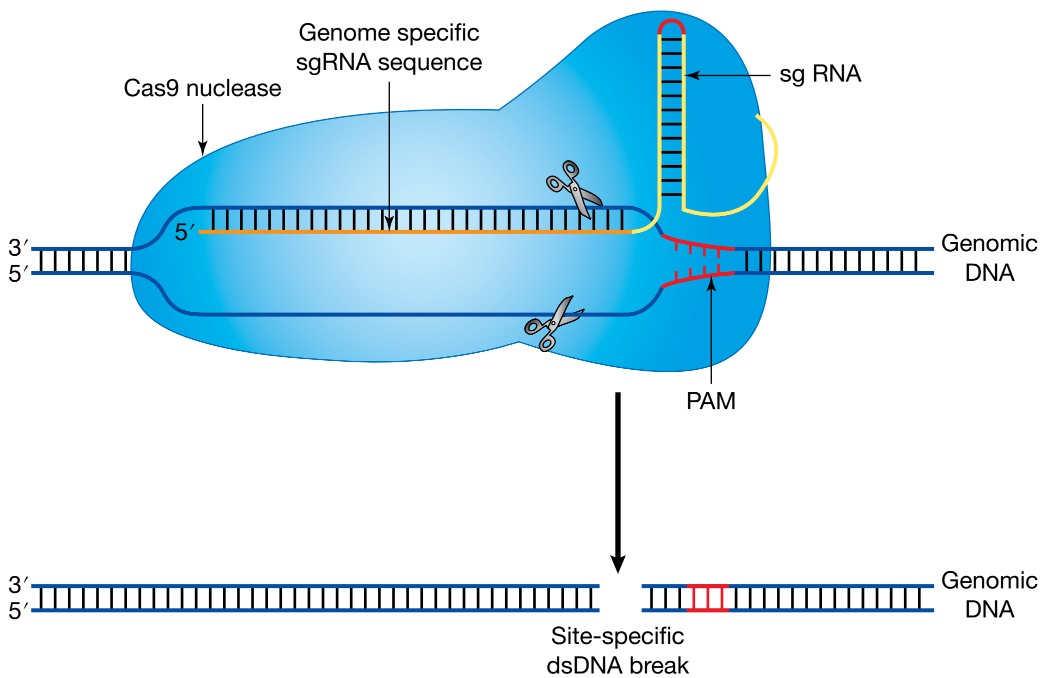



**FIGURE 13.3** Stylised diagram showing the components of CRISPR-Cas9 technology. Note the single-stranded guide RNA (sgRNA) with its leading segment (shown in orange) that is complementary to the DNA of interest. The sgRNA enables the Cas9 nuclease enzyme (shown as the large blue shape) to be positioned to cut both strands of the DNA close to the PAM site (shown in red). The two pairs of scissors indicate the DNA sites cut by the Cas9 nuclease.

After the Cas9 nuclease enzyme has cut both strands of the DNA of interest, the breaks can be repaired and, in this process, two kinds of gene editing can occur.

1. In one type of repair process, called a gene 'knock in', a specially designed DNA sequence is inserted into a precise location in the genome (see figure 13.4). The aim of this type of repair is to edit faulty alleles and restore their normal function. For example, using this process, the gene mutation responsible for cystic fibrosis has been corrected in cultured stem cells taken from individuals with cystic fibrosis.

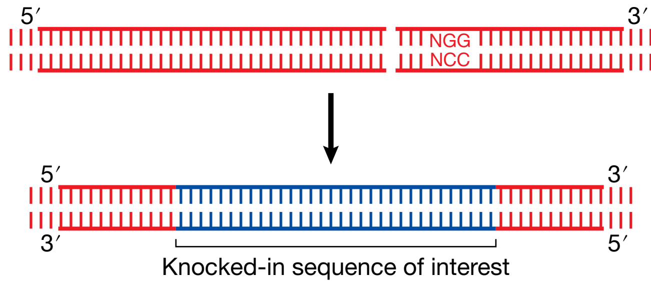



**FIGURE 13.4** Diagram showing a 'knock in' type of repair after the Cas9 nuclease has produced a double-stranded break. The faulty gene is edited and its normal function restored by inserting a specially designed segment of DNA that corrects the defect. In this case, the defect was a missing segment of DNA.

2. In the second type of repair, called a gene 'knock out', the re-joining that repairs the break in the DNA involves a process that is subject to error. This process can result in a random insertion or deletion (indel) of one or two bases, producing a frameshift mutation. A frameshift mutation can either disable a gene or can produce a STOP signal. So, the aim of this type of repair is to disable or silence a gene (see figure 13.5).

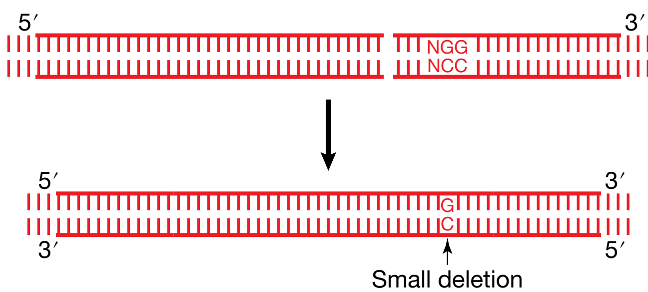



**FIGURE 13.5** Diagram showing a 'knock out' type of repair in which the gene function is disrupted either as a result of a frameshift or as a result of creating a premature stop triplet.

## What about the future?

CRISPR-Cas9 technology has extraordinary potential for widespread use in clinical, agricultural, and research settings, for example:

- for a myriad of research purposes:
  - to 'knock out' genes, one at a time, in order to identify their function
  - to introduce specific mutations in a DNA sequence.
- to edit a faulty allele of a gene in a person with a severe inherited disease
- to snip out the faulty segment of a gene and replace it with a working copy
- to activate or to repress a gene
- to add a new gene to the genome.

This potential is highlighted by the fact that major pharmaceutical companies have made significant investments in CRISPR-Cas9 technologies. For example, in October 2015, Vertex Pharmaceuticals paid in excess of US$100 million to use the gene-editing technology of CRISPR Therapeutics. Their intention is to develop treatments for cystic fibrosis and sickle-cell diseases. More recently, Bayer Pharmaceuticals and CRISPR Therapeutics set up a joint venture aimed at developing therapies to treat blood disorders, blindness and congenital heart disease.

**eBook*plus***

**Weblink**
CRISPR

**eBook*plus***

**Weblink**
CRISPR in the media

In a special review in the international journal *Nature* (25 May 2015), CRISPR technology was introduced as follows:

### CRISPR: THE GOOD, THE BAD AND THE UNKNOWN

'A DNA-editing technology called CRISPR has rapidly become one of the most popular ways to alter genomes. Concerns about its risks temper excitement about its usefulness. It has already been used to modify human embryos, and the technology could alter wild animal populations; it works in everything from wheat to mice . . .'

CRISPR-Cas9 technology itself is neutral — it is a cut and paste or cut and replace technique that can operate on the genomic DNA of a human cell or a mouse cell or a mosquito cell or a wheat cell or a yeast cell, and the list goes on . . .

What about the good, the bad and the unknown? Like any technology, CRISPR will be identified as good or bad depending on the uses to which it is put and their immediate and longer term consequences. As the technology is applied more widely, more unknowns will emerge.

In April 2015, a group of Chinese scientists reported the results of their use of CRISPR technology on human embryos in an attempt to edit a faulty gene. (The embryos came from an in-vitro clinic and, because each was the product of fertilisation by two sperm, these embryos were not capable of completing development.) The results of their experiments appear to have raised many questions and few answers.

Issues being debated include: *Should gene editing be allowed on human cells, such as eggs, sperm and embryos, whose DNA can pass to future generations?* Or, *should gene editing be restricted to somatic cells where that DNA is not passed to the next generation?* These and related ethical issues will continue to engage scientific communities, governments and the public.

In late 2015, the first *International Summit on Human Gene Editing* was held in the United States in Washington, DC. The following headline from a Canadian newspaper that reported on this meeting encapsulates the situation: **CRISPR gene-editing tool has scientists thrilled — but nervous.** (Source of headline: Crowe, Kelly 2015 'CRISPR gene-editing tool has scientists thrilled — but nervous', CBC News, 30 November.)

CRISPR-Cas9 is one technique that involves the manipulation of DNA at the level of genes or even the genome. In the following sections, we will explore some of the other tools that can be used to manipulate DNA.

### KEY IDEAS

- CRISPR-Cas9 technology is an efficient, easy-to-use, and inexpensive tool for precise gene editing.
- CRISPR-Cas9 technology can be applied to every eukaryotic species, including humans.
- The key components of CRISPR-Cas9 technology are a guide RNA and an endonuclease enzyme, Cas9.
- Early gene therapy techniques added DNA segments to genomes but at unpredictable locations.
- CRISPR-Cas9 technology enables precise genomic editing.
- CRISPR technology can not only add DNA but can also edit, modify or disable or delete DNA from the genome.
- Like any new technology with so many potential applications, aspects of the use of CRISPR technology raise ethical considerations.

## QUICK CHECK

1 Identify one capability provided by CRISPR-Cas9 technology that was not possible with earlier gene therapy techniques.
2 Identify the following statements as true or false:
  a CRISPR technology can only be applied to certain species.
  b When CRISPR is used to add DNA to a genome, it is added in a random location.
  c CRISPR can be used to correct faulty genes.
  d CRISPR can be used to silence a gene.
  e The scissors in CRISPR technology consist of a bacterial endonuclease enzyme.

# Tools to manipulate DNA

Carpenters work with wood and sculptors work with clay. Scientists who use gene manipulation technology are sometimes called 'genetic engineers'. Genetic engineers work with the genetic material DNA.

Each operator — carpenter, sculptor or genetic engineer — requires tools to manipulate the raw material. To construct a bookcase, a carpenter needs tools such as a ruler, saw and hammer so that the wood may be measured, cut and joined. Similarly, the genetic engineer requires tools to cut, join, copy and separate DNA. The following section describes the tools and techniques that, when combined, enable genes to be manipulated in various ways.

Table 13.1 shows some of the 'tools' used in gene manipulation. The same tools can be used regardless of the source of the genetic material.

**TABLE 13.1** Tools for gene manipulation.

| Action | 'Tool' |
|---|---|
| • *Cut* DNA into fragments at precise locations | restriction enzymes (see Cutting DNA into fragments) |
| • *Separate* fragments by size | electrophoresis (see Sorting DNA fragments) |
| • *Find* particular DNA fragments | probes (see Finding DNA fragments of interest) |
| • *Join* DNA fragments | ligase enzyme (see Joining DNA fragments) |
| • *Make* specific DNA fragments | DNA synthesiser & reverse transcriptase (see Making DNA) |
| • *Amplify* DNA | Polymerase chain reaction (PCR) (see Amplifying DNA) |

## Cutting DNA into fragments

One class of enzymes can cut double-stranded DNA into a reproducible number of fragments. These enzymes are called **restriction enzymes** or cutting enzymes. These cutting enzymes occur naturally in microbes, mainly in bacteria, but also in some archaea.

The main feature of cutting enzymes is that they do not snip the two strands of a DNA molecule at random. Each cutting enzyme snips DNA at certain positions only, and is *restricted* to cutting at specific locations only. (This is why enzymes in this class are called *restriction* enzymes.)

Each cutting enzyme has its own name. These are strange-looking names such as Bgl I, Hap II, and Taq I, which are pronounced as 'Buggle one', Hap two' and 'Tack one'. The box below explains how the names of cutting enzymes are organised.

## ODD FACT

The first restriction enzyme was discovered at Johns Hopkins Medical School in 1970.

What is significant about restriction enzymes is that their actions on DNA are predictable as they cut DNA only at their specific restriction sites, and they act in the same way on every occasion. Before the first restriction enzyme was discovered in 1970, it was not possible to cut DNA into fragments in predictable and reproducible ways.

## *ECO* RI — WHAT'S IN A NAME?

The names of cutting (restriction) enzymes look very strange indeed . . .
Hap II, Eco RI, Bgl II, Not I, Pst I Hin dIII

The name of a restriction enzyme has two parts. The first part is a set of three letters, such as Eco or Bgl or Pst. The three letters are derived from the scientific name of the microbe in which the restriction enzyme occurs.

The first letter is the first letter of the genus and the next two letters are the leading letters of the trivial (specific) name:

- Eco is a restriction enzyme found in *Escherichia coli*.
- Bgl is a restriction enzyme found in *Bacillus globigii*.
- Hin is a restriction enzyme found in the Haemophilus influenza.

The second part of the name is a Roman numeral such as I or II or III. What does this mean? A particular species of microbe may contain several restriction enzymes. Each different restriction enzyme from the same microbe is given a Roman numeral as follows: I for the first enzyme isolated from that species, II for the second, and so on.

In some cases an additional letter is added, such as in Hin dII, and the letter 'd' denotes a particular strain or type (serotype) of the bacterium.

The position where a cutting enzyme can snip double-stranded DNA is the recognition sequence of that restriction enzyme. A restriction site is a particular order of nucleotides (see figure 13.6). Some restriction enzymes cut the two strands of a DNA molecule at points directly opposite each other to produce cut ends that are 'blunt'. Other cutting enzymes cut one strand at one point, but cut the second strand at a point that is not directly opposite. The overhanging cut ends made by these cutting enzymes are called 'sticky' or 'protruding'. These sticky ends are complementary.

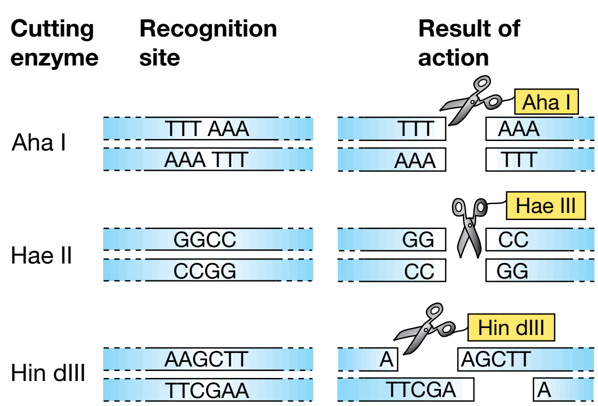



**FIGURE 13.6** Cutting by Aha I, Hae III and Hin dIII. Which restriction enzymes produce 'sticky ends'? Are the sticky ends produced by Hin dIII complementary?

The expected distance between cut sites of a restriction enzyme is given by the formula $4^n$, where n is the number of bps in the restriction site.

The number of nucleotide base pairs (bps) in a restriction site varies from four to about eight base pairs (bps). This length determines how frequently a restriction enzyme is likely to cut a random sequence of DNA, or in other words, the average distance between cuts. For example, restriction enzymes with a 4-bp recognition sequence, such as Hae III, are expected on average to


**FIGURE 13.7** Vial of restriction enzyme, EcoRI. (Reprinted from www.neb.com (2016) with permission from New England Biolabs, Inc.)

studyon

Unit 4 | AOS 2 | Topic 1 | Concept 1

**Restriction enzymes**
Summary screen and practice questions

cut about every 256 bps along a sequence. In contrast, a restriction enzyme with a 6-bp recognition sequence, such as Aha I, will on average cut about every 4096 bps.

Several thousand restriction enzymes have now been isolated, mainly from bacteria, but some from archaea. About 300 restriction enzymes are available from commercial suppliers. Genetic engineers simply buy a quantity of the required enzyme in highly purified form from a commercial supplier (see figure 13.7). Table 13.2 below shows some commonly used restriction enzymes.

**TABLE 13.2** Restriction enzymes in common use and their recognition sites. The arrowheads indicate the site of the cut in each strand of the double-stranded DNA molecule.

| Enzyme | Source | Recognition site |
|---|---|---|
| Alu I | *Arthrobacter luteus* | AGˇCT<br>TCˆGA |
| Bam HI | *Bacillus amyloliquefaciens H* | GˇGATC C<br>C CTAGˆG |
| Bcl I | *Bacillus caldolyticus* | TˇGATC A<br>A CTAGˆT |
| Bgl II | *Bacillus globigii* | AˇGATC T<br>T CTAGˆA |
| Eco RI | *Escherichia coli R factor* | GˇAATT C<br>C TTAAˆG |
| Hae III | *Hemophilus aegyptus* | GGˇCC<br>CCˆGG |
| Hin dIII | *Hemophilus influenzae Rd* | AˇAGCT T<br>T TCGAˆA |
| Kpn I | *Klebsiella pneumoniae* | G GTACˇC<br>CˆCATG G |
| Not I | *Norcadia otitidis-caviarum* | GCˇGGCC GC<br>CG CCGGˆCG |
| Pst I | *Providencia stuartii* | C TGCAˇG<br>GˆACGT C |
| Sac I | *Streptomyces achromogenes* | GAGˇCTC<br>CTCˆGAG |
| Taq I | *Thermus aquaticus* | TˇCG A<br>A GCˆT |

studyon

Unit 4 | AOS 2 | Topic 1 | Concept 1

**See more**
Understanding restriction enzymes

For cutting, a sample of DNA is dissolved and the particular restriction enzyme is added. Provided recognition sites are present in the DNA sample, the DNA molecules will be cut into two or more fragments. The lengths of the fragments depend on the relative positions of the recognition sites (see figure 13.8). In some cases, two different restriction enzymes may be added to a DNA sample. How, if at all, would the resulting fragments be expected to change when two restriction enzymes are used, as compared to just one?

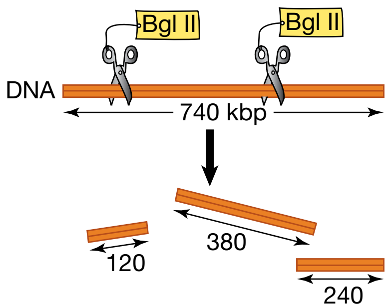



**FIGURE 13.8** Each cutting enzyme snips DNA at a particular site. Would a different cutting enzyme, say Aha I, also be expected to produce three fragments of the same sizes as those produced by Bgl II? (Note that kbp = kilobase pairs.)

## Sorting DNA fragments

OK, so now we have a solution containing DNA that has been cut into a number of predictable fragments using a restriction enzyme. How can we sort out this mixture of DNA fragments? Sorting — no problem: use electrophoresis.

DNA fragments can be sorted using the technique of **electrophoresis**. Electrophoresis sorts DNA fragments according to their lengths, that is, their sizes.

To separate a mixture of dissolved DNA fragments, the mixture is placed in a slot at one end of a slab of jelly-like supporting material, known as a gel, that is immersed in a buffer solution. The gel, which is made of agarose, is then exposed to an electric field, with the positive (+) pole at the far end and the negative (−) pole at the starting end or origin.

Look at figure 13.9, which shows a researcher using a micropipette to load a sample of DNA into one slot at the cathode end of a gel that is covered by a buffer solution. Typically eight to ten different samples of DNA can be loaded onto one gel, with each sample being loaded into a different slot. If you look carefully, you will see blue colouring, which indicates that DNA samples have been loaded into almost all the slots. The colour is added simply to assist the researcher to load the DNA samples into the gel slots.

**FIGURE 13.9** Apparatus for separating DNA fragments by electrophoresis. Notice the red (+) anode at the far end and the black (–) cathode at the origin where a DNA sample is being loaded into the last slot in the gel.

All DNA fragments have a net negative charge, because of the phosphate groups in their sugar–phosphate backbone. As a result, DNA fragments will move towards the positive pole when placed in an electric field. The shortest DNA fragments move most quickly and the longest fragments move most slowly. Fragments of the same size move at the same rate. The end result of electrophoresis is a series of parallel bands of DNA fragments at differing distances down the gel. (see figure 13.10). Each band can contain thousands or millions of DNA molecules of the same size.

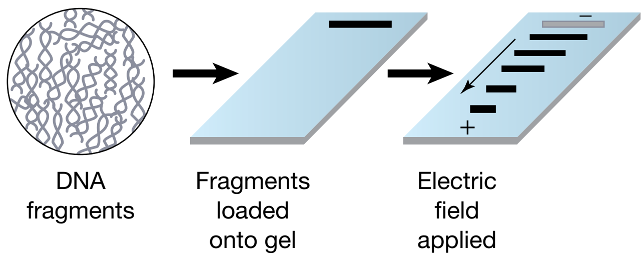



**FIGURE 13.10** Diagram showing the result of electrophoresis of a mixture of DNA fragments loaded into one slot on a gel. The fragments are separated according to their sizes. Which fragments move furthest from the origin — the longest or the shortest?

DNA fragments are invisible so, after the gel run is complete, the separated DNA bands must be made visible either through the use of a dye or a labelled probe. One technique makes use of the dye ethidium bromide (EtBr), which binds to the major groove of DNA molecules. When illuminated by ultraviolet light, the DNA bound to EtBr fluoresces pale pink (see figure 13.11).

A standard ladder of DNA fragments of known sizes is usually run through the gel at the same time as the unknown DNA samples. The sizes of unknown DNA fragments can be approximated by comparing the positions of their bands with those of the known standards. Figure 13.11 shows a ladder of DNA markers of known sizes in lane 1, at the left-hand side of the gel.

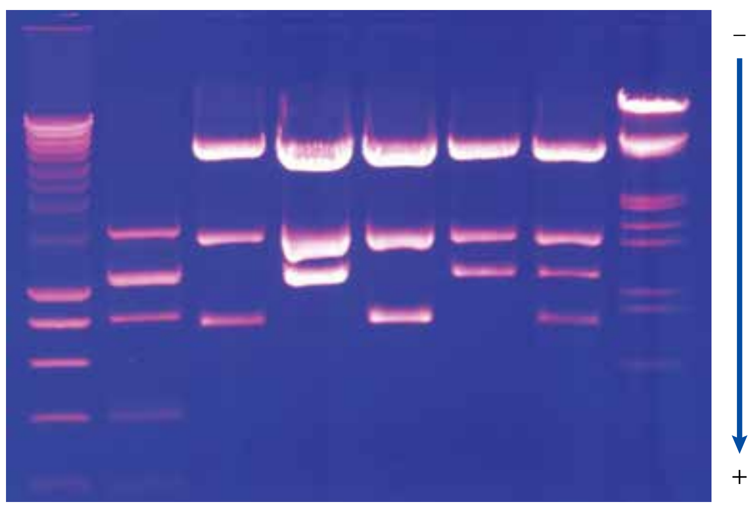



**FIGURE 13.11** Gel stained with ethidium bromide after completion of electrophoresis. This image shows DNA bound to EtBr in bands revealed by pink fluorescence under ultraviolet illumination. Lane 1 (at left) contains DNA markers of known sizes and lanes 2 through 8 contain DNA fragments of unknown sizes.

## Interpreting gel runs

Let's look at examples of electrophoresis in action:

**Example 1:** One of the main reasons for using electrophoresis to separate DNA fragments is to obtain estimates of the size of the DNA fragments in the mixture. Remember that electrophoresis separates DNA fragments according to their sizes. DNA fragments of unknown sizes can be separated by electrophoresis and then their sizes can be approximated by comparing them with DNA molecules of known size separated on the same gel and under the same conditions.

Examine figure 13.12, which shows a gel run of a mixture of DNA fragments of unknown sizes from the RH slot, and samples of DNA markers of known size from the LH slot. Note that the unknown sample of DNA has resolved into four bands.

The ladder of marker DNAs of known sizes shows values ranging from 23 130 base pairs to 564 base pairs. Estimate the size of the DNA fragments that have separated into the four bands. Remember that this is just an approximation.

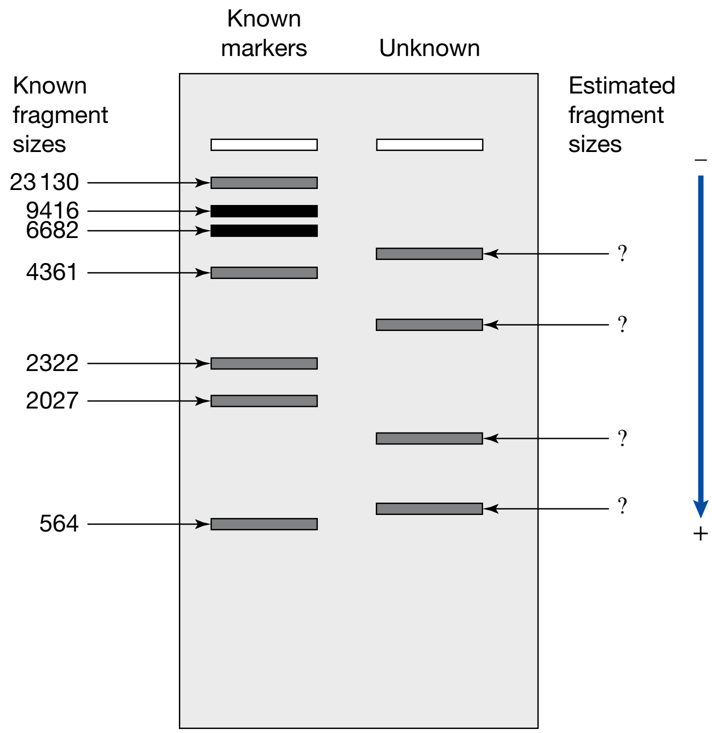



**FIGURE 13.12** Diagram showing the separation of DNA fragments of unknown sizes on a gel (at RH side) alongside a separation of DNA markers of known sizes (at LH side). This technique is used to obtain estimates of the sizes of unknown DNA fragments in a band by comparing their positions with those of bands of DNA markers of known sizes.

**Example 2:** Samples of the DNA of the **F8C** gene, which controls blood clotting, were collected from several different people, two males (J and T) and one female (K).

The F8C gene is located on the X chromosome and has two alleles: the ***C*** allele that determines normal blood clotting, and the ***c*** allele that determines defective blood clotting (haemophilia).

The DNA of each sample was treated with the restriction enzyme **Bcl I**.

The result of Bcl I treatment is shown in figure 13.13a.

The DNA samples were then subjected to electrophoresis and the positions of the bands were identified using a fluorescent probe. The result is shown in figure 13.13b.

Let's consider the results. Remember that the only bands visible are those labelled by the fluorescent probe. Remember that females have two X chromosomes, but males have only a single X chromosome.

You can estimate the sizes of the fragments by comparing their positions with those of the known standards.

How many fragments would result from Bcl I digestion of the ***C*** allele? (Answer: 1)

How many fragments would be produced by Bcl I treatment of the ***c*** allele? (Answer: 2)

What are their expected sizes? (Answer: 879 and 286 bp)

What has happened to the smaller 286 bp fragment? (Answer: The fragment is there, but it is not visible because the probe does not pair with it.)

**study on**

Unit 4
AOS 2
Topic 1
Concept 3

**Gel electrophoresis**
Summary screen and practice questions

**study on**

Unit 4
AOS 2
Topic 1
Concept 3

**See more**
Gel electrophoresis

Note that person K, a female, has both the ***C*** allele and ***c*** allele, as shown by the presence of two fragments.

Would all females be expected to show this pattern? (No. Most females would be expected to be genotype ***CC*** and show just only one band, the 879 band.)

Person J, a male, has the ***C*** allele only. However, person T, a male, shows the presence of the relatively shorter ***c*** allele and so has the X-linked disorder, haemophilia.

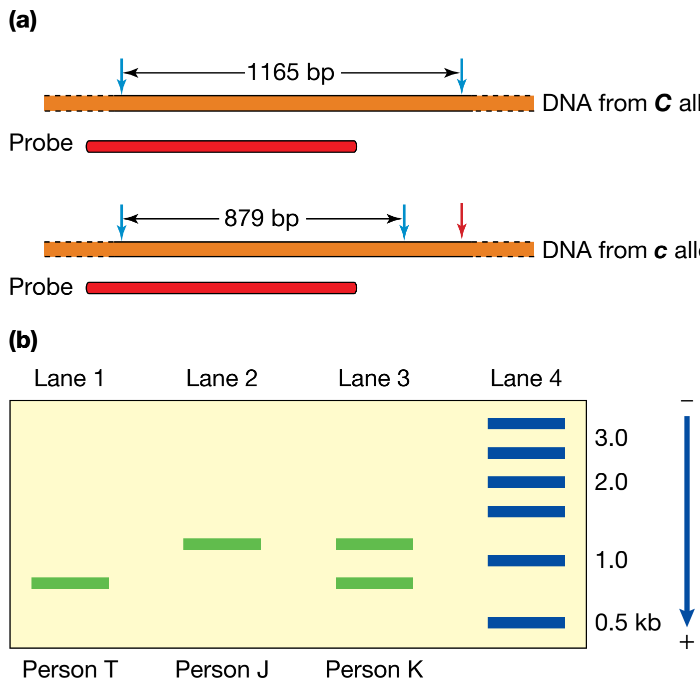



**FIGURE 13.13** **(a)** Diagram showing the DNA of the two alleles of the F8C gene and their cut sites for the restriction enzyme Bcl I. Note that the ***C*** allele has two restriction sites, while the ***c*** allele has three restriction sites. **(b)** Result of electrophoretic run of the three different DNA samples.

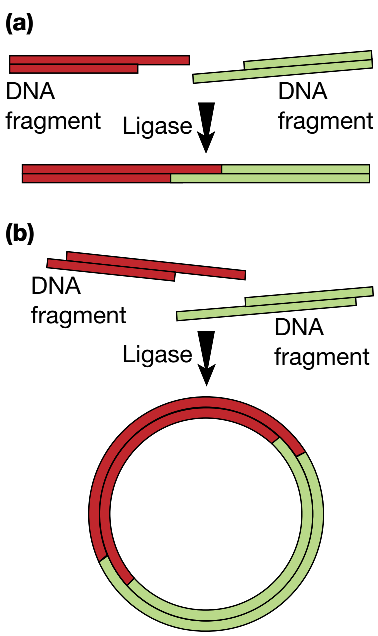



**FIGURE 13.14** In both **(a)** and **(b)**, the DNA fragments have sticky ends created with the same cutting enzyme. Ligase enzyme forms strong covalent bonds so that the fragments are joined permanently. Why is the result a circular DNA molecule in **(b)**?

## Joining DNA fragments

In gene manipulation, the joining of pieces of DNA is sometimes required. An enzyme known as **ligase** catalyses the joining of pieces of double-stranded DNA at their sugar–phosphate backbones. The bonds that form in this case are strong (**covalent**) bonds. The joining can produce one longer piece of DNA (see figure 13.14a) or it can produce a circular molecule of DNA if the fragments have complementary 'sticky ends' at both ends (see figure 13.14b).

### KEY IDEAS

- DNA can be manipulated in various ways, including by being cut into fragments, and DNA fragments can be sorted and can be joined.
- Cutting or restriction enzymes cut DNA into fragments at specific recognition sites.
- Cuts of double-stranded DNA by some restriction enzymes produce 'sticky' ends, and some produce blunt ends.
- Electrophoresis sorts DNA fragments according to their lengths, measured in base pairs.
- DNA fragments can be joined through the action of the enzyme ligase.

## QUICK CHECK

**3** Identify the following statements as true or false:

- **a** Restriction enzymes cut DNA molecules at random sites along their length.
- **b** The restriction enzyme Aha I produces DNA fragments with sticky ends.
- **c** Electrophoresis separates DNA fragments according to their different charges.
- **d** In electrophoresis, DNA fragments travel from slots at the origin of a gel to the cathode.
- **e** A restriction enzyme that can act on human DNA would not be able to act on mouse DNA.
- **f** In a long DNA sequence, it is reasonable to predict that there would be more cut sites for Hae III than for Hin dIII.
- **g** In electrophoresis, shorter fragments of DNA travel farther than longer fragments.

# More tools for manipulating DNA

Let's now identify other tools and procedures that are used in DNA manipulation.

## Finding DNA fragments of interest

Electrophoresis can separate a mixture of DNA fragments according to differences in their sizes, but how can one particular set of DNA fragments be picked out from a large number of DNA fragments? This is a bit like finding a needle in a haystack. However, this can be done using a **probe**.

You might use a person's dog to locate that person in a crowd. In this case, the person is the target object and the dog is the 'probe' that locates the target. You could use a Geiger counter to find the source of a radioactive emission. In this case the probe is the Geiger counter and the target that it will locate is the source of a radioactive emission. In both cases, the probe will home in on the target. Follow the probe and you have found your target! Likewise, as people pass through security checks at airports, both their bodies and their luggage are 'probed' by scanners that detect targets, including prohibited metal objects, such as knives or guns. In this case, the 'probe' does not physically attach to the target but identifies it on a screen display (see figure 13.15).

**FIGURE 13.15** The luggage scanner is like a probe that can home in on particular prohibited targets without making physical contact with the targets.

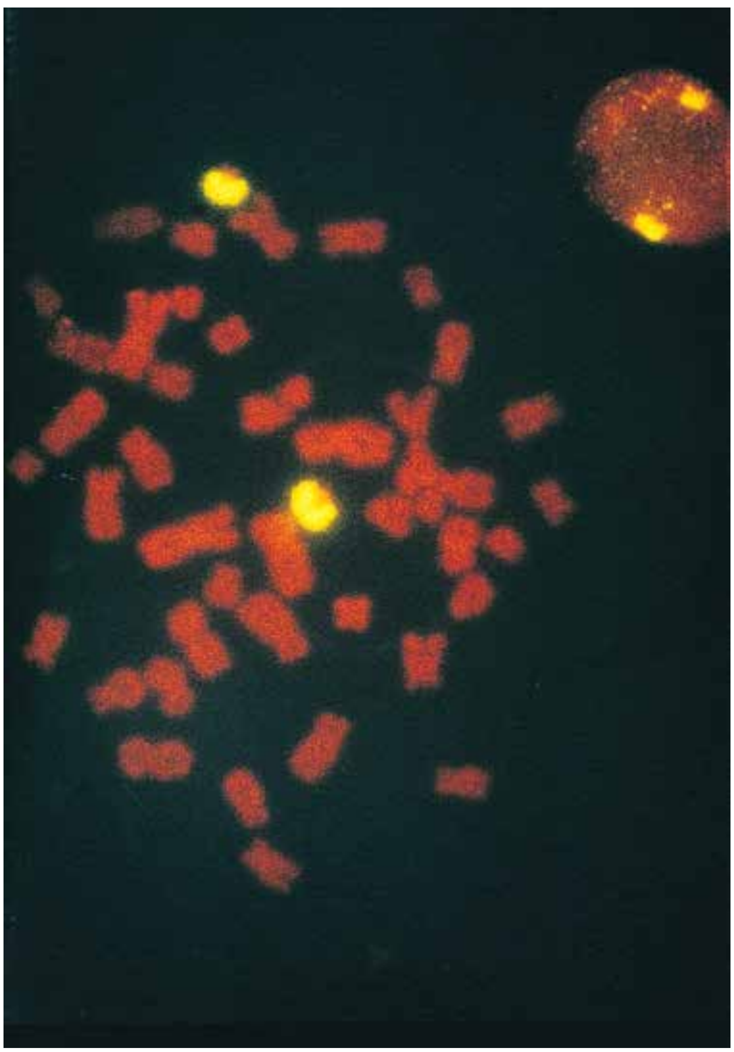
Likewise, a molecular probe can locate particular target DNA fragments from among a mixture of many DNA fragments. This kind of probe is commonly a piece of single-stranded nucleic acid, either DNA or RNA, with a base sequence that is complementary to part of the base sequence of one of the strands of the target DNA. The key to the use of probes is their ability to hybridise with the complementary sequences on their single-stranded target DNA.

To locate the DNA fragments of interest, the target DNA must first be denatured to separate its two strands, otherwise the probe cannot home in and pair with it. Figure 13.16 shows a probe and part of the target it can locate. Because the probe is complementary to part of the target, it is able to pair with it.


**FIGURE 13.16** The base sequence of the probe is complementary to part of its target. A probe carries a radioactive or fluorescent label, represented by the flag, so that it (and its target) can be located.

**ODD FACT**

When *in situ* hybridisation is carried out using a probe with a fluorescent label, the technique is referred to as **FISH** (fluorescence *in situ* hybridisation).



A probe must be labelled in some way so that it can be easily located. This label is typically a fluorescent marker or it may be a radioactive marker. The position of the probe and the target that it has latched onto are signalled by the fluorescent label on the probe.

Probes can be used to locate specific DNA segments, either within a cell or outside cells, such as on a gel.

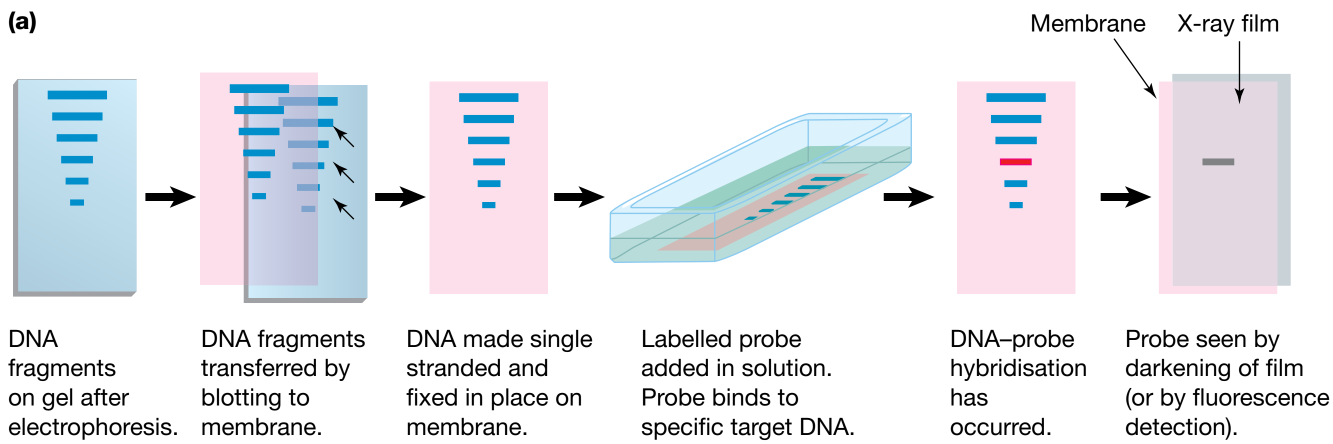
Figure 13.17 shows the chromosomes of a human cell that have been exposed to a probe that binds specifically to a unique sequence of DNA in part of the human Y chromosome. This probe has a label that fluoresces bright yellow when exposed to ultraviolet light. Notice the two bright yellow spots that indicate that the probe has hybridised at two sites, signalling the presence of two Y chromosomes in the cell. This process is known as ***in situ* hybridisation**. The term *in situ* means 'in place' and identifies the fact that the DNA is in its normal place in the chromosomes. In fact, these cells came from a male who has 47 chromosomes, including one X and two Y chromosomes, in each of his somatic cells.

**FIGURE 13.17** A probe that fluoresces bright yellow has been mixed with chromosomes. The probe is specific for the Y chromosome. At how many sites has the Y-specific probe become bound? Notice that the second cell (upper right) is at the interphase stage of the cell cycle. (Image courtesy of Dr Arabella Smith.)

Probes can also be used to locate a particular DNA fragment from a mixture of fragments that has been sorted by electrophoresis. Figure 13.18a outlines the procedure called 'Southern blotting' in which the DNA fragments are transferred from the gel to a nylon membrane.

The steps in this procedure are as follows:

- The gel is treated with a denaturing agent to make the DNA single stranded.
- The gel is transferred to a salt solution, and a nylon membrane is placed over the gel.
- Layers of absorbent paper, weighted on top, are placed over the membrane, and the salt solution is drawn through the gel, carrying the DNA that is absorbed by the nylon membrane (see figure 13.18b for details). This procedure typically lasts for about 18 hours and, by the end of this time, all the DNA from the gel has been transferred to the nylon membrane.
- The membrane is then soaked in a solution containing a labelled probe that consists of a single-stranded nucleic acid probe that binds specifically to the intended DNA target.
- The probe, and hence the DNA of interest, is located, either by use of X-ray film in the case of a radioactively labelled probe, or by fluorescence in the case of a probe with a fluorescent label.



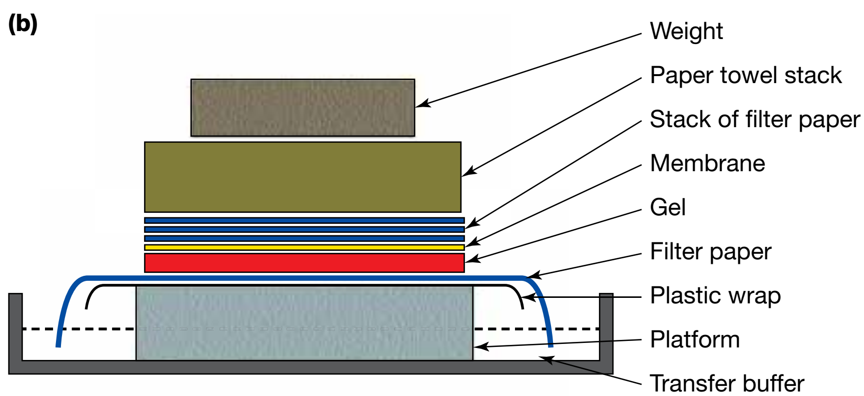




**FIGURE 13.18** **(a)** Steps in the use of a probe to locate a particular DNA target after electrophoresis. **(b)** Details of arrangement for the transfer of denatured (single-stranded) DNA from the gel to a nylon or nitrocellulose membrane.

## Making DNA fragments

It is easy to isolate the total DNA from human somatic cells. This, however, consists of the DNA of two sets of 23 chromosomes, each set containing about 3000 million base pairs encoding about 21 000 genes, as well as very long sequences of noncoding DNA. Some scientists may want short lengths of DNA to act as probes. Other scientists might want copies of just one specific

**eBook*plus***

**eLesson**
Somatic cell nuclear transfer
**eles-2465**

DNA base sequence as part of their research. How can particular segments of DNA be obtained for these various purposes?

Two methods for making specific DNA fragments are outlined below:

1. *Synthesise DNA from nucleotide building blocks*. This method makes use of an instrument called a DNA synthesiser (see figure 13.19 for one example). To use this method, the base sequence of the required DNA must be known. Instruments called DNA synthesisers can join nucleotide sub-units in a predefined order to produce DNA segments with lengths greater than 100 bases. The chemical synthesis of DNA does not require a template strand nor does it require the enzyme DNA polymerase. The products can be used as primers or as probes.

**FIGURE 13.19** A scientist working in a DNA laboratory. DNA synthesisers can generate primers and probes as well as lengths of DNA up to about 100 bps. DNA synthesisers work in conjunction with a computer and DNA synthesis software.

**ODD FACT**

In 1972, Khorana and his team reported the first gene to be artificially synthesised; this was the gene that encodes the information for a transfer RNA (tRNA) in yeast. In 1977, the first protein-coding gene was artificially synthesised.

2. *Make a copy of DNA using an mRNA template*. To do this, mRNA is isolated from the specific cells in which the gene concerned is active. The enzyme **reverse transcriptase** uses the mRNA as a template to build a single-stranded DNA with a complementary base sequence. This DNA is known as **complementary DNA (cDNA)**. The reverse transcriptase procedure has been used successfully to make copies of the human growth hormone gene and also the gene for tissue plasminogen activator (t-PA), a blood clot-dissolving enzyme used in the treatment of some forms of heart attack.

Figure 13.20a outlines a simplified version of the process to produce single-stranded cDNA using reverse transcriptase enzyme and an mRNA template. Commercial kits are available for the use of reverse transcriptase to produce cDNA (see figure 13.20b).

To make this DNA double stranded, the enzyme DNA polymerase is then used. Note that this method will work only for coding DNA: that is, genes that produce mRNA. This method cannot be used for the bulk of the genome that consists of noncoding DNA.

The name of the enzyme reverse transcriptase describes its action. 'Reverse' and 'transcriptase' refer to the fact that the process is the exact opposite of transcription — normally in cells, DNA directs the manufacture of RNA.

**(a)**

mRNA isolated from cystol of specific cells.

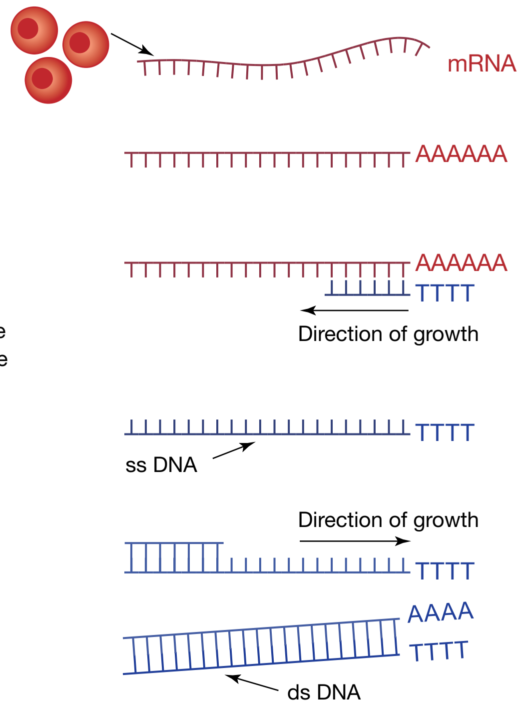
Poly-A tail added to this mRNA. (This segment provides an anchor to which a '**primer**' can attach.)

Oligo-dT primer added and binds to poly-A mRNA tail. Reverse transcriptase enzyme added. DNA lengthens by addition of nucleotides. The order is controlled by the sequence in the mRNA.

When DNA chain is complete, mRNA is removed by alkali treatment.

Polymerase enzyme added. This catalyses the building of complementary DNA strand.

Final double-stranded DNA product (cDNA).




**(b)**


**FIGURE 13.20** **(a)** Simplified version of the steps involved in making complementary DNA (cDNA) using reverse transcriptase enzyme and an mRNA template (ss = single stranded). Note that the end product is single-stranded DNA. How might it be made double stranded? **(b)** One of the commercial kits that provides everything needed to reverse transcribe RNA to single-stranded cDNA. (Image (b) courtesy of Thermo Fisher Scientific.)

## Amplifying DNA fragments

In forensic investigations, critical evidence can come from trace amounts of DNA obtained from a single root hair, a licked stamp, a dried blood stain, a discarded cigarette butt, a piece of chewing gum, drinking glasses, a soiled tissue or a trace of semen. For medical purposes, a few fetal cells obtained from amniotic fluid may be the source of DNA that can reveal if a fetus has an inherited disease. These DNA traces become useful when they are amplified to millions of copies using the polymerase chain reaction (PCR).

Figure 13.21 shows an instrument that automates the PCR procedure. Minute quantities of DNA provide the starting point for PCR. From a few starting copies of DNA, millions of copies can be quickly produced.


**FIGURE 13.21** A Veriti® 96-Well Fast Thermal Cycler, which can automatically take up to 96 samples through a programmed polymerase chain reaction. (Image courtesy of Thermo Fisher Scientific.)

**ODD FACT**

Tiny bone and tissue fragments on ancient stone tools are a source of DNA. This DNA has been amplified using PCR to identify the species from which it came.

**ODD FACT**

Kerry Mullis shared the 1993 Nobel Prize in Chemistry for his development of the use of the heat-tolerant enzyme *Taq* polymerase to make multiple copies of genes via PCR.

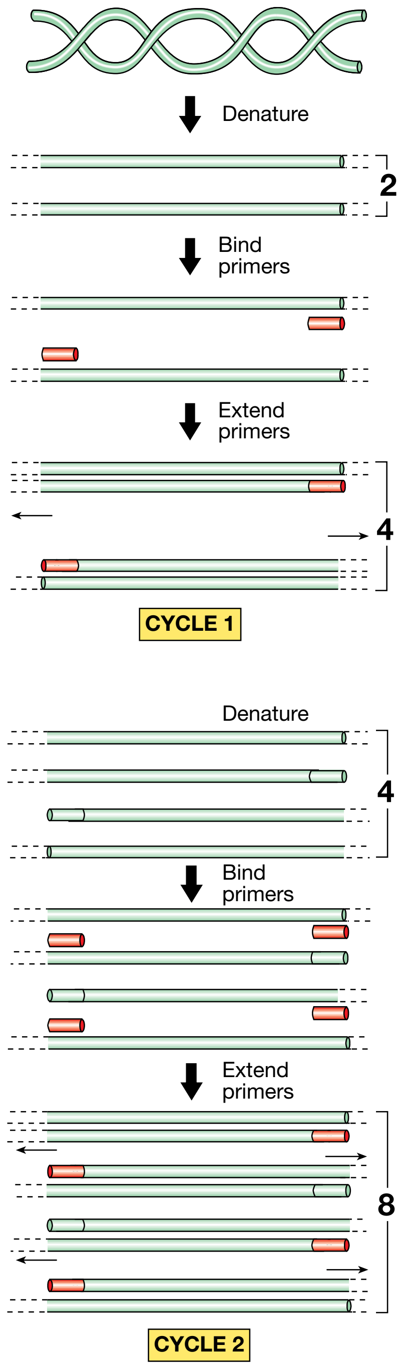



**FIGURE 13.22** Polymerase chain reaction (PCR) was developed in 1984. Primers are short single-stranded sequences of DNA (20–25 bp) complementary to regions at the boundaries of the DNA to be amplified. Notice that four DNA strands are present after one cycle, and eight strands after two cycles. How many copies would be present after eight cycles?

The polymerase chain reaction (see figure 13.22) depends on the enzyme DNA polymerase to amplify or make multiple copies of a sample of DNA. The polymerase enzyme comes from a bacterial species (*Thermus aquaticus*) that lives in hot springs. It is known as the *Taq* polymerase.

The procedure is carried out in a test tube (*in vitro*) as follows:

1. *Denature:* The DNA sample is denatured by heating so that it dissociates to single strands (94 °C for two minutes).
2. *Bind primers:* Short segments of single-stranded DNA, known as primers, are added; these primers pair with regions at either end of the DNA region of interest (55 °C for two minutes).
3. *Extend primers:* The polymerase enzyme uses the primers as a starting point and extends them so that two complete double strands are formed. For this step, a supply of nucleotides must be available (72 °C for one minute).

About five minutes is required for each cycle of 'denature — bind primers — extend primers'. Over time, an exponential growth of the DNA from the two original strands occurs: 4, 8, 16, 32 . . . , so that after 20 cycles about one million copies of the DNA exist.

The amplified DNA produced in a few hours by PCR is enough to provide material for a DNA profile in a forensic investigation, or to identify whether an embryo or fetus has inherited a genotype that causes a serious disorder, such as sickle-cell anaemia. In contrast, other **prenatal** procedures require several days or longer to obtain enough DNA by cell reproduction.

## KEY IDEAS

- Probes typically consist of short single-stranded nucleic acids, either DNA or RNA, that can locate target DNA through complementary base pairing.
- Probes carry a label that allows them (and their targets) to be located either inside or outside cells.
- DNA fragments can be artificially made using a DNA synthesiser that requires a known DNA base sequence.
- Reverse transcriptase enzyme can use an mRNA template to build a single strand of complementary DNA (cDNA).
- The polymerase chain reaction (PCR) can amplify minute amounts of DNA.

study on

Unit 4
AOS 2
Topic 1
Concept 2

**DNA amplification**
Summary screen and practice questions

## QUICK CHECK

4 Identify the following statements as true or false:
   a Target DNA must be denatured before it can be located with a probe.
   b The base sequence of a probe is complementary to part of the base sequence of its DNA target.
   c The reverse transcriptase process can be used to produce cDNA copies of noncoding DNA.
   d A DNA synthesiser does not require enzymes or a template strand.
5 What is meant by the phrase 'denature — bind primers — extend primers'?

study on

Unit 4
AOS 2
Topic 1
Concept 2

**See more**
DNA amplification

# Transforming bacteria

In this section we will examine how DNA can be transported into bacterial cells resulting in their transformation, and see how recombinant plasmids are involved in this story.

## Transporting DNA into cells

Suppose that a team of genetic engineers wants to insert copies of a DNA fragment isolated from plant cells into bacterial cells, or they wish to insert copies of a human gene into bacterial cells. Special carriers or vectors can be used to transport DNA from the external source into the bacterial cells. We can give the label of 'foreign DNA' to any DNA that is transported from the external environment into cells using a vector. Very often, in the laboratory setting, **plasmids** are the vectors used to transport foreign DNA into bacterial cells.

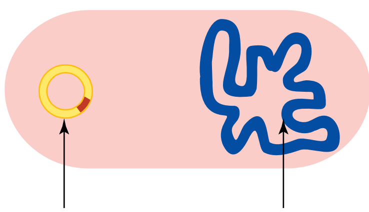



**FIGURE 13.23** Stylised diagram of one bacterial cell, showing its main chromosome and one plasmid. Plasmids are small circular DNA molecules, additional to and separate from the main genetic information held in the main bacterial chromosome. Multiple copies of plasmids may be present in one bacterial cell.

### Plasmids: vectors of DNA

What are plasmids? In addition to their main chromosome, most bacterial cells contain one or more plasmids. Plasmids are small circular pieces of double-stranded DNA ranging in length from about 2000 to 10 000 base pairs (see figure 13.23) found mainly in bacterial cells. All plasmids are self-replicating because they contain an **origin of replication (ORI)** that is a specific DNA base sequence where DNA **replication** of a plasmid begins. As a result, plasmids can replicate independently of the main bacterial chromosome.

In addition to naturally occurring plasmids, plasmids can also be constructed so that they include certain features in addition to an origin of replication. For example, plasmids that are used to transfer DNA include several features, including:

1. genes for one or more traits, including a **selectable marker gene**
   The marker gene is often a gene governing antibiotic resistance, with the particular form of the gene being the allele for resistance to an antibiotic, such as ampicillin ($Amp^R$). Bacteria containing the plasmid with the marker can be selected from among other bacteria by exposure to the antibiotic. Bacteria that have taken up the plasmid with the marker are unaffected by the antibiotic, but bacteria lacking that particular plasmid will be killed by the antibiotic.
2. a **multiple cloning site (MCS)**, also known as a **polylinker**
   This is a region of the plasmid DNA that contains the recognition sites for several restriction enzymes that are ones producing 'sticky' ends. These are sites where it is possible to insert foreign DNA. Treatment of plasmids with the specific restriction enzyme cuts the plasmid DNA open, allowing the incorporation of foreign DNA with complementary sticky ends.

Figure 13.24 shows two of many plasmids that can be used as vectors to transport foreign DNA into bacterial cells to transform them. Both plasmids have the required origin of replication (ORI) and also have a marker ($Amp^R$) for the antibiotic ampicillin. In addition, the pUC19 plasmid has a second marker ($Tet^R$) for a second antibiotic, tetracycline.

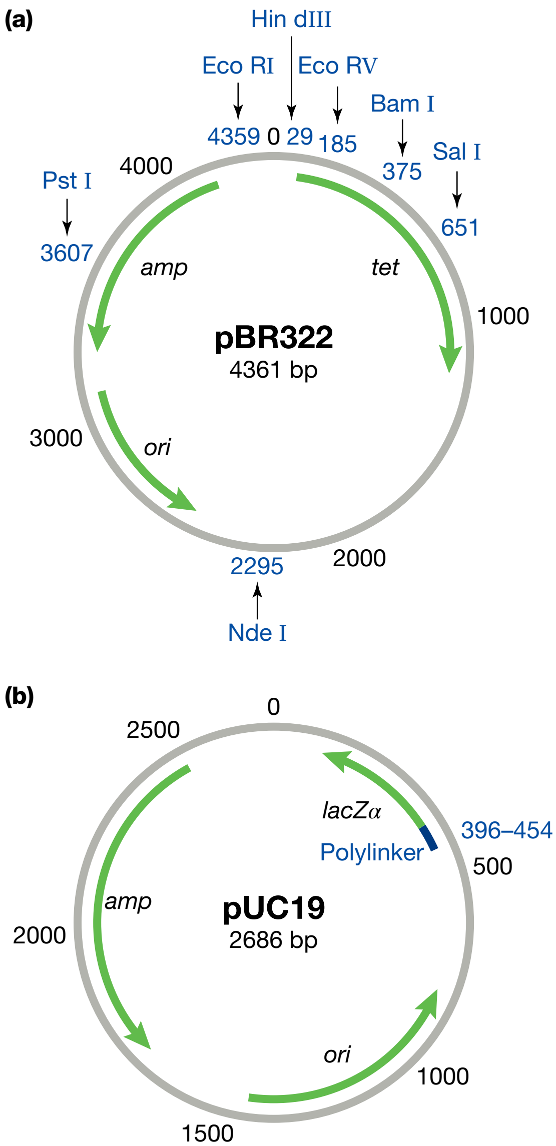



**FIGURE 13.24** Diagrams of plasmids. **(a)** Plasmid pBR322 with many cloning sites present throughout the plasmid. **(b)** Plasmid pUC19 has a cluster of many restriction enzyme recognition sites in the blue sector of the plasmid labelled as polylinker (or multiple cloning site). Note the presence in both plasmids of an origin of replication (*ori*) and a marker gene for ampicillin resistance (*Amp*$^R$).

**ODD FACT**

The allele that encodes antibiotic resistance (*Amp*$^R$) produces the protein beta-lactamase, which inactivates ampicillin.

### Bacterial transformation

In the laboratory, scientists commonly use a plasmid as the carrier or **vector** to introduce new genes into bacterial cells. When foreign DNA is directly taken up from an external source by bacterial cells, those bacteria are said to be transformed. Evidence that **bacterial transformation** has occurred is seen in the expression of the foreign marker gene in the phenotype of the bacteria. For example, some bacteria are susceptible to a particular antibiotic, such as ampicillin. If these bacteria acquire a plasmid carrying a gene for ampicillin resistance, these bacteria are transformed and become ampicillin resistant.

Bacterial transformation may occur naturally, as when bacteria swap plasmids among themselves. Naturally occurring transformation of bacterial species by plasmid transfer has produced increasing numbers of strains of bacteria that are resistant to multiple antibiotics. Bacterial transformation may also occur artificially, as a deliberate process carried out by microbiologists in laboratories.

Bacterial transformation can involve the direct uptake by bacteria of naked fragments of DNA from the environment, such as DNA fragments from dead bacteria. Very commonly, however, transformation is achieved when vectors, such as plasmids, introduce foreign DNA into bacterial cells.

### First knowledge of bacterial transformation

Transformation of bacteria by natural processes has probably been taking place for at least a thousand million years or so. However, it was not until 1928 that it was recognised — this by an English scientist, Frederick Griffith (1877–1941). Griffith experimented with pneumococcus bacteria; the smooth variety of these bacteria causes pneumonia, but the rough variety is harmless. Griffith showed that when the contents of dead smooth bacteria cells were added to a culture of living rough bacteria cells, some of the rough cells were transformed into pneumonia-causing smooth cells that, when injected into mice, killed them. At that time, Griffith called the agent responsible for this change the 'transforming factor' but its identity was not known. It was not until 1944 that Oswald Avery (1877–1955) demonstrated that the transforming factor was DNA.

## Making recombinant plasmids

Recombinant plasmids are plasmids that carry foreign DNA. Making recombinant plasmids involves the following steps:

1. The DNA of the plasmid is cut at one point using a specific cutting enzyme that creates sticky ends. This changes the form of the plasmid from circular to linear.
2. The foreign DNA fragments are prepared using the same cutting enzyme so that the foreign DNA has sticky ends that match those of the cut plasmid.

   The foreign DNA fragments and the plasmids are mixed, and, in some cases, their 'sticky ends' pair by using weak hydrogen bonds. These are **recombinant plasmids**.

   Other pairings will also occur, such as cut plasmids resealing themselves so that they are not recombinant plasmids.
3. The joining enzyme, ligase, is added and this makes the joins permanent through covalent bonding.

Figure 13.25 shows the process by which recombinant plasmids are formed.

**ODD FACT**

Smooth pneumococcus bacteria have an external capsule that prevents their destruction by phagocytic cells of the immune system. Infection by a single smooth cell can lead to the death of a mouse by pneumonia, while an infection by hundreds of rough cells that lack capsules is harmless.

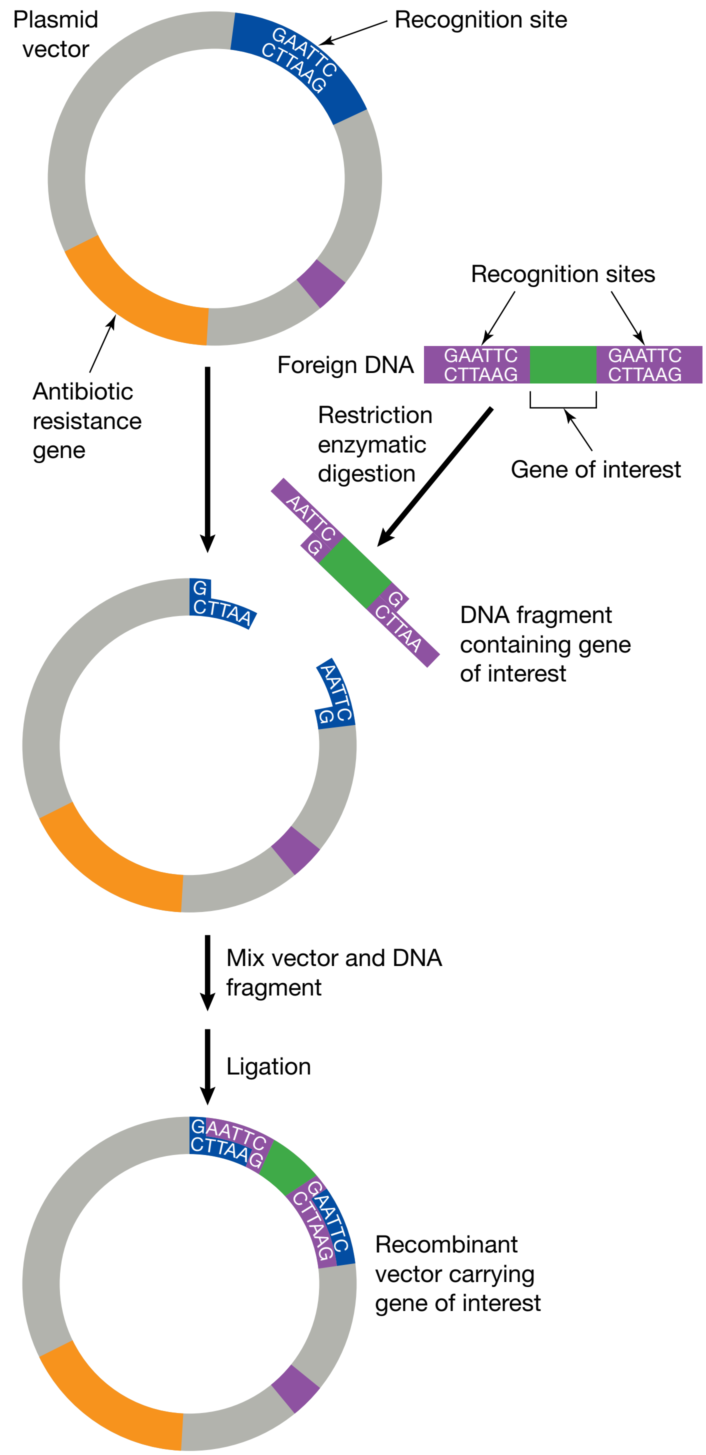



**FIGURE 13.25** Diagram showing the steps in the formation of a recombinant plasmid. The formation of complementary sticky ends in both the foreign DNA fragment and the cut plasmid means that pairing between the plasmid and the foreign DNA can occur.

**FIGURE 13.26** The PlasmidFactory company logo. (Image courtesy of PlasmidFactory.)

In reality, commercial companies exist that will produce DNA plasmids at the highest quality to meet the design requirements of research groups, as, for example, PlasmidFactory®, a German company founded in 2000 (see figure 13.26).

## Getting plasmids into bacterial cells

Once recombinant plasmids have been created, the challenge is to transfer them into bacterial cells. In general, the success rate for the transfer of recombinant plasmids into bacterial cells is low, but the success rate can be increased through various techniques:

- One technique is termed **electroporation**. In this method, cells are briefly placed in an electric field that shocks them and appears to create holes in their plasma membranes so that plasmid entry is facilitated.
- Another method is to heat shock the bacterial cells by suspending them in an ice-cold salt solution, and then transferring them to 42 °C for less than one minute. This treatment appears to increase the fluidity of the plasma membranes of the bacterial cells and increases the chance of uptake of plasmids.

An example of the heat-shock treatment is shown in figure 13.27. The starting culture of bacterial cells is sensitive to the antibiotic tetracycline, and is denoted $Tet^S$.

1. The bacterial culture is placed in an ice bath and chilled.
2. Recombinant plasmids with the tetracycline resistance allele $Tet^R$ are added to the bacterial culture and chilled.
3. The bacteria and plasmid mix is placed in hot water at 42 °C for 50 seconds, producing a heat shock. This is the stage when the plasma membranes of the bacterial cells are altered, increasing the chance of uptake of plasmids by the cells.
4. The mix is returned to an ice bath for two minutes.
5. The bacteria are plated on an agar plate containing the antibiotic tetracycline and incubated at 38 °C overnight. Bacteria that have not taken up the plasmids are killed by the tetracycline. Bacterial cells that have taken up the plasmids will be selected; these bacteria have been transformed and will replicate.

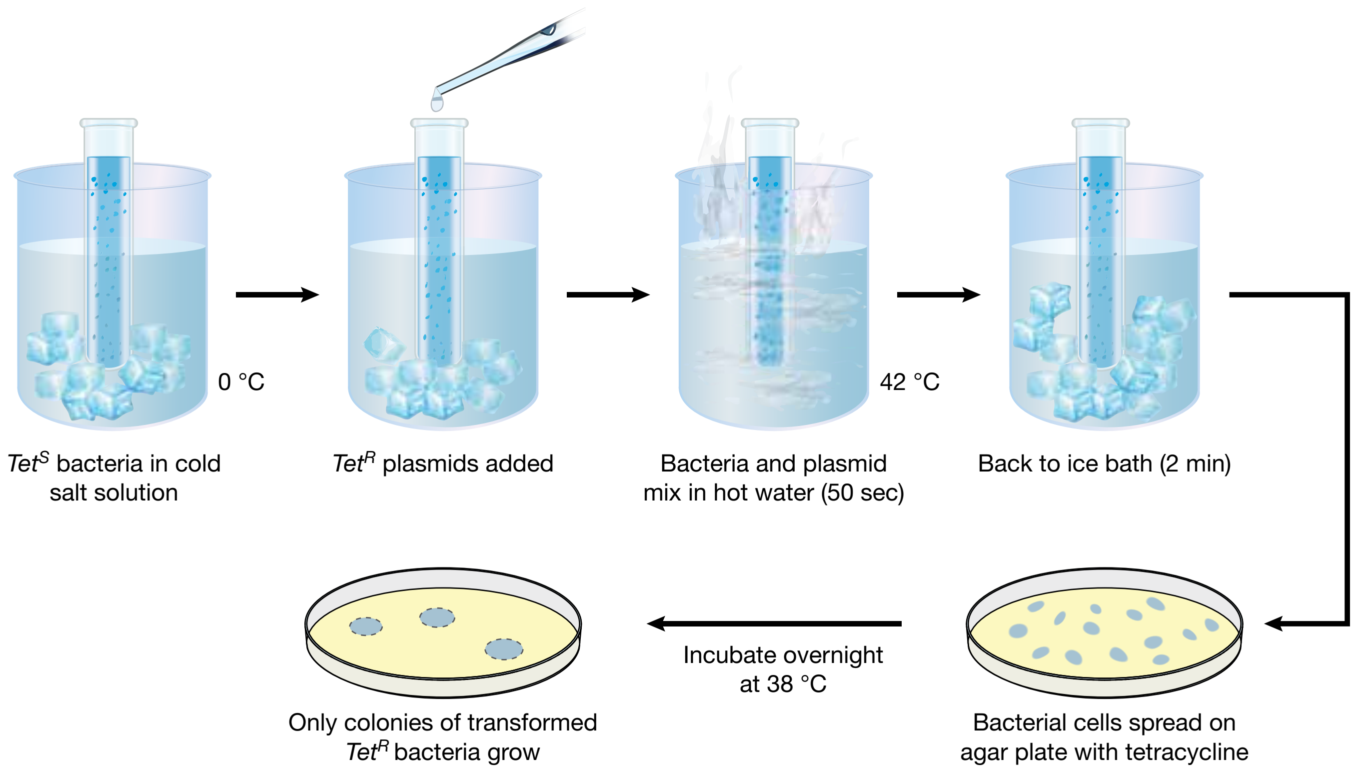



**FIGURE 13.27** Stages in the uptake of plasmids by bacterial calls, resulting in their transformation.

studyon

Unit 4 | AOS 2 | Topic 1 | Concept 4

**Recombinant DNA**
Summary screen and practice questions

The importance of recombinant plasmids is that the foreign DNA they incorporate can come from any source: it may be a human gene, a plant gene, a jellyfish gene, a yeast gene, and so on. The importance of transformed bacteria is that they express the foreign gene in their phenotype. We will see in chapter 14 how these transformed bacteria can be factories to produce the protein products of these foreign genes, such as human insulin.

Now that you have met plasmids, let's look at how they can be used to make multiple copies of DNA.

## Making multiple copies of DNA

It is possible to obtain multiple copies of DNA using the polymerase chain reaction (PCR) as described earlier in the chapter. However, the PCR technique is used to amplify DNA only when very tiny amounts of DNA are available, such as in forensic investigations.

In contrast, if DNA is available in relatively large amounts, recombinant plasmids can be used to make multiple copies of large fragments of DNA or even entire genes. This technique is called **gene cloning** and we will explore this technique in detail in chapter 14 (see page 652). Note that gene cloning using recombinant plasmids uses the fact that plasmids are self-replicating and can multiply independently of the main bacterial chromosome.

A simplified diagram showing how plasmids can make multiple copies of foreign DNA that they carry is shown in figure 13.28. This multiplication is achieved through (i) replication of the plasmids and (ii) binary fission (division) of the bacterial cells in which the plasmids are located.

In summary, the DNA to be copied (cloned) is joined to a bacterial plasmid and the plasmid is transferred into a bacterial cell. Once inside the bacterial cell, the plasmid and its foreign DNA multiply to produce up to 20 copies. Within about 20 minutes, the bacterial cell itself divides by binary fission to produce two daughter cells. Within these cells, the plasmids again multiply and then the bacterial cells divide again. Over a period of several hours, millions of bacterial cells and even larger numbers of copies of the foreign DNA will be produced.

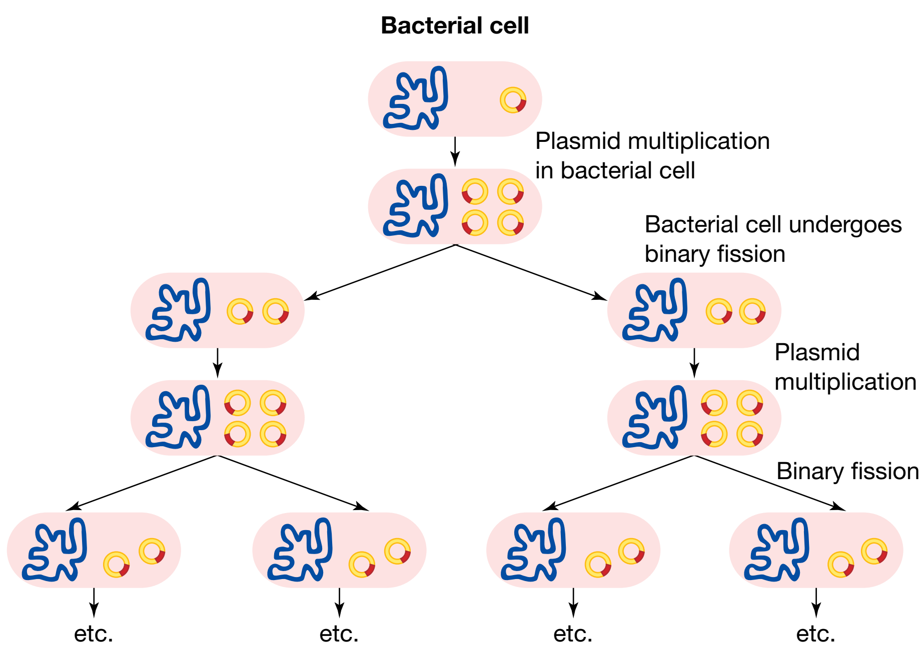



**FIGURE 13.28** Gene cloning involves the formation of many copies of a particular fragment of DNA. In this case, the red region of the recombinant plasmid is the foreign DNA fragment. Multiplication of this foreign DNA involves both replication of the plasmid's DNA within cells and binary fission of the bacterial cells. In this diagram, after one binary division of the bacterial cell and plasmid replication, how many copies of the original plasmid are present?

**BIOLOGIST AT WORK**

### Steven Mileto — PhD student and microbiologist

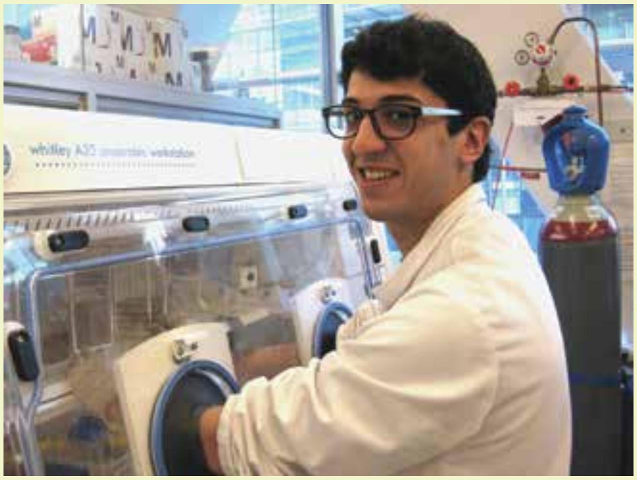


**FIGURE 13.29** Steven Mileto is a microbiologist and a PhD student at Monash University. (Image courtesy of Steven Mileto.)

'When I tell someone without a science background that I am a microbiologist they often ask, 'How do you know what you're working on? I thought you always needed a microscope to see and work on microorganisms'.

It's always fascinated me that something so small, something that as a single organism cannot been seen by the naked eye, can be so beneficial to our existence, yet can also result in so much harm to people, animals and plants. Microorganisms aid in our breakdown of nutrients, assist in the production of renewable energy sources, life-saving antibiotics, chemicals, compounds, and enzymes and many food products that we adore. But we often focus on the negative aspects of microorganisms, particularly their ability to cause life-threatening and debilitating disease, their ability to develop resistance to antibiotics and the many widespread plagues and epidemics that arise due to pathogenic (disease-causing) microorganisms.

As a microbiology PhD student, my thesis and lab research focuses on one particular pathogenic bacterium, *Clostridium difficile*. *Clostridium difficile* is the most common cause of antibiotic-associated hospital-acquired diarrhoea and can result in a spectrum of disease ranging from mild diarrhoea to severe and life-threatening diarrhoea, which may lead to more serious and fatal complications. Disease is mediated by two large toxins produced by *C. difficile*, which lead to cell death and inflammation. My research specifically focuses on how *C. difficile* interacts with the host during infection. From my research we hope to better understand how the host responds to infection with *C. difficile* at the immunological level, as well as understanding other systemic disease complications that may arise during severe infection. To do so, we use models of infection to mimic human disease and analyse the pathologies, and immunological and host responses, induced by infection with *C. difficile*. We then compare this to infection with genetically engineered *C. difficile*, which lacks the ability to produce the disease-inducing toxins, in the hope of identifying key differences in host responses between both groups. Medical research is, however, not a one-man job and relies on our team of students and staff working and collaborating together to achieve our research outcomes. By working together on various aspects of *C. difficile* disease, we aim to pinpoint key pathways involved in pathology and disease progression during *C. difficile* infection, and hope to identify more suitable treatment options and markers for disease severity. Hopefully this will translate to better disease outcome and fewer fatalities from *C. difficile* infection.'

**KEY IDEAS**

- Plasmids are small circular fragments of double-stranded DNA occurring naturally in bacteria.
- Plasmids are vectors that can carry foreign DNA from any source.
- Plasmids can be constructed to carry selectable marker genes, such as antibiotic resistance.
- Transformation of bacteria occurs when these cells take up DNA from an external source and express it.
- Recombinant plasmids can be used to make multiple copies of foreign DNA fragments or even genes.

## QUICK CHECK

6 Identify the following statements as true or false:
  - **a** Plasmids are the main chromosome of a bacterial cell.
  - **b** Plasmids are self-replicating entities.
  - **c** When bacterial cells and plasmids are mixed, all bacterial cells will be transformed by taking up the plasmids.
  - **d** Bacteria that are ampicillin sensitive will be killed by exposure to this antibiotic.
  - **e** If ampicillin-sensitive bacteria take up plasmids that contain *Amp*$^R$, these bacteria will be transformed and will become ampicillin resistant.
  - **f** The foreign DNA in a recombinant plasmid could come from mice or men.

7 Identify two procedures that can increase the chance of uptake of plasmids by bacteria.

# BIOCHALLENGE

The plasmid known as pAMP contains 4539 base pairs of DNA. These include an origin of replication (ORI), the ***Amp**$^R$* allele of the gene conferring ampicillin resistance.

In addition, this plasmid includes a single occurrence of each of the following sequences:

| | | |
|---|---|---|
| **GGATCC** | **and** | **AAGCTT** |
| **CCTAGG** | | **TTCGAA** |

The pAMP plasmid was treated with a mixture of the restriction enzymes Bam HI and Hin dIII.

The result of this treatment was two fragments as follows:

- a fragment of 784 base pairs
- a fragment of 3755 base pairs that included the ORI and the ***Amp**$^R$* gene sequence.

1. Use the above information to draw a rough sketch of the pAMP plasmid and show the positions of the two recognition sites.
2. What kinds of ends are produced by treatment of these plasmids:
   a. with Bam HI
   b. with Hin dIII?
3. What is the function of the ORI?

   Various samples were subjected to electrophoresis in an agarose gel (at 1%) as follows:

   Lane 1: A set of DNA markers of known sizes, 5000 bp, 4000 bp, 3000 bp, 2000 bp, 1000 bp, and 500 bp

   Lane 2: pAMP plasmids after treatment with Bam HI and Hin dIII

   Lane 3: pAMP plasmids after treatment with Hin dIII only

   Lane 4: pAMP plasmids after treatment with Bam HI only
4. Your task is to draw a gel with four lanes showing the relative positions of the bands after the electrophoresis is completed.

   Indicate the origin of your gel that marks the position where the samples are loaded.

   Number each lane.

   At the side of your gel, include an arrow with a positive (+) and a negative (–) sign to show the direction of migration of the various DNA fragments.

   First, draw the bands for the sample of known markers in lane 1 and show their sizes at the LH side of the gel.

   Then draw the bands for the other samples and show their sizes at the RH side of the gel.
5. Fragments of foreign DNA with a size of 500 base pairs have ends resulting from Bam HI treatment. Copies of this foreign DNA were mixed with pAMP plasmids that had also been treated with Bam HI only. Some recombinant plasmids were formed.
   a. Make a rough sketch of the recombinant pAMP plasmid.
   b. Only some, but not all, plasmids became recombinant. Suggest a possible reason for this.
   c. The enzyme ligase is involved in the formation of recombinant plasmids. Identify its role.
6. Recombinant plasmids were mixed with recipient bacterial cells that were ampicillin sensitive (***Amp**$^S$*).

   Some bacterial cells took up recombinant plasmids, but others failed to do so.
   a. Which bacteria were transformed?
   b. What evidence would indicate that these bacteria were transformed?
   c. Briefly explain how you could separate the transformed bacteria from those that were not transformed.

# Chapter review

studyon
Unit 4
AOS 2
Topic 1
DNA manipulation
Practice questions

## Key words

bacterial transformation
Bcl I
Cas9 nuclease
complementary DNA (cDNA)
covalent
CRISPR-Cas9
DNA editing
electrophoresis
electroporation
F8C
gene cloning
gene editing
genome editing
guide RNA
*in situ* hybridisation
ligase
multiple cloning site (MCS)
origin of replication (ORI)
plasmids
polylinker
prenatal
probe
recombinant plasmids
replication
restriction enzymes
reverse transcriptase
selectable marker gene
vector

## Questions

**1 Recognising similarities →** Copy and complete the following:
*Example:*
Petrol : (is to) engine :: (as) food : (is to) animal
Saw : carpenter :: ................ : genetic engineer
Needle and thread : fashion designer :: ................ : genetic engineer

**2 Recognising similarities →** An analogy is defined as a comparison between two different things on the basis of some similarity. Statements such as 'The circulatory system is like an irrigation system,' and 'The cell membrane is like the outer wall of a castle' are examples of analogies.

- **a** Identify a tool of the genetic engineer that might be seen as analogous to one of the following objects. Give a reason for each of your choices.
  ................ is like a photocopier because ..............
  ................ is like adhesive tape because ...............
  ................. is like a scalpel because ......................
- **b** Things compared in an analogy have some similarities, but they are not similar in every way. For each analogy, describe one important way in which the items compared differ.

**3 Demonstrating knowledge →** A solution contains many copies of a piece of DNA with three cutting sites for a particular restriction enzyme. This solution of DNA is treated with the restriction enzyme under controlled conditions.

- **a** What is the expected result of this treatment?
- **b** Is this change in DNA an example of a chemical or a physical change?

**4 Applying knowledge and understanding →** The action of the cutting enzyme Pst I is as follows:

```
C T G C A G              C T G C A          G
             ────────→
G A C G T C              G          A C G T C
```

- **a** What kinds of ends does this enzyme produce?
  Copies of one long fragment of DNA were treated with Pst I and three smaller fragments were produced:

| | |
|---|---|
| fragment 1 | 2.3 kbp |
| fragment 2 | 8.1 kbp |
| fragment 3 | 4.4 kbp |

  (One kbp denotes one kilo (thousand) base pairs.)
- **b** How long was the original DNA fragment?
- **c** How many recognition sites for this enzyme were present?
- **d** If the three fragments were subjected to electrophoresis, draw the expected outcome, showing the + and – electrodes.

**5 Demonstrating knowledge and understanding →** A young woman from a family with a history of haemophilia had a test that revealed that she was a carrier of the haemophilia ***c*** allele. This woman later became pregnant and found out that her fetus was male. She wished to know if her fetus would be affected by haemophilia.

- **a** Carefully explain if and how an answer might be given to her.
  One DNA test for haemophilia involves the use of the restriction enzyme Bcl I.
- **b** For this test, could Bcl I be replaced with another restriction enzyme, such as Sac I? Explain your decision.

**6 Analysing data and drawing conclusions →** A piece of double-stranded DNA is shown below.

← 8.4 kbp →

A sample of this DNA was treated with the cutting enzyme Alu I. The cutting site for this enzyme is AGCT.

- **a** After this treatment, the sample was subjected to electrophoresis. The result is shown in figure 13.30. The numbers denote the sizes of the fragments in kbp. What conclusion may be drawn?

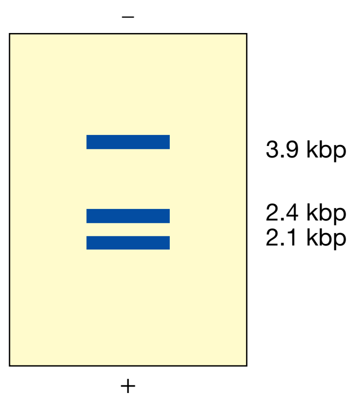



FIGURE 13.30

**b** A second sample of this DNA was treated with a different cutting enzyme, Hin dIII. The cutting site for this enzyme is AAGCTT. The result after electrophoresis revealed two DNA fragments of sizes 5.4 kbp and 3.0 kbp. What conclusion may be drawn?

**c** A third sample of this DNA was treated with yet another cutting enzyme, Not I. The result after electrophoresis revealed one fragment of 8.4 kbp. What conclusion may be drawn?

**7** **Demonstrating understanding →** One sample (S1) of DNA was cut into pieces using a cutting enzyme that produces blunt ends. A second sample (S2) of DNA was cut using a cutting enzyme that produces sticky ends. In both cases, the sizes of the fragments produced were similar.

Scientist AA wanted to join the DNA pieces in sample S1. Scientist BB wanted to join the DNA pieces in sample S2. Both scientists carried out their experiments with the DNA in solution under identical conditions.

**a** Explain the terms *sticky end* and *blunt end*.

**b** What enzyme would the scientists use to catalyse the joining?

**c** The scientists observed that the pieces of DNA from one sample joined more quickly than those in the other sample. For which sample would the faster rate of joining be observed? Explain your choice.

**8** **Demonstrating knowledge →**

**a** Describe one means by which a scientist can obtain a copy of a gene.

**b** Having obtained a copy, describe one means by which the scientist can obtain millions of copies of the gene.

**9** **Analysing data and applying principles →** Examine figure 13.31 and answer the following questions:

**a** Identify the lane that contains:

i the shortest fragments of DNA

ii DNA fragments that include one restriction site for Bgl II and were treated with that restriction enzyme

iii DNA fragments with an approximate size of 1500 bps

iv DNA fragments with a size in excess of 3000 bps.

**b** If this same DNA that was treated with Bgl II had instead been treated with Hin dIII, would the pattern of bands obtained in electrophoresis have been the same? Explain.

**c** A different sample of DNA was treated with a particular restriction enzyme that produced three fragments of sizes 400, 750 and 2500 bps. Consider that this DNA mixture was loaded into lane 5 of the gel shown in figure 13.31. Draw the expected pattern on completion of electrophoresis of this DNA in lane 5.

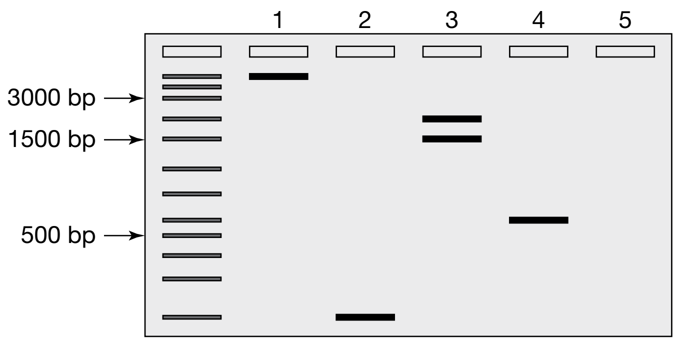



FIGURE 13.31

**10** **Demonstrating knowledge and understanding →** Complete the following sentences:

**a** In bacterial transformation, bacteria take up foreign ________.

**b** A vector commonly used in transformation experiments is a ________.

**c** For transformation to occur, the DNA taken up by bacteria must be ________ in their phenotypes.

**d** Sticky ends join temporarily through the formation of ________.

**e** Temporary joins can be made permanent through the use of ________.

**f** The replication of plasmids is ________ of the replication of the main bacterial chromosome.

**11** **Analysing data and applying principles →** Examine figure 13.32, which shows a pAES30 plasmid and the location of recognition sites for various restriction enzymes. This plasmid comprises 4715 bps. Note that $Km^R$ is the DNA of the allele for resistance to the antibiotic kanamycin, and MCS is a multiple cloning site for the restriction enzymes shown alongside it.

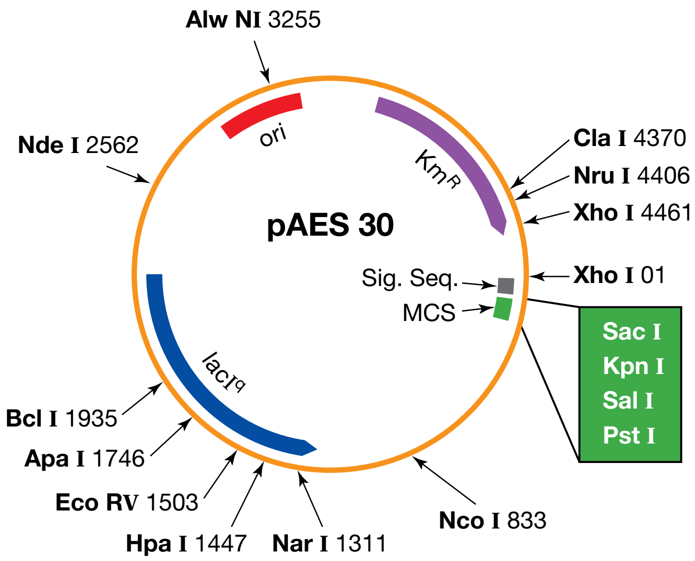



**FIGURE 13.32**

**a** What would be the effect of such a plasmid in terms of the number of fragments and sizes, if it was subjected to the following treatments:

i treatment with Nde I
ii treatment with Nde I plus Sac I
iii treatment with Nde I plus Sac I plus Cla I.

These plasmids are to be made recombinant by incorporating a fragment of foreign DNA. The decision needs to be made about which restriction enzyme to use for this purpose.

Student P says: '*Use any one of them. It makes no difference*.' Student Q says: '*No way. There are some we should definitely not use*!'

**b** With which student do you agree? Explain your choice.

**12 Discussion question →**

CRISPR-Cas9 technology allows for gene editing. An issue that is the subject of debate is:

*Should gene editing be allowed on human cells, such as eggs, sperm and embryos, whose DNA can pass to future generations?* Or, *should gene editing be restricted to somatic cells where that DNA is not passed to the next generation?*

Discuss this issue with your classmates.

CHAPTER

# 14 DNA applications in society

**FIGURE 14.1** This new cemetery in France is the last resting place of some Australian soldiers killed in the tragic battle of Fromelles in July 1916 and whose remains were unaccounted for until 2009. In that year, the remains of 211 soldiers were recovered from an unmarked mass grave. Since then, many have been identified, partly with the assistance of DNA technology. In this chapter, we will explore several gene technologies developed because of our understanding of the workings of DNA and consider some social and ethical implications associated with their use. (Image courtesy of Marjory Martin.)

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- understand the processes involved in various gene technologies, in particular, gene cloning and DNA profiling
- become aware of a range of applications of gene technologies
- recognise that the use of gene technologies may raise ethical and social issues.

# Insulin: a life saver

A young teenage boy lay dying in the Toronto General Hospital in Canada in January, 1922. The cause of his impending death was the fact that his blood glucose levels were out of control and had risen to dangerously high levels. At the same time, in spite of the excessive glucose in his blood, the boy's cells were starved of the glucose they needed to generate the energy for their life-sustaining metabolic reactions. With no treatment available, his urine was loaded with glucose, he had lost weight, was physically weak and his organs had been damaged. It was just a matter of time until he fell into a coma and died. The boy was suffering from what today we call type 1 diabetes, but it was out of control.

The boy's imminent fate was not unusual at that time. With no effective medical treatment available, anyone with type 1 diabetes was faced with the prospect of a greatly shortened lifespan. At best, the lifespan of affected young people might be prolonged, perhaps by up to a year, by placing them on a near-starvation diet with virtually no intake of carbohydrates. (The major product of digestion of carbohydrates is glucose, so eating carbohydrates would produce a further increase in already dangerously high levels of blood glucose.) This intervention, however, simply delayed the inevitable death of an affected person.

However, the young Canadian boy was to cheat death on this occasion. He became the first person ever to receive an injection of insulin, a pancreatic hormone that regulates blood glucose levels. This injection reduced the boy's blood and urinary glucose concentrations to normal levels and enabled his cells to take up glucose. He gained weight, his appetite was restored and, by May 1923, the boy was in good health and was discharged from hospital.

## Finding the elusive substance

In the first part of the twentieth century, it had been known for some years that particular cells of the pancreas, called the islets of Langerhans, produced a mysterious substance that kept blood glucose levels within narrow limits. In a vain attempt to alleviate the symptoms of diabetes, people were fed or injected with pancreatic tissue, but this had no effect. Many attempts to isolate this elusive substance failed.

However, in July 1921, after many trials, Frederick Banting (1891–1941) and Charles Best (1899–1978), working under the direction of John Macleod (1876–1935) at the University of Toronto in Canada, isolated a crude substance that could alleviate the diabetic symptoms of dogs that had their pancreases removed (see figure 14.2).

In December 1921, biochemist James Collip (1892–1965) purified the crude pancreatic extract and made it suitable for human use. This substance was later named insulin by them.

**FIGURE 14.2** Charles Best and Frederick Banting, co-discoverers of insulin, photographed with Marjorie, the first diabetic dog to be brought out of a coma using insulin.

**ODD FACT**

At a scientific meeting in May 1923, John Macleod presented a paper titled *'The Effects Produced on Diabetes by Extracts of Pancreas'*, which was received with great excitement. The authors of the paper were Banting, Best, Collip, Campbell, Fletcher, Macleod and Noble. Campbell and Fletcher were the physicians who injected the dying Canadian boy with insulin extracted from cattle pancreases.

**ODD FACT**

Almost 1000 kilograms of pig pancreases were required to produce about 100 grams of pure insulin.

In January 1922, the insulin isolated by Banting and Best and purified by Collip was injected into the dying teenager in the Toronto General Hospital, and it saved his life. By 1923, large-scale production of insulin by drug companies made this life-saving drug widely available, saving millions of lives. In the same year, the Nobel Prize in Physiology or Medicine was awarded jointly to Banting and Macleod 'for the discovery of insulin,' and they shared their prizes with Best and Collip.

## Early commercial production of insulin

The starting points for the commercial production of insulin by drug companies were the pancreases of cattle and pigs obtained from abattoirs. For nearly six decades, the insulin used for injections of people with type 1 diabetes came from these sources.

This insulin effectively controlled blood sugar levels, but some people suffered local reactions to this insulin, such as rashes and swellings at the injection sites. This may have been because the product contained trace amounts of impurities. Human insulin is a protein hormone built from a total of 51 amino acids organised as two chains: an amino acid A chain containing 21 amino acids, and an amino acid B chain containing 30 amino acids (see figure 14.3). In comparison, insulin from cattle differs in three amino acids, and there is a single amino acid difference between pig insulin and human insulin. Because insulin from pigs and cattle is not a perfect match to human insulin, there is also the risk of allergic reactions occurring in some people.

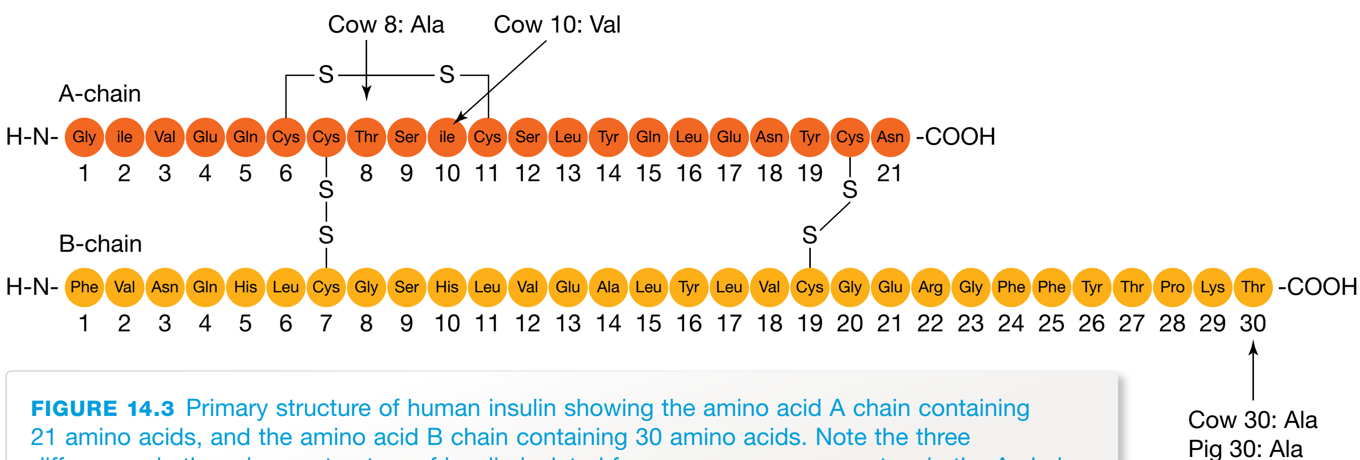



**FIGURE 14.3** Primary structure of human insulin showing the amino acid A chain containing 21 amino acids, and the amino acid B chain containing 30 amino acids. Note the three differences in the primary structure of insulin isolated from a cow pancreas — two in the A chain at amino acid residues 8 and 10, and one in the B chain at amino acid residue 30. Pig insulin has a single difference at position 30 in the B chain.

## Then came recombinant insulin

First, some definitions:

- Gene cloning is the process by which a *gene* of interest is located and cloned (copied) to produce multiple copies.
- Recombinant DNA is DNA that is formed by combining DNA from different sources, often from different kinds of organisms.
- Recombinant proteins are proteins that are expressed by recombinant DNA present in an organism. A protein obtained by introducing recombinant DNA into a host (such as a bacterial cell or yeast cell) produces the gene product. A cloned gene encodes the amino acid sequence of such a protein.

In the 1980s, a major breakthrough occurred when the American drug company Eli Lilly used gene cloning and expression of the copied genes to produce recombinant (rec) human insulin, known as Humulin®


**FIGURE 14.4** When beta cells of the human pancreas fail to produce insulin, recombinant human insulin made by gene manipulation techniques can substitute for the missing hormone. Shown here is Humulin® or recombinant (rec) insulin produced by gene cloning.

(see figure 14.4). In late October 1982, the Food and Drug Administration (FDA) of the United States approved Humulin® for human use, and it has now replaced the insulin prepared from non-human sources, such as from the pancreatic tissue of cows and pigs.

Today, if the beta cells of the human pancreas fail to produce insulin, recombinant human insulin can substitute for the missing hormone. Following the development of Humulin®, other recombinant insulins have been developed with slightly different properties, including rapid-acting forms (e.g. generic name: Insulin lispro) and long-acting forms (e.g. generic name: Insulin glargine). Figure 14.5 shows the time course of action of these various forms of insulin on blood glucose levels. Insulin is now supplied in vials, as pumps and as pens (see figure 14.6).

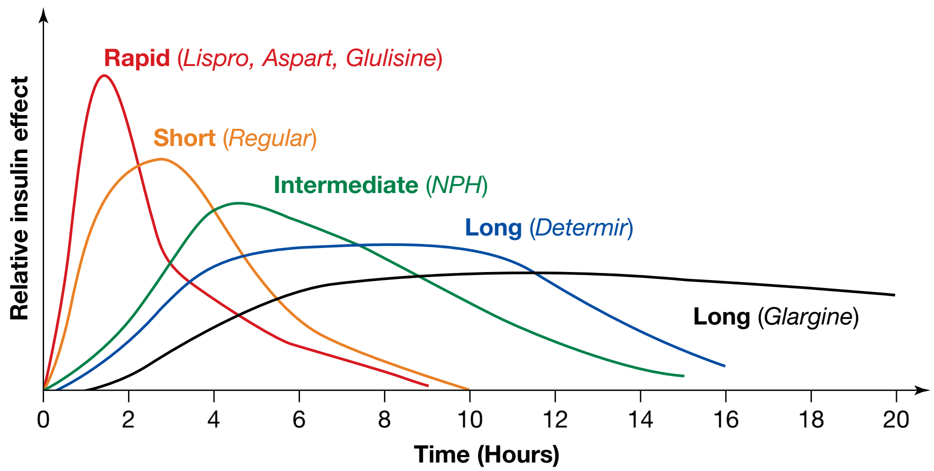



**FIGURE 14.5** Time course of action of various forms of recombinant insulin on blood glucose levels. Which type would be expected to be the best to manage the surge in blood glucose levels soon after eating a meal?


**FIGURE 14.6** An insulin pen that can be conveniently carried and used to administer insulin.

## The list of recombinant proteins expands

The first recombinant protein to be developed and made available commercially was Humulin®. Since 1982, the expression of cloned genes has been used to produce many other recombinant proteins, both hormones and enzymes, that are used clinically. Table 14.1 identifies just some of these recombinant proteins.

**TABLE 14.1 Some recombinant DNA technology products. The host cell is the 'cellular factory' in which the genes concerned are cloned and then expressed as a particular protein product. (Humulin®, Puregon®, etc. are registered trademarks for products, just like Weeties and Vegemite.)**

| Example of rec protein | Host cell type | Use of rec protein |
|---|---|---|
| Humulin® (rec insulin) | *E. coli* bacteria or yeast cells | used in treatment of type 1 diabetes |
| Humatrope® (rec human growth hormone) | *E. coli* bacteria or yeast cells | used in treatment of growth deficiency in children |
| Puregon® (rec human follicle-stimulating hormone) | hamster ovary cells | promotes ovulation and is used in the treatment of infertility |
| Recombinate® (rec factor 8) | hamster ovary cells | blood clotting factor used to control bleeding episodes in haemophilia |
| Erythropoietin (rec hEPO) | hamster ovary cells | used in treatment of severe anaemia |
| Interferons (recIF) | *E. coli* bacteria | used as an anti-viral treatment |
| Hepatitis B vaccine (rec HBsAg) | yeast cells or hamster ovary cells | used in immunisation against hepatitis B infection |

Some of these recombinant proteins, such as Humulin®, replaced products that were previously extracted from tissues of other species. Other recombinant proteins, such as human growth hormone (rec hGH), replaced products that were extracted from tissues of dead people (cadavers).

Other recombinant proteins replaced products that were extracted from tissues of living people, such as follicle stimulating hormone (FSH), which was previously extracted from the urine of post-menopausal women, and hepatitis B vaccine, which was previously based on antibodies extracted from the plasma of people with chronic hepatitis infection.

Finally, other recombinant proteins are new products that have only ever been produced through expression of cloned genes — these include proteins involved in aspects of immunity, including monoclonal antibodies, (rec Mabs), colony stimulating factor (rec CSF), and interferon (rec IF).

Advantages of producing proteins through expression of cloned genes as compared to extraction from other biological sources include:

- high levels of purity
- reliability of supply
- consistency of quality between batches.

The following section identifies a tragic problem that emerged when the only source of one hormone for clinical use was by extraction from human tissue sourced from a very large numbers of cadavers.

## The growth hormone story

Human growth hormone (hGH), secreted by the pituitary gland, stimulates the growth of all body tissues, including bones. hGH was first isolated from human pituitary glands in 1956. By the 1960s, hGH was in clinical use in the treatment of children and adolescents with a particular growth deficiency. This treatment often involved the subcutaneous injection of growth hormone on a daily basis.

A growth deficiency in a child or an adolescent is typically identified when he or she is much shorter than most others of the same age and gender.

Throughout the 1960s, 1970s and early 1980s, the only source of human growth hormone was from pituitary glands recovered from cadavers. To produce sufficient growth hormone for use, very large numbers of pituitary glands from cadavers were required. Each batch of hGH was derived from a pool of pituitary glands from about 16 000 cadavers. During a year, patients received their hGH injections from two or three different batches. (A batch is a quantity of a medicine or drug that is manufactured under specific conditions during a specific period.)

In the mid 1980s, a small number of young adults were diagnosed with Creutzfeldt-Jakob Disease (CJD). This disease has a long incubation period but, once the disease symptoms appear, it progresses rapidly and is fatal, typically within a period of 8 to 18 months. CJD is the result of a slow-acting abnormal prion infection (refer back to chapter 6, page 228).

The appearance of CJD in young adults was unheard of and highly unexpected. Investigation revealed that, in every case, all these affected young adults had been treated over an extended period in childhood and adolescence with daily injections of hGH extracted from pooled human cadaver pituitary glands. It was concluded that one or more of the pituitary glands used to prepare growth hormone during some period of time came from people with CJD and was the source of the abnormal prions.

**ODD FACT**

An iatrogenic disease is one that is the result of a medical or clinical procedure. The first ever case of iatrogenic CJD was reported in 1974 and involved the recipient of a corneal transplant from an infected cadaver donor.

The distribution and clinical use of cadaver-derived hGH was stopped. By 2012, worldwide, 226 cases of this iatrogenic CJD resulting from cadaver-derived growth hormone had been reported, with most cases occurring in France, the United Kingdom and the United States.

In 1985, recombinant hGH became available — this was produced using gene cloning and expression of the cloned genes, a process that removed the risk of prion infection. This rec hGH is now the standard treatment for children and adolescents with growth deficiencies arising from a lack of this hormone.

In the upcoming box, you can read how the islet cells of your pancreas produce inactive pre-proinsulin, which is transformed in these cells to proinsulin and then to active insulin, which is secreted into the bloodstream.

**KEY IDEAS**

- The pancreatic hormone, insulin, was first isolated and purified for human use as a result of the combined efforts of four Canadian scientists.
- Insulin derived from cattle pancreases was first used clinically in 1923.
- For more than five decades, insulin for clinical use was extracted from the pancreases of cattle and pigs.
- Recombinant human insulin became available in 1982 using the expression of cloned genes.
- The range of recombinant polypeptides and proteins continues to increase.
- Growth hormone derived from pituitary glands from cadavers was a source of prions that resulted in deaths from CJD of some young adults treated in childhood and adolescence with this hormone.

## FROM PRE-PROINSULIN TO INSULIN

The **INS** insulin gene is translated in the beta cells of the islets of Langerhans of the pancreas on the ribosomes.

The product of translation is an inactive polypeptide consisting of 106 amino acids known as pre-proinsulin.

Pre-proinsulin moves from the ribosomes into the channels of the rough endoplasmic reticulum where it is acted on by a protease enzyme that removes the leading 24 amino acids of this molecule, a signal sequence. This leaves an inactive precursor molecule of 82 amino acids called **proinsulin**.

Proinsulin is transported to the Golgi complex where it undergoes further modification. In the Golgi complex, proinsulin molecules are folded, covalent bonds form and the molecules are packaged into membrane-bound secretory granules.

Within the secretory granules, the proinsulin is cleaved by an enzyme that removes a segment of 31 amino acids, called the C chain (or C peptide). The remainder forms active insulin molecules, each composed of an A chain and a B chain held together by strong covalent disulfide **- S – S -** bonds. Active insulin is then exported from the Golgi complex into the bloodstream as required. The signal for its release is an elevation in the normal blood glucose levels.

Figure 14.7 gives a simplified summary of this process of conversion of pre-proinsulin to insulin.

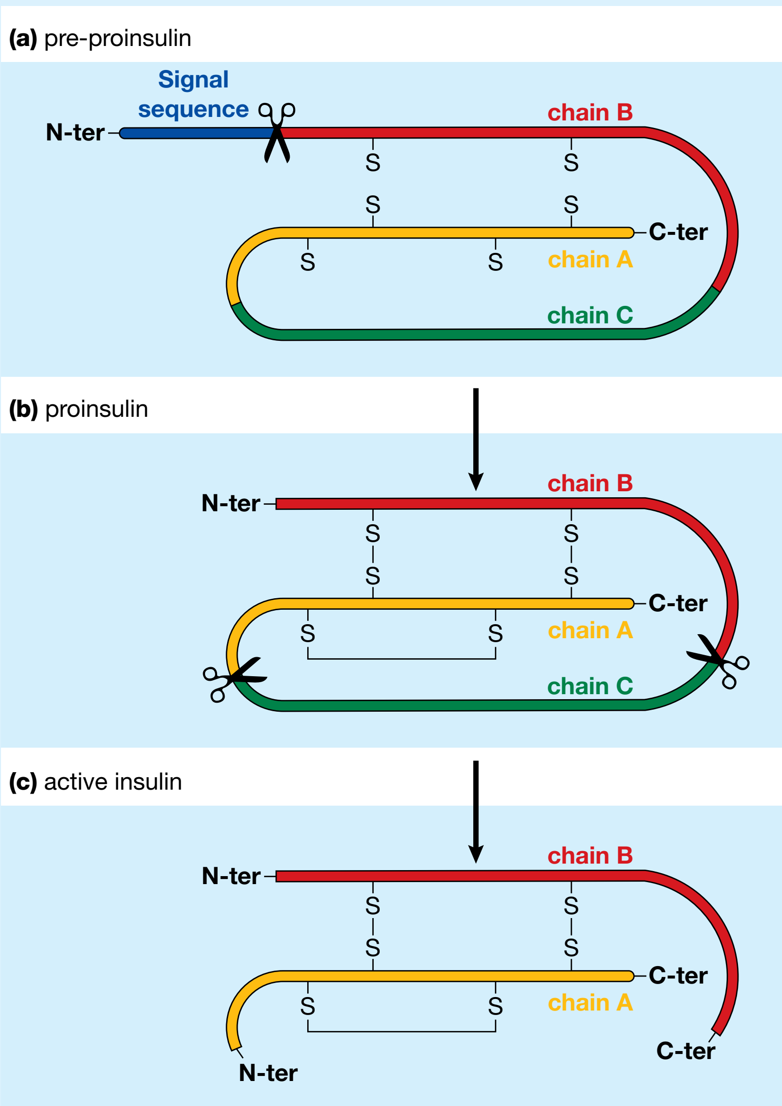



**FIGURE 14.7** Diagram showing the formation of active insulin from pre-proinsulin. **(a)** Pre-proinsulin, the immediate product of gene translation, is quickly modified to proinsulin by removal of the leading signal sequence. **(b)** Proinsulin, the inactive precursor of insulin, is modified to active insulin by removal of the C chain. **(c)** Active insulin is formed by the removal of the C fragment, leaving the A and B chains held together by covalent disulfide bonds.

## QUICK CHECK

**1** Briefly explain the following terms:
 **a** 'recombinant protein'
 **b** host cell.
**2** Why was insulin called 'a miracle drug' when it first became available?
**3** Identify the following statements as true or false:
 **a** All recombinant proteins are replacements for proteins that were extracted from mammalian, including human, tissues in the past.
 **b** The first recombinant protein to be produced for clinical use was insulin.
 **c** The past source of a vaccine for hepatitis B made use of antibodies from the plasma of people with a chronic hepatitis B infection.
**4** What was the cause of CJD in young adults from the 1970s to the mid 1980s?

# Making recombinant insulin using cloned genes

To make recombinant insulin or any other protein, some important preliminary steps must be carried out:
1. Get copies of the required DNA to transfer to plasmids — to clone a gene, you must first find it or make it.
2. Choose a suitable plasmid to carry this foreign DNA — to clone a gene, it needs a vector or carrier.
3. Select the host cell for the recombinant plasmids — to clone a gene, the plasmid must transfer it to a host cell that supplies the environment in which cloning can take place.

### Getting the gene of interest

The DNA of the gene of interest can be obtained, either by making it using a DNA synthesiser, or by using its mRNA template to make a cDNA copy of the gene. In the case of insulin, we can start with the DNA that encodes proinsulin, the precursor of insulin.

As an alternative to starting with the gene for proinsulin, recombinant insulin could also be made by cloning the DNA sequences that encode the A chain and the B chain of insulin.

When the amino acid sequence of a gene of interest is known, the DNA sequence that encodes it can be inferred by reference to the genetic code polypeptides. For example, proinsulin includes the amino acids *gly-ile-val-* in order, so a DNA sequence that can encode these three amino acids will comprise the base sequence: **CCT – TAT – CAC –**. If the gene is not too large, the required sequences can then be made using a DNA synthesiser that has an output of double-stranded (duplex) DNA. Otherwise, a cDNA copy must be made (refer back to chapter 13, page 626).

Each sequence must also include a recognition site for a restriction enzyme at its leading and trailing ends (see figure 14.8). The restriction enzyme selected is one that has a recognition site in the plasmids that will be used to transfer the gene of interest to the host cells. The restriction enzyme, Bam HI, meets this requirement.

Let's outline the procedure to clone the DNA that encodes proinsulin. This is our gene of interest. Figure 14.8 shows a diagram of our gene of interest showing the recognition sites for the Bam HI restriction enzyme at each end of this DNA sequence. Note that the DNA of the plasmid contains several thousand base pairs, the gene of interest consists of several hundred base pairs, and the restriction site is a DNA segment of just six base pairs.

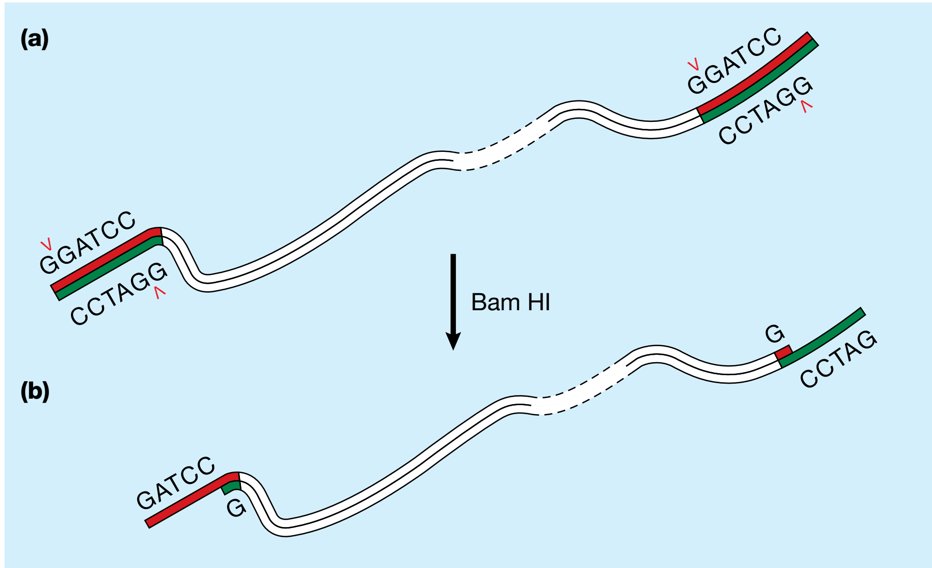



**FIGURE 14.8** **(a)** Diagram showing a representation of the DNA sequence of our gene of interest. The sequence of this double-stranded DNA includes a Bam HI recognition site at its leading and trailing ends. The red symbols denote the cut points in these recognition sites. **(b)** Diagram showing this DNA sequence after treatment with the restriction enzyme Bam HI. Note the sticky ends that are produced by Bam HI.

## Choosing a plasmid to be the vector

Before a gene can be cloned, it must first be transferred to a suitable vector, such as a plasmid. We will use plasmid **pUC19** as our vector. This plasmid has a recognition site for the same restriction enzyme, Bam HI, that created the 'sticky ends' on our gene of interest.

The role of the plasmid is twofold:

(i) The plasmid must carry the gene of interest into a living cell and, once there, the plasmid must replicate itself and produce multiple copies of or clone the inserted gene. Any plasmid that carries the gene of interest is termed a **recombinant plasmid** because it contains DNA from at least two different sources.

(ii) A plasmid must advertise if it is recombinant with the gene of interest inserted into its structure. Two markers present in the plasmid will enable us to recognise recombinant plasmids as well as host cells that possess these recombinant plasmids. These markers are a selectable antibiotic resistance marker, ***Amp$^R$***, and a screening marker, ***lacZ.***

The DNA of the *lacZ* screening marker encodes an enzyme that acts on a colourless substrate (X-gal) and produces an intracellular blue product:

$$\text{colourless X-gal substrate} \xrightarrow[\textit{enzyme}]{\textit{galactosidase}} \text{blue product}$$

The *lacZ* marker of the pUC18 plasmid contains the recognition site for the Bam HI restriction enzyme, so if this gene has been incorporated into this site, the enzyme cannot be produced and no blue product can appear.

Figure 14.9 shows the features of our pUC18 plasmid.

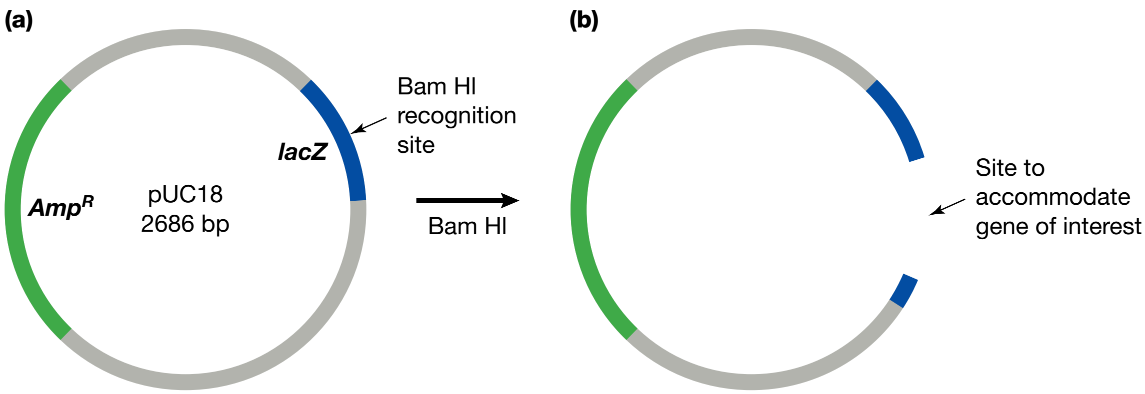



**FIGURE 14.9** **(a)** Simple diagram of plasmid pUC18 showing its antibiotic resistance markers, ***Amp***$^{R}$, and its colour marker, ***lacZ***, before treatment with Bam HI and **(b)** after treatment with Bam HI, which cuts open the plasmid, forming sticky ends. Note that the treatment with Bam HI interrupts the *LacZ* DNA sequence. If the gene of interest is incorporated here, the enzyme that is responsible for the colour marker cannot be formed.

Because of these markers, recombinant plasmids can easily be identified, as well as cells that carry these plasmids (see figure 14.10). If the gene of interest is incorporated into the plasmid at the Bam HI recognition site, the enzyme that makes the blue product is lost because the DNA of this segment is disrupted. If the gene of interest is not present, the enzyme is produced and the blue colour can be produced.

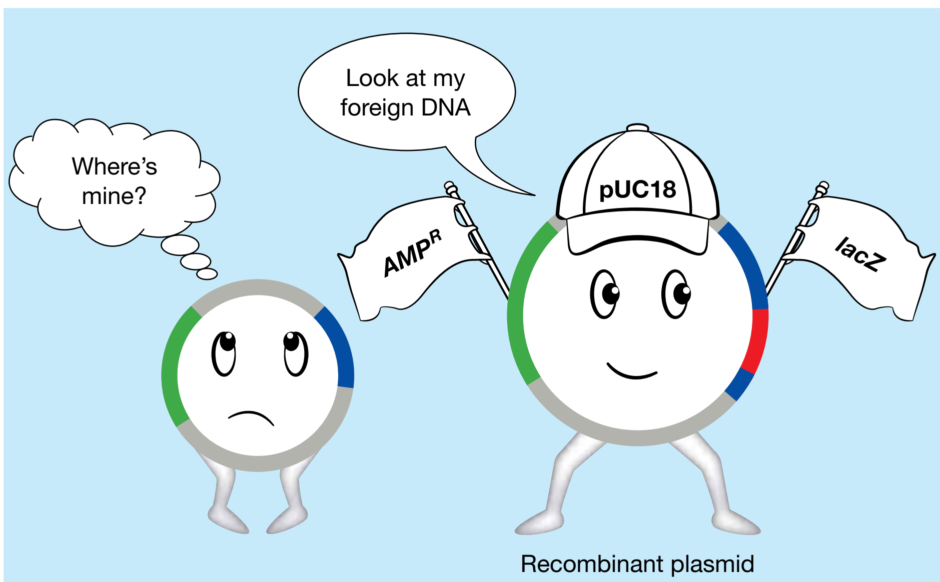



**FIGURE 14.10** Cartoon figure representing a plasmid (at right) with a gene of interest (shown in red) inserted into the *lacZ* gene (shown in blue). This insertion means that the *lacZ* gene cannot be expressed. Such a plasmid is said to be recombinant, in contrast to the LH plasmid that is non-recombinant. Which plasmid can multiply to produce a colony of blue cells?

It is also possible to use a *lacZ* screening marker to identify host cells that possess recombinant plasmids. These cells multiply in culture and produce white colonies, while all colonies of non-recombinant host cells are blue. Go to the upcoming box and read how the *lacZ* screening marker can show which bacterial cells are recombinant and which are not.

## *LacZ:* AN EASY WAY TO SCREEN FOR RECOMBINANT BACTERIA

Consider the plasmid shown in figure 14.9, which has a *lacZ* DNA sequence that includes a recognition site for a particular restriction enzyme. The *lacZ* acts as a screening marker that can indicate which bacterial cells have incorporated recombinant plasmids carrying foreign DNA and which bacterial cells are non-recombinant.

How does this work?

The DNA of the *lacZ* screening marker encodes an enzyme, galactosidase, that converts a particular substrate to a blue-coloured product. The neat trick with the *lacZ* screening is that if foreign DNA is incorporated at the recognition site, the DNA coding sequence of *lacZ* is disrupted so that no functional enzyme is formed. No enzyme: no blue colour!

On the other hand, if foreign DNA is not incorporated at this site, the *lacZ* DNA can produce a functional enzyme. Enzyme present: blue colour formed.

$$\text{colourless substrate (X-gal)} \xrightarrow[\text{enzyme}]{\text{galactosidase}} \text{blue product}$$

After recombinant plasmids are mixed with host bacterial cells and subjected to a heat shock (or to electroporation), some bacterial cells will take the recombinant plasmids that carry the foreign DNA. Because this foreign DNA disrupts the *lacZ* DNA sequence, no *lacZ* enzyme is formed. However, other bacterial cells will take up non-recombinant plasmids. These bacteria can produce the *lacZ* enzyme and it catalyses the conversion of a particular substrate to a blue-coloured product.

Next, a sample of the mixture of bacterial cells is grown on agar plates that contain the substrate for the *lacZ* enzyme. The bacteria that have taken up recombinant plasmids will reproduce, but will form white colonies. In contrast, bacteria that lack recombinant plasmids will reproduce and form blue colonies. Figure 14.11 shows an agar plate with bacterial colonies growing on it. Note that some colonies are blue and some are white. Which colonies are composed of recombinant bacteria — blue or white?

**FIGURE 14.11** Photo of an agar plate with bacterial colonies. The agar includes a substrate that forms a blue product in the presence of the *lacZ* enzyme. Which colonies contain bacteria with an active *lacZ* enzyme? Why are some colonies white?

### Choosing host recipient cells for the plasmids

A harmless strain of the bacterial species *Escherichia coli* is commonly selected for cloning genes and expressing the proteins encoded by these genes. The cells can replicate rapidly, doubling their numbers about every 20 minutes. In addition, any recombinant plasmids they possess replicate independently and increase at an even faster rate (refer back to chapter 13, page 633). These *E. coli* host cells have special nutritional requirements so that they cannot survive outside the laboratory. We will use these weakened *E. coli* as our host cells.

Yeast cells and other eukaryotic cells can also be used as host cells. Yeast cells or hamster ovary cells must be used when the final functional protein is a glycoprotein, that is, a protein with sugar residues attached to many of its amino acids. Bacterial cells cannot carry out this process of sugar attachment (glycolation).

So, we now have our copies of the gene of interest — the proinsulin gene, our plasmids and our host cells. Let's see how we can clone our gene of interest and let it express the protein that the gene encodes.

**ODD FACT**

*E. coli* host cells produce their recombinant proteins as intracellular inclusion bodies. In contrast, yeast cells and other eukaryotic host cells produce soluble recombinant proteins that they secrete into their culture medium.

**ODD FACT**

The recombinant protein erythropoietin (rec hEPO) is a glycoprotein. It contains 40 per cent by weight of sugars, and so, is a very highly glycolated protein. What kind of host cell would be needed to produce rec hEPO?

## Cloning a gene and expressing the protein it encodes

**Step 1: Inserting the gene of interest into plasmids.** The plasmids and the copies of the gene of interest, both with their sticky ends from treatment with Bam HI, are mixed together. What will happen? Because of the matching sticky ends, many plasmids will join with the DNA of the gene of interest to form recombinant plasmids (that is, with DNA from different sources). However, some plasmids simply re-join with themselves or join with other plasmids, and some DNA fragments will join with other foreign DNA fragments with matching sticky ends. When ligase enzyme is added to the mixture, any joins are made permanent.

Figure 14.12 shows some of the next steps that we need to complete in order to use cloned recombinant cells as factories to produce insulin.

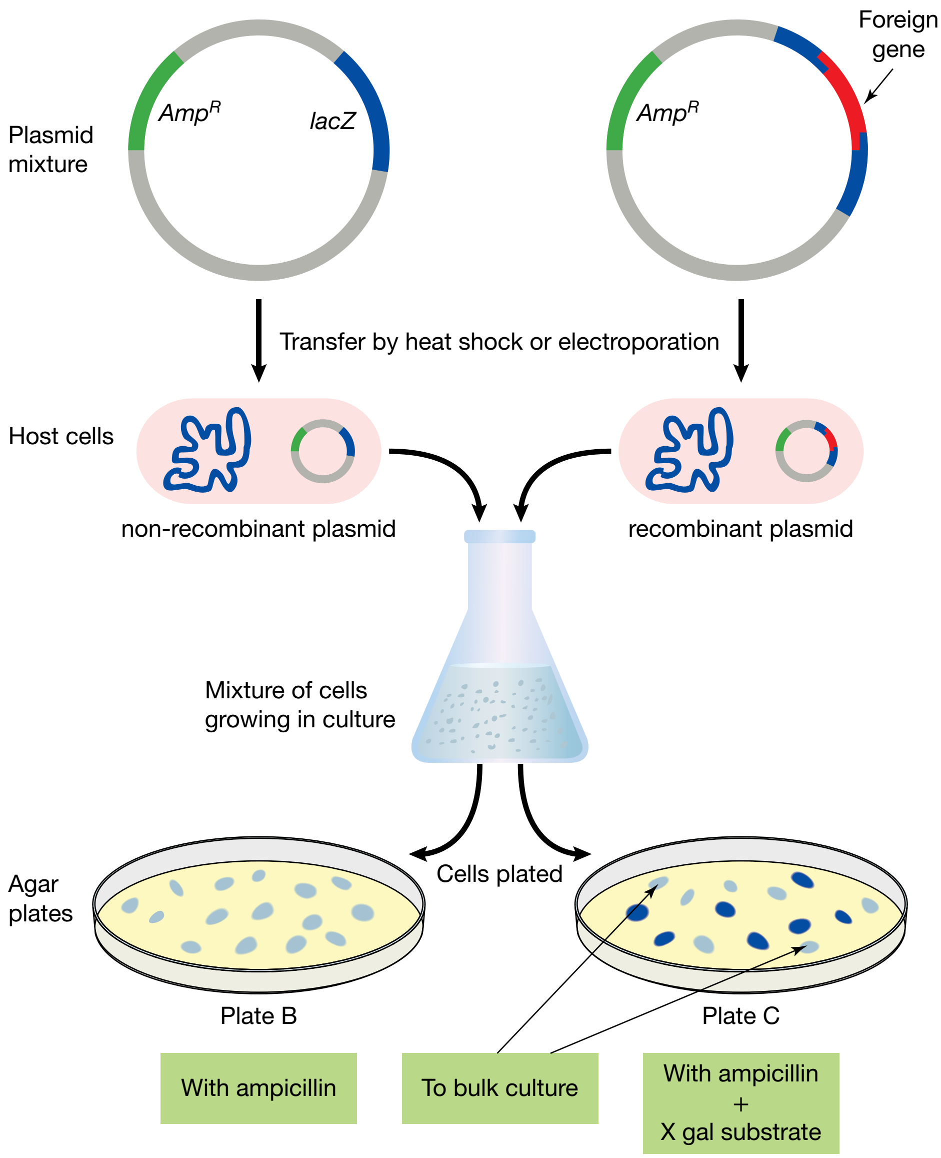



**FIGURE 14.12** Outline of some of the steps to produce insulin using cloned genes carried in recombinant plasmids in host *E.coli* cells. More detail is provided in the steps outlined.

**Step 2: Transferring plasmids into host cells.** The plasmid mixture is added to a culture of host bacterial cells that are ampicillin sensitive. The host cells and plasmid mixture is then subjected to either heat shock or electroporation to increase the chance of uptake of plasmids by the bacterial cells. The bacterial culture is incubated overnight, during which period the cells divide by binary fission and plasmids replicate (refer back to figure 13.28 on page 633).

**Step 3: Identifying host cells with recombinant plasmids.** Bacterial cells are spread over culture plates as described below. What will happen?

***Plate A:*** nutrient agar. Any viable bacterial cells will grow on this plate.

***Plate B:*** nutrient agar plus ampicillin. Only cells with the pUC18 plasmid will grow on this plate, both cells with recombinant plasmids and cells with non-recombinant plasmids. Cells without this plasmid are ampicillin sensitive and are killed by this antibiotic.

***Plate C:*** nutrient agar plus ampicillin plus X-gal substrate. Cells that possess recombinant plasmids with the proinsulin gene inserted in the *lacZ* segment of the plasmid will grow on this agar, but the colonies that they form are white. Cells with non-recombinant plasmids have an intact *lacZ* segment and produce the enzyme that converts the colourless X-gal substrate to a blue product. The colonies that these cells form are blue.

**FIGURE 14.13** A small-scale bioreactor used for growth of recombinant cells and for production of recombinant proteins. Testing on a small scale enables the identification of the optimal culture conditions and provides a check that the recombinant cells are expressing the cloned gene.

**Step 4a: Small-scale culturing of recombinant cells.** Samples of recombinant cells from the white colonies on plate C are tested on a small scale for their ability to express the proinsulin gene that they carry and to fine-tune the conditions under which maximal growth can be achieved (see figure 14.13).

**Step 4b: Mass culturing of recombinant *E. coli* cells.** Once a cell line has been identified in small-scale tests as having a high level of expression of the proinsulin gene, it is moved to large-scale mass culturing.

The recombinant bacteria are cultured in large stainless steel fermentation tanks (see figure 14.14) that have a capacity of up to 50 000 litres. The culture medium, kept at optimal temperature, supplies all the requirements required for bacterial cell growth, with the bacteria replicating about every 20 minutes. Each of the millions of bacterial cells in the mass culture is like a mini proinsulin-producing factory as it expresses the cloned proinsulin genes carried in its recombinant plasmids.

**FIGURE 14.14** Stainless steel fermentation vats in which mass culturing of recombinant *E. coli* bacterial cells takes place and where those cells express the proinsulin gene.

**Step 5: Isolation and purification of insulin.** The proinsulin molecules produced by the bacterial cell factories are retained within the cells as inclusion bodies. These inclusion bodies are freed from the cells by use of enzymes that break down bacterial cell walls. Following this, an elaborate procedure of purification occurs. Finally, proinsulin is converted to active insulin by removal of the C peptide (C chain ), leaving active insulin comprising the A and B chains linked by disulfide bonds.

**study on**

| Unit 4 | **Transporting DNA into cells** Summary screen and practice questions |
|---|---|
| AOS 2 | |
| Topic 1 | |
| Concept 5 | |

**study on**

| Unit 4 | **Gene cloning** Summary screen and practice questions |
|---|---|
| AOS 2 | |
| Topic 2 | |
| Concept 1 | |

**study on**

| Unit 4 | **See more** Gene cloning |
|---|---|
| AOS 2 | |
| Topic 2 | |
| Concept 1 | |

Purification of recombinant proteins from the host cells in which they are prepared is achieved in various ways. Figure 14.15 shows the use of affinity chromatography in purifying soluble recombinant protein prepared in *E. coli* cells.

**FIGURE 14.15** A scientist purifying a sample of recombinant protein using affinity chromatography.

## KEY IDEAS

- Recombinant proteins can be made through the expression of cloned genes.
- To express a cloned gene, it must be transferred into a host cell using a carrier or vector, such as a bacterial plasmid.
- Plasmids carrying cloned foreign genes are said to be recombinant.
- A plasmid used to carry foreign DNA such as a cloned gene will have several features, including a drug resistance marker, a screening marker, and will have a recognition site for a suitable restriction enzyme.
- Host cells with plasmids can be differentiated from cells lacking plasmids by selecting cells for resistance to an antibiotic, such as ampicillin.
- Host cells with recombinant plasmids can be differentiated from cells with non-recombinant plasmids using *lacZ* colour screening.
- Recombinant cells in large-scale culture express the protein product of the gene of interest.

## QUICK CHECK

5 In a sentence, briefly explain:
   a why the same recognition sites must be present on the ends of the gene of interest and at a known site in the selected plasmid
   b why the blue colour fails to appear in colonies of cells derived from cells possessing recombinant plasmids, where those plasmids have a *lacZ* screening marker that includes the recognition site for the restriction enzyme used to cut plasmids.

6 Identify the following statements as true or false:
   a One role of a plasmid in gene cloning is to incorporate a gene of interest into its structure for transfer to a host cell.
   b One role of a recombinant plasmid is to clone the gene of interest after the plasmid has been taken up by a host cell.
   c Isolated plasmids can replicate.
   d Host cells used in gene cloning can only be bacterial cells.
   e A gene of interest to be cloned can be produced by chemical synthesis.
   f If an ampicillin-sensitive cell takes up a plasmid with an $Amp^R$ selectable marker, that cell will become ampicillin resistant.
   g In a mixture of plasmids and foreign DNA, both with matching sticky ends, it would be expected that the plasmids would all join with the foreign DNA fragments.

# Testing to identify genetic status

**Genetic screening** most commonly refers to an organised program of testing of groups of a population in order to achieve early detection or exclusion of inherited disorders. In genetic screening, testing is available to all members of a population group even if they have no obvious signs of any disorder. In Australia, for example, a genetic screening program exists that tests all newborn babies for several inherited disorders (see below).

Another definition is that genetic screening is the testing of a population to identify individuals who are at risk of a *genetic* disease or of transmitting a *gene* for a *genetic* disease.

**Genetic testing** refers to the screening of an individual, most commonly an 'at-risk' person. The purpose of **genetic** testing is to confirm or rule out the presence in a person of a defective allele that is involved in an inherited disorder or that increases the person's susceptibility to a particular disease. Genetic testing may be sought by an individual or may be offered to an individual for many different reasons (see below).

Let's look first at genetic screening.

## Genetic screening

studyon

Unit 4 | AOS 2 | Topic 2 | Concept 2

**Genetic screening** Summary screen and practice questions

Any population-screening program may have potential damaging outcomes; for example, a false positive test result for a newborn baby would likely be the cause of significant worry and distress for the parents. For this reason, genetic screening of a population or sub-group within a population is only undertaken when the benefits clearly outweigh any potential harm and when certain conditions are met:

1. The screening program must relate to a significant inherited disorder either in terms of the number of people affected or the severity of the disorder.
2. A suitable test for the disorder must exist that is reliable, has proven predictive value, has a minimal rate of false positive results, and is cost-effective.
3. An effective intervention or treatment for the disorder must exist to address the condition in those people found to have the disorder.

In the newborn screening programs in Australia, a blood sample is taken from babies a few days after their birth. This blood sample is sent to state laboratories where tests are carried out for several inherited disorders and for a congenital condition, as shown in table 14.2. Although there is no cure for cystic fibrosis, early diagnosis and treatment can ease symptoms and reduce complications. Figure 14.16 on page 657 shows the genetic defect in galactosemia that, if left untreated, has drastic effects on an affected baby.

**TABLE 14.2** **Genetic screening program of newborn babies in Australia tests for inherited disorders and a common congenital condition.**

| Disorder | Frequency | If undetected and untreated | Intervention |
|---|---|---|---|
| phenylketonuria | 1 in 10 000 | brain damage from build-up of phenylalanine in blood | low phenylalanine diet |
| galactosemia | 1 in 40 000 | organ damage from build-up of galactose in blood | diet of lactose-free milk; no dairy products |
| congenital hypothyroidism | 1 in 3500 | impaired growth and brain development | daily tablet of thyroid hormone |
| cystic fibrosis | 1 in 2500 | impaired function of lungs and digestion | drugs to prevent lung infections, open airways and thin the mucus |

## Genetic testing

For more detail of the procedure for screening of newborn babies, refer to *Nature of Biology Book 1 Fifth edition*, page 590.

There are many reasons an individual person might seek or be offered genetic testing in order to identify whether they possess a variant (allele) of a particular gene that causes a genetic disorder, or significantly increases their susceptibility to a disease, or would allow them to benefit from a particular drug. Some reasons for seeking or being offered genetic testing are identified in figure 14.17 on page 658.

Types of genetic testing include the following:
- adult screening for increased risk of a disease
- carrier detection
- prenatal screening and diagnosis
- predictive screening or presymptomatic testing
- embryo biopsy or pre-implantation genetic screening.

### Adult screening for increased risk of disease

Ellen, a young woman in good health makes an informed decision to have a double mastectomy or breast removal. Why? Her family medical history shows that her maternal grandmother, her mother and two of her mother's sisters died of breast cancer and a first cousin on her mother's side died of ovarian cancer.

Only a few cancers have an inherited genetic component and one of these is breast cancer. About 5 to 10 per cent of women who are diagnosed with breast cancer have a hereditary form of that disease. About 80 per cent of women with a heritable form of breast cancer have a mutation in one of two genes, either **BRCA1** or **BRCA2**, that makes them far more susceptible to developing breast and ovarian cancers.

The **BRCA1** (**BR**east **CA**ncer 1) gene is located on the number-17 chromosome while the **BRCA2** (**BR**east **CA**ncer 2) gene is located on the number-13 chromosome (see figure 14.18 on page 658). A person with a mutation of either gene is said to be BRCA positive. Females who are BRCA1 positive or BRCA2 positive have a much higher than average probability of developing breast cancer in their lifetimes (40 to 85% versus 12% for the general female population). BRCA positive women also have a high risk of developing ovarian cancer (6% to 44%, versus just under 2% for the general female population). Men who are BRCA positive have a higher than average risk of developing prostate cancer and, if they are BRCA2 positive, they are at higher than average risk of developing breast cancer.

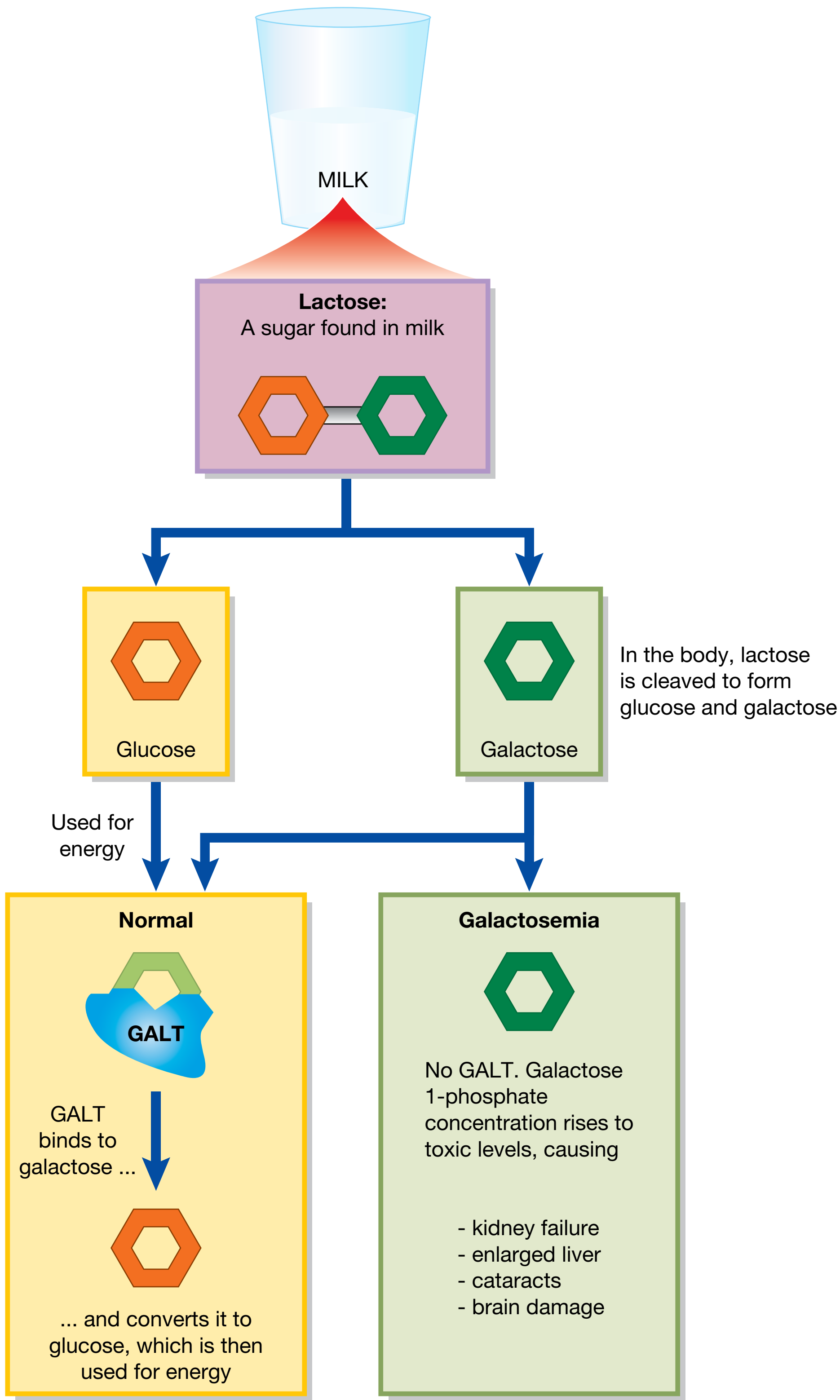



**FIGURE 14.16** Normal metabolism of the galactose sub-unit of lactose (at LH side) compared with the inherited disorder galactosemia, which is due to the absence of the enzyme GALT, encoded by a gene on chromosome 9. (GALT = Galactose-1-phosphate uridylyltransferase.)

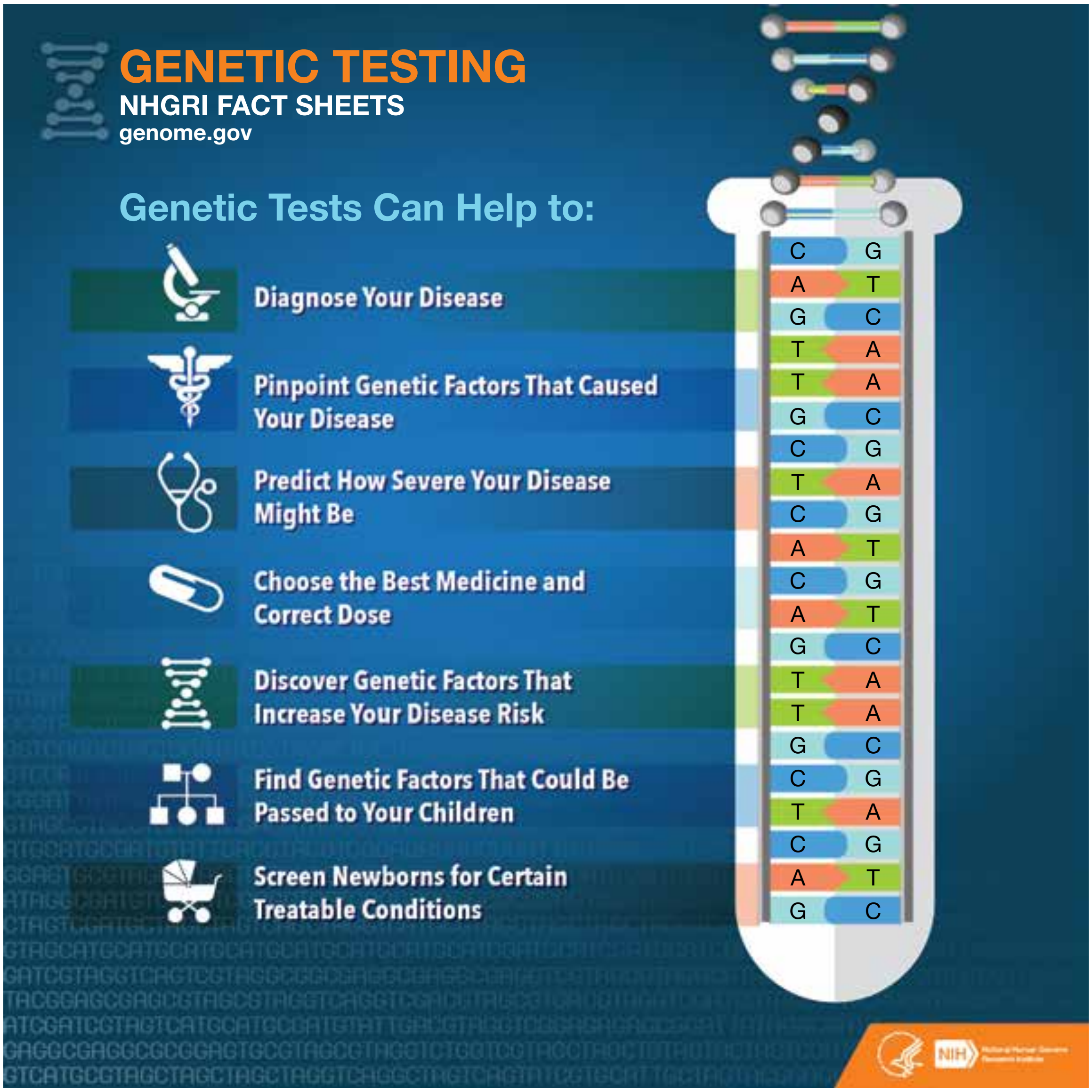



**FIGURE 14.17** Some of the many reasons an individual may seek or may be offered genetic testing. (Image courtesy of the National Human Genome Research Institute.)

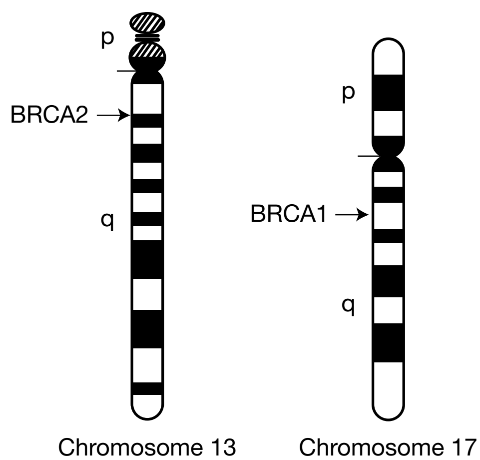



**FIGURE 14.18** Human chromosomes showing the **BRCA2** and **BRCA1** gene loci.

Genetic testing can be done for women who, based on their family medical history, may be at a significantly increased risk of breast cancer. In Ellen's case, her family background put her in the high-risk category. At the time, Ellen did not have breast cancer but the genetic screening test on a sample of her blood showed that she had inherited the mutant form of the **BRCA1** gene from her mother, which put her at a high risk of developing the breast and ovarian cancers that killed so many members of her family.

Publicity about and general awareness of the risks associated with BRCA positive gene variants reached a high level in May 2013. This was the result of the announcement by Angelina Jolie, high-profile film actress, film director and UN envoy for refugees, stating that she had both her breasts surgically removed (see figure 14.19). Jolie decided on this course of action after discovering she was BRCA1 positive, and, in light of the fact that her family history includes many female members who died of cancer, including her mother, who died of ovarian cancer aged just 56 in 2007, her grandmother, aged 45, and her great-grandmother, aged 53, her great aunt and a female cousin. In March 2015, Jolie announced that she had had her ovaries and fallopian tubes removed.

**FIGURE 14.19** Angelina Jolie on set during the filming of her first film as director. Jolie said that she wrote about her double mastectomy '*because* [she] *hope*[s] *that other women can benefit from* [her] *experience*'.

The view of some is that, because screening for breast cancer is available and reliable, a pre-emptive double mastectomy before cancer is detected is not necessarily an essential preventive measure. This view would argue that a woman could retain her breasts and undergo more frequent cancer screenings in an effort to catch cancer early, should it ever develop, and make use of the breast cancer prevention drugs tamoxifen and raloxifene. If any sign of cancer appeared, a lumpectomy could be performed in which the tumour and a small amount of surrounding tissue are removed. However, for a BRCA positive woman at high risk of breast cancer because of her family history, the decision to have both breasts removed is a personal one that she has the right to make after receiving expert advice and evaluating all the options.

There appears to be no argument that the pre-emptive surgical removal of ovaries from a BRCA positive woman after she's finished having children is a desirable preventive measure. Why? There is no reliable screening test for ovarian cancer and often, by the time the cancer recognised, it is well advanced and the prognosis is poor.

**ODD FACT**

One study reported that Jolie's public statements led to a general increase in awareness by women of issues and options relating to breast cancer and ovarian cancer. In the UK, it led to a doubling of referrals for genetic tests of breast cancer over an extended period.

## Carrier testing

**Carrier testing** is used to identify people who 'carry' a mutant allele of a gene that is linked to an inherited disorder, and to inform them of the probability of transmission of their gene defect to their children. While carriers of recessive alleles will not show clinical signs of the disorder, they can pass the faulty allele to their children. Carrier testing is usually offered to or sought by people who have a family history of a particular inherited disease and who are interested in the risk to their children.

Tom has the inherited X-linked disorder, haemophilia. Haemophilia is a condition in which factor 8, one of the factors required for normal blood clotting, is missing. Because it is an X-linked disorder, haemophilia is more common in males than in females. People with uncontrolled haemophilia have prolonged periods of bleeding after an injury, spontaneous bruising and internal bleeding.

Tom's sister, Jane, wishes to know her status for the **F8C** gene that encodes production of factor 8. She is aware that she may be a heterozygous carrier of the defective allele, with genotype $X^HX^h$, or she may be homozygous unaffected, with genotype $X^HX^H$. Heterozygous carriers of the defective allele have reduced levels of factor 8 but show no clinical signs of haemophilia. Jane is not concerned for herself but for the future children she may have, particularly sons.

Jane was tested and the result showed that she was a heterozygous carrier, genotype $X^HX^h$. This test result gave Jane the knowledge that she had a 1 in 2 chance of passing the defective allele to a son who would be affected by haemophilia. She and her partner chose to undertake prenatal testing during their first pregnancy as they did not want to have a boy who would experience the pain and difficulties that she saw her brother, Tom, experience. Happily, their first pregnancy was of a male fetus that had not inherited her X chromosome that carries the $X^h$ allele.

**eBook*plus***

**eLesson**
Autosomal recessive disorders
**eles-2462**

Carrier testing is also carried out on individuals who belong to an ethnic group who have a higher than average risk of such a disease, such as beta thalassaemia in people from southern Mediterranean countries. Beta thalassaemia is an autosomal recessive disorder, and Cyprus has one of the highest carrier rates in the world, estimated at 1 in 7. Greek–Cypriot couples who wish to marry in the Cypriot Orthodox Church must undergo premarital screening to identify whether or not they are carriers. Regardless of their test results, two people identified as carriers are still free to marry. In most countries, carrier testing for beta thalassaemia is voluntary, but premarital testing is mandated in several countries, including Iran, Saudi Arabia and Cyprus.

## Prenatal screening and diagnosis

Optional prenatal screening may be offered to a woman during the first or second trimester of her pregnancy. Such prenatal tests include blood tests, ultrasound screening (see figure 14.20), and cell-free DNA screening, which is a technique in which fetal DNA present in the mother's bloodstream is screened for chromosomal abnormalities such as Down syndrome.

Prenatal testing can indicate the genetic status of an embryo or a fetus where a specific genetic disorder is suspected to be present, or in some cases it may signal possible birth defects. It should be noted that prenatal screening tests of the types identified above are indicative but not diagnostic. If an abnormality is indicated, an invasive diagnostic test can be carried out, such as chorionic villus sampling, which takes a sample of placental villi, or amniocentesis, which takes a sample of amniotic fluid from the uterus. These invasive tests have a slight risk of a miscarriage: for amniocentesis, the probability is 0.01 to 0.5 per cent, and for chorionic villus sampling, the probability is 1.0 per cent.

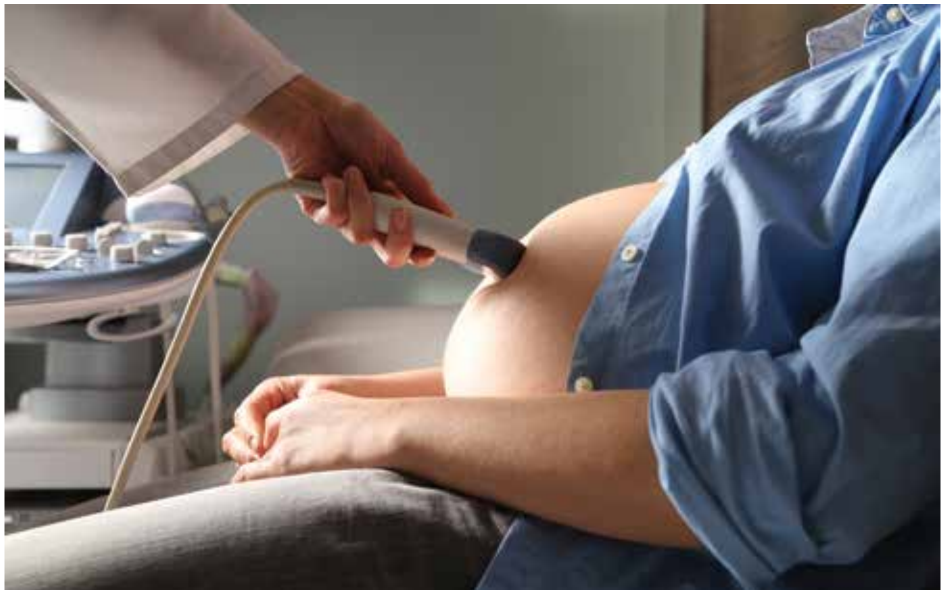


**FIGURE 14.20** Ultrasound examination of the fetus is used to screen for physical abnormalities.

Normal results from prenatal testing will be a source of relief and happiness for the parents-to-be. However, results that indicate the presence of a genetic disorder that cannot be treated, or a chromosomal abnormality that causes serious mental and physical disabilities, make parents-to-be face the difficult decision about whether or not to continue with the pregnancy. Some parents-to-be will choose not to have prenatal testing because they would wish to continue with their pregnancy regardless of any test results.

If an abnormality is detected, the result can provide parents-to-be with the information that they need to plan for the after-birth care of their affected baby. In yet other cases, prenatal testing can alert parents and medical personnel to the presence of a condition that can be treated and that requires treatment of the baby to start immediately after birth.

## Predictive or presymptomatic testing

**Predictive testing** is a form of genetic testing. It is also known as **presymptomatic testing**. This type of testing is used to detect gene mutations associated with disorders that appear after birth, often later in life. (Definition from Human Genetics of Australasia.)

Predictive testing is often requested by individuals 'at risk' of inheriting, from an affected parent, the defective ***H*** allele that encodes Huntington's disease (HD). (You first met HD in chapter 9 (see page 405) as an example of a founder effect that has produced a high incidence of HD in Tasmania, with the next highest in Victoria.)

Let's look at how predictive or presymptomatic testing is done in the case of HD.

**Huntington's disease (HD)** is a dominant inherited disease that usually first appears in an affected person in the age range of 30 to 50 years. The HD gene is located on the number-4 chromosome, and a single copy of the ***H*** allele is sufficient for the HD phenotype to appear. This progressive disorder affects the brain. People with Huntington's disease show emotional disturbances, personality changes and loss of control of motor function. At present, there is no cure for this devastating disease.

Children of a parent with Huntington's disease have a 50 : 50 chance of inheriting the defective ***H*** allele from that parent and are said to be 'at risk'. Typically, those children who inherit the ***H*** allele show the first signs of this disorder only in mid adulthood. By then, they have often started their own families and may also have transmitted the defective ***H*** allele to some of their own children. Through DNA technology, the genetic status of people at risk of Huntington's disease can be identified before they show the disease symptoms. This type of predictive testing can be carried out, not only on an at-risk adult, but also on the fetus of an at-risk person, and can even be performed on a single cell taken from an early human embryo (see embryo biopsy below).

The ***H*** allele that causes Huntington's disease can be distinguished from the normal allele (***h***) by the presence in the coding region of the gene, typically, of 40 or more copies of a trinucleotide repeat, CAG. In people who are *not* at risk of HD, the number of repeats of this trinucleotide is in the range of 6 to 35. The **HD** gene contains the instructions for the protein huntingtin. The presence of the extra trinucleotide repeats in the ***H*** allele means that this protein has additional *gln* amino acids in its structure, and this abnormal huntingtin protein produces neurological symptoms, typically in middle age.

The predictive test involves using polymerase chain reaction (PCR) to amplify the region of the **HD** gene that includes the repeats. This region of amplified DNA from each subject is separated by electrophoresis (see figure 14.21), which reveals the number of CAG repeats. The box at the end of this section outlines the current protocol for predictive testing of people at risk of inheriting the ***H*** allele, with special reference to the pre- and post-test consultations and counselling.

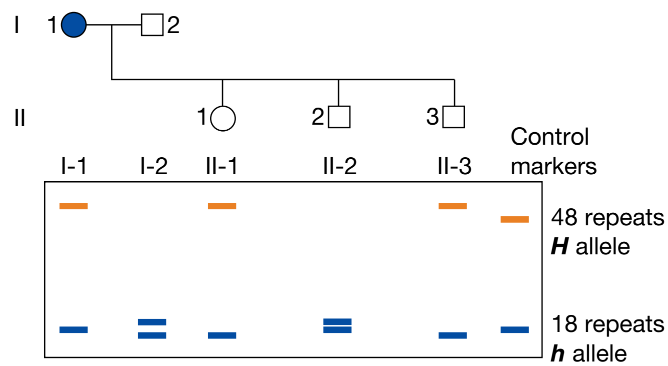



**FIGURE 14.21** A family showing segregation for the two alleles of the **HD** gene. Some members are homozygous for the normal *h* allele, and show 16 to 20 CAG repeats (shown in blue). The ***H*** allele in this family has 52 CAG repeats (shown in orange) and members with this allele include the mother (I-1), who already shows severe symptoms of HD, and two children (II-1 and II-3), who at the time of the test showed no symptoms.

Questions that may be asked include:

With whom, if anyone, should a symptomless person whose test results show that she or he possesses the ***H*** allele share the test result: a fiancé, an employer, an insurance agent, a living parent who does not show the HD symptoms?

What psychological support is available to symptomless individuals who test positive for the ***H*** allele?

Should the age limit for presymptomatic testing be lowered or left as it is?

## Embryo biopsy or pre-implantation genetic diagnosis

Embryo biopsy or pre-implantation genetic diagnosis (PGD) is a test that involves the removal of a single cell from an embryo conceived by in-vitro fertilisation (IVF) and the testing of that cell to determine whether or not the embryo will be affected by a specific inherited disease.

**Pre-implantation genetic diagnosis (PGD)** involves identifying, in the genotypes of embryos created through IVF, the presence of one or two allele(s) that cause particular inherited diseases. The purpose behind this procedure is to allow the couple involved to have a child that is free of the disease and, therefore, cannot transmit the allele responsible to the next generation. Figure 14.22 shows a simplified outline of the PGD procedure.

One inherited disease for which PGD has been used is Huntington's disease (HD). Where one partner in a couple is known to have inherited the ***H*** allele responsible for HD, the use of pre-implantation genetic diagnosis after IVF allows the couple to have a child who is free of the possibility of inheriting the ***H*** allele that causes HD.

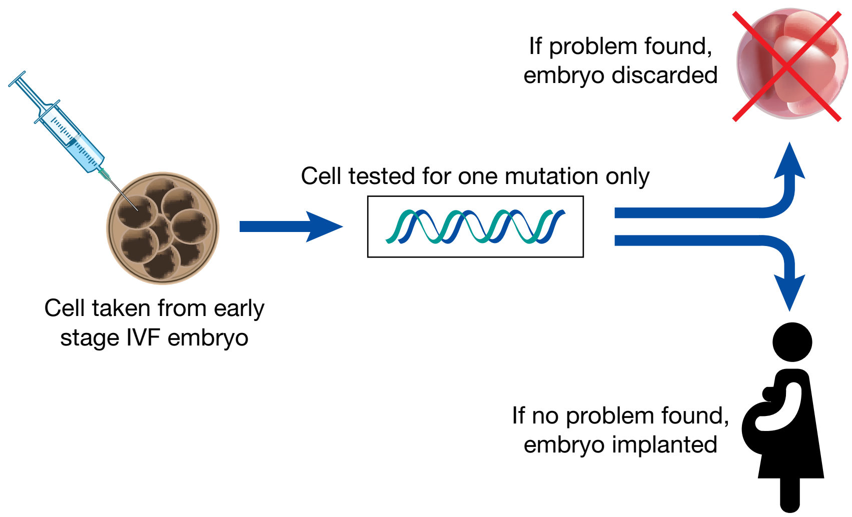



**FIGURE 14.22** Highly simplified diagram showing pre-implantation genetic diagnosis.

**ODD FACT**

The emerging field of pharmacogenetics is the study of inherited genetic differences in drug metabolic pathways that can affect a person's responses to particular drugs, both their beneficial effects as well as their adverse effects. Pharmacogenomic testing identifies why drugs work in some patients but not in others because of differences in their genetic make-up.

Other inherited conditions where PGD has been used include autosomal recessive conditions such as beta thalassaemia and cystic fibrosis, and X-linked condition such as fragile-X, Duchenne muscular dystrophy and haemophilia A.

Some people are opposed to PGD; their comments include:
'I am horrified to think of these people sitting in judgment on these embryos and saying who should live and who should die.'

'Who is going to make the decision about who should and should not live? We believe all babies have an equal right to life.'

'My only concern is that we are constantly messing around with nature.'

Others strongly support PGD:
'Any technique which has the potential to reduce the risk of serious, debilitating and potentially life-threatening disease has to be greeted with some enthusiasm.'

## Genetic testing: pros and cons

Genetic testing provides results that can be used for many different positive purposes, as for example:

- People who know from carrier testing that they are at risk of a particular cancer can ensure that their health status is regularly monitored, and they can make any relevant lifestyle changes.
- People who know from carrier testing that they can pass a defective allele to their children can plan their pregnancy and use prenatal testing or pre-implantation diagnosis to ensure that they have a baby free of the specific disease.
- Results of prenatal testing and diagnosis mean that parents have adequate time to make decisions about the future of the pregnancy.
- Presymptomatic testing may remove the anxiety and stress of not knowing whether or not a late-onset inherited disease will develop.

In contrast, genetic testing may be the source of negative outcomes, as for example:

- Presymptomatic testing may be a source of anxiety and stress, such as in the case of a person testing positive for a defective allele that is the cause of an inherited late-onset disease, such as HD, for which no treatment is presently possible.
- The result of antenatal testing and diagnosis may place an emotional burden on prospective parents faced with making a decision about the fate of a pregnancy.

- Misuse of test results, such as their disclosure to third parties, may discriminate against the person tested.
- Disclosure of information about family members who have not consented to testing would be an invasion of privacy; misuse of the information based on the test results after disclosure to third parties, such as insurers and employers, would be a form of discrimination.
- Test results may be inconclusive or difficult to interpret, creating uncertainty for the individual tested.

## PROTOCOL FOR PREDICTIVE TESTING FOR HD

Predictive testing for those at risk of developing HD is available in every state in Australia. The Northern Territory and Tasmania derive support from interstate services. A team consisting of test program coordinator/genetic counsellor, neuropsychiatrist or neurologist, and a geneticist provide specialised consultation, education and counselling support for those undergoing testing.

The person requesting testing (known as the 'consultand') must:

- be at risk of HD, with the diagnosis of HD having been confirmed in a parent (or grandparent) or sibling
- be 18 years of age or older. In rare circumstances, a younger teenager who demonstrates capacity to give informed consent may enter discussions with a view to testing
- be free from severe depression or psychotic disorder
- agree to participate in counselling sessions, including result giving (disclosure) and post-test follow-up
- preferably be accompanied by a partner or supportive friend, not another 'at-risk' person.

'Predictive testing' means testing for the presence or absence of an abnormal gene in someone who is presently showing no signs of the disease in question. Normally, there are anywhere from about 6 to 35 CAG repeats in a section of the DNA of the huntingtin gene. In HD, there are typically 40–50 repeats, occasionally more. The increased number of glutamines may in some way affect cell function, in particular populations of neurons within the brain so that, over time, through a process that is not completely understood, cell death occurs very gradually in several centres of the brain.

In the laboratory, the CAG repeat sequence of the gene is examined and, according to its length, will fall into one of three areas: the normal range, the 'grey zone' or the affected range. With a result in the normal range, the consultand can be reassured that he/she will never develop HD symptoms and his/her offspring face no risk of HD. A result in the affected range means that HD will develop in that person's lifetime. That person also carries a 50 per cent risk of passing the affected gene on to any child.

Grey zone results are complicated to interpret. The consultand with a grey zone result carries the risk of passing on an abnormal gene to their child, possibly with the CAG expansion having increased up to the size of the typical affected HD range. That is to say, the gene can expand as it is transmitted from the parent to the child — a phenomenon known as anticipation. This is particularly evident when the abnormal allele is passed on from a father to his children. Some consultands with a result in the grey zone have never developed symptoms of HD, and some have developed milder symptoms at a later age than is usual. In practice, grey zone results are not commonly encountered.

The current Predictive Testing Program in Victoria — other states differ a little — takes around six to eight weeks. The consultand is required to attend a number of sessions that usually include:

- gathering the consultand's family history or pedigree
- discussions about:
  - motivation and timing of testing (ensuring no other major stressors exist for the client at that time)
  - personal experience of HD
  - the genetic nature of HD
  - confidentiality and legal implications of testing.
- consultation with neuropsychiatrist/neurologist; this provides an opportunity for the consultand to gain further understanding of the clinical implications of HD, treatment options and research. The consultand may choose to undergo a physical examination to confirm absence of symptoms; the counsellor attends this session
- blood collection
- discussion about the impact of testing — how either a positive or negative result will impact on their life and that of those around them

*(continued)*

- reproductive options if relevant
- disclosure of the result
- post-test follow-up — initially with the counsellor and, in Victoria, longer term by staff of the Huntington Disease Association.

The consultand is free to initiate contact with the counsellor at any time following the program or to withdraw at any time during it.

In summary, the Predictive Testing Program provides information about HD and the gene test. Counselling provides opportunities to explore in a supportive climate the implications of testing for the consultand's relationships, especially within the family context, at the time of the test process, and projecting into the future. Consideration is also given to legal, employment and other long-term ramifications of the gene test.

**Reproductive choices**

Many couples who have the HD genetic mutation have children without any genetic testing of the fetus. However, there are some choices for couples who do not want to transmit the risk of the illness to their children. Options are best explored before a pregnancy occurs, and supportive counselling and care are provided, whichever course a couple chooses.

Prenatal testing involves chorionic villus sampling (CVS) at 11 weeks gestation. This test is performed under ultrasound guidance, and tissue from the developing placenta is sampled. Chorionic villi develop from the conceptus and are fetal in origin, so these cells represent the fetal genotype. A direct DNA test is performed on these cells. If the result shows that the fetus has inherited the HD mutation, the parents may request termination of the pregnancy.

Pre-implantation genetic diagnosis (PGD) uses in-vitro fertilisation (IVF) techniques combined with genetic analysis of the embryo. IVF can now be accessed by a couple where one is known to have (or be at risk of having) the HD mutation. The embryo is created in the laboratory with the ova and sperm from the couple. Approximately 48 hours after conception, a single cell is drawn from the developing embryo and genetic analysis is performed. Only an embryo without the mutation is transferred to the woman for implantation and development into a baby. PGD is a complex process with many technical and medical considerations.

Other options for couples may include use of donor sperm and donor embryo. In Victoria, these services can be accessed directly through the Royal Women's Hospital or Monash Medical Centre, or by inquiring through Genetic Health Services Victoria. Similar services are available in other states. Couples with a family history of HD are not eligible to adopt.

**KEY IDEAS**

- Genetic screening of populations or groups within a population is aimed at identifying individuals who have specific genetic or congenital disorders for which treatment exists.
- Genetic testing involves testing for the presence of a faulty allele that is the cause of an inherited human disorder.
- Genetic testing is carried out for various purposes and can use embryonic, fetal or adult cells as a source of DNA.
- Use of genetic testing raises ethical issues.

**QUICK CHECK**

**7** Which of the following statements best describes genetic screening?
- **a** Genetic screening cures genetic disorders.
- **b** Genetic screening detects genetic disorders.
- **c** Genetic screening detects the presence of alleles that produce genetic disorders.

**8** How does an embryo biopsy differ from prenatal screening?

**9** Identify the following statements as true or false:
- **a** Population screening is for genetic disorders that are incurable.
- **b** Population screening requires a testing procedure that is reliable, has proven predictive value and has a minimal rate of false positive results.
- **c** Presymptomatic testing involves identifying the carrier status of an individual.
- **d** A woman who is BRCA1 positive will develop breast and ovarian cancer.

# Use of DNA in identification

**Each of us is genetically unique, with the exception of identical (monozygotic) siblings.** While phenotypic differences are apparent among us, the most fundamental expression of our uniqueness is in our genetic material, DNA. Headlines (see figure 14.23) remind us of this uniqueness.

Today, individuals can be identified through a technique known as **DNA profiling**. The amount of DNA needed for DNA profiling is very small because DNA can be amplified through the polymerase chain reaction (PCR) (refer back to Amplifying traces of DNA on page 627 in chapter 13). One person's DNA profile is constant, regardless of the type of cell used to prepare the profile. A DNA profile prepared from a person's white blood cells is identical to that prepared from the same person's skin cells or other somatic cells.

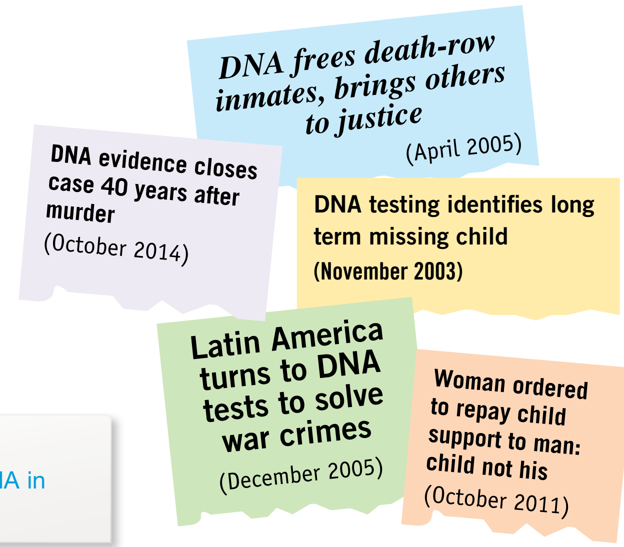



**FIGURE 14.23** Headlines signalling the power of DNA in identification.

## DNA identification helps solve crimes

Because DNA molecules are only slowly degraded, DNA profiling can be carried out on biological samples from crime scenes from years ago, and this profiling has led to the solution of many 'cold cases' worldwide. For example, in 1980 in Tasmania, the body of a 22-year-old female bank clerk was recovered from the Derwent River. On several occasions, police interviewed a cab driver, Gerald Hyland, because bloodstains were found in his cab and these were identified as having come from two different people. While one blood sample corresponded to the blood type of the bank clerk, the tests available at that time could not match this blood exclusively to her. Hyland refused to give a blood sample and, because he was just a suspect, he could not be compelled by law to do so. Another stain found on the back seat of the cab was identified as semen, but, at that time, there was no way of matching this semen to one particular person. Because of the lack of conclusive evidence, Hyland could not be charged.

Thirteen years later, in 1993, the preserved blood samples from the cab were retested using new DNA technology and one sample was shown to have come from the murdered bank clerk. Based on this evidence, Gerald Hyland was arrested and a blood sample from Hyland was then able to be taken. Through examination of Hyland's blood sample using DNA technology, the semen and the second blood sample from the cab were matched exclusively to him. Hyland pleaded guilty to the rape and manslaughter of the bank clerk.

**ODD FACT**

The oldest 'cold case' solved using DNA profiling was a murder in 1968. Semen traces from the crime scene were kept on file and, in 2003, were used to produce a DNA profile. This led to the arrest of a suspect, William Speer. In November 2005, he was found guilty of murder.

In 1973, two teenage girls were murdered near Swansea, Wales. A suspect in this case was Joe Kappen, a nightclub doorman, who died in 1990. Later, when the technology became available, the killer's DNA profile was established using DNA samples kept from the crime scene. Relatives of Kappen provided samples from which scientists were able to infer Kappen's DNA profile and it was found to match the killer's profile. This gave police grounds for gaining permission to exhume Kappen's body for DNA profiling. In June 2002, police announced that Kappen's DNA profile matched that from the murder scene and this case was closed 29 years after the murders.

In December 2005 in the Northern Territory, Bradley J. Murdoch was found guilty of the murder in July 2001 of British tourist Peter Falconio. Critical evidence presented at Murdoch's trial included DNA analysis of a bloodstain found on the sleeve of a T-shirt worn at the time by Joanne Lees, Mr Falconio's travelling companion (see figure 14.24). The DNA profile of this blood matched that of Murdoch. Blood samples from a pool of blood found on the Stuart Highway and from nearby rocks were found to match Peter Falconio's DNA profile. His body has not been found.

**FIGURE 14.24** Senior biologist Kamen Eckhoff (at left) holds up the T-shirt that Joanne Lees was wearing when she was abducted; at right, Assistant Police Commissioner John Daulby. In the background are images associated with the case, including a photofit image of the suspect, a mock-up of the utility that the suspect was driving and the plastic ties used by Murdoch to bind the wrists of Joanne Lees. (Image courtesy of Newspix.)

DNA evidence makes it more likely that a person guilty of a crime will confess and plead guilty. In the town of Wee Waa in outback New South Wales in January 1999, an elderly woman was bashed and sexually assaulted. A police investigation continued for more than a year, but no arrest was made. Fifteen months after the attack, police asked local men aged between 18 and 45 to volunteer to give a saliva swab for DNA testing. Samples were taken from about 500 men. However, before the DNA samples were processed, a local man went to the police and confessed to this crime.

DNA evidence can clear innocent people. Kirk Bloodwood, convicted of murder in 1984 in the United States, was the first person to be released from death row on the basis of post-conviction DNA testing, in 1993.

**ODD FACT**

Several years after the release from prison of Kirk Bloodwood on DNA evidence, the DNA profile from this murder was matched to another person, who pleaded guilty in May 2004.

## The lost soldiers of the battle of Fromelles

Identification using DNA is not always used in criminal investigations. In some cases, DNA is used for identification where the purpose is to give closure to families and the knowledge that a missing chapter of their family history has been made known. In these cases, mitochondrial DNA (mtDNA) is often used.

DNA samples recovered from the remains of 250 Allied soldiers found near Fromelles in France in 2008 became part of the evidence that enabled their identification. Many of these remains were those of Australian soldiers who were killed in the Battle of Fromelles in August 1916 and buried by German troops in unmarked mass graves behind German front lines. It was the detective work of an amateur WWI historian, Lambis Englezos, a Greek-born Melbourne art teacher who gathered the evidence, that led to the discovery of the mass graves.

In May 2008, six mass graves were partially excavated and conclusive evidence of human remains was found. Artefacts, such as badges and uniform buttons, indicated that the remains were those of Australian and British soldiers. In May 2009, excavation of the mass graves began and, within a few months, the remains of 250 soldiers had been recovered. In January 2010, reburials of the remains began at the new Fromelles Military Cemetery (refer back to figure 14.1). The remains of 219 Australian soldiers, 75 yet to be identified, have been reburied with full military honours (see figure 14.25). The remains of each soldier now lie in an individual grave with a headstone that, in most cases, bears the name of the soldier. The process of identification continues with six more Australian soldiers identified in 2016.

**FIGURE 14.25** A solemn military procession transporting the remains of a casualty of the disastrous Battle of Fromelles (1916), for re-interment to a final resting place at the new Fromelles Military Cemetery.

Examination of ancient DNA has also been used to answer questions in many areas, including British history, diseases of ancient Egyptians, and human ancestry. Ancient DNA is DNA isolated from long-dead objects, such as mummies, museum specimens and long-buried bones. For example, the mtDNA extracted in 2013 from a skeleton buried under a carpark in Leicester, England, assisted in confirming that the skeleton was that of Richard III, King of England, who died in 1485 (see figure 14.26). Even older samples of DNA have been obtained from fossils of Neanderthal humans who lived tens of thousands years ago (refer back to figure 12.63 on page 589).

**FIGURE 14.26** Model of the bones of Richard III in front of an MRI machine. Magnetic resonance imaging was one of the techniques used to examine the original bones. (Image courtesy of Judith Kinnear.)

In summary, identification using DNA is a powerful tool that can be applied in many situations, including:

- forensic investigations
  - Can the DNA found at a crime scene be matched to a person on the national DNA database?
  - Is this blood spot from the victim or from the possible assailant?
  - In a rape case, is this semen from a previously convicted rapist?
- mass disasters, such as passenger aircraft crashes, the 9/11 terrorist attacks, the Bali bombings
  - Can the various remains that have been recovered be matched to a particular person known to have been at the site of the disaster?
- identification of human remains
  - Are these remains those of a particular missing person?
  - Who was the unknown child, tagged as body number 4, recovered after the sinking of the RMS *Titanic* in 1912?

## Which DNA is used for identification?

Depending on the purpose and circumstances of the identification, the DNA used comes from either the chromosomes (nuclear DNA) or from mitochondria (mtDNA). Some important differences between nuclear DNA and mtDNA include:

1. mtDNA is exclusively inherited via the maternal line — that is, it is solely matrilineal; nuclear DNA is inherited from both parents.
2. mtDNA does not undergo recombination and any changes in mtDNA must arise by mutation; nuclear DNA undergoes recombination, meaning that it is shuffled around in each generation by crossing over during meiosis.
3. mtDNA in each somatic cell is present in high numbers of copies; nuclear DNA is present as two copies only. For this reason, mtDNA is the DNA that can be recovered from ancient specimens, such as bones and teeth.

In both nuclear DNA and mtDNA, however, their use in identification depends on the existence of segments of DNA that vary greatly between individuals. Such regions of DNA are termed hypervariable.

**ODD FACT**

In 2001, mtDNA was recovered from teeth and bone taken from the grave of an unknown child (body number 4) who died when RMS *Titanic* sank in 1912. Comparisons with females from various families who were possible relatives revealed that the unknown child was 13-month-old Eino Panula from Finland, who died along with his mother and four brothers.

**ODD FACT**

Even seemingly unrelated people can have the same mtDNA profiles because they shared the same female ancestor several generations ago.

Hypervariable regions of DNA that are currently used for identification are:

- **short tandem repeats** (STRs) in the nuclear DNA, also known as microsatellites. A large number of STRs are present on different human chromosomes. On its own, DNA from STRs can identify one person uniquely (apart from identical siblings). DNA samples from relatives are *not* required. So, when there is a need to match a DNA sample from a crime scene to just one particular person, STRs are used.
- **hypervariable regions** (HVRs) in the noncoding region of mtDNA. It should be noted that mtDNA identification is less precise because individuals from the same maternal line have identical mtDNA profiles. mtDNA is used, for example, when nuclear DNA cannot be recovered or when nuclear DNA is degraded because of age.

Identification using mtDNA is mainly applied either (1) to identify victims of mass disasters where the names of the victims are known but where identification of the remains by conventional means, such as visual inspection or dental records, cannot be done or (2) to identify decomposed remains when the identity is thought to be one of a few specific missing individuals. In 2009, mtDNA was important in identifying the Australian soldiers killed at Fromelles in 1916. **In cases of identification using mtDNA, there must be living relatives on the maternal line to provide mtDNA for comparison with the mtDNA from the remains.**

## First there was DNA fingerprinting

The original technique of identification using DNA was called **DNA fingerprinting**. DNA fingerprinting was developed as an identification tool in 1985 by Professor Sir Alec Jeffreys (see figure 14.27). This technique used hypervariable regions of DNA, known as minisatellites, that are located near the ends of chromosomes. Minisatellites are chromosomal regions where sequences of 9 to 80 base pairs are repeated tens or hundreds of times.

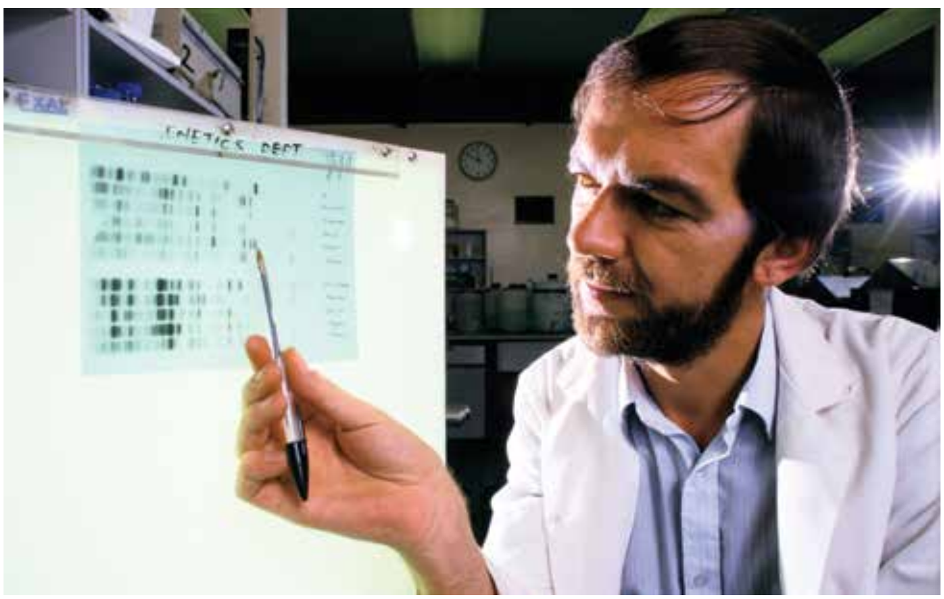


**FIGURE 14.27** Professor Sir Alec Jeffreys developed the original technique of DNA fingerprinting, based on hypervariable regions of DNA known as minisatellites. Sir Alec is seen here examining an X-ray film that shows the DNA fingerprints of various individuals. Because of variations in the numbers of repeats in different minisatellites, each individual has a unique pattern that appears as a series of bands in a vertical column. This pattern is the so-called DNA fingerprint.

The DNA fingerprinting technique was first used in a criminal investigation in England in 1986. The case involved the rape and murder in November 1983 of a 15-year-old schoolgirl in the village of Narborough, England. Almost three years later, a rape and murder of a second teenage girl occurred in a nearby village. At that time, semen stains from both victims could reveal only that the murderer had blood group A and had one of 10 variants of a particular enzyme — a description that fitted about 10 per cent of the male population! This finding, however, suggested that the same assailant was probably involved in both cases.

Police asked Professor Jeffreys to use his DNA fingerprinting technique to assist in the investigation. Blood and saliva samples were collected on a voluntary basis from 4583 men aged between 16 and 34 living in and around the two villages. DNA fingerprints were obtained only from those samples with the same blood type and enzyme status as the murderer. However, none of these DNA fingerprints, including one from a disturbed young hospital porter who confessed to the second murder, matched that of the scene-of-crime DNA.

Police were later informed that a man was overheard in a hotel boasting that he gave a blood sample for a friend. When questioned, this man admitted that he called himself Colin Pitchfork when he gave his blood sample. Pitchfork was arrested, a blood sample was taken from him and his DNA fingerprint was found to match that of the semen recovered from the crime scene. Pitchfork confessed, pleaded guilty to the murders and, in January 1988, was sentenced to life imprisonment. This was the first time that the genetic evidence of DNA fingerprinting was used to assist in convicting a murder suspect, and also in exonerating the hospital porter who confessed to the crime.

**ODD FACT**

In Australia, DNA evidence was first used in 1989 in a court case in Canberra. The defendant, Desmond Applebee, who was charged with sexual assault, first claimed no involvement. When evidence was given that DNA from semen found on the victims matched his DNA profile, although Applebee then claimed consensual sex, he was convicted.

## Then came DNA profiling

During the late 1980s and early 1990s, two breakthroughs occurred that led to the development of DNA profiling, which is now the technique of choice for identification today:

1. The minisatellites used by Sir Alex Jeffreys were replaced by the use of a different class of hypervariable DNA segments termed **short tandem repeats (STRs)**.
2. New and highly sensitive methods of visualising DNA were developed that used fluorescent labelling of DNA.

These two developments, combined with the use of the polymerase chain reaction (PCR) to amplify DNA, produced the powerful technique of DNA profiling that has applications in so many areas, including forensics and medicine.

**ODD FACT**

Because of PCR, even the handling of a mobile phone by a person can generate sufficient material from which that person's DNA profile can be produced.

### What is an STR?

STRs are hypervariable regions of chromosomes where sequences of just two to five base pairs are repeated over and over. These regions are very common and hundreds are scattered throughout the human chromosomes. STRs are termed 'short' because the repeat sequences are only two to five base pairs long, and 'tandem' because the repeats occur one after the other (see figure 14.28). However, the number of repeats of an STR marker can vary between people and each variation is a distinct allele. The number of repeats of a 4-base pair sequence of one STR marker on the number-5 chromosome ranges from 7 to 15. In most cases, the alleles at an STR locus on a human chromosome are named according to the number of repeats and so are identified as allele 7, allele 8 and so on.

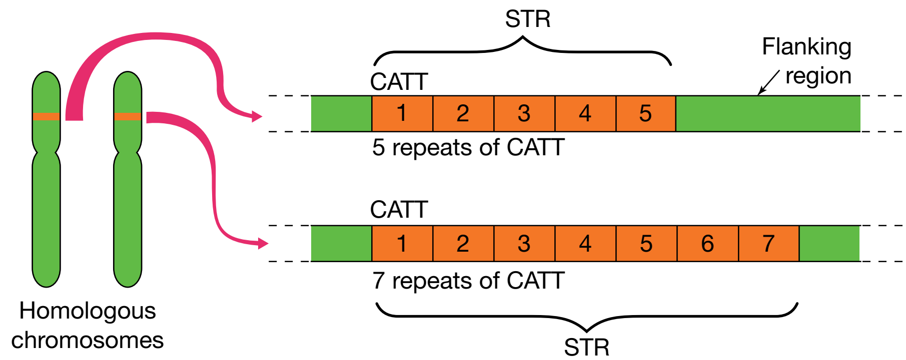
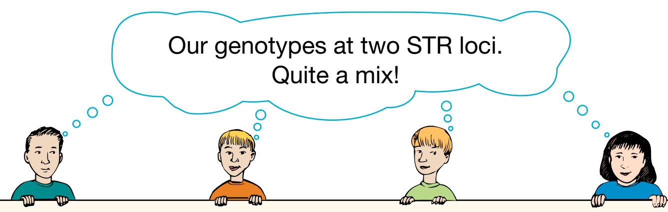



**FIGURE 14.28** Diagrammatic representation of an STR showing one allele with five repeats of a four-base sequence and the other allele with seven repeats on homologous chromosomes. How would these alleles be named? While the STR regions are variable, the flanking regions on either side of the STR are constant.

At each STR locus, an individual is either homozygous or heterozygous and so can have a maximum of just two different alleles. These alleles are inherited in a Mendelian fashion, and are co-dominant; that is, where two different alleles of the same STR marker are present, both can be detected. Figure 14.28 shows a person who is heterozygous 5/7 at one particular STR locus and has one allele with 5 repeats and the other allele with 7 repeats. Within the gene pool of a population, however, many different alleles can exist at each STR locus (see figure 14.29).




| | | | | |
|---|---|---|---|---|
| STR locus 1 | 7, 10 | 9, 15 | 11, 16 | 10, 12 |
| STR locus 2 | 4, 8 | 6, 7 | 5, 9 | 7, 10 |

**FIGURE 14.29** Alleles at the same STR loci can vary between different people. What is the maximum number of different alleles at one STR locus that can exist in one person? In a group of four people?

Note that sometimes a repeat of a 4-base STR is not complete and may be missing the final base or the final two bases. In that case, an allele might be identified as 9.2 or 30.3. The notation 9.2 means that this allele has 9 repeats followed by the first 2 bases of the next repeat. Likewise, 30.3 denotes an allele with 30 repeats plus the first 3 bases of the next repeat.

Table 14.3 shows the frequencies observed in three different ethnic groups in the Australian population for the alleles at one particular STR locus on the number-5 chromosome, known as D5S818. The number of repeats of this 4-base STR marker ranges from 7 to 15. Note that the allele frequencies vary *within* a population, with allele 11 being far more common than allele 15. Note also that the frequencies vary *between* populations, with allele 7 being about 20 times more common in Asian populations than in the other two populations.

**TABLE 14.3** Frequencies of the alleles at the D5 STR locus on the number-5 chromosome for three sample populations in Australia. Which allele has the lowest number of base pairs? For a person who is homozygous 13/13 at this locus, how many repeats of the STR are present on each of her number-5 chromosomes?

| | Frequency in sample populations | | |
|---|---|---|---|
| **Alleles at D5 STR locus** | **Aboriginal** | **Asian** | **Caucasian** |
| allele 7 | **0.0010** | **0.0244** | **0.0010** |
| allele 8 | **0.0016** | **0.0014** | **0.0021** |
| allele 9 | **0.0683** | **0.0551** | **0.0364** |
| allele 10 | **0.2107** | **0.2252** | **0.0644** |
| allele 11 | **0.2903** | **0.2999** | **0.3739** |
| allele 12 | **0.2862** | **0.2343** | **0.3547** |
| allele 13 | **0.1267** | **0.1499** | **0.1552** |
| allele 14 | **0.0139** | **0.0091** | **0.0104** |
| allele 15 | **0.0013** | **0.0007** | **0.0019** |

(*Source:* Data from Weir, B.S., September 2004, 'Matching and partially matching DNA profiles,' *Journal of Forensic Science,* Vol. 49, Issue 5.)

## Why use STRs rather than minisatellites?

Compared with DNA fingerprinting, DNA profiling:

- is far more sensitive and requires smaller quantities of DNA (even a pinhead-sized spot of blood can provide sufficient DNA) and the STRs can be amplified by the polymerase chain reaction (PCR)
- is based on alleles whose sizes allow fragments differing by just one base pair to be distinguished
- is carried out in a much shorter time — hours rather than days
- uses several single-locus probes rather than one multilocus probe as occurs in DNA fingerprinting
- uses coloured fluorescent labels to visualise the STRs, rather than radioactive labels, so that each different STR allele can be identified by colour as well as by size
- produces less complex patterns that are more easily interpreted.

In addition, unlike minisatellites, population data on allele frequencies of STR alleles can be obtained.

In the following section, we will examine the technique of DNA profiling.

### KEY IDEAS

- Apart from the DNA of identical siblings, the DNA of each person is unique.
- Depending on the purpose, identification can involve either chromosomal (nuclear) DNA or mitochondrial DNA.
- Hypervariable regions present in DNA include minisatellites and short tandem repeats (STRs).
- Minisatellites were the basis for identification of the now superseded technique known as DNA fingerprinting.
- The current technique of identification using DNA is known as DNA profiling and is based on STRs.

QUICK CHECK

**10** Identify the following statements as true or false:
  **a** The current technique of identification by DNA is known as DNA fingerprinting.
  **b** A minisatellite is also known as a short tandem repeat.
  **c** The length of the repeated unit in an STR is from two to five base pairs.
  **d** Chromosomal DNA gives more precise identification than mtDNA.
  **e** One person could have 10 different alleles at one STR locus.
**11** List two advantages of DNA profiling as compared to DNA fingerprinting.
**12** Who invented the technique of DNA fingerprinting?

# DNA profiling in Australia

Where DNA profiling is carried out for forensic investigations and the results are to be used in court, the biological samples must be collected according to a strict protocol. Laboratories handling and testing the samples must use standardised procedures under controlled conditions. Samples for DNA analysis must be safeguarded and accounted for at every point, from their collection at the scene of a crime to their delivery at the laboratory. Why is this unbroken chain of custody important? Only certain laboratories are accredited to carry out DNA profiling for forensic purposes.

From 1995 up to late 2012, all Australian states used a common method of DNA profiling for forensic purposes that involved nine STRs from different human chromosomes. These STR markers were chosen because they are reproducible, robust, easy to score and are highly informative, and have low mutation rates. In addition, a tenth marker (that is not an STR) was used to identify the gender of the individual. This gender marker is the *Amel* locus, which is present on both the X chromosome and the Y chromosome. The Amel gene on the X chromosome is just 107 base pairs long while that on the Y chromosome contains 113 base pairs. As a result, the gender of a person can be identified from this marker.

Table 14.4 shows the nine STR loci used prior to 2013 for DNA profiling in Australia and the *Amel* locus used as a gender identifier. Figure 14.30 shows the size range of all the alleles of each STR that can be present in human populations. Where an overlap occurs in the size ranges of the alleles of different STRs, they can be distinguished on the basis of coloured fluorescent labels.

**ODD FACT**

In Victoria, DNA profiling is carried out at the Victoria Police Forensic Services Department at Macleod.

**TABLE 14.4** The nine STR markers used in DNA profiling in Australia prior to 2013. For simplicity, STR markers that start with the letter D are identified by their chromosomal location only. In fact, the naming of STR markers is more complex; for example, the full designation of D13 is D13S317 and D7 is D7S820.

| STR locus | Chromosome location | Number of repeats of 4-base sequence | Number of alleles* |
|---|---|---|---|
| **D3** | 3 | 8–20 | 11 |
| **vWA** | 12 | 10–25 | 12 |
| **FGA** | 4 | 12–51 | 27 |
| **D8** | 8 | 7–19 | 11 |
| **D21** | 21 | 24–38 | 30 |
| **D18** | 18 | 7–27 | 21 |

**ODD FACT**

In 1997, the United States identified 13 STRs and the *Amel* gender marker for use in forensic DNA profiling. In 2011, the FBI proposed that this be expanded to a set of 20 STRs and the *Amel* gender marker. This has now happened.

**TABLE 14.4** (*continued*)

| STR locus | Chromosome location | Number of repeats of 4-base sequence | Number of alleles* |
|---|---|---|---|
| **D5** | 5 | 7–16 | 9 |
| **D13** | 13 | 5–16 | 10 |
| **D7** | 7 | 5–15 | 14 |
| ***Gender marker*** **Amel** | X<br>Y | | 1 (size 107)<br>1 (size 113) |

*Number of alleles detected in Caucasian Australian population (data from B.S. Weir, op. cit.). In fact, other alleles also exist in other populations.

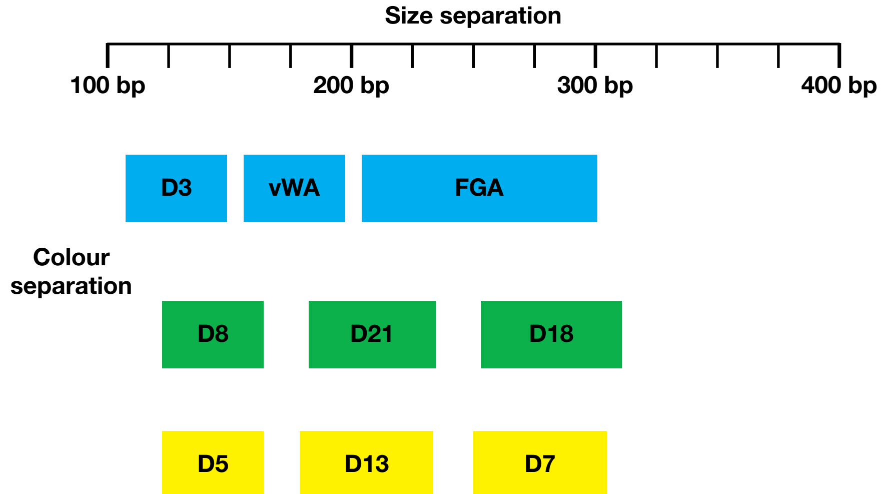



**FIGURE 14.30** Alleles of the different STR markers used in DNA profiling differ in their sizes and can also be distinguished by the colours of the fluorescent labels that they carry. Since late 2012, DNA profiling laboratories in Australia are using an expanded set of 20 STRs and the *Amel* gender marker to generate DNA profiles.

To produce a DNA profile, the DNA of the STR markers is amplified using the polymerase chain reaction (refer back to Amplifying traces of DNA, page 627 in chapter 13), and the various alleles of each STR are then separated and made visible with fluorescent dyes.

The resulting DNA profile is a series of coloured peaks at different locations, with each peak being one allele of one specific STR. The location of each peak indicates the size of the allele and hence the number of repeats. Where sizes overlap, alleles of different STRs are distinguished by fluorescent labels of different colours. A person shows either one or two peaks at each STR locus where a peak corresponds to an allele, depending on whether the individual is homozygous or heterozygous at that locus. For the *Amel* gender marker, if just a single peak with a size of 107 base pairs appears on the profile, the person is female; if two peaks are detected, one at 107 and the second at 113 base pairs, then the person is male.

Figure 14.31a shows the DNA profile of an individual based on nine STR loci and the gender marker. What is the gender of this person? Figure 14.31b shows a scientist examining a DNA profile on a computer screen.

(a)

(b)

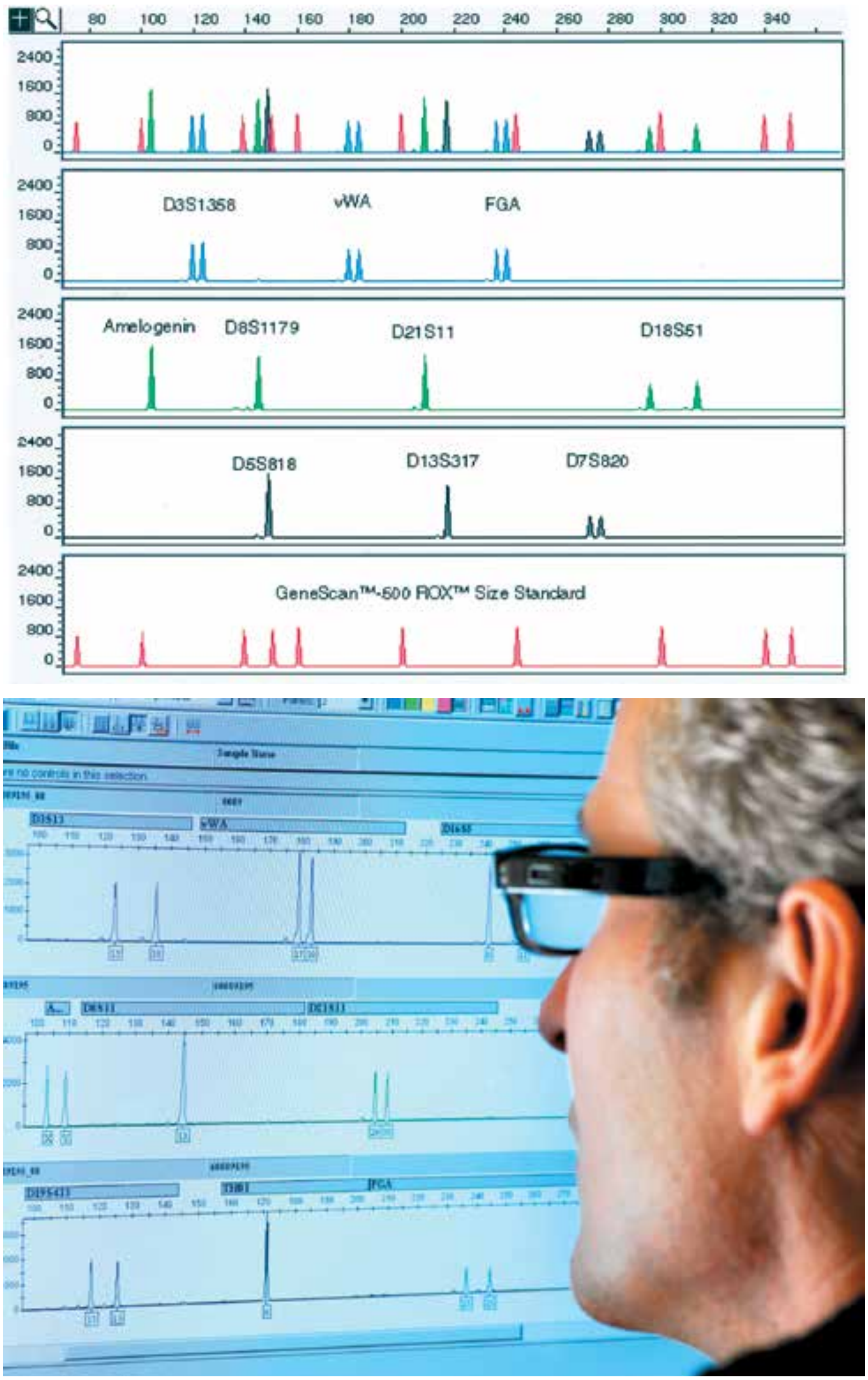



**FIGURE 14.31** **(a)** DNA profiles of a person that show the results for nine STR markers plus the *Amel* gender marker. The horizontal scale at the top identifies the sizes of the various STR alleles. The top row of peaks is the combined DNA profile. The lower rows show the alleles at the various STR loci that are distinguished by different fluorescent labels, blue, green and yellow (for clarity, the yellow is shown here as black). (Image (a) courtesy of Applied Biosystems.) **(b)** A scientist examining a DNA profile. What is the maximum number of peaks that could appear in the DNA profile of one person, based on the nine STR markers shown in **(a)**? What is the minimum number?

The expanded DNA profiling system operating in Australia since 2013 consists of 18 STRs and the *Amel* gender marker, and generates DNA profiles, each comprising a maximum of 40 peaks plus the gender indicator. (Compare this with the number of peaks in figure 14.31a.) This expanded system greatly increases the discriminatory power of DNA profiling, reducing the chance of any mismatches.

These STR markers are included among the 24 STR markers shown in figure 14.32. This expanded set includes additional STR markers that are used worldwide to construct databases.

The box at the end of this section introduces a forensic scientist who works at the Victorian Institute of Forensic Medicine.

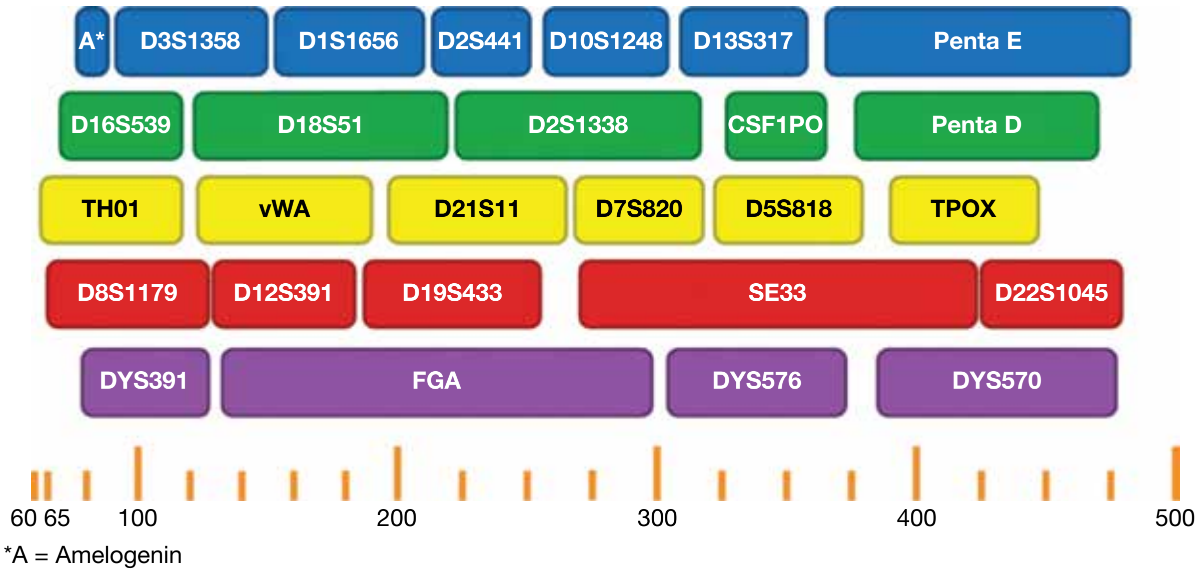



**FIGURE 14.32** An expanded STR DNA profiling system consisting of 24 STR markers and an Amelogenin gender identifier. This system includes the 18 STR markers used in Australia. '(Image courtesy of Promega Corporation 2016, Madison, Wisconsin, USA.)

## STR loci generate many different genotypes (profiles)

For one gene locus with n different alleles, the number of different genotypes possible is:

$[n \times (n + 1)] / 2$. This means that an STR locus with 7 different alleles can produce $[7 \times 6] / 2 = 21$ different genotypes.

So, an STR locus with 14 alleles can produce 105 different genotypes in a population, and a different STR locus with 9 alleles can have 45 different genotypes (see note below). Together, these two STR loci produce $105 \times 45 = 4725$ different genotypes.

As the number of STR loci increases, the number of different genotypic combinations in a population increases enormously. As a result, a DNA profile based on 18 STRs is a unique combination that allows a person to be identified with an incredibly high level of probability. The chance that the DNA profile of one person is identical to that of another person (except for an identical sibling) is one in many billions.

## What is meant by a DNA profile match?

In a criminal investigation, there are two DNA profiles to be considered: the DNA profile from the crime scene and that of the person who has been charged with the crime. What is the chance that these two profiles could match purely by chance, meaning that an innocent person has been accused of the crime?

In other words, what is the probability of a match between the DNA from the crime scene and the DNA of some unknown person, unrelated to the accused?

Using profiles based on 10 STR markers, the UK Forensic Service answers this question as follows: If the DNA from the crime scene came from some unknown person, unrelated to the defendant, the probability of a match would be of the order of one in a billion.

### WHAT'S MY DNA PROFILE?

Most people are interested in their family histories and the names and places of birth of their various ancestors. Today, many people are also interested in their genetic history as expressed in their personal DNA profiles. Commercial firms exist that can provide people with their DNA profiles for their personal interest or for other purposes.

A DNA identification (ID) card identifies a person not only by name and date of birth, but more significantly by the right thumbprint and a unique DNA profile based on a set of STRs, commonly nine, but sometimes more. STR alleles are inherited in a Mendelian fashion and so, at each STR locus, the person will have inherited one of the alleles from their mother and one from their father. While the individual's appearance will change as she or he ages, their physical fingerprint pattern and DNA profile will remain unchanged.

## National DNA databanks

The power of DNA profiling in forensic investigations has led to the setting up of national DNA databanks in many countries. These databases hold the DNA profiles of people convicted of serious crimes and the DNA profiles from scenes of crime. This resource can be used as evidence to link people to scenes of crime, to link one crime scene to the scene of another crime, to clear innocent people and to assist in convicting the guilty.

Perhaps the most spectacular conviction that resulted from the use of national DNA databases relates to the rape and murder, in a youth hostel in Brittany, France, of a 13-year-old British schoolgirl, Caroline Dickinson, on a school excursion in July 1996. DNA samples from her five male school companions, two male teachers and the coach driver excluded them, and the case remained unsolved. One possible suspect was Francisco Montes, formerly from Spain, but he had disappeared. In 2001, an immigration official in the United States read about this case and discovered that Montes had immigrated to America. Just a few weeks earlier, Montes had been arrested for trying to break into a youth hostel in Florida. Because of this, his DNA profile was on file in the American national DNA database (CODIS) and his DNA profile was found to match that from the murder scene in Brittany. Montes was arrested, extradited to France and, in June 2004, was sentenced to 30 years imprisonment for Caroline's murder.

Some national DNA databases have been established as follows:

- April 1995: The UK DNA database (NDNAD) was the first National DNA Databank to be set up in the world. As at October 2013, NDNAD held the DNA profiles of more than 5.8 million people plus more than 400 000 unknown scene-of-crime profiles. DNA profiles are held in the database as a series of numbers and a gender identifier, either XX or XY, as, for example:

  **14,18; 30,31.2; 16,17; 13,14; 28,11; 12,14; 19,23; 6,7; 12,14; 21,23; . . . . X,Y**

- October 1998: The Federal Bureau of Investigation (FBI) in the United States set up its **Co**mbined **D**NA **I**ndex **S**ystem (CODIS). As at December 2011, this national database held more than 110 million profiles.

- July 2000: Australia's CrimTrac database set up by the federal government included a national criminal DNA database, a national automated fingerprint identification system (NAFIS) and a national child sex offender system. In July 2016, CrimTrac ceased to exist and its functions were absorbed by the newly established Australian Criminal Intelligence Commission (ACIC), which has a broader scope of operation involving investigative, research and information delivery services (see figure 14.33).

In Australia, the collection of DNA samples is done in accordance with relevant state and territory legislation. DNA profiles are prepared using the agreed standard STR loci at accredited forensic laboratories. DNA profiles are then supplied to the National Criminal Investigation DNA Database (NCIDD) where they are held in digital form. When a DNA profile is supplied to the NCIDD, the name of the person from whom the sample comes does not form part of the data — this is one of the privacy safeguards. Only the state police authorities know the names of the people who provide the DNA profiles. So, when a match is made, details are returned to the police service concerned, where the identification is made.

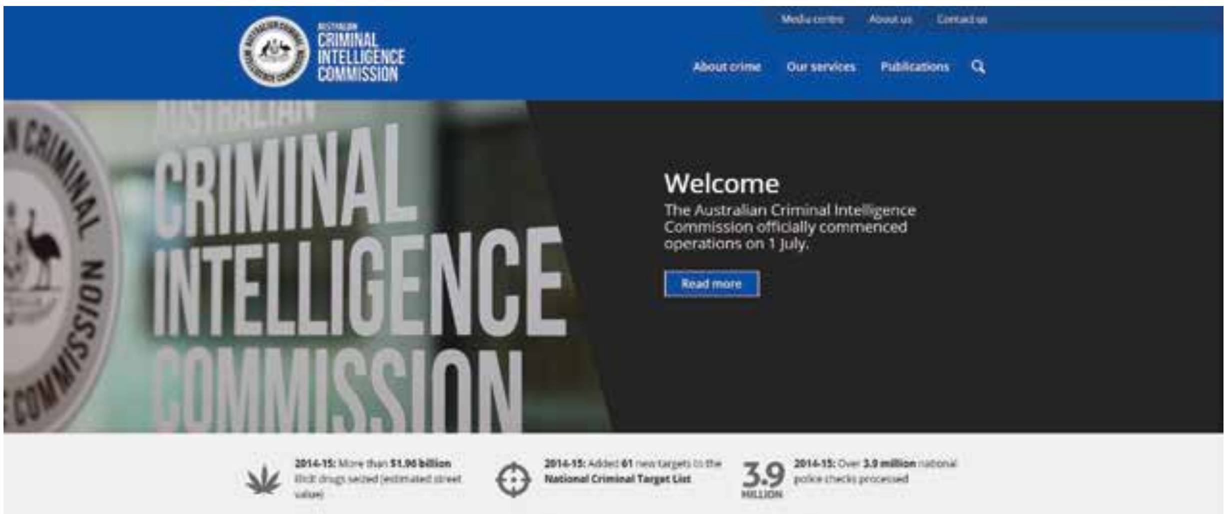



**FIGURE 14.33** The Australian Criminal Intelligence Commission (ACIC) delivers key law enforcement information and services to police Australia-wide to equip them to investigate, solve and prevent crime. Included in the many functions of ACIC is the maintenance of the National Criminal Investigation DNA Database (NCIDD). (Image courtesy of the ACIC.)

## Using the DNA database

DNA profiles from crime scenes are searched against the DNA profiles from individuals:

- A full match occurs when all the numbers (and gender marker) representing an individual's DNA are an exact match to the numbers representing the DNA left at the crime scene (see figure 14.34a).
- A partial match occurs when some data points are missing from the scene-of-crime DNA sample (see figure 14.34b), perhaps because the criminal tried to remove evidence from the scene of the crime. Partial matches can still provide useful information for police investigators.
- Another kind of database search is a **familial search**. This is used to look for possible close relatives of the suspected offender in cases where the police have found the offender's DNA at the crime scene but the suspected offender does not have a profile in the databank. In this case, the search is for individuals with close familial links to the suspected offender. From this information, the DNA profile of the suspect can be inferred, and, depending on the outcome, the suspect will either be arrested or will be cleared of any suspicion.

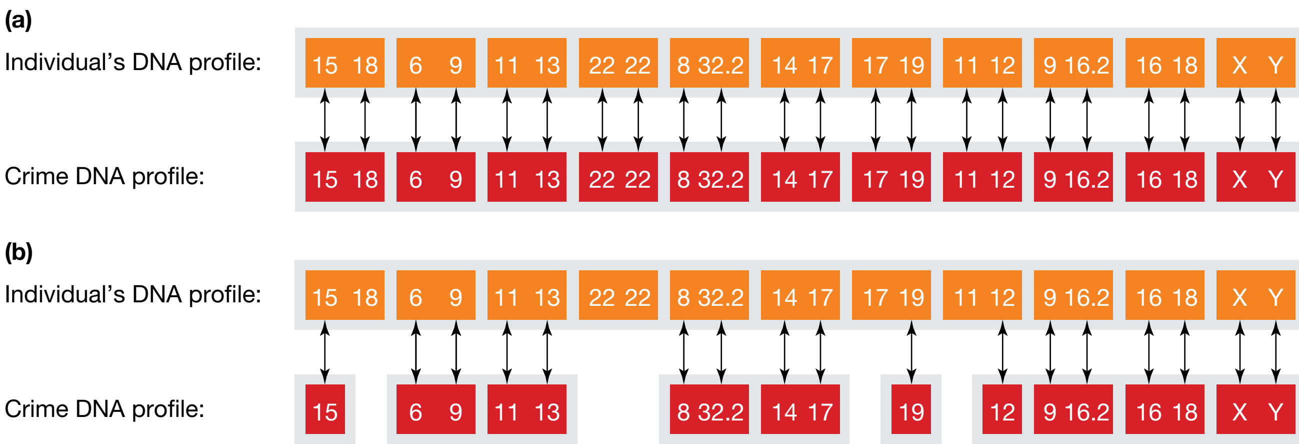



**FIGURE 14.34** Matches between an individual's DNA profile and a crime scene DNA profile may be **(a)** a full match or **(b)** a partial match. (*Source* for both (a) and (b): National DNA Database Strategy Board Annual Report 2012–13, United Kingdom.)

studyon

Unit 4 | AOS 2 | Topic 2 | Concept 3

**DNA profiling**
Summary screen and practice questions

## Issues in DNA profiling

1. **DNA databanks**: Debate exists in terms of the balance between human rights and protection of the community, as, for example:
   - Whose DNA profiles should be added to the national databank — only individuals convicted of serious crimes or should individuals convicted of minor crimes be included too? Or should individuals accused of crime but not convicted, that is, innocent people, also be included?
   - Should DNA profiles be held in databanks permanently or should they be removed after a given period?
2. **Familial searching**: This type of search involves the comparison of a crime scene profile with suspect and convicted offender profiles on a DNA database, looking for relatives to the crime scene profile. The result of a kinship search is a list of candidate relatives in decreasing order of being the likely source of the crime scene DNA sample. Such a search will return many possible candidates that have similarities to the unknown profile. This is very different from using the databank to find a direct full match between a crime scene DNA sample and a single databank DNA profile.

   A familial search of the Californian DNA database found a partial match between a younger man's profile and that of the DNA of crime scenes related to the murders of African-American women over a period of more than 20 years. This finding led police to the father of the younger man who was apprehended and charged.

   In spite of its success in solving cold cases, familial searching has raised some legal and ethical concerns. Concerns have been expressed that wider implementation of familial searching will cause investigators to rely too much on genetic leads at the expense of other potentially more useful lines of investigation. Some argue that familial searching should only be used to protect public safety in those cases where there is a high probability that an unknown perpetrator of a serious crime will strike again.

   The use of familial searching also raises the issue of whether relatives of convicted criminals whose DNA is in databanks and who can be reached through familial searching are being unfairly discriminated against. This is as compared to relatives of individuals whose DNA is not in the databank. Some people are concerned that familial searching identifies a person not because of what they are known to have done but because they have criminal relatives.

3. **Predictive testing**: Predictive testing refers to the ability to use a sample of a person's DNA to predict aspects of that person's phenotype. As studies of the human genome and the related areas of proteomics and metabolomics have expanded, this is becoming less fictional and more possible. Already, a few phenotypic traits such as blue eye colour and red hair colour can be predicted with reasonable accuracy. In addition, DNA can give an indication of the geographic ancestral origin of individuals through identification of their DNA haplotypes.

   Concerns have been raised about predictive testing of DNA samples to identify aspects of a person's appearance because presently this is not a mature science but a science under development and, in addition, a phenotype is not determined exclusively by one's DNA but is also affected by many environmental factors, including childhood infections, quality of diet, accidents and plastic surgery, to name a few.

   Other issues raised relate to the possibility of reinforcing prejudices against certain groups within the community and promoting racial stereotypes if DNA samples are used to determine genetic ancestry and whether this information is made public in relation to individuals suspected of crime.

## BIOLOGIST AT WORK

Dr Dadna Hartman is a forensic scientist with the Victorian Institute of Forensic Medicine.

My scientific journey started as a postgraduate student in the Department of Biochemistry at La Trobe University where I first discovered the wonders of DNA. At the time I was fascinated by the advent of PCR and DNA sequencing and the power these techniques promised. I could only imagine what one could achieve with technologies capable of deciphering the genetic code locked away in the genome of a cell; as they say, I was hooked. After the completion of my PhD studies, I commenced my postdoctoral career as a Howard Hughes Associate at UT Southwestern Medical Center in Dallas, Texas. Early in my postdoctoral career, I came to the conclusion that I wanted to do applied science, rather than pure research. My scientific training in molecular biology enabled me to transfer my skill set to varied fields of application, taking me from cancer research to animal genetics and genomics and ultimately leading me to forensics (where I would like to say I found my calling).

The Victorian Institute of Forensic Medicine (VIFM), Melbourne, Victoria, is charged with the responsibility of assisting the Coroners Court of Victoria with the medico-legal death investigation of reportable deaths. Part of this process is the identification of deceased persons in cases where identification can only be achieved through scientific approaches such as DNA analysis. In these instances, a DNA profile is obtained from the deceased person and matched to the DNA profile obtained from family members to verify a genetic relationship. Although I enjoyed shows like CSI and NCIS as much as the next person, I did not truly appreciate the complexity of the work undertaken as part of a forensic investigation until I started working at the VIFM as the Chief Molecular Biologist.

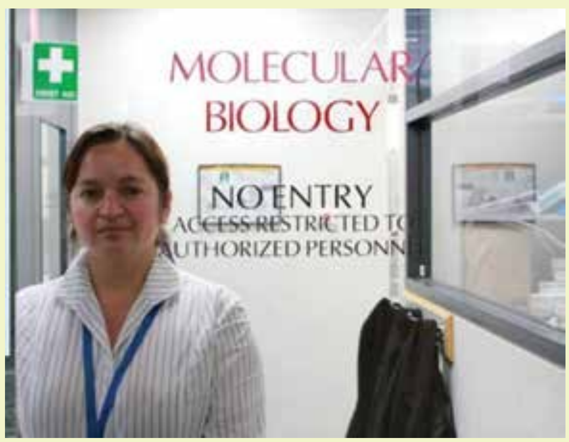



**FIGURE 14.35** Dr Dadna Hartman, Manager Molecular Biology and Chief Molecular Biologist, Victorian Institute of Forensic Medicine. (Image courtesy of Dr Dadna Hartman.)

*(continued)*

At the VIFM, my team routinely works with challenging samples that are compromised with respect to the quality and quantity of recoverable DNA. We routinely work with a variety of different samples that each have their own nuances: blood, bone, teeth, hair, and even toenail clippings. To ensure that we obtain unadulterated DNA profiles, we work in a laboratory specifically designed to minimise the risk of contamination. Additionally, all of our methods are performed under NATA accreditation so our stakeholders can have confidence in our findings. Our casework includes the identification of long-term missing persons, as well as assisting our Victorian and interstate police colleagues in cold case and homicide investigations. Through the provision of DNA analysis, particularly of mitochondrial DNA, we have assisted in solving crimes that can date back more than 20 years. As our findings generally form part of an evidence brief to either the Coroners Court of Victoria or a police agency, we are from time to time required to provide evidence in court. In this capacity, I am required to explain complex scientific procedures to lay members of a jury and to withstand the pressures of cross-examination.

An important aspect of our work centres on Disaster Victim Identification (DVI). My team, like many around the country, are trained and ready to deploy should we be required to assist in the identification of victims in a mass disaster. This was poignantly brought home to us during the 2009 Victorian Bushfires, where we assisted in the identification of those who perished in the fires. During this time, we completed a year's worth of DNA analysis in three months.

Although our work is service-delivery focused, we are constantly investigating new ways to improve our DNA analysis capabilities through research activities aimed at improving our extraction and profiling techniques, such as massive parallel sequencing and SNP analysis. To this end, the VIFM runs an active research program at the Department of Forensic Medicine, Monash University. Through this avenue, we offer honours-year research projects aimed at improving our DNA analysis capabilities.

In conclusion, although our work is challenging at times, particularly when working on complex or high-profile cases, there is an immense sense of satisfaction in helping families at a time when they are particularly vulnerable. I come to work with the knowledge that I am doing something the community values and that I enjoy immensely.


**FIGURE 14.36** Molecular Biologist extracting DNA from a forensic sample. Note: The scientist is wearing appropriate PPE and working within a Lamina Flow cabinet to safeguard the sample from potential contamination from exogenous DNA. (Image courtesy of Judith Kinnear.)

## KEY IDEAS

- DNA profiling in Australia is based on 18 short tandem repeat (STR) loci plus a gender marker.
- DNA profiling has played a role in criminal convictions.
- National DNA databanks have been set up in many countries, including Australia.
- Extension of the use of DNA databanks for purposes such as familial searching and predictive testing has raised debate on some issues.

## QUICK CHECK

**13** Identify the following statements as true or false:
  **a** The gender marker used in DNA profiling is an STR.
  **b** The national DNA databank in the United States is known as CODIS.
  **c** DNA profiling in Australia uses a number of different STR loci in different states.
  **d** The greater the number of alleles of an STR, the higher the number of possible different genotypic combinations in a population.
  **e** It is now possible to give an accurate description of a crime suspect based on DNA from the crime scene.

**14** What is the name of the national DNA databank in Australia?

**15** A DNA profile shows two peaks for the gender marker. What is the gender of the person concerned?

**16** Give one example of the use of DNA profiling in assisting in a criminal conviction.

# Identification in mass disasters

In mass disasters, such as aircraft crashes or terrorist bombings, it is not always possible to identify victims by conventional methods, such as visual identification, fingerprints, dental records, X-rays and tattoos. In the terrorist attack on the US World Trade Center twin towers on 11 September 2001 (known as 9/11), nearly 2800 people died. Because of the intense heat and the violent collapse of the towers, fewer than 300 intact bodies were recovered. Of the victims, about 1600 have been identified and, for more than half of these, the means of identification was DNA analysis. In these cases, DNA profiles of body parts were matched with DNA profiles from victims' personal effects, such as toothbrushes, combs, hairbrushes and razors. In addition, reference DNA samples were supplied by family members.

On 12 October 2002, terrorist bombings in Bali caused the deaths of more than 200 people, including 88 Australians (see figure 14.37). A team of Australian experts in forensic identification travelled to Bali to identify the victims. Of these, some were identified from dental records, some by fingerprints, but some could be identified only using DNA. In the box below, you can read the remarkable story of Linzi Wilson-Wilde, a DNA analysis expert, who was a key figure in the Bali victim identifications. Linzi received a Medal of the Order of Australia for this work. She is now Director of the National Institute of Forensic Science at the Australia New Zealand Policing Advisory Agency (ANZPAA NIFS). Linzi is recognised as the most senior forensic scientist in Australia.

### ODD FACT

All United States troops have a drop of blood on file at the Armed Forces Identification Library in Rockville, Maryland, for identification purposes in the event that military personnel are killed.

**FIGURE 14.37** Australian forensic officers from the Western Australian police force search the Bali bombing site for evidence in October 2002. The bombs killed more than 200 people, including 88 Australians. Of these victims, some were identified from dental records, some by fingerprints, but some could be identified only using DNA.

# DNA and paternity testing

Are they my children? This question can be answered by DNA profiling since the alleles at each STR locus behave in a Mendelian fashion, with half the alleles of a child being contributed by the mother and half by the father. For each test locus, a child inherits one allele from the mother and the other from the father.

If the paternity of a child is in question, a DNA test can be carried out on the child and on the man in question to resolve the issue. If the child carries paternal alleles that are not present in the man, the man in question can be excluded as the biological father. If, on the other hand, the paternal alleles in the child match those present in the man, the man cannot be excluded as the biological father.

Table 14.5 shows the DNA profiles, based on just four DNA markers, of two parents and three children. A question has been raised about the paternity of child 3. What conclusion can be reasonably drawn about child 3?

studyon

Unit 4
AOS 2
Topic 2
Concept 3

**Do more**
DNA profiling
(Family Ties)

**TABLE 14.5** DNA profiles of two parents and three children.

| STR locus | Mother | Father | Child 1 | Child 2 | Child 3 |
|---|---|---|---|---|---|
| **D3** | 14, 18 | 16, 17 | 14, 17 | 17, 18 | 14, 16 |
| **D8** | 10, 15 | 12, 13 | 13, 15 | 10, 13 | 10, 15 |
| **D21** | 28, 30 | 31, 32 | 28, 31 | 28, 32 | 30, 30 |
| **D7** | 8, 10 | 7, 13 | 7, 8 | 10, 13 | 7, 10 |

DNA identification has played a role in reuniting children who have been separated from their families deliberately, such as being abandoned at birth, or by accident, such as displacement during war. DNA identification has been used, for example, in reconstructing the family histories of Aboriginal children forcibly separated from their families.

## BIOLOGIST AT WORK

### Dr Linzi Wilson-Wilde OAM — forensic scientist

As a forensic scientist, your role is to examine evidence, formulate opinions and present those opinions in courts of law. Although this can be very confronting, it is a very rewarding career.

I became interested in forensic science during my undergraduate years at La Trobe University, Victoria, where I was given a forensic-related genetics project. I gained a Bachelor of Science degree in 1995 and went on to complete a Postgraduate Diploma in Science in population genetics and forensic analysis in collaboration with the Victoria Police Forensic Services Department (VPFSD) and La Trobe University, where I gained my passion for forensic science.

After graduating, my first position was with the Biology Division of the VPFSD, where I received all of my training and was later involved in the establishment of the DNA analysis method. While at the VPFSD, and as part of a cooperation exercise with the Vietnamese Police, I was deployed to Vietnam to train local scientists in the method and use of DNA profiling in the establishment of a government DNA laboratory.

In 2000 I was employed as a forensic DNA specialist with the Forensic Services Group of the New South Wales Police Service. While there, I worked on the investigation of a number of 'cold case' murders and a contemporary triple murder. I was also involved in Australia's first mass DNA screen in the town of Wee Waa, where the voluntary DNA screening of 496 men was conducted. This led to the identification and conviction of a rapist. I also served on the Working Party on Law Enforcement and Evidence for the Australian Law Reform Commission Report into the Protection of Human Genetic Information, released in 2003.

I later joined the Biological Criminalistics Team at the Australian Federal Police (AFP) as the Team Leader. The team was responsible for providing biological and DNA analysis services for the AFP in the Australian Capital Territory and for federal and overseas investigations. My role involved examining items, such as clothing and weapons, for the presence of biological material and then conducting DNA analysis on the material found. I was also trained in bloodstain pattern analysis, which looks at bloodstain patterns at crime scenes to determine the sequence of events. Results of examinations and opinions drawn are then presented in courts of law.

My career highlight has been my work on Disaster Victim Identification (DVI) and criminal investigation of the Bali bombings in 2002, for which I was awarded a Medal in the Order of Australia. For five months, initially working 18 hours a day, seven days a week, I was responsible for up to 50 scientists analysing samples from remains and missing persons' reference samples for DNA comparison. Over 3000 DNA tests were performed, which helped to identify 199 victims; only three bodies remain unidentified. The team also analysed crime scene samples, gathering evidence that contributed to the conviction of the bombers. I was also integrally involved in the drafting and passage of appropriate legislation that allowed the matching of DNA results for the DVI process. To facilitate this, I sat on the floor in both the House of Representatives and the Senate while the amendment to the legislation was tabled and debated — an exciting experience.

I have since returned to Melbourne where I became the General Manager of the National Institute of Forensic Science at the Australia New Zealand Policing Advisory Agency (ANZPAA NIFS). During this time I completed my PhD in the investigation of wildlife crime using molecular DNA analysis at the Melbourne Museum and with the University of Canberra. I was also inducted into the Victorian Honour Roll of Women in 2014.

I am now the Director of ANZPAA NIFS, which involves running the agency for Australia and New Zealand. We are working on facilitating national standardisation of forensic analysis, research and innovation, education and training, quality programs and national information exchange and collaboration. It is a wonderful opportunity to give back to the field I love. In life you never stop learning. Whatever you pursue in life, whatever career you choose, always remain focused and committed.

**FIGURE 14.38** Dr Linzi Wilson-Wilde OAM. (Image courtesy of Dr Linzi Wilson-Wilde.)

## KEY IDEAS

- DNA identification has played a major role in identifying victims of mass disasters where other methods cannot be used.
- DNA testing is used in paternity testing and in reuniting children separated from their parents.

**QUICK CHECK**

**17** Identify the following as true or false:
- **a** Alleles at each STR locus behave in a Mendelian fashion.
- **b** Boys inherit all their STR alleles from their fathers.
- **c** DNA testing can be used to exclude a man as the biological father of a child.

**18** True or false? A mother with alleles 10 and 15 and a man with alleles 12 and 14 at the D8 STR locus could have a child with the alleles 14 and 15 at the same locus.

# DNA profiles are not just for people!

**ODD FACT**

In Canada, DNA databases for moose (*Alces alces*), caribou (*Rangifer tarandus*) and black bear (*Ursus americanus*) enable species to be identified from biological traces. These databases have been used to assist in prosecuting people for illegally poaching bears in Ontario, where they are protected, and in disproving the false claim that the bears were legally hunted in the province of Manitoba.

Like human chromosomes, the chromosomes of other mammals, including domestic pets, contain many short tandem repeats (STRs) that can be used to create DNA profiles. Just as people can be identified by their DNA profiles, so too can domesticated mammals, such as horses, cattle, sheep, dogs and cats (see figure 14.39), as well as wildlife. In addition, the mtDNA of these species can be used to show pedigree relationships.

DNA profiling of mammals can be used for various purposes, such as prosecuting people who are trading illegally in wildlife, confirming the parentage of valuable breeding stock or, in the case of a dog attack, matching saliva traces left on a victim of a dog attack to a particular dog. Domestic pets leave traces, such as hair, saliva and blood, on clothing, carpets and furnishings. DNA from these traces can provide forensic evidence to link human suspects to the scene of a crime or to a victim. As with human DNA profiles, it is possible to identify individual animals with a very high degree of certainty and to identify with certainty whether an individual animal is the offspring of particular parents.

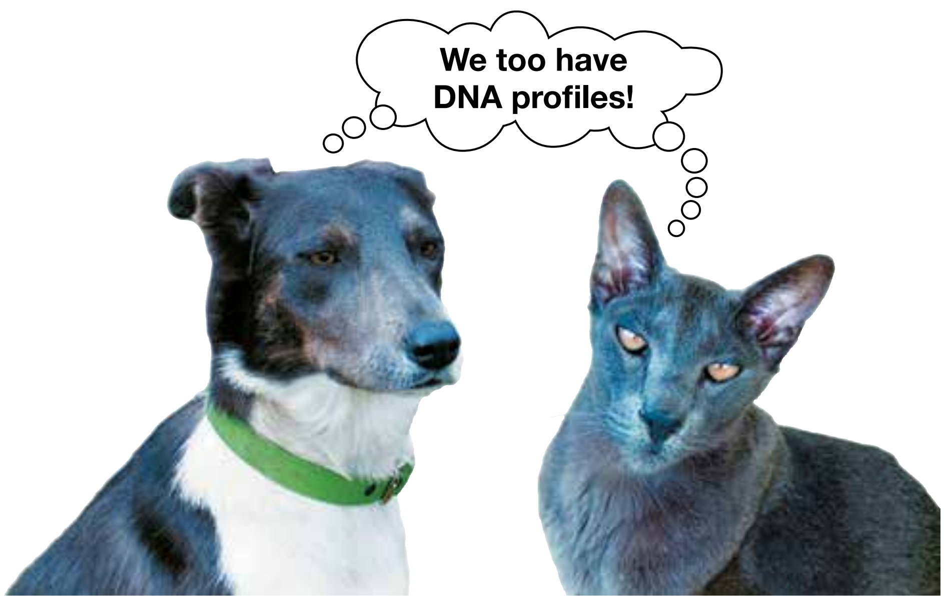



**FIGURE 14.39** DNA is the common genetic material of all animals, including mammals. As a result, DNA profiles can be prepared not only from people but also from domestic pets and other mammals.

## Protecting Australia's wildlife using DNA

Australia's rare native birds, such as the red-tailed black cockatoo (*Calyptorhynchus banksii*), and native reptiles, such as thorny devils (*Moloch horridus*), attract high prices in Australia and even higher prices when smuggled overseas (see figure 14.40).

**FIGURE 14.40** Red-tailed black cockatoo, a native Australian bird. How are gene manipulation techniques assisting in protecting these and other species of native birds in their natural habitats?

Our native wildlife is protected by legislation and cannot be collected in the wild. Several laboratories in Australian states have the expertise to perform DNA fingerprinting and mtDNA analysis of wildlife. These laboratories work with government agencies to protect our wildlife from exploitation such as unlawful collecting and unlawful export.

For example, registered breeders hold licences from state wildlife authorities that allow them to breed rare native birds and reptiles using approved captive stock. However, many of these species are very difficult to breed in captivity. DNA profiling has been used to test pedigree claims and distinguish between rare native birds produced through legal captive breeding and those obtained by illegal nest robbing. Illegally obtained birds can be readily identified because their DNA profiles do not match those expected in the offspring of the alleged parent birds.

DNA technology has also been used to test the validity of claims that native wildlife specimens in the possession of collectors were obtained legally by purchasing them from approved breeders. In a case in NSW in 2010, DNA evidence was used for the first time to secure the conviction of a man who had illegally collected pygmy spiny-tailed skinks (*Egernia depressa*) from their habitat in Western Australia (see figure 14.41). His claim that he bought the skinks from a registered breeder was disproved when the DNA profiles of the skinks failed to match those expected to be produced by the alleged parents.

**FIGURE 14.41** The pygmy spiny-tailed skink (*Egernia depressa*). When illegally exported from Australia and sold on the black market, these reptiles are worth many thousands of dollars each. Taking these skinks from their habitat involves not only their loss from the native populations, but also the destruction of the rocks in which they shelter.

## Cats have DNA profiles

**ODD FACT**

The likelihood that the DNA profile of the cat hair found on the bloodstained jacket would match Snowball's DNA profile by chance was estimated at 1 in 45 million.

Snowball the cat has the distinction of being the first cat whose DNA profile led to a conviction of a person for murder — this was the first occasion on which STR genotyping of pet animal hairs was admitted in evidence in a criminal trial, in Canada.

In October 1994, a woman disappeared from her home. In woodlands near to where the missing woman's car was found, a man's bloodstained leather jacket was found and this blood matched that of the missing woman. Several white hairs were found on the lining of the jacket — not human hairs but hairs from a white cat. In May 1995, Canadian police found the woman's body in a shallow grave. Her estranged husband, Douglas Beamish, was charged with her murder. The leather jacket could be identified only as being like one owned by Beamish. However, police recalled seeing a white cat at the home of Beamish's parents where Beamish was living. DNA analysis showed that the white hairs came from this white cat, Snowball. The presence of Snowball's hairs at the crime scene implicated Beamish in this murder. This DNA evidence was a major factor in convicting Douglas Beamish of the murder of his wife.

Hairs, blood and buccal (cheek) swabs obtained from cats can be used to generate DNA profiles. The identification of domestic cats (*Felis catus*) for forensic purposes is now typically based on 11 STRs plus one gender identifier. Table 14.6 and figure 14.42 show the STR loci used in cat DNA profiling, the number of different alleles identified for each STR and the chromosome where the STR is located.

**ODD FACT**

As for human DNA profiles, the STR loci in cats used for forensic identification contain variable numbers of four-base repeats.

**TABLE 14.6** Features of the STRs used in forensic identification with cat samples. Which STR has the highest number of alleles?

| STR locus | No. of alleles | Chromosome |
|---|---|---|
| **FCA441** | 8 | **D3** |
| **FCA723** | 20 | **A1** |
| **FCA731** | 6 | **B1** |
| **FCA733** | 16 | **B2** |
| **FCA736** | 23 | **B4** |
| **FCA740** | 7 | **C1** |
| **FCA742** | 15 | **D4** |
| **FCA749** | 14 | **F2** |
| **F53** | 11 | **A1** |
| **F85** | 32 | **B1** |
| **F124** | 20 | **E1** |
| ***Gender identifier*** **SYR** | | **Y** |

(*Source:* Menotti-Raymond et al., 2005, *Journal of Forensic Sciences*, Vol. 50, p. 1061.)

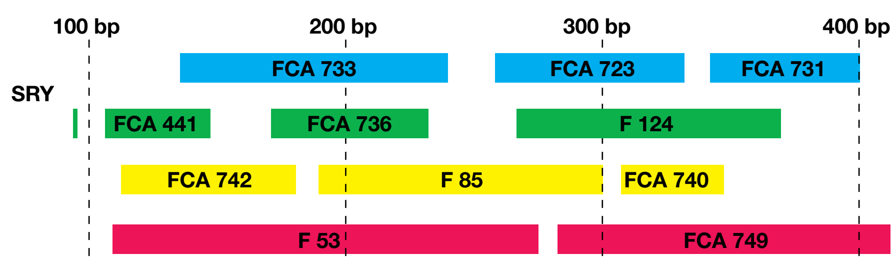



**FIGURE 14.42** Size ranges of STR sequences used for cat samples, and the different fluorescent dyes used to distinguish them in cases of size overlap (data from Menotti-Raymond et al., 2005, *Journal of Forensic Sciences*, Vol. 50, p. 1061).

Figure 14.43 shows the DNA profiles of two domestic cats, one male and one female. Note that only the male shows the presence of the gender identifier, at the extreme lower right-hand corner.

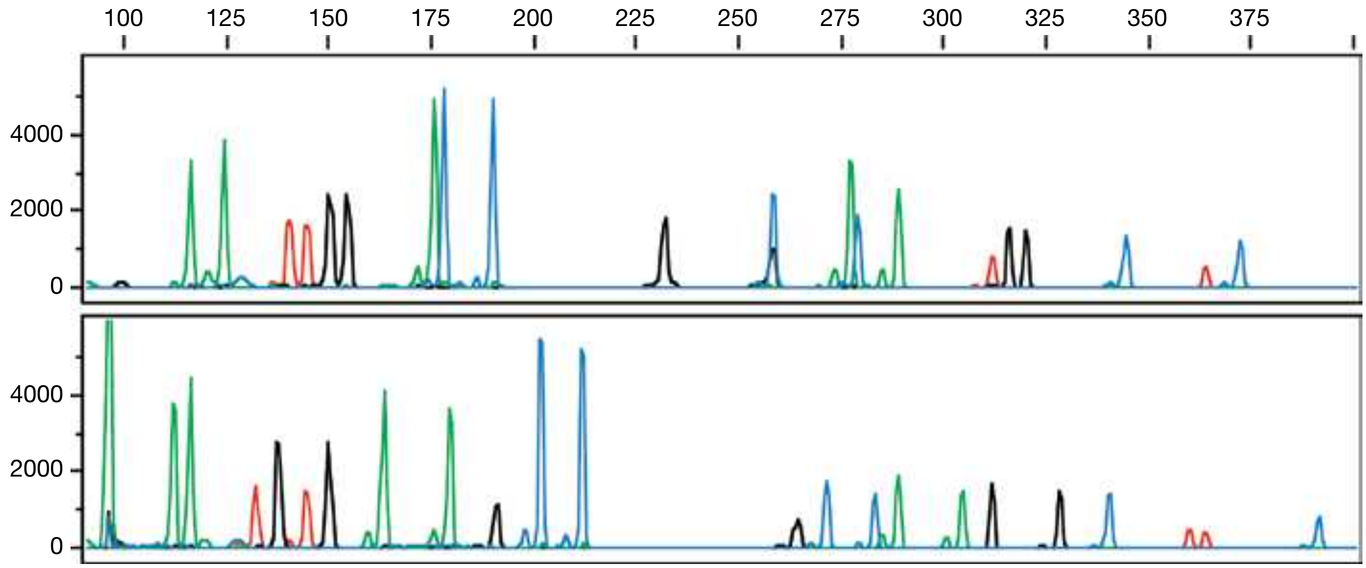



**FIGURE 14.43** DNA profiles of two domestic cats, one female at top and one male at bottom. The scale at the top shows the size of the various alleles and the vertical scale at the side shows the fluorescence intensity. For clarity in these profiles, peaks with a yellow fluorescent label are shown in black. (Reprinted with permission from the *Journal of Forensic Sciences*, Vol. 50, Issue 5, copyright ASTM International.)

## And dogs are crimebusters too!

In 1998 in the United States, critical evidence against two people accused of murder proved to be bloodstains on their clothing that came from the victim's dog, which had also been shot. The DNA profile of these bloodstains matched that of a blood sample taken from the dog and linked the accused to the crime scene.

Like people, DNA profiling of dogs is based on 10 STR markers. Figure 14.44 shows a sample of an ID card produced in the United States for dogs.

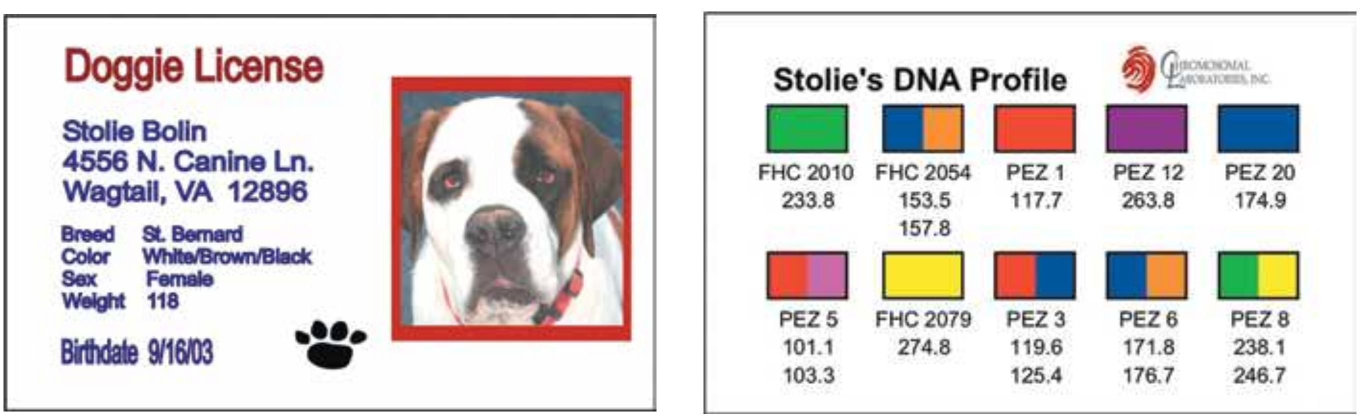



**FIGURE 14.44** Here we see a 'doggie licence' for Stolie, a St Bernard dog. On the back of this licence is Stolie's DNA profile based on 10 STR loci. Can you identify an STR locus for which Stolie is homozygous? At how many loci is Stolie heterozygous? (Image courtesy of Genomic Diagnostics.)

In several countries, dog breeders who wish to be accredited with the relevant national body, such as The Kennel Club (UK), are required to have their breeding stock DNA profiled. Discussion is underway about possible compulsory DNA testing of breeding dogs in order for them to be registered with the Australian National Kennel Council (ANKC). Each dog, identified by its microchip or tattoo, would be required to have a DNA test based on material obtained from a buccal (cheek) swab.

Figure 14.45 shows a dog having a buccal swab taken for DNA testing. This test uses 11 STR markers as well as a gender identifier.

**FIGURE 14.45** Ricky, a golden retriever, having a buccal (cheek) swab taken for DNA testing. (Image courtesy of Genomic Diagnostics.)

## ODD FACT

For any greyhound to be raced in Australia or New Zealand, Greyhounds Australasia requires that the sire (father) and the dam (mother) of the greyhound be DNA-certified. This certificate, combined with a microchip, guarantees accurate identification of each racing greyhound.

Canine DNA profiles, like human DNA profiles, are a series of peaks, with each peak corresponding to one allele at an STR locus. Figure 14.46 shows the DNA profile of Oscar the dog.

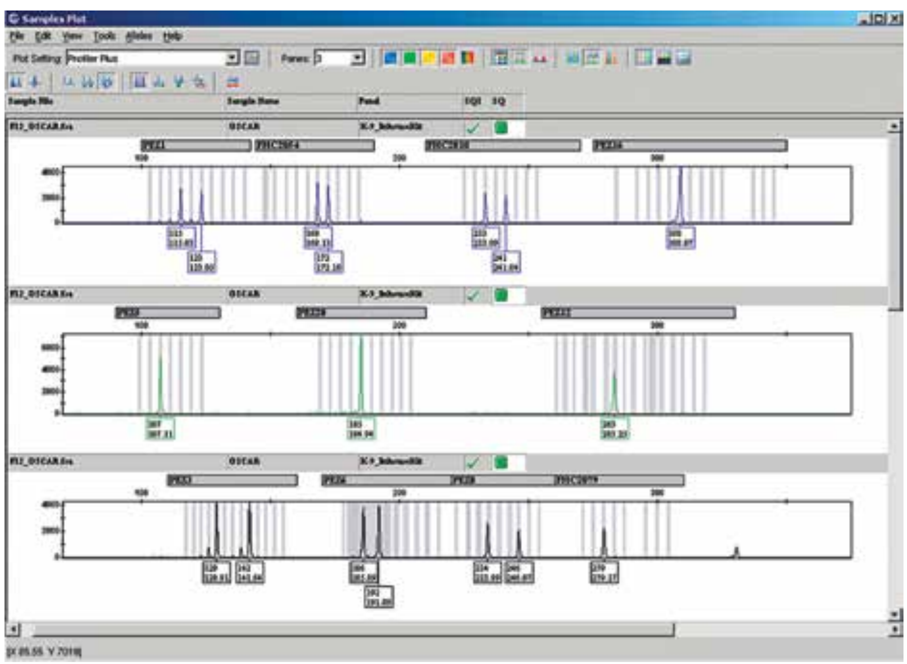


**FIGURE 14.46** DNA profile of Oscar the dog. The name of each STR locus is shown above the relevant peaks. At which STR locus is Oscar homozygous? (Image courtesy of Genomic Diagnostics.)

DNA profiling is also carried out on thoroughbred horses (see figure 14.47). Since 2003, registration in the Australian Thoroughbred Stud Book requires that all thoroughbred foals born in Australia have a DNA test using material from the roots of hairs taken from their manes. In the case of horses, a set of 17 particular STRs is used to create a unique DNA profile.

**FIGURE 14.47** A thoroughbred mare and her colt. (Image courtesy of Judith Kinnear.)

DNA profiling is also done on cattle. Figure 14.48 shows samples of hairs (with roots) taken from cattle. Each sample is then sent to a laboratory for DNA profiling and this profile, plus an electronic tag, forms a unique identifier for each animal. This means that animals can be tracked at any stage, and that a meat sample taken from a carcass can be matched back to a particular animal.


**FIGURE 14.48** Hair samples from cattle ready to be sent for DNA profiling. (Image courtesy of Zoetis.)

Breeders of pedigreed animals such as cattle and horses need to know the sire (father) and dam (mother) of the animals that they breed. Mistakes in assigning parents can occur, owing to cross-suckling, semen and embryo mix-ups during artificial insemination or embryo transfer programs, mistakes in record keeping or mixed-sire matings such as occur when a gate is left open or an animal jumps the fence. The error rate in assigning parentage in one New Zealand study was found to be 19 per cent — this means that the parents of nearly one animal in five had been misassigned. Wrong parent assignment might later mean legal action!

## KEY IDEAS

- **DNA profiles can be prepared from biological material from mammals, including domestic pets.**
- **As for humans, DNA profiling of other mammalian species involves the use of short tandem repeats (STRs).**
- **DNA profiling of domestic pets has been used as evidence in courts of law.**
- **Paternity testing based on DNA profiles is used in pedigreed animals.**

## QUICK CHECK

**19** List two non-human species for which DNA profiling has been done and, in each case, identify one reason why it was done.

**20** Refer to figure 14.42 and identify:

- **a** the approximate size of the gender marker for cats
- **b** the size range for the FCA740 marker in the cat population.

**21** Identify one means by which tissue is obtained from dogs for DNA profiling.

# BIOCHALLENGE

1 The following table shows the DNA profiles of the members of the Romanov family murdered in 1918 in Ekaterinberg, based on a set of 15 autosomal STR markers. Samples 4.3, 7.4, 3.46, 5.21 and 6.14 were obtained from skeletal remains recovered from the grave opened in 1991. Samples 147 and 146.1 were from skeletal remains from the grave opened in 2007.

| | Sample 4.3 | Sample 7.4 | Sample 3.46 | Sample 5.21 | Sample 6.14 | Sample 147 | Sample 146.1 |
|---|---|---|---|---|---|---|---|
| **Marker** | **Tsar Nicholas II** | **Tsarina Alexandra** | **Olga** | **Tatiana** | **Maria or Anastasia** | **Anastasia or Maria** | **Alexei** |
| Amelog | X, Y | X, X | X, X | X, X | X, X | X, X | X, Y |
| D3S1358 | 14, 17 | 16, 18 | 17, 18 | 17, 18 | 16, 17 | 17, 18 | 14, 18 |
| TH01 | 7, 9.3 | 8, 8 | 8, 9.3 | 7, 8 | 8, 9.3 | 7, 8 | 8, 9.3 |
| D21S11 | 32.2, 33.2 | 30, 32.2 | 30, 33.2 | 32.2, 33.2 | 30, 33.2 | 30, 33.2 | 32.2, 33.2 |
| D18S51 | 12, 17 | 12, 13 | 12, 12 | 12, 12 | 13, 17 | 12, 17 | 12, 17 |
| D5S818 | 12, 12 | 12, 12 | 12, 12 | 12, 12 | 12, 12 | 12, 12 | 12, 12 |
| D13S317 | 11, 12 | 11, 11 | 11, 11 | 11, 11 | 11, 11 | 11, 11 | 11, 12 |
| D7S820 | 12, 12 | 10, 12 | 12, 12 | 10, 12 | 12, 12 | 10, 12 | 12, 12 |
| D16S539 | 11, 14 | 9, 11 | 11, 11 | 11, 11 | 11, 14 | 9, 11 | 11, 14 |
| CSF1PO | 10, 12 | 11, 12 | 11, 12 | 11, 12 | 10, 11 | 10, 12 | 10, 12 |
| D2S1338 | 17, 25 | 19, 23 | 17, 19 | 23, 25 | 17, 19 | 17, 23 | 23, 25 |
| vWA | 15, 16 | 15, 16 | 15, 16 | 15, 16 | 15, 16 | 15, 16 | 15, 16 |
| D8S1179 | 13, 15 | 16, 16 | 13, 16 | 15, 16 | 13, 16 | 15, 16 | 15, 16 |
| TPOX | 8, 8 | 8, 8 | 8, 8 | 8, 8 | 8, 8 | 8, 8 | 8, 8 |
| FGA | 20, 22 | 20, 20 | 20, 22 | 20, 20 | 20, 22 | 20, 22 | 20, 22 |
| D19S433 | 13, 13.2 | 13, 16.2 | 13.2, 16.2 | 13.2, 16.2 | 13, 16.2 | 13, 13 | 13, 13.2 |

(*Source:* Coble, M.D. *et al.*, 2009. Mystery solved: the identification of the two missing Romanov children using DNA analysis, *PLoS ONE* Vol. 4, No. 3.)

a Suggest why samples 6.14 and 147 are identified as *either* Maria *or* Anastasia.
b What role does the Amelog marker play in a profile?
c What is an STR?
d A student stated: 'Why use 11 STR markers rather than just 2? Surely that would make the situation far less complicated'. Do you agree with the student?
e The entries in the columns include numbers such as 12, 13 and 30, 30.2. What do these numbers denote?
f For the **D18S51** marker, which parent contributed the 12 allele to Alexei?
g For the **D21S11** marker, which parent contributed the 32.2 allele to the person represented by sample 147?
h Are all possible genotypes represented in the four children of Tsar Nicholas and the Tsarina?
i If the genotype of sample 146.1 for the **FGA** marker had been 22, 22, would this have made a difference?

2 In September 2011, the announcement was made that skeletal remains (minus the head) exhumed from Pentridge Prison in Melbourne were those of the bushranger Ned Kelly, who was hanged at the Old

Ned Kelly.

Melbourne Gaol in 1880. The Victorian Institute of Forensic Medicine (VIFM) explored several lines of evidence in the process of identifying these remains. This evidence included X-rays, CT scans, skeletal measurements and historical records. In addition, mtDNA was successfully isolated from the skeletal fragments by a DNA laboratory in Argentina.

a Suggest how historical records might contribute to the identification process.

b Suggest how skeletal measurements and other skeletal features might contribute to the identification process.

c Leigh Olver, a teacher in Melbourne, is the great-grandson of Ellen, one of Ned's sisters. Leigh was a key player in the identification process because the mtDNA from the remains was found to match Leigh's mtDNA.

In figure 14.49, which of the following pedigrees (A, B or C) shows the inheritance of mtDNA from Ned's mother to Ned and then, over several generations, to Leigh Olver? Briefly explain your decision.

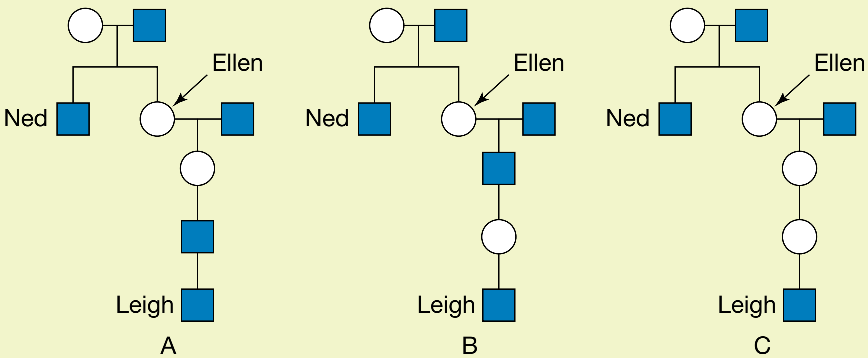



FIGURE 14.49

# Chapter review

studyon

Unit 4 | DNA manipulation
AOS 2 | Biological knowledge and society
Topics 1&2 | Practice questions

## Key words

carrier testing
DNA fingerprinting
DNA profiling
familial search
genetic
genetic screening
genetic testing
Huntington's disease (HD)
predictive testing
pre-implantation genetic diagnosis (PGD)
presymptomatic testing
short tandem repeats (STRs)

## Questions

**1 Making connections →**

a Construct a concept map for 'DNA profiling' using any of the key words above. Add other terms as needed.

b Construct a second concept map for 'recombinant proteins' using any of the chapter keywords.

**2 Demonstrating and communicating understanding →** *E. coli* cells carrying the foreign gene for human growth hormone are to be grown in mass culture in 50 000 litre fermenting tanks over 2 days. The recombinant plasmid present in these *E. coli* cells is pUC18.

a Person A suggested that 0.5 gram of these bacteria would be sufficient to start this mass culture, but person B said 'No way! You will need lots more'. Indicate with which person you agree and give a reason for your choice.

b Outline a procedure by which the *E. coli* bacteria with recombinant plasmids might be separated from bacteria not possessing recombinant plasmids.

c The culture fluid in which the *E. coli* bacteria cells will grow provides the water, glucose, salts and nitrates required for their growth, but it will also support the growth of many other kinds of bacteria. Identify a possible measure that could be taken to prevent the growth of contaminating microbes.

**3 Considering possibilities →** Detection and amplification of DNA has become much more sensitive so that DNA can now be isolated from just a small number of cells, estimated at 5 to 20 cells. This fact has given rise to the concept of '**touch DNA**', that is DNA that may be recovered from skin cells left behind on any object that a person has touched or handled.

a Use the web to identify estimates of the numbers of skin cells that are lost daily from the skin of an adult. Any object that is a probable source of touch DNA at a crime scene, such as a door handle or a knife handle, or a glass would be swabbed with a sterile, DNA-free cotton swab (see figure 14.50).

b Suggest why the technician in this figure is gloved and masked.

c Suggest three other possible sources of touch DNA at a crime scene.

The DNA profile from touch DNA found on a door handle at a crime scene was found to match that of the next-door neighbour of the victim.

d Would this evidence alone justify the immediate arrest of that neighbour for the crime concerned? Explain your answer.

**FIGURE 14.50** Sterile, DNA-free cotton swab being used to swab the rim of a wine glass, which is a possible source of touch DNA.

**4 Applying principles in a new situation →** A phenomenon termed secondary transfer of DNA is known to occur. Secondary transfer is defined as the transfer of the DNA of one person to an object by a second person after skin-to-skin contact has been made, such as during a handshake between two individuals (see figure 14.51).

**FIGURE 14.51** Could this handshake transfer the DNA of the person on the right to the person on the left and/or vice versa?

Consider the following facts:

i L.A. was charged with a murder because his DNA matched a sample of DNA obtained from the victim's hands.
ii L.A. does not have an identical sibling.
iii Hospital records proved that before the murder took place, L.A. was highly intoxicated and unconscious on the floor of a liquor store where he was treated by paramedics before being transported to the hospital. As a result, L.A. was released.

**a** Given that the DNA match was real, suggest two ways in which this situation might have occurred:
  i one way involving a secondary transfer of DNA
  ii some other explanation.

**b** Is a secondary transfer explanation more or less likely if you are told that the paramedics who treated L.A. in the liquor store were the first emergency responders to the murder scene an hour or so later.

**5 Demonstrating knowledge and understanding ➜** Suppose that you are required to produce a recombinant protein.

Some of the steps that you will perform are identified as follows, but they are not in the correct order:

*Copy recombinant DNA, identify relevant gene, screen transformed cells, transfer recombinant DNA into host cells, construct recombinant DNA vector*.

**a** List these steps in the appropriate order.
**b** For each step, briefly describe its outcome.

**6 Evaluating the limitations of DNA identification evidence ➜** Throughout this chapter, reference was made to DNA identification in the following terms:

i 'DNA samples recovered from the remains . . . became part of the evidence that enabled their identification . . .'
ii ' . . . the mtDNA extracted from a skeleton . . . assisted in confirming that the skeleton was that of Richard III.'
iii '. . . the first time that the genetic evidence of DNA fingerprinting was used to assist in convicting a murder suspect.'

Review these statements and try to identify a key point about the use of DNA identification that is embedded in them.

**7 Evaluating the limitations of DNA identification evidence ➜** Consider the following information: Daniel Fitzgerald was convicted of a murder after DNA evidence linked him to a didgeridoo found at a crime scene. There was no other evidence linking this man to the crime.

**a** If you had been a member of the jury involved in this case, would you have agreed with the verdict? Give a reason for your decision.

In 2013, Fitzgerald appealed against his conviction to the Court of Criminal Appeal, but this appeal was unsuccessful. However, he was given leave to appeal to the High Court. In 2014, this appeal went forward.

**b** If you were one of the High Court judges hearing this appeal, would you uphold the appeal against the conviction or reject the appeal?

The High Court judges ruled that the recovery of the appellant's DNA from the didgeridoo did not raise any inference about the *time when* or *circumstances in which the DNA was deposited* there.

**c** Suggest two possible ways in which Fitzgerald's DNA might have been transferred to the didgeridoo.

The High Court upheld the appeal, overturning the conviction on the grounds that the DNA evidence was not sufficient to prove guilt because the means by which the DNA was transferred to the didgeridoo was not known. The key issue is that no other evidence existed that linked the convicted man to the crime.

**8 Discussion question ➜** In 1997, the author of a paper in the *Journal of Pediatrics* included the following comments about the use of cadaver-derived human growth hormone (hGH) in the United States:

'*Patients . . . were subjects in clinical studies for part of the first year of treatment, after which the limited availability of GH dictated treatment for only part of the year and caps on final heights.*

*By 1984 treatment was year round and the height caps largely unenforced. In the last year of its distribution GH was used in 2450 patients in the United States. . . . And then, in 1985, came Creutzfeldt-Jakob disease. While the not-so-good old days are gone and need not be lamented, there remains virtue in a conservative therapeutic philosophy. If anything can be learned from the use of pituitary GH in children, it is a healthy respect for the law of unintended consequences.'*

Review the growth hormone story on page 646 and consider the above statement.

Discuss the following:

- **a** The availability of cadaver-derived hGH until 1984 was limited. How did this impact the treatment of patients?
- **b** How did treatment of patients change from 1984 when the supplies of hGH became more plentiful?
- **c** What is meant by 'a conservative therapeutic philosophy'?
- **d** The author states that the lesson to be learned from the GH story is 'a healthy respect for the law of unintended consequences'. What does the author mean? Do you agree?

Some people obtain injectable hGH through internet pharmacies and other websites and use it in the hope of improving their athletic performance. However, at present, the effect of hGH on athletic performance is unknown. Some people purchase sprays and pills containing hGH in the hope of slowing or stopping age-related processes in the body. These claims are also unproven.

- **e** Do you consider these to be harmless practices?
- **f** Given that hGH is a protein comprising a chain of 191 amino acids, what is the expected effect on GH if ingested by mouth?

CHAPTER

# 15 Manipulating genes in organisms

**FIGURE 15.1** Photograph of GloFish® taken under blue actinic lighting. Short-wave blue light is absorbed by a protein in the muscle tissue of these transgenic fish and is then emitted as longer wavelength bright fluorescent red, green, yellow or orange light. In this chapter, we will explore transgenic organisms and genetically modified organisms (GMOs), and examine their uses in agriculture. (Images courtesy of Glofish.)

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- recognise the distinction between genetically modified organisms (GMOs) and transgenic organisms
- give examples of transgenic plants and animals
- gain knowledge and understanding of how GMOs are applied in agriculture
- understand the techniques used to produce transgenic organisms
- consider some issues related to GMOs.

# 1970s: Genetic engineering begins

For millennia, human beings have used selective breeding to alter the characteristics of domesticated plant and animal species (refer back to chapter 9, page 423). Selective breeding involves manipulation at the level of the phenotype, as, for example, by selecting for breeding only those sheep with fleece with the thinnest (finest) fibres, and culling sheep with undesirable features. In the past, genetic material could only be transferred between members of the same species (or very rarely between very closely related species). The restrictions that normally prevent gene transfer between different species is termed the 'species barrier' and it includes many factors, such as the inability of gametes of one species to fertilise those of a different species.

In the early 1970s, however, all this changed when scientists gained the power to directly manipulate DNA, the genetic material of all organisms. Through direct manipulation of DNA, genes can be transferred between organisms from totally unrelated species, extra copies of genes can be added to a genome, and genes can be silenced. Because DNA is the universal genetic material and because the genetic code in organisms is closely related, foreign genes might theoretically come from any kind of living organism — from an orchid, an oyster, a dragonfly, a moss, a monkey, a squid and so on.

Gene manipulation and genetic engineering techniques have made the 'species barrier' irrelevant, and enable genes to be transferred between any two species — plant cells can incorporate bacterial genes, bacterial cells can incorporate human genes, and fish can incorporate sea anemone genes. Because of its scope and power, genetic manipulation or genetic engineering of species raises various legal, ethical, economic and ecological issues.

Genetic engineering (or gene manipulation or biotechnology) differs from the conventional breeding programs that were once the only means by which improvements could be achieved in farm animals and agricultural crops. Some key differences between genetic engineering techniques and conventional breeding are shown in table 15.1.

**TABLE 15.1** Differences between conventional breeding and genetic engineering.

| | Genetic engineering | Conventional breeding |
|---|---|---|
| ***Source of genes*** | Genes can be introduced from almost any species, closely or remotely related | Genes can only come from the same or a very closely related species |
| ***Time scale*** | Improvement or other intended outcome can be achieved in much shorter time | Time required to achieve improvement typically requires many generations |
| ***Minimum number of genes inserted*** | Just a single gene of interest can be transferred from a source organism to the host organism | An entire single genome must be acquired that may include undesirable genes as well as the desired gene |
| ***Chance of outcome*** | Higher probability of achieving the desired genetic outcome | Little guarantee that a specific gene combination can be achieved |
| ***Silence genes?*** | Yes, silencing of genes is possible | No, silencing of genes is not possible |

However, let's meet the fish with the genes of a sea anemone . . .

## Fluorescent fish of many colours

In 2003, scientists at the University of Singapore developed transgenic zebra fish (*Danio rerio*) that expressed in their muscle tissue a red fluorescent protein encoded by a gene transferred from a sea anemone. The incorporation of the sea anemone gene into the fish genome was stable so that these fish could pass the genetic information for the fluorescence trait to their offspring.

**ODD FACT**

Natural fluorescence (bio-fluorescence) is common in some animal groups, such as corals and jellyfish, and it also occurs in butterflies and parrots. A 2014 publication reports on the discovery of natural fluorescence in more than 180 species of fish.

**eBook*plus***

**Weblink**
Natural fluorescence in fish

In Australia, it is illegal to import or possess genetically modified or transgenic fish without authorisation under the *Gene Technology Act 2000* and from the Department of Agriculture, Fisheries and Forestry.

Since then, genes for green and yellow fluorescent proteins have been introduced into stable transgenic lines of other ornamental fish, such as tetras (*Paracheirodon ternetzi*) and tiger barbs (*Puntius tetrazona*). The fluorescence of these various transgenic fish species is most vividly expressed when the fish are illuminated by blue actinic lighting (420–460 nm). The fluorescent proteins absorb short wavelength blue light and re-emit it as longer wavelength fluorescent light of colours such as red, green, and yellow (see figure 15.1).

The rights to the international marketing of these so-called GloFish® were purchased from the University of Singapore by Yorktown Technologies, an American company. The various fluorescent colours of these fish have registered descriptive labels, including Starfire Red™, Electric Green™, Moonrise Pink™ and Sunburst Orange™. Glofish are the first transgenic organisms to be sold as pets in the United States.

Genetic engineering techniques combined with the use of fluorescent protein markers are valuable tools that enable scientists to observe biological processes in living organisms, such as events in mammalian embryonic development. In the case of the fluorescent fish, the Singaporean scientists developed them, not for ornamental purposes, but for research purposes — in particular, in order to develop fish that would fluoresce in the presence of specific environmental pollutants.

## Making fluorescent fish

Scientists at the National University of Singapore created the transgenic, fluorescent zebra fish by micro-injecting the DNA of the genes for red, green, and orange fluorescent proteins into fish eggs. This micro-injection takes place soon after fertilisation, but before the sperm and egg pronuclei have fused. The source of the genes is either from a sea anemone (for red fluorescence) or from a jellyfish (for green fluorescence).

Provided the injected DNA inserts itself into the genome of the fertilised egg at an appropriate site, this new feature becomes an inherited trait. The mature fish can pass the genes encoding the fluorescence traits to their offspring, and this process can continue, generation after generation. Once fluorescent fish have been produced using genetic engineering techniques, no further genetic engineering is required to build the fish stocks. The new generations of the various fluorescent fish are simply produced though natural breeding.

## GMOs include transgenic organisms

**Genetically modified organisms (GMOs)** are those organisms whose genomes have been altered using genetic engineering technology. GMOs have novel combinations of genes that do not occur in natural populations of the species concerned but are created through the means of genetic engineering. The genetic alteration achieved through genetic engineering may involve:

- the addition of a gene or a segment of DNA
- the silencing of a gene so that its function is lost.

For an organism to be identified as a GMO, any gene or DNA segment that is added through genetic engineering should be heritable, that is, should be able to be passed on to the next generation.

**Transgenic organisms** comprise a subgroup of GMOs that includes those GMOs in which the alteration to the genome involves the genetic material from a different species. So, it follows that all transgenic organisms are GMOs, but not all GMOs are transgenic.

The question to ask in order to identify a transgenic organism is: *'Does the gene that is added to the genome come from the same species or does it come from a different species?'* If the gene is from the same species, the organism is termed a GMO. If the gene comes from a different species, the recipient organism is described as being a transgenic organism and, at the same time, a GMO. For example, GloFish that have incorporated genes from a sea anemone into their genomes are transgenic as well as GMOs. An organism whose genome has *not* been subjected to change by genetic engineering is termed a **non-GMO**.

One of the early transgenic mammals produced was a mouse that, in 1983, had a copy of the human growth hormone gene (**GH1**) inserted into its genome. This transgenic mouse, when compared with its non-GMO littermates, showed a much greater growth rate (see figure 15.2). The **GH1** gene, plus a promoter sequence, was inserted via a plasmid.

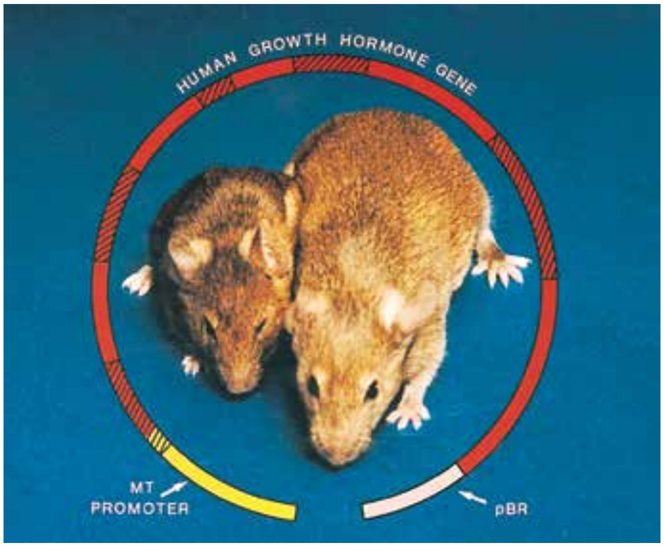



**FIGURE 15.2** Mice from the same litter, aged about 24 weeks. A transgenic mouse (at right) carries an inserted copy of the human **GH1** gene in its genome and weighs 59 grams. Its non-GM littermate (at left) weighs just 28 grams. The background shows the plasmid involved with the **GH1** gene (shown in red) with its promoter sequence (shown in yellow).

(Source: Brinsters, R.L., School of Veterinary Medicine, University of Pennsylvania.)

Since the early transgenic mice appeared in the 1980s, many transgenic organisms, both plant and animal, have been produced. The ability to introduce foreign genes and produce transgenic organisms with phenotypic traits not previously present in the species has been used on several occasions to 'rescue' a species in a local region. One example may be seen in Hawaii relating to the local papaya-growing industry. In the 1990s, a papaya ringspot virus (PRSV) infection severely impacted on the commercial growing of papaya fruits (*Carica papaya*) (see figure 15.3), with annual production falling from 56 million in 1992 to 34 million in 1998. The development and release of transgenic papaya plants resistant to PRSV after 1998 allowed papaya harvests in the virus-infection region to increase. This measure, combined with the cultivation of non-GM papaya in new quarantined regions free of the virus, enabled the Hawaiian papaya industry to recover to its previous levels of fruit production.

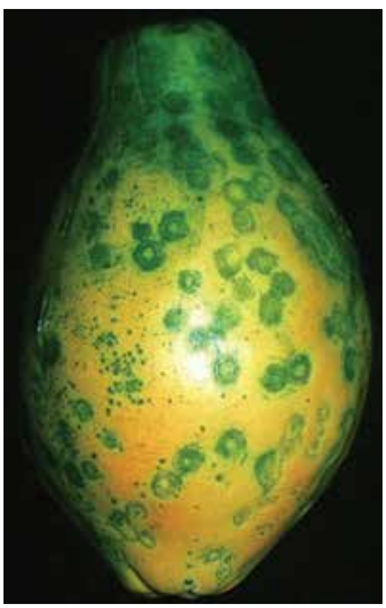


**FIGURE 15.3** A papaya fruit showing the signs of papaya ringspot virus (PRSV) infection. As well as affecting the fruit, this virus also causes leaf damage.

## How are genes added to or 'removed' from GMOs?

Let's now look briefly at the production of transgenic mammals, such as mice. Genetic engineering has produced many different transgenic mouse stocks that carry copies of various human genes responsible for inherited diseases, including Huntington's disease, cystic fibrosis, thalassaemia, gout, and retinoblastoma, an inherited cancer. Transgenic mice play an important role in research into these human inherited diseases as they provide models that can be used for testing drugs, for studying fine details of the development of these diseases, and for identifying the role of various genes in the progress of these diseases.

### Making transgenic mice

Various genetic engineering techniques may be used to produce transgenic mice. These techniques, as outlined below, include: 1. Micro-injecting DNA into fertilised eggs, and 2. Modifying embryonic stem cells.

1. **Micro-injecting DNA into fertilised eggs:** DNA micro-injection of eggs is used when the aim is to add a gene to create GM mice. In this procedure, the DNA of the gene concerned is injected into the fertilised egg of a mouse at a stage when the male and female pronuclei have not yet fused. Figure 15.4 shows a diagrammatic summary of this step of the process. The step involves the use of a very fine micropipette (at right). A much larger pipette (at left) holds the egg cell in place.

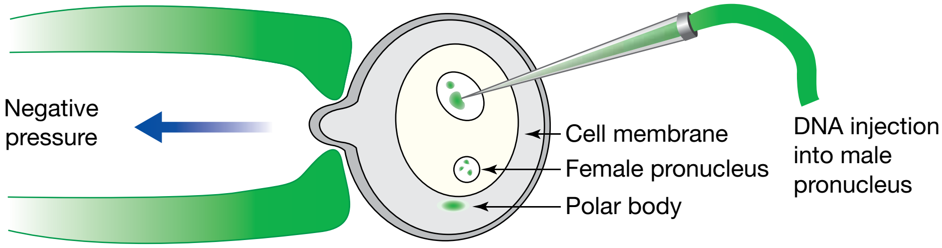



**FIGURE 15.4** Micro-injection of naked DNA into the male pronucleus of a fertilised egg prior to fusion of the male and female pronuclei to form the nucleus of the zygote.

**ODD FACT**

The yield of viable transgenic eggs from the micro-injection procedure is low because there is no control over the site in the genome where the foreign gene inserts itself into the egg nucleus, and there is no control over how many copies of the foreign gene are inserted into the genome.

This step produces some mouse eggs with one or more copies of the injected DNA that has randomly integrated into their genomes. A number of eggs are implanted in a surrogate female mouse, and live mouse pups will be born 19 days later. These mouse pups will be screened to select those that have integrated the foreign DNA and so are transgenic.

Figure 15.5 is a diagram showing a highly simplified version of the overall process of producing transgenic mice for research into inherited human diseases.

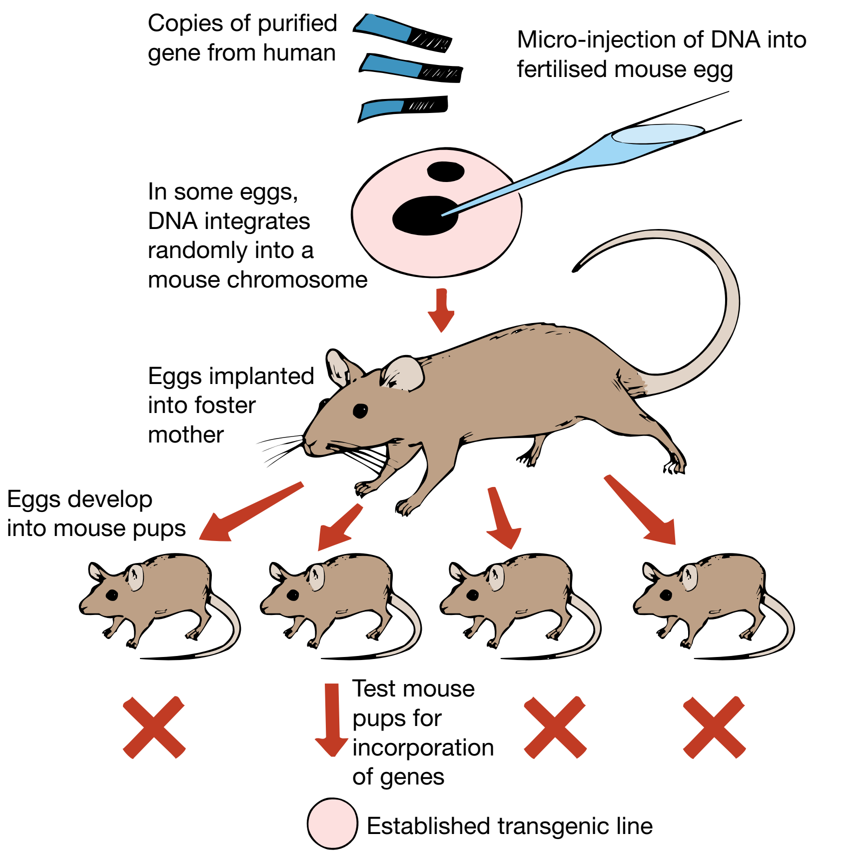



**FIGURE 15.5** Producing a transgenic mouse. The DNA is not added to a particular point of the mouse chromosome, but is incorporated at random. What problem would arise if the added gene integrated in the middle of a gene controlling production of an essential enzyme?

**ODD FACT**

Mice that that are engineered to lack a specific functional gene are termed 'knock-out' mice. This is in contrast to transgenic mice engineered to carry an extra gene. Predictably, these are called 'knock-in' mice.

2. **Modifying embryonic stem cells:** This procedure is used to obtain the answer to the question: *'What happens when a particular gene is silenced?'* The aim of this procedure is to silence or disable a specific target gene in order to identify its function by seeing what happens when the gene is inactivated. The disabling is achieved by creating a gene that is homologous to the target gene but which carries a defect in its base sequence that makes it non-functional. This defective DNA sequence is injected into an embryonic stem cell and, in some cells, the defective gene replaces the functional target gene by recombination. If this happens, the target gene in those cells is 'knocked out' or silenced.

studyon

| | |
|---|---|
| Unit 4 | **Genetically modified organisms** |
| AOS 2 | Summary screen and practice questions |
| Topic 2 | |
| Concept 4 | |

## Making transgenic plants

Several methods exist for making transgenic plants, and examples of commonly used methods are outlined below.

**1. Physical uptake of DNA:**

Several methods of transferring DNA into plant cells have been used.

- One method is the so-called 'gene gun'. In this technique, DNA-coated gold particles, just 1 to 3 micrometres in diameter, are accelerated at high speed into plant tissue by using compressed helium. Once inside the cell, the DNA diffuses from the surface of the gold particles and some DNA integrates into the chromosomal DNA of the plant. In this technique, the site of insertion of the foreign DNA is random.
- Plant cells can be treated with enzymes to remove their cell walls. Note that a plant cell minus its cell wall is termed a **protoplast**. Protoplasts can be placed in a buffer solution with the foreign gene of interest. The mixture is subjected to electroporation, a treatment that temporarily increases the permeability of the plasma membrane to macromolecules, such as DNA. If the foreign DNA is taken up and inserts into the chromosomal DNA of the protoplasts, the protoplasts can be placed in tissue culture and then grown into transgenic plants.

**2. *Agrobacterium*-mediated uptake of DNA:**

The transfer of foreign genes into higher plants is commonly carried out using soil bacteria of the genus *Agrobacterium*. In nature, these bacteria infect plants, causing diseases such as crown gall disease (see figure 15.6a). These diseases develop when the bacteria insert their Ti plasmids into plant cells and, once there, the plasmids transfer a segment of their DNA, called T-DNA (transfer DNA), into the plant genome. Figure 15.6b shows a diagram of the Ti plasmid of *Agrobacterium* bacteria. Note the presence in the plasmid of the T-DNA region.

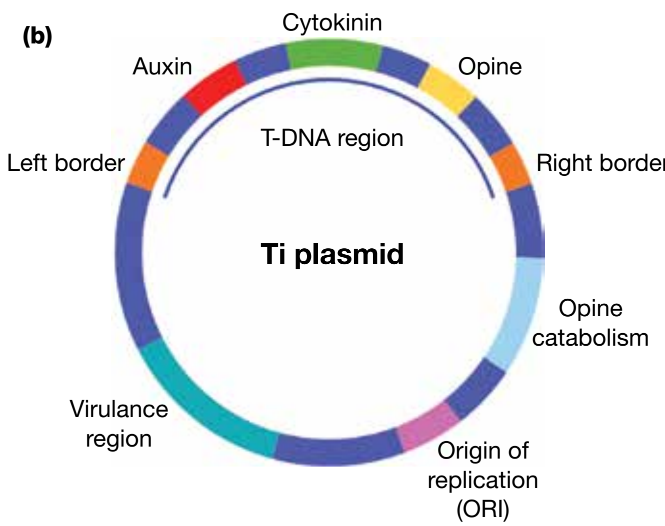




**FIGURE 15.6** **(a)** Crown gall disease in various plant species is caused by the soil bacterium species *Agrobacterium tumefaciens.* Galls are dark masses of undifferentiated tissue that typically form on the crown of the plant. The crown is the point at the soil line where the stem is joined by the main roots. **(b)** Ti plasmid of *Agrobacterium tumefaciens.* The Ti plasmid produces crown gall tumours in plants when they are infected by these bacteria.

This natural behaviour of *Agrobacterium* bacteria of inserting their T-DNA into the genome of plant cells is used by biotechnologists. This behaviour of *Agrobacterium* can be used to transfer foreign DNA into plant cells to develop genetically modified plant crops. The foreign DNA must be inserted into the plasmid within the T-DNA region. However, before they are used to transform plant cells, the *Agrobacterium* bacteria are modified to make them harmless and non-tumour producing. Importantly however, these bacteria still retain their ability to transfer the T-DNA segments of their plasmids into host plant cells along with any foreign DNA. These foreign genes can then be transferred by *Agrobacterium* cells into host plant cells by growing pieces of plant tissue in culture with the harmless Ti plasmids carrying the foreign gene.

The steps in transforming plant cells are outlined below and shown in figure 15.7.

**Step 1:** Recombinant *Agrobacterium* plasmids carrying the DNA sequence of the foreign gene are formed.

**Step 2:** *Agrobacterium* cells and plasmids, both recombinant and non-recombinant, are mixed, and some cells take up recombinant plasmids.

**Step 3:** *Agrobacterium* cells are grown on agar, and cultures of recombinant cells with the plasmids that carry the foreign gene are identified.

**Step 4:** Plant cells are infected by these bacteria, and the recombinant plasmids transfer the foreign genes into the plant cells where these foreign genes insert into the genetic material of the plant chromosomes.

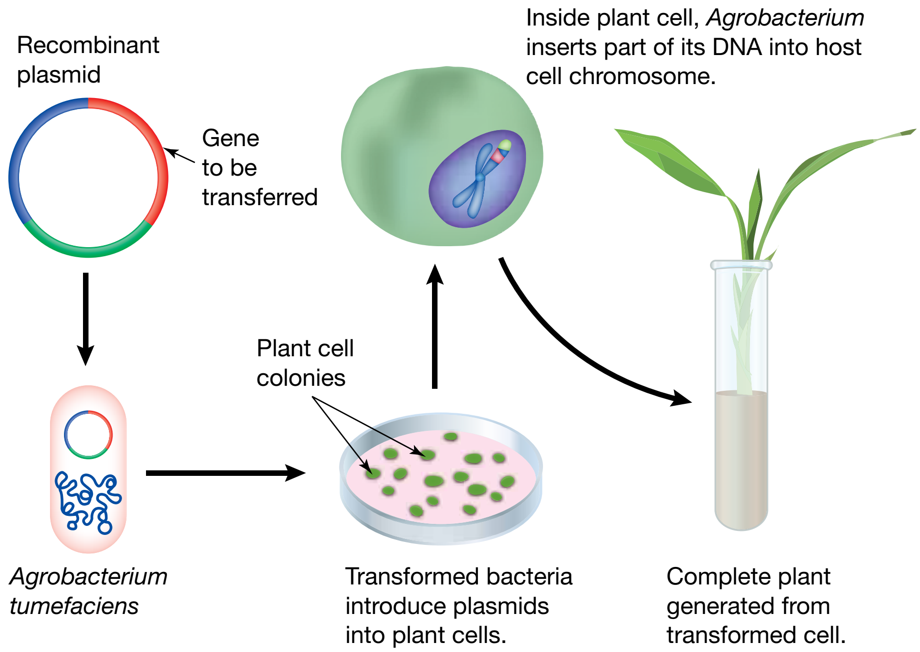



**FIGURE 15.7** Diagram showing steps in the formation of a genetically modified plant where the foreign gene is transferred to the plant cell via the recombinant plasmids of *Agrobacterium* bacterial cells.

## KEY IDEAS

- Gene manipulation displays many differences from conventional breeding.
- A genetically modified organism (GMO) is one that has had its genome altered by genetic engineering technology, with the alteration being permanent and heritable.
- A transgenic organism is also a GMO, but with the extra proviso that the alteration to its genome involves DNA from a foreign species.
- A number of different methods exist for producing genetically modified organisms.
- Many transgenic mice stocks have been developed that carry various human genes, and they are used in research on human inherited diseases.

**QUICK CHECK**

1 Identify the following statements as true or false:
  a GloFish were created by incorporating fluorescent protein into fish genomes.
  b The gene gun is one technique for creating transgenic plants.
  c It is theoretically possible to produce a transgenic plant that has been transformed by a salmon gene.
  d While GloFish are examples of genetically modified organisms, they are not transgenic organisms.
  e When micro-injection is used, the site of insertion of a foreign gene into a mouse egg cell can be predicted with accuracy.
  f *Agrobacterium* species can be used as vectors to transfer foreign genes into plant cells.
2 Could you buy GloFish at your local pet shop? Explain.
3 List two differences between genetic engineering and conventional breeding.
4 Identify two different methods of introducing foreign genes into plant cells.

# GMOs in agriculture

Foreign genes that confer resistance to insect pests and viruses, and that confer tolerance to various herbicides, have been engineered into major food crops with high market value. Worldwide, nearly 180 million hectares of GM crops were planted in 2015, including pest-resistant maize, cotton, potato and rice, virus-resistant squash and papaya, and herbicide-tolerant cotton, maize, soybean and canola. In addition, some crops, such as corn and cotton, have been engineered to carry the combined ('stacked') traits of insect resistance and herbicide tolerance.

Table 15.2 gives some examples of the purposes of genetic manipulation of various plant crops. In some cases, outcomes have been achieved with GM crops receiving regulatory approval and entering the commercial production stage in one or more countries. In other cases, projects are still at the experimental stage or at field trial stage.

**TABLE 15.2** **Examples of intended purposes of genetic manipulation of plant crops.**

| Transgenic trait | Crops |
|---|---|
| insect (pest) resistance | cotton, corn, potato, tomato |
| herbicide tolerance | canola, cotton, corn, rice, flax, sugar beet |
| virus resistance | papaya, squash, potato |
| delayed ripening | tomato |
| tolerance to environmental stress<br>• drought tolerance<br>• flood tolerance<br>• salt tolerance | <br>rice<br>rice<br>rice, wheat, barley, tomato |
| altered oil composition | canola, soybean |
| enhanced nutritional value | rice, wheat |
| improved post-harvest shelf life | tomato, papaya, mango, pineapple |

## GM crops worldwide

Worldwide in 1996, an initial area of 1.7 million hectares was planted with GM crops. This increased by 2015 to nearly 180 million hectares, with the main GM crops under commercial cultivation being herbicide-tolerant and insect-resistant cotton, canola, corn and soybean. Figure 15.8 shows the growth in area under cultivation with GM crops for the industrial (developed) countries versus developing countries over the period 1996 to 2015.

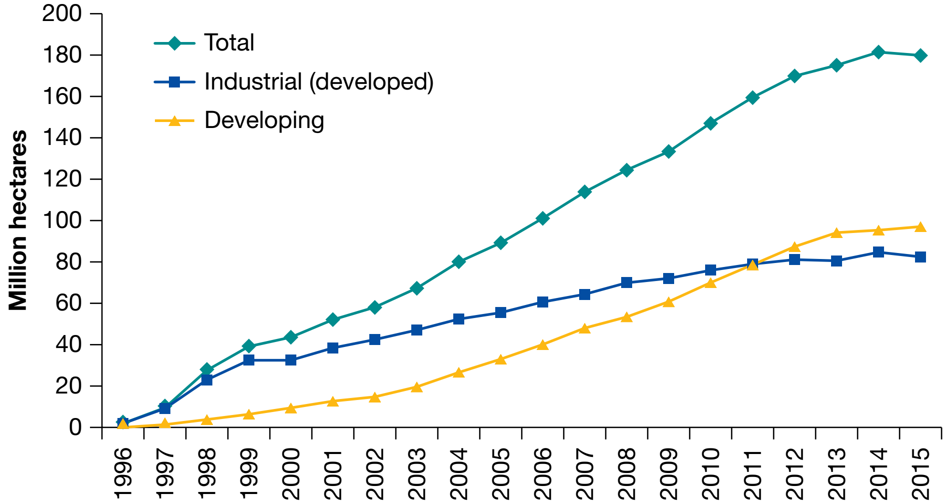



**FIGURE 15.8** Graph showing the increase in area under cultivation with GM crops in the period 1996–2015 and the breakdown for industrial (developed) and developing countries. (*Source:* ISAAA, 2015.)

In total, about 30 countries have GM crops under cultivation. However, in 2014 just 10 of these countries accounted for 98 per cent of the total GM area under cultivation. Countries with the largest areas under cultivation with GM crops are shown in figure 15.9. USA has the largest area, then Brazil and Argentina. Other countries with significant GM crop plantings are India and Canada. The group identified as 'Others' includes Australia and eight European Union nations, including Sweden and Germany, that have planted GM potatoes engineered to have altered starch content for use as stock food and industrial use.

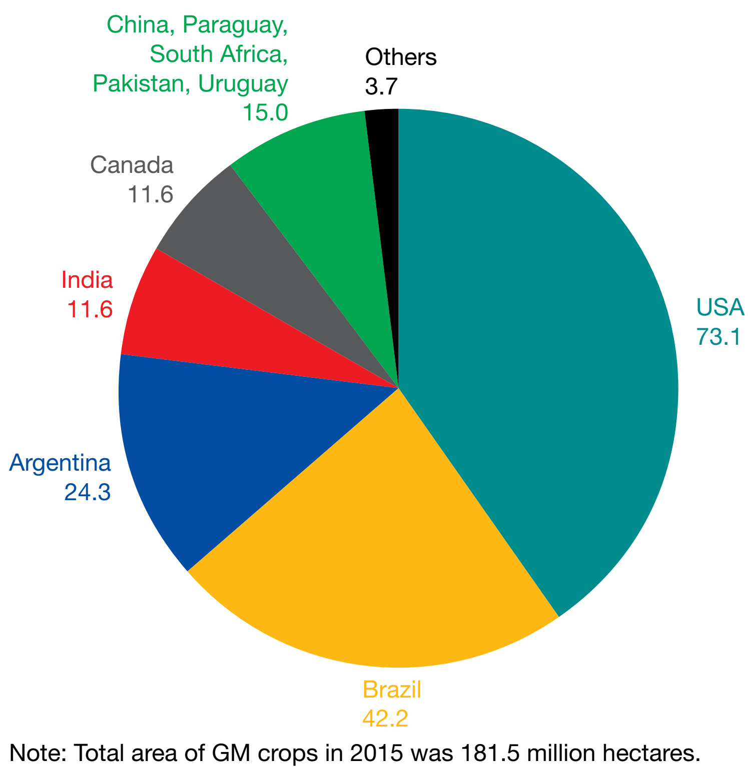



**FIGURE 15.9** Pie chart showing the worldwide areas of cultivation of GM crops in 2015. Note that the United States has the largest area under cultivation. (Source: ISAAA, 2015.)

As well as the largest area of GM plantings, the United States has the widest range of GM crops, either in commercial plantings or under development. The commercial plantings of approved GM food crops in the USA include soybean, corn, cotton, apple, canola, chicory, flax and linseed, melon, papaya, potato, rice, squash, sugar beet, tomato and tobacco.

Worldwide, however, the GM crops under cultivation representing about 99 per cent of the total area of GM plantings include just four crops: soybean (51%), maize (corn) (30%), cotton 13% and canola (5%). All other GM crops constitute just 1 per cent of the total GM plantings (see figure 15.10).

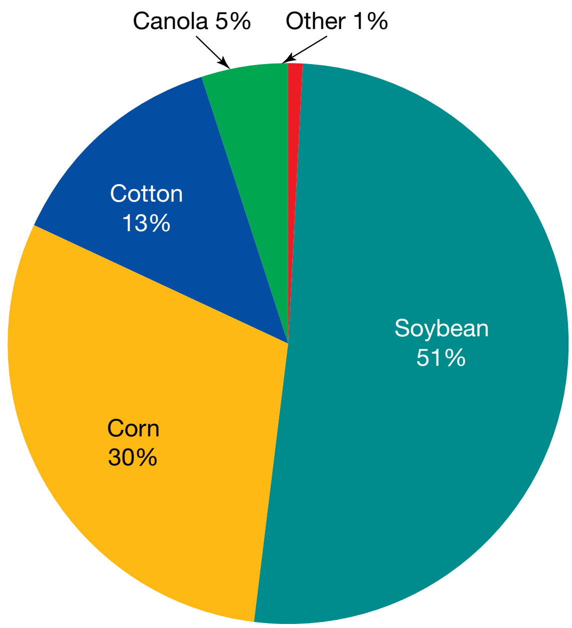



**FIGURE 15.10** GM crops as a percentage of total area under GM cultivation. Just four crops dominate worldwide. Plantings of all other GM crops constitute just one per cent (Source: CBAN, 2016.)

## GMOs in Australia

In Australia, two genetically engineered crops grown on a large scale are cotton (*Gossypium hirsutum*) and canola (*Brassica napus*).

1. **GM Cotton:** The fruit of the cotton plant is a structure called the cotton boll. Inside each boll is the lint that forms the cotton fibres that are spun into the cotton fabrics of clothing, towels, furniture covers and the like. The cotton boll also contains a number of seeds, which are a source of cottonseed oil.

**FIGURE 15.11** Crop of cotton plants, with fruits ready for harvesting in autumn. Cotton is one of the two approved GM plants that can be grown commercially in Australia.

**FIGURE 15.12** Larva of *Helicoverpa armigera*, known as the bollworm, eating a cotton boll. *H. armigera* and other *Helicoverpa* species are the most serious insect pests of cotton and are also pests of other crops, including corn. (Courtesy of Judith Kinnear.)

In 1996, **Ingard®cotton** became the first genetically modified field crop to be grown in Australia. These cotton plants were genetically engineered to include the **cry1Ac** gene from the bacterial species *Bacillus thuringiensis*. The **cry1Ac** gene encodes a protein, known as **Bt**, that acts as an insecticide. Bt insecticides are highly effective against the leaf-feeding larvae (caterpillars) of the cotton bollworm (*Helicoverpa armigera*), which are major pests of cotton crops (see figure 15.12).

In 2003, **Bollgard II®cotton** was developed by CSIRO in collaboration with the US-based multinational company Monsanto, and was released for commercial use. A second gene, the **cry2Ab** gene, that encodes a toxin that acts against different insect pests was engineered into Bollgard II cotton plants.

Many insecticides, such as endosulfan and broad-spectrum organophosphates, are highly toxic to people, pets, livestock on farms, wildlife, and they kill beneficial insects as well as their major target pests. In contrast, Bt insecticide is non-toxic, has a narrow target range, and is not a contact insecticide (one that kills on direct contact). Instead, Bt must be ingested by caterpillars to take effect.

Note: In April 2016, the Australian Office of the Gene Technology Regulator (OGTR) received an application from Monsanto Australia Limited for the commercial release of two new types of GM cotton. These are Bollgard® 3 XtendFlex™ cotton and XtendFlex™ cotton. XtendFlex™ cotton contains three genes that confer tolerance to the herbicides glyphosate, glufosinate and dicamba. Bollgard® 3 XtendFlex™ cotton contains the same three introduced genes for herbicide tolerance and three introduced genes that confer insect resistance. Monsanto is seeking approval to grow these GM cottons commercially in all cotton-growing areas of Australia. The GM cottons and their products would enter general commerce, including their use in human food (cottonseed oil) and animal feed. (Source: OGTR Application DIR145.)

**ODD FACT**

In 2008, GM canola was first grown commercially in Victoria.

2. **Canola:** In spring, crops of canola form vivid displays of bright yellow flowers (see figure 15.13) that are a valuable source of oilseed.

**FIGURE 15.13** Canola plants (*Brassica napus*) in flower. Canola seeds are a source of canola oil, which is used in many processed food products.


**FIGURE 15.14** Guess who has the Roundup-resistant gene?

Two GM varieties of canola have been developed for use in Australia:

(i) **Roundup Ready® canola**, developed by Monsanto, was approved for commercial production in Australia in 2003. The **gox** gene, from the bacterial species *Ochrobactrum anthropic*, was engineered into canola plants using *Agrobacterium* as a vector. The gox gene encodes an enzyme that breaks down glyphosate, the active ingredient of Roundup® herbicide, into a harmless product. Spraying canola crops with this herbicide kills weeds and other plants that lack the gox gene, but leaves GM canola plants unaffected (see figure 15.14).

(ii) **InVigor® canola**, developed by Bayer CropSciences Pty Ltd, has been engineered to include a gene that confers tolerance to the herbicide glufosinate. The gene that confers this tolerance is the **bar** gene, which comes from soil bacteria of the genus *Streptomyces*.

## Getting approval to manipulate genes

In Australia, the production of genetically modified (GM) crops and other organisms is controlled at both the federal and state levels. At Commonwealth level, experiments involving GMOs are governed by the *Gene Technology Act 2000*. Under this Act, a gene technology regulator, based in the Office of the Gene Technology Regulator (OGTR), administers this legislation and makes decisions on applications to release GMOs.

Several committees provide advice to the OGTR on various matters as follows:

- Gene Technology Technical Advisory Committee (GTTAC) — technical and biosafety matters
- Gene Technology Ethics Committee (GTEC) — ethical matters
- Gene Technology Community Consultative Committee (GTCCC) — issues of community concern.

**eBook*plus***

**Weblink**
OGTR website

The object of the national scheme is to 'protect the health and safety of people, and to protect the environment, by identifying risks posed by or as a result of gene technology, and managing those risks by imposing conditions on the licences required for dealings with GM organisms (GMOs).'

### GM foods approved for Australia

Where edible, products from GM crops are termed **genetically modified (GM) foods.** In Australia, GM products cannot be purchased as fresh foods, but GM products can be used in processed foods.

The bi-national government agency Food Standards Australia New Zealand (FSANZ) has approved the use of specific GM products in the manufacturing of foods that can be sold in Australia. These GM products, many of which are imported, include the following: oil from GM canola; oil, flour, syrup and sugar from GM corn; cottonseed oil from GM cotton; oil and protein meal from GM soybean; sugar from GM sugar beet; GM potato products, and GM lucerne.

Some of these GM products are used as animal feeds, such as GM soybean meal and lucerne.

Novel DNA or protein is defined as DNA or a protein that, because of the use of gene technology, differs in chemical sequence or structure from the DNA or protein present in the counterpart non-GM food.

**Labelling of GM foods:** FSANZ requires that GM whole foods and ingredients in processed foods that *contain novel DNA or protein* must be labelled with the words 'genetically modified.' Exceptions are as follows:

- foods in which the GM ingredients are highly refined, such as cooking oils, margarine, sugars, starches, chocolate and baked goods
- foods made at bakeries, restaurants and takeaways
- foods from animals that are fed GM feed.

### CRISPR-Cas9 technology: to be or not to be regulated?

All plants and animals genetically engineered by a range of techniques, including those described earlier in this chapter, are subject to regulatory approval before development and release. This applies to animals that are genetically engineered through micro-injection of DNA or modification of embryonic stem cells, and also to plants genetically engineered using gene guns or by *Agrobacterium*-mediated gene transfer.

The CRISPR-Cas9 technology (refer back to chapter 13, page 613), however, has created new and simple ways of modifying the genome of organisms, and without introducing any genetic material from either the same or a different species. Because CRISPR is a gene-editing technique, it might be expected that new genetic variants produced by this technology would not be regarded as GMOs and, if so, would not come under GM legislation.

In the United States, the regulator of GM plants and fungi is the US Department of Agriculture (USDA). In April 2016, the USDA stated that it will not impose its GM regulation requirements on common white mushrooms (*Agaricus bisporus*) that have been genetically modified using CRISPR-Cas9. The gene editing carried out on these mushrooms deleted a few base pairs from one of the genes encoding an enzyme that causes browning of the mushrooms — browning that had been reducing the market value of the mushrooms. This decision means that, in the United States, these mushrooms can immediately be cultivated and sold.

**KEY IDEAS**

- **The uptake of GM plants in commercial agriculture worldwide has increased since the first GM plants were approved.**
- **Worldwide, about 30 different countries have GM plantings, but 10 countries have 98 per cent of the total area under GM plantings.**
- **As at 2015, varieties of only two plant species, cotton and canola, have been approved for large-scale cultivation in Australia.**
- **In Australia, production of GM organisms is governed by Commonwealth and state legislation.**
- **Edible GM products cannot be purchased as fresh foods in Australia, but specific GM products have been approved for use in manufactured foods that are sold here.**
- **In the USA, gene-edited organisms produced using CRISPR-Cas9 technology do not come under GM regulations.**

**QUICK CHECK**

5 Identify the following statements as true or false:
  a Worldwide, GM crops are dominated by just four crops.
  b The major GM crop worldwide in terms of area under cultivation is soybean.
  c Australia is one of the top ten countries in terms of area of GM crops under cultivation.
  d The GM crops approved for large-scale cultivation in Australia are cotton and corn.
  e 'Stacked' refers to a GM organism with two or more genetically engineered traits.

6 Who is responsible for the administration of the *Gene Technology Act 2000*?

## Other GM plant crops

Examples of GM crops are outlined below and, in some cases, an outline of the method of creating the GM plant is given. GM crops are developed for various reasons. In the previous section, we saw that the major reason was pest control. In this section, you will see different reasons for the genetic modification of plants, such as enhanced nutritional value in the case of Golden Rice, improved shelf life in the case of tomatoes, and the creation of new ornamental flowers in the case of blue carnations.

1. **GM tomatoes:** In May 1994, the first genetically engineered food licensed for human consumption came on the market in the United States. This was a genetically engineered tomato developed by the Calgene company in which the **PG** gene encoding the enzyme polygalacturonase was silenced. This enzyme, also known as pectin depolymerase, degrades the pectin in plant cell walls, causing the tomatoes to become soft as they ripen. Typically, tomatoes are harvested as hard green fruit and are ripened artificially. In contrast, these GM tomatoes could be picked after ripening on the vine, had longer shelf life and had a flavour that reflected the on-vine ripening of these tomatoes.

The silencing of the PG gene was achieved by inserting another gene that produced mRNA that was complementary to the mRNA produced by the PG gene — this mRNA is termed **antisense mRNA**. Because they are complementary, the antisense mRNA pairs with the normal mRNA and stops it from being translated, stopping production of the softening enzyme (see figure 15.15).

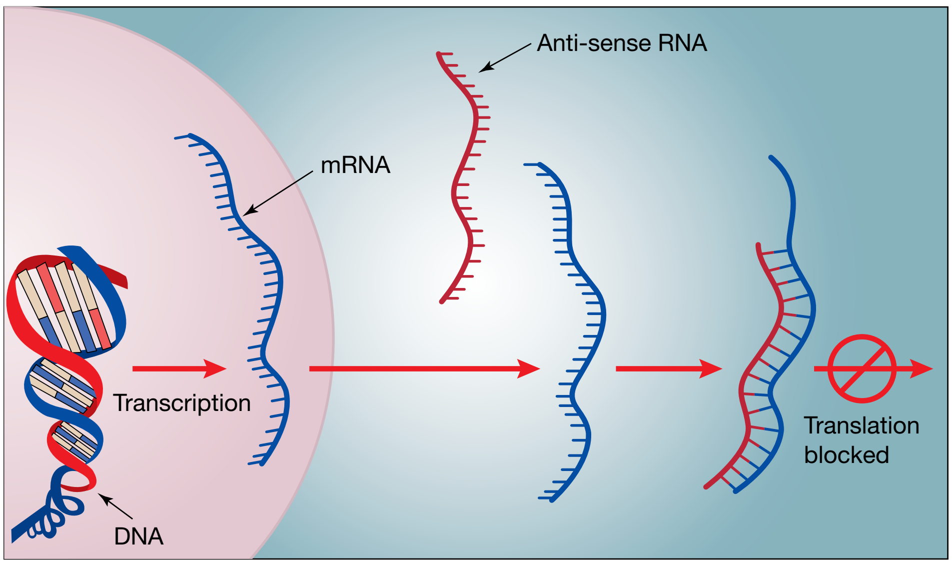



**FIGURE 15.15** Diagram showing a simplified version of gene silencing through the use of antisense mRNA.

Marketed under the label of FlavrSavr, these GM tomatoes were generally positively received by most US consumers, but production and distribution costs exceeded income and the venture was not a commercial success. In 1996, the Calgene company was bought out by Monsanto and the FlavrSavr tomato program was closed down.

Genetic engineering of tomatoes has more recently been directed at silencing genes that are induced by the ripening hormone ethylene. Ethylene is the signalling molecule that initiates the expression of various genes involved in the ripening of many kinds of fruit. Over-ripening and the associated softening of these fruits causes loss after they are harvested, and limits their shelf life. Suppression of these genes using a different technique known as RNA interference (RNAi) has been shown to delay ripening, reduce the rate of softening and extend shelf life. See the box on the next page to check out a technique of gene silencing by RNA interference (RNAi).

## SILENCING GENES THROUGH RNA INTERFERENCE

The first clue that it might be possible to switch off or turn down genes came in the 1980s. Scientists tried to intensify the natural purple colour in petunias by adding double-stranded RNA (dsRNA) to petunia cells where one strand of this RNA matched the mRNA transcribed by the pigment-controlling petunia gene. To the scientists' surprise, instead of becoming deeper purple, most of the petunias treated in this way turned white. At that time, this result could not be explained.

In 1998, researchers discovered that adding double-stranded RNA (dsRNA) to cells caused some genes to be silenced. The silenced genes were those whose mRNA transcript matched one of the strands of the added dsRNA. This finding gave biologists a means of switching off one or more specific genes in cells. This method of silencing genes is called **RNA interference** (RNAi).

RNA interference does *not* act directly on the DNA of genes but works by breaking down the mRNA produced by one specific gene, leaving all other genes unaffected. RNA interference acts through short fragments of RNA, called small interfering RNA (siRNA).

### How does RNA interference work?

The process of switching genes off by RNA interference (see figure 15.16a) involves the following:

1. Long synthetic double-stranded RNA (dsRNA) that is complementary to the mRNA to be silenced is added to a cell.
2. An enzyme, known as Dicer, cuts this dsRNA into short RNA fragments called **small interfering RNA** (siRNA).
3. These siRNA fragments combine with a complex of cellular proteins to form an RNA-induced silencing complex (RISC). At this point, the siRNA is separated into its two strands, and only the guide strand is retained.
4. The guide strand locates the mRNA target. The specific targets recognised by the RISC complex are any mRNA molecules with a base sequence that is complementary to the guide strand of the siRNA.
5. The target mRNA is broken down by the RISC complex. Degrading this mRNA means that the protein encoded by the gene concerned cannot be produced and so that gene is effectively silenced or knocked out.

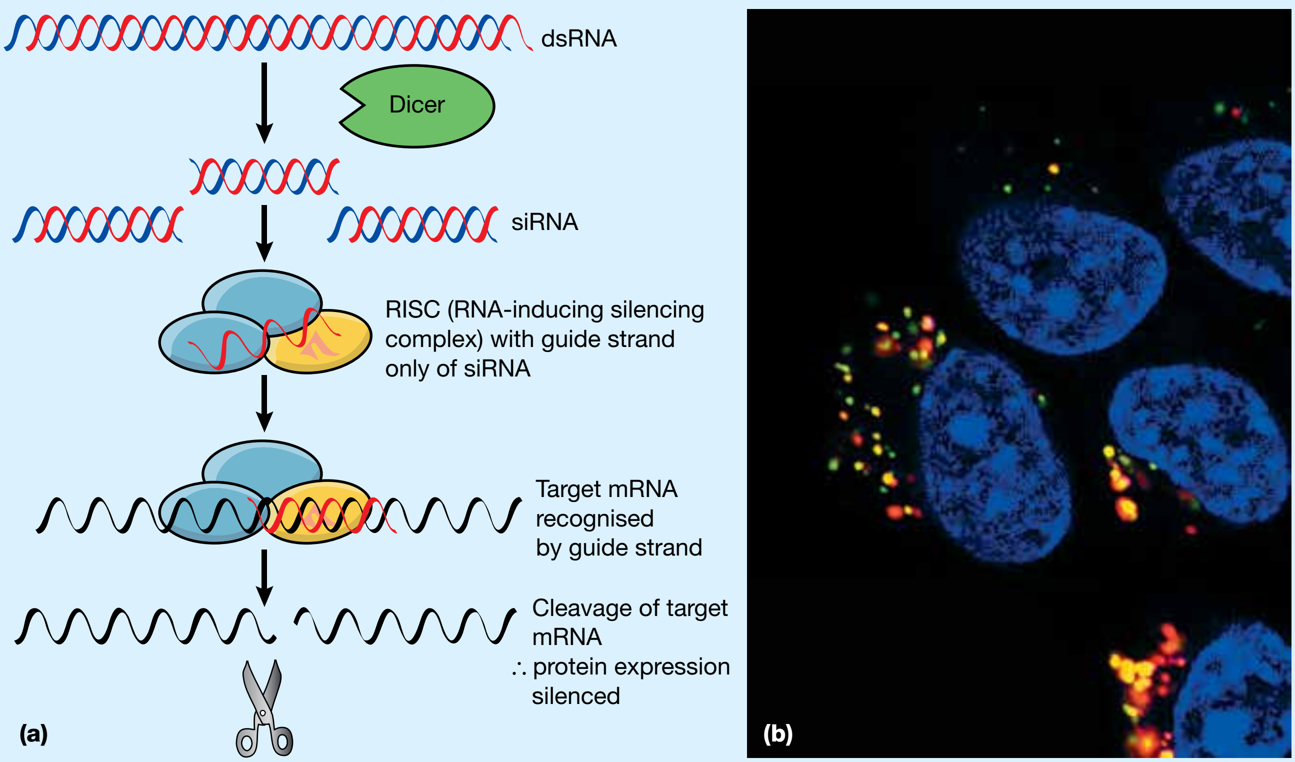



**FIGURE 15.16** **(a)** Diagram showing a simplified version of RNA interference (RNAi). **(b)** RNAi in human cells. The nucleus is shown by its blue fluorescent label. siRNAs are revealed by fluorescent labels, with the guide strand of each siRNA carrying a red label and the other strand carrying a green label. (Where the two strands overlap, a yellow signal results.) The red guide binds to the target mRNA by complementary base pairing, enabling the RISC to destroy the mRNA. (Image (b) courtesy of Tariq M. Rana, PhD, University of California, San Diego.)

**FIGURE 15.17** Carnations *(Dianthus caryophyllus)* with blue colour resulting from gene manipulation.

**2. Pineapples:**
In Australia, when the growing temperature of pineapples (*Ananas comosus*) falls below 25 °C by day and below 20 °C at night, a condition known as blackheart develops, in which the centre of the pineapple blackens. This condition can also arise post-harvest if pineapples are exposed to low temperatures during storage, handling and transportation. Pineapples with blackheart appear normal externally, and the condition is not detected until they are cut open for processing or for consumption. Blackheart is a cause of major economic loss to growers and to food processors.

Blackheart is due to the increased action of the enzyme polyphenol oxidase (PPO), which catalyses the production of quinones. The quinones can then react further to produce black pigments. One line of research on eliminating blackheart has involved the silencing of the **PPO** gene using RNA interference. An Australian company, Benitec, produced pineapple cells in which the gene that encodes this enzyme was silenced using RNAi and these cells were cloned to produce pineapples that do not blacken.

**3. Blue carnations:**
Blue is not a colour that is naturally produced in carnations. In 1995, the first GM blue carnations were approved for growth in Australia.

More recently, in October 2015, the Office of the Gene Technology Regulator (OGTR) granted a licence to International Flower Developments Pty Ltd, giving approval for importing three more types of GM blue carnations to Australia: Moonaqua™, Moonberry™ and Moonvelvet™. The licence does not permit the growing of these GM carnations in Australia. Figure 15.17 shows GM colour-modified carnations.

The colours seen in the petals of various species of flowering plants are due in part to the presence of the red and blue anthocyanin pigments, such as the red and pink cyanidin pigment and the blue delphinidin pigment. Another group of pigments prominent in some flowers are the yellow and orange carotenoids.

A simplified version of the pathway for production of the two anthocyanin pigments is shown in figure 15.18. While carnations have the A and C genes needed to produce red pigment and its precursor, they lack the B gene, which encodes the enzyme needed to produce the precursor of the blue pigment.

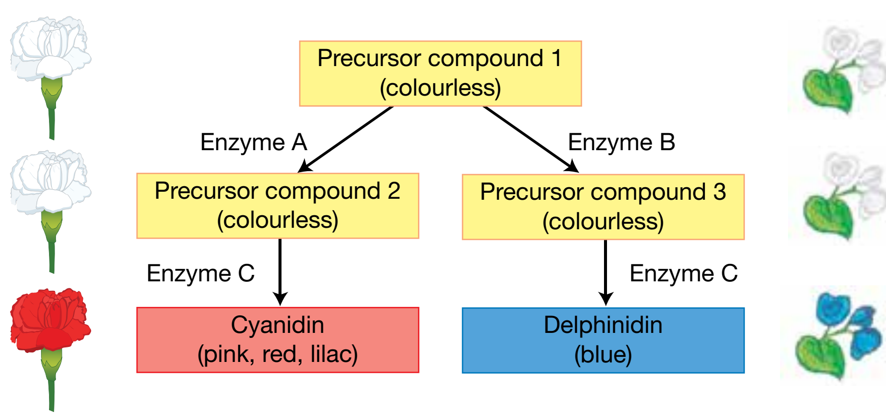



**FIGURE 15.18** Part of the pathway for production of two of the common flower pigments. Each step is catalysed by an enzyme whose production is controlled by a specific gene. Carnations lack the gene that encodes enzyme B.

**ODD FACT**

Anthocyanin pigments (red, pink, mauve and blue) are found in solution in the vacuoles of petal cells. A second major group of flower pigments are carotenoids (yellow, orange), which are found in special plastids, known as chromoplasts.

One way of producing a blue carnation is shown in figure 15.19 below:

1. Start with a red carnation.
2. Silence the gene responsible for producing enzyme C using RNA interference (RNAi).
   A carnation with this gene silenced is white. (Why?)
3. Add the pansy gene responsible for producing enzyme B. Even with this gene, the carnation will still be white. (Why?)
4. Add the iris gene encoding enzyme C that converts precursor compound 3 to blue delphinidin pigment. (Because the iris gene is slightly different from the homologous gene, it is not affected by the RNAi.)

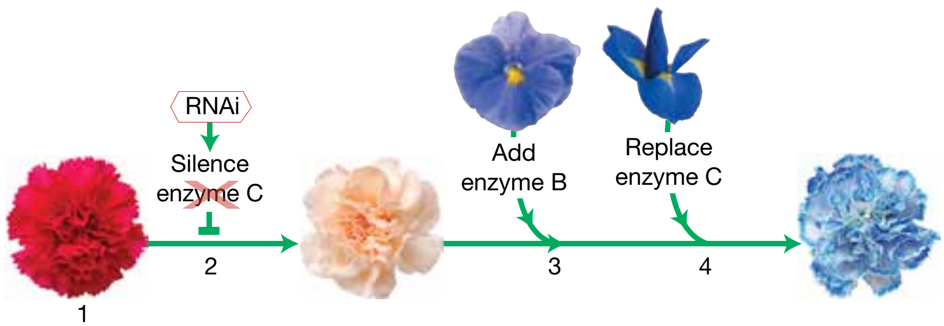



**FIGURE 15.19** One method of producing a blue carnation.

### 4. Golden Rice:

The World Health Organization (WHO) has identified vitamin A deficiency as one of the major causes of preventable blindness in children. Vitamin A deficiency also increases the risk of disease and death from infections, particularly in the poorest segments of populations in low- and middle-income countries. Vitamin A deficiency is the lack of an adequate intake of vitamin A in the diet. The direct sources of vitamin A in the human diet are animal products, such as egg yolk, butter, cream, cod liver oil and liver. However, plants, such as carrots, leafy green vegetables, sweet potatoes and cantaloupes, are also the source of beta-carotene, which, in the human body, is converted to vitamin A.

The normal role of the carotenoids is that of accessory photosynthetic pigments (refer back to *Nature of Biology Book 1 Fifth Edition,* page 100) found in the leaves of green plants. Rice plants produce beta-carotene in their photosynthetic leaf cells, but the genes concerned are not active in the cells of rice seeds (grains). This means that a serve of rice does not provide any vitamin A precursor to the person eating it.

**Golden Rice 2** is a strain of rice (*Oryza sativa*) that has been genetically modified to produce beta-carotene, a precursor of vitamin A, in the endosperm of the rice grain. Golden Rice is so named because it has a distinctive golden colour that is due to the presence of high levels of beta-carotene (see figure 15.20). Eating a serve of Golden Rice 2 provides beta-carotene, which can be converted to vitamin A. The development of Golden Rice does not involve any multinational companies; it has been developed as a public-good project.

The genome of the Golden Rice variety has been genetically engineered by the insertion, into the rice genome, of two additional genes in the pathway for the biosynthesis of carotene and the addition of a promoter sequence that keeps this pathway active in the cells of the rice endosperm. The two genes added are:

1. the **PSY** gene from daffodil (*Narcissus pseudonarcissus*), which encodes the enzyme phytoene synthase
2. the **CRTI** gene from the soil bacterium *Pantoea ananatis*, which encodes the enzyme phytoene desaturase; this enzyme catalyses multiple steps in the synthesis of carotenoids up to lycopene.

**FIGURE 15.20** Standard rice (left) and Golden Rice (at right). The golden colour is due to the presence of high levels of beta-carotene: the more intense the golden colour, the higher the concentration of beta-carotene. Each grain is a seed that contains the rice embryo (rice germ), but the bulk of the rice seed is the endosperm, a food store for the developing embryo. Once in the body, beta-carotene is converted to vitamin A. (Images courtesy of the Golden Rice Humanitarian Board.)

**ODD FACT**

Chinese researchers have shown that one bowl of cooked Golden Rice (50 g dry weight) provides about 60 per cent of the Chinese Recommended Nutrient Intake of vitamin A for children aged six to eight years.

Lycopene is then converted to beta-carotene by another enzyme, lycopene cyclase, that is already present and in the rice. Figure 15.21 shows a simplified outline of the pathway for the synthesis of beta-carotene in rice endosperm. What happens to beta-carotene when this rice is eaten as part of a meal?

The Golden Rice 2 variety holds significant potential to assist in reducing the incidence of vitamin A deficiency in some population groups. However, Golden Rice 2 is not presently under cultivation because of regulatory issues and, in some countries, its adoption faces strong public opposition.

**20-C precursor**

↓ ***Phytoene synthase*** encoded by PSY daffodil gene

**phytoene**

↓ ***Phytoene desaturase*** encoded by CRTI bacterial gene

**zeta-carotene**

↓ ***Phytoene desaturase*** encoded by CRTI bacterial gene

**lycopene**

↓ ***Lycopene cyclase*** encoded by rice gene

**beta-carotene**

**FIGURE 15.21** Diagram showing a simplified account of the activated beta-carotene synthesis pathway in the GM rice known as Golden Rice 2. Substrates are shown in black and enzymes in red italics.

## BIOLOGIST AT WORK

### Dr Alex Johnson — Plant Molecular Biologist

Dr Alex Johnson is a research scientist and Senior Lecturer in the School of BioSciences at The University of Melbourne. Alex writes:

'Rice and wheat provide 20 per cent and 19 per cent of the world's dietary energy supply respectively, yet low concentrations of iron in these widely consumed cereal grains frequently cause debilitating malnutrition disorders in humans. Over two billion people, roughly 30 per cent of the world's population, are affected by iron deficiency with symptoms ranging from poor mental development in children to iron deficiency anaemia. Many of these people live in developing countries and rely on cereals as staple foods. However, iron deficiency anaemia is also a major public health problem in developed countries such as Australia, where an estimated 8 per cent of preschool children, 12 per cent of pregnant women and 15 per cent of non-pregnant women of reproductive age are affected. The development of new cereal varieties containing increased concentrations of iron and other essential micronutrients in the grain, an approach known as biofortification, is an economical and sustainable approach to increase micronutrient intakes for humans around the world.

Genetic modification (GM) technology will be needed to develop most, if not all, of the world's biofortified cereals due to inherently low levels of several micronutrients, such as iron and zinc, in cereal grain and the complete absence of others (such as pro-vitamin A). For many years, my research team at Melbourne has been using GM technology to increase the activity of a group of rice genes called nicotianamine synthase genes. This approach has led to striking results — we have produced GM varieties of rice that have fourfold more iron and twofold more zinc in the grain. The iron- and zinc-biofortified rice plants also grow vigorously on calcareous (high pH) soils where micronutrients such as iron and zinc are normally limiting for plant growth. We are now working to produce iron and zinc-biofortified wheat plants using similar GM technology. A number of promising wheat lines have been developed and we are conducting field trials of the GM wheat plants at several sites across Australia.

I have been fascinated by plants and how they grow for as long as I can remember. One of my earliest memories is of me following my mum around in her garden while she planted bean, watermelon and other seeds, and me digging up the seeds each day to check on them! As a result my mum often had terrible-looking gardens but my interest in plant biology only grew. I started my PhD in the States in the late 1990s when the first GM crops were being released and I knew from that moment that I wanted to work in this field. Twenty years on and I am still happy with my decision!'

**FIGURE 15.22** Alex Johnson in a field trial of biofortified wheat plants in South Australia. (Image courtesy of Dr Alex Johnson.)

**studyon**

| Unit 4 | **Transgenic organisms** |
|---|---|
| AOS 2 | Summary screen and practice questions |
| Topic 2 | |
| Concept 5 | |

## Animal GMOs

Genetic manipulation is being applied to animal species as well as to plants. Among projects underway are those aimed at improving yields and, hence, reducing costs, such as the AquAdvantage® salmon.

### 1. GM salmon

On 19 November 2015, the United States Food and Drug Administration (FDA) gave approval for the first genetically engineered animal to be sold as food. The GE animal approved for human consumption is a transgenic Atlantic salmon (*Salmo salvar*), designated as AquAdvantage® salmon.

The transgenic features of this salmon were the result of a DNA construct that was inserted into the salmon genome. This DNA construct included (i) a gene encoding growth hormone from the Chinook salmon *Oncorhynchus tshawytscha*, the largest of the various species of Pacific salmon, and (ii) a promoter gene and a terminator sequence from an eel-like fish, the ocean pout (*Zoarces americanus*) (see figure 15.23). The promoter gene acts on the growth hormone gene so that it is expressed continuously. (In contrast, the Atlantic salmon's own promoter is only expressed intermittently in response to signals such as day length and temperature.)

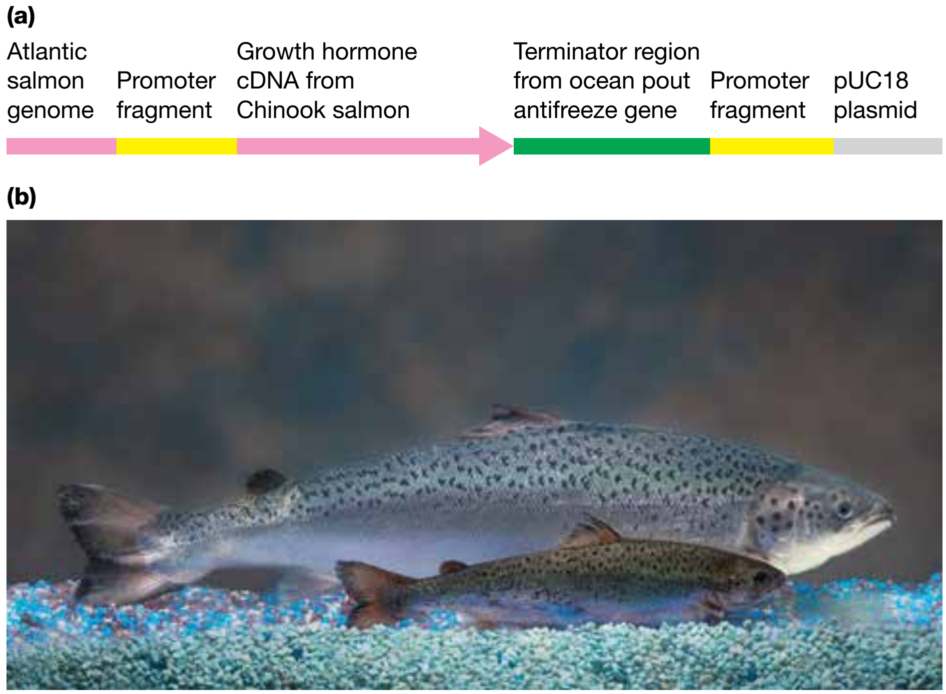



**FIGURE 15.23** **(a)** Diagram (not to scale) showing the arrangement of the DNA construct that was inserted into the genome of Atlantic salmon eggs to start the AquAdvantage® GM salmon. The short segment of the pUC18 plasmid included was non-coding DNA. **(b)** Size comparison of an AquAdvantage GM Salmon (background) vs a non-transgenic Atlantic salmon sibling (foreground) of the same age. Both fish reach the same size at maturity but the non-transgenic salmon will take twice as long to grow to the mature size. (Image courtesy of AquaBounty Technologies.)

Other modifications were made to these GM salmon to prevent any possibility of their interbreeding with native populations of wild salmon. These features included making the GM eggs from which GM salmon develop: 1. all-female, 2. Triploid, with three copies of each chromosome, and 3. Sterile, because meiosis cannot produce balanced gametes in a triploid organism. A physical measure to prevent any possibility of escape of these GM salmon to the ocean was imposed by the FDA. A condition of FDA approval was that the GM salmon be grown in contained land-based facilities, not in the net-fenced pens that are commonly used in coastal waters. This land-based contained setting provides an opportunity for the salmon to be reared in waters that are continuously cleaned and kept free of environmental contaminants.

Data of comparative growth rates for the GM and non-GM salmon are shown in table 15.3. Note that the GM salmon reaches maturity more quickly than the non-GM salmon and so can be harvested sooner (see figure 15.24). It is estimated that the time to market was reduced from three years to 18 months.

**ODD FACT**

Triploid salmon are produced by exposing salmon eggs to high pressure, a treatment that causes the offspring to receive the usual one set of chromosomes via the sperm but two from its mother's egg because the polar body of the egg is retained rather than being ejected.

## 'Pharming' to make clinical proteins

Human genes have been engineered into mammalian cells, such as hamster cells and mouse cells. Human genes have also been engineered into mammalian clones, such as cattle, sheep and goats. These gene transfers are not possible under normal evolutionary processes. Why has this been done? This was done with the aim of using the transgenic mammalian cells or the transgenic mammals to make the protein products of these genes. Let's look at two case studies.

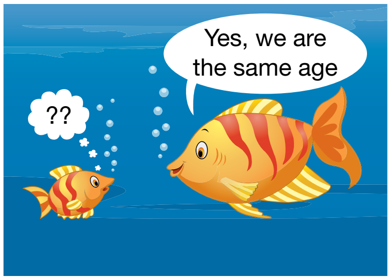



FIGURE 15.24

**TABLE 15.3** Table showing the comparison of the average times taken to reach a given weight (in kilograms) for GM salmon and non-GM salmon. These data show the significant reduction in time taken to reach market size for the GM salmon. Could you show these data in a graph?

| Approx. time (days) to reach weight | | | |
|---|---|---|---|
| Weight | GM fish | non-GM fish | Difference |
| 1 kg | 350 | 570 | 220 |
| 2 kg | 490 | 670 | 180 |
| 3 kg | 550 | 760 | 210 |
| 4 kg | 600 | 840 | 240 |

## Case study 1: transgenic hamster cells as factories

Andrew and Diane were keen to start a family but, after 18 months of trying without success to conceive, the young couple sought medical advice. Andrew was found to have a normal sperm count and sperm motility but Diane's tests showed that she was not producing adequate levels of follicle-stimulating hormone (FSH) and so was not ovulating.

Decades ago, Diane would have been treated with hormone isolated from the urine of menopausal women. However, in 1994, FSH that was produced using recombinant DNA techniques was approved for use in several countries. In Australia, it is marketed as Puregon®. Recombinant FSH (recFSH) is available in larger quantities, its supply is more reliable and it is purer than the urine-derived product because it does not contain the extraneous materials that were often present in the latter.

The production of human recFSH occurs in transgenic hamster cells. These cells have been genetically engineered so that they contain the human gene that controls production of follicle-stimulating hormone.

**ODD FACT**

In the United States on 2 October 1992, twins Karine and Cedric were born. This birth was the first ever to a woman who had been treated with recFSH.

## Case study 2: transgenic mammals as factories

Some research has involved the production of transgenic cows and goats whose cells carry human genes for specific protein products. If these genes are expressed in the milk produced by these transgenic mammals, they function as mini-pharmaceutical factories to produce particular protein products.

George and Charlie are cloned transgenic cows with the gene for a human protein known as serum albumin. These transgenic cows have expressed this gene in their milk. It has been estimated that one transgenic cow can produce about 80 kilograms of this human protein per year, and so herds of transgenic cows could become an important source of this protein, which would otherwise have to be obtained from human blood donations.

Likewise, the cloned goats Mira, Mira and Mira have produced milk containing a human protein that prevents clotting, which is used in patients undergoing heart surgery. Figure 15.25 shows an outline of the procedure for producing a human protein from transgenic goats. Note that the human gene of interest to be transferred is not inserted as a 'naked' structural gene. Instead, the structural gene is transferred as part of a DNA construct that includes promoter, terminator and other regulator DNA sequences.

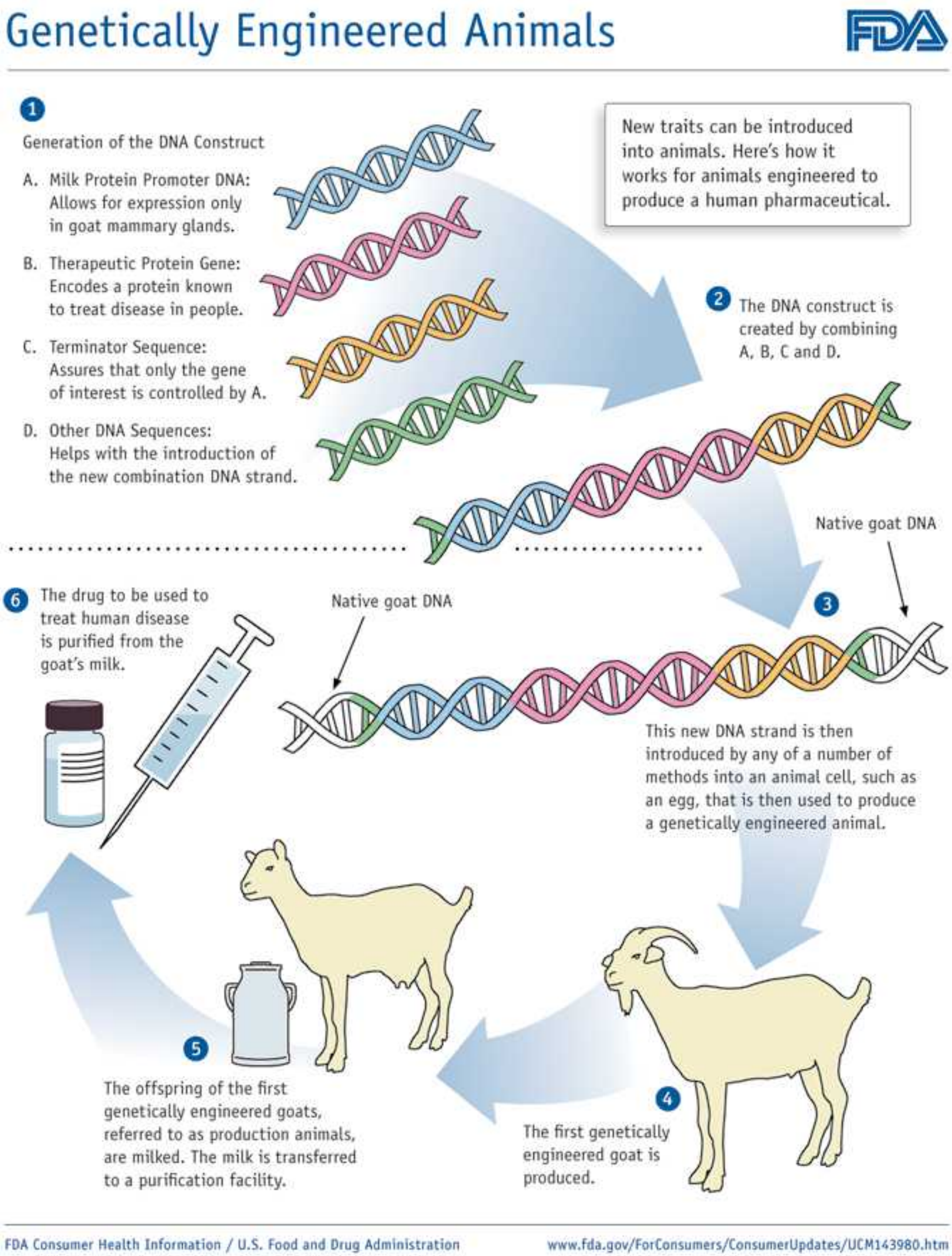



**FIGURE 15.25** (Image courtesy of the US FDA.)

# Debate about GMOs continues

Genetic engineering continues to be the subject of public debate, centring around ethical concerns, ecological concerns, and economic issues raised by the fact that GM techniques and GM organisms are subject to intellectual property law. The various interest groups with a voice in this debate include consumers, farmers, scientists, biotechnology companies, food processors and retailers, government regulators and non-government organisations. Each group will have its own perspective on the various issues, which may not necessarily be shared by other groups. For consumers, an important issue has been the need for explicit labelling of GM foods, while other groups may regard this as an unnecessary imposition.

In regard to GM plants in agriculture, debate has focused on various issues:

### Ecological issues

- Is there the possibility that the insecticides genetically engineered into plants will kill beneficial non-target insects as well as the target pest insects? It is noted that the **Bt** insecticide must be ingested for its effect to be realised. However, as more insecticides are added to GM plants, will these insecticides also require ingestion to realise their effects?

  Although the Bt pesticide can protect crops against unwanted insects, it may have the potential to have unintentional effects on neutral or even beneficial species.
- Is there the possibility that insect pests will become resistant to the Bt insecticide over time, leading to the need for the use of stronger and more persistent chemicals for pest control in the future? Evidence is already emerging that some of the insect pests targeted by the Bt insecticide are developing resistance to it.

  It may be noted that GM plants are increasingly being stacked with multiple traits, and these new varieties have replaced the GM plants with single traits, such as pesticide resistance or herbicide tolerance.
- Are there issues relating to gene flow between species? The potential for unintended gene flow between GM crops and related plants in the wild is a matter of concern to ecologists. In many cases, it is possible for transplanted genes to spread through the normal cross-pollination of plants. Should this occur, cross-fertilisation could result in the creation of herbicide-resistant so-called 'superweeds.'

### Social impacts

- What are the long-term social impacts of increasing ownership of GM seeds and patents for genes by a small number of global companies? Large biotechnology companies, such as Monsanto, sell the herbicide-tolerant seeds and also own and sell the herbicide concerned. Farmers who buy Monsanto seeds cannot collect seeds for the following year's crop, but must buy more seeds. Does this place the benefits of GM beyond the reach of poor farmers in developing countries? Does the large-scale adoption of a few crops pose new risks to food security?

### Ownership of intellectual property issues

- In the past, some developing nations accused major biotechnical companies of genetic 'piracy,' claiming that company employees collected genetic material in these countries without permission, took the material back to their laboratories in industrialised countries, and genetically modified and patented the genetic material. In response to concerns related to intellectual property rights, the international Convention on Biological Diversity was negotiated in 1992. This convention legally guarantees that all nations, including developing countries, have full legal control of 'indigenous germplasm.'

### Health and safety issues

- The need for thorough and rigorous testing is a key issue for consumers.
- GMOs give the possibility of introducing additional nutrients into foods, as well as antibiotics and vaccines. This technology can provide nutrition and disease resistance to those countries that don't have the means to provide these otherwise. The distribution of these foods is more feasible than mass inoculations for current diseases. However, this promise carries with it potential negative effects such as the creation of antibiotic-resistant microbes.

### Cultural issues

- These have been prominent in the debate relating to Golden Rice 2 in the Philippines. The question asked by the public is 'Why impose a GM rice on us when we have plenty of natural fresh foods that are sources of vitamin A?' The public fears that large-scale adoption of GM crops to address dietary deficiencies poses a threat to biodiversity. Alternative views are that Golden Rice 2 can prevent major incidence of childhood blindness and so this rice should be widely cultivated.

These issues and others relating to the uses of GM technology will continue to be discussed. What are your views on these issues?

## KEY IDEAS

- **Gene manipulation may be carried out for a variety of purposes.**
- **GM plants containing particular genes that are silenced have been produced using antisense mRNA and RNA interference techniques.**
- **GM animals have been produced through the insertion of DNA constructs carrying a structural gene of interest and relevant promoter genes.**
- **Issues, such as social, ecological, and legal ones, have arisen in relation to various aspects of genetic engineering.**

## QUICK CHECK

**7** Identify the following statements as true or false:

**a** The use of gene technology raises debate, including ecological, legal and cultural issues.

**b** Genes can be silenced and GM organisms produced through the use of techniques involving antisense mRNA and RNA interference (RNAi).

**c** DNA constructs are used to knock out growth hormone genes in GM Atlantic salmon.

**d** The first GM animal has been approved as a food for human consumption in the USA and is a transgenic Atlantic salmon.

# BIOCHALLENGE

**1 Looking at GMOs worldwide:**

**eBook*plus***

**Weblink**
Database of GM approvals worldwide.

Use an internet search engine to research the International Service for the Acquisition of Agri-biotech Applications (ISAAA), in particular the potato: New Leaf™ Y Russet Burbank potato (*Solanum tuberosum*). Find the **RBMT15-101** entry, which details the genetic engineering of this potato variety.

- a What traits were engineered into this variety of potato?
- b How were the foreign genes concerned transferred into the potato cells?

One of the genes transferred to the potato cells was the **pvy cp** gene.

- c What was the source of this gene?
- d What is the function of this gene?

Now, research ISAAA to find the **EMBRAPA 5.1** entry to see the details of the genetic engineering of the bean (*Phaseolus vulgaris*).

- e What trait was engineered into the beans?
- f How was the foreign gene concerned transferred into the bean cells?

The gene transferred into the bean cells was the **ac1** gene.

- g What was the source of this gene?
- h What is the function of this gene?

**2 Looking at GM products in Australia:** The only genetically modified food crops grown in Australia are cotton and canola. Food Standards Australia New Zealand (FSANZ) allows manufacturers to use a wide range of GM food imported from overseas as ingredients in packaged foods. These include specific GM varieties of soybean, corn, potatoes and sugar beet. For example, **GM soybean products** are used as ingredients in many processed foods, including pastries, chocolates, potato chips, margarine and mayonnaise.

The list of food produced using gene technology and approved for use under the *Food Standards Australia New Zealand Act 1991* can be seen by searching online for the 'Record of GM Product dealings' at the OGTR (Office of the Gene Technology Regulator).

- a Identify the various processed food products from GM soybean that are approved for sale in Australia.
- b What condition was placed by FSANZ on the use of high oleic soybean oil or protein meal?

Check out the three approved GM potatoes (*Solanum tuberosum*).

- c What features have been genetically engineered into these potatoes?
- d Could you buy these potatoes at your local greengrocer or find them in the vegetable section of a supermarket?

Now, compare this with the three approved GM canola (*Brassica napus*) lines.

- e What features were genetically engineered into the GM canola plants from which the listed canola products are derived?

Check out the corn (*Zea mays*) products.

- f What is the difference between Yieldgard corn and Bt-11 corn?

# Chapter review

studyon

Unit 4 | AOS 2 | Topic 2

Biological knowledge and society

Practice questions

## Key words

antisense mRNA
bar
Bollgard II®cotton
Bt
CRTI
cry1Ac
cry2Ab
genetically modified (GM) foods
genetically modified organisms (GMOs)
Golden Rice 2
gox
Ingard®cotton
InVigor® canola
non-GMO
PG
PPO
protoplast
PSY
RNA interference
Roundup Ready® canola
small interfering RNA
transgenic organisms

## Questions

1 **Making connections** → Construct a concept map for 'genetically modified organisms' using any of the key words above. Add other terms as needed.

2 **Applying your understanding** →
  a Identify the essential characteristic of a transgenic organism.
  b Of the following, which are examples of transgenic organisms or cells:
    i mammals produced by cloning
    ii plants produced by artificial pollination
    iii wheat cells with a mouse gene
    iv cow cells with a gene mutation?

3 **Interpreting graphical data** → Examine figure 15.26, which provides data for four crops. The total area under cultivation worldwide in 2015 for each crop is shown grouped into two categories: (i) under cultivation with GM plants and (ii) under cultivation with non-GM (conventional) plants.

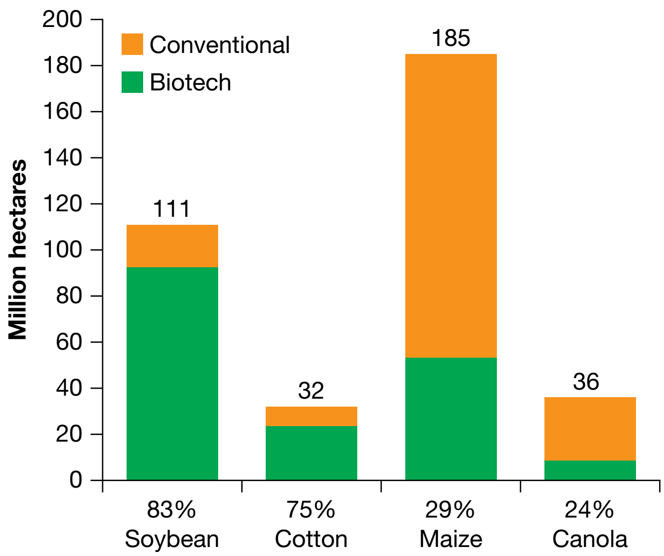



**FIGURE 15.26** Total area worldwide under cultivation in 2015 for four crops and the percentage of each crop that comprises biotech (GM) plants.
(Source of data: ISAAA.)

  a Which crop covers the largest area worldwide?
  b Which crop covers the largest area of GM plants?
  c Which crop has the highest percentage of GM plants?
  d What is the area in hectares under cultivation with GM soybean crops?
  e Given that the area of the playing surface at the Melbourne Cricket Ground (MCG) is 7.5 hectares, approximately how many MCG playing fields are under cultivation with GM canola crops worldwide?

4 **Developing logical explanations** → Consider the following observations and give an explanation to account for each:
  a Inserting a foreign gene into the genome of an organism may unintentionally cause unexpected harmful effects.
  b Individuals who were allergic to Brazil nuts suffered an allergic reaction when they ate GE soybeans into which a Brazil nut gene had been incorporated.
  c GM crops are more extensively tested than non-GM varieties before release, both for their environmental effects and as foods.
  d DNA constructs used to genetically engineer an organism would be expected to include the gene encoding the protein of interest, a promoter and a terminator.
  e In the USA, organisms genetically modified using the CRISPR-Cas9 technology are not subject to the GM regulation requirements of the US Department of Agriculture (USDA).

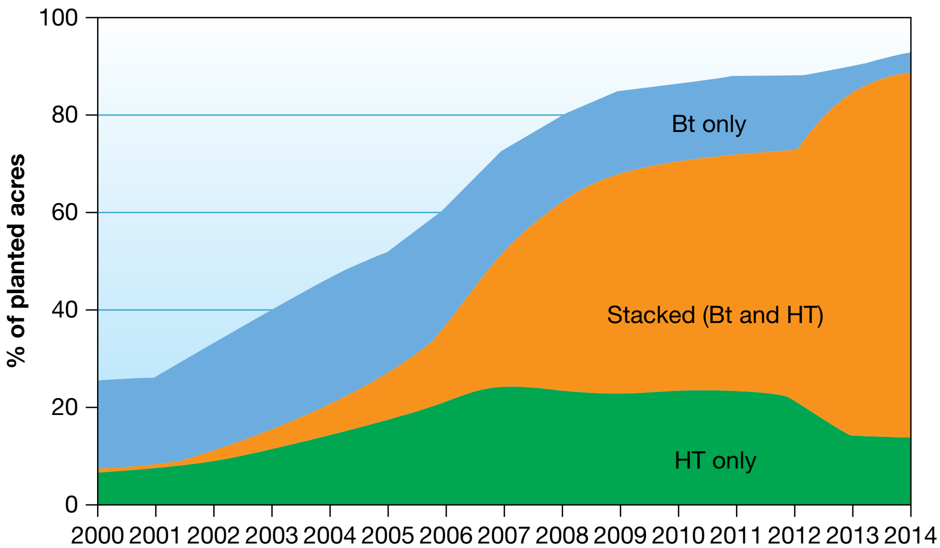
5 **Identifying critical differences** → What is the essential difference between the members of the following pairs:
  a a transgenic organism and a gene-manipulated organism
  b physical uptake of DNA and *Agrobacterium*-mediated uptake of DNA
  c insect resistance and herbicide tolerance
  d Ingard®cotton and Bollgard II®cotton?

**6 Demonstrating and communicating knowledge →** Briefly explain the role of the following:
- **a** the T-DNA segment of the *Agrobacterium* plasmid
- **b** the **gox** gene in Roundup-tolerant plants
- **c** the **Bt** protein in Ingard®cotton
- **d** RNA interference (RNAi)
- **e** the DNA construct in AquAdvantage® Atlantic salmon.

**7 Demonstrating knowledge →** Which of the following options could produce useful quantities of protein through genetic engineering?
- **a** bacteria containing recombinant plasmids
- **b** yeast carrying foreign genes
- **c** transgenic plants
- **d** transgenic mammals producing substances in their milk
- **e** all of the above.

**8 Interpreting data →**
- **a** What story is told by figure 15.27, which refers to the cultivation of GM corn in the United States?
- **b** Check back on Bollgard® 3 XtendFlex™ cotton. What level of stacking exists in this cotton variant?

**9 Discussion question →** Discuss the following statements with your classmates. Indicate whether, as a group, you agree or disagree with each statement, giving valid reason(s) for your decisions:
- **a** 'The use of genetically engineered crops reduces food security in the long run by decreasing biodiversity.' (Careful! Could the same be true of non-GM crops?)
- **b** 'Genetic engineering is a precise and predictable science.'
- **c** 'Genetic engineering is a minor variation on traditional breeding practices.'
- **d** 'Genetic engineering will cure world hunger.'




(Source: USDA, Economic Research Service using data from USDA, National Agricultural Statistics Service, June Agricultural Survey.)

**FIGURE 15.27**

CHAPTER

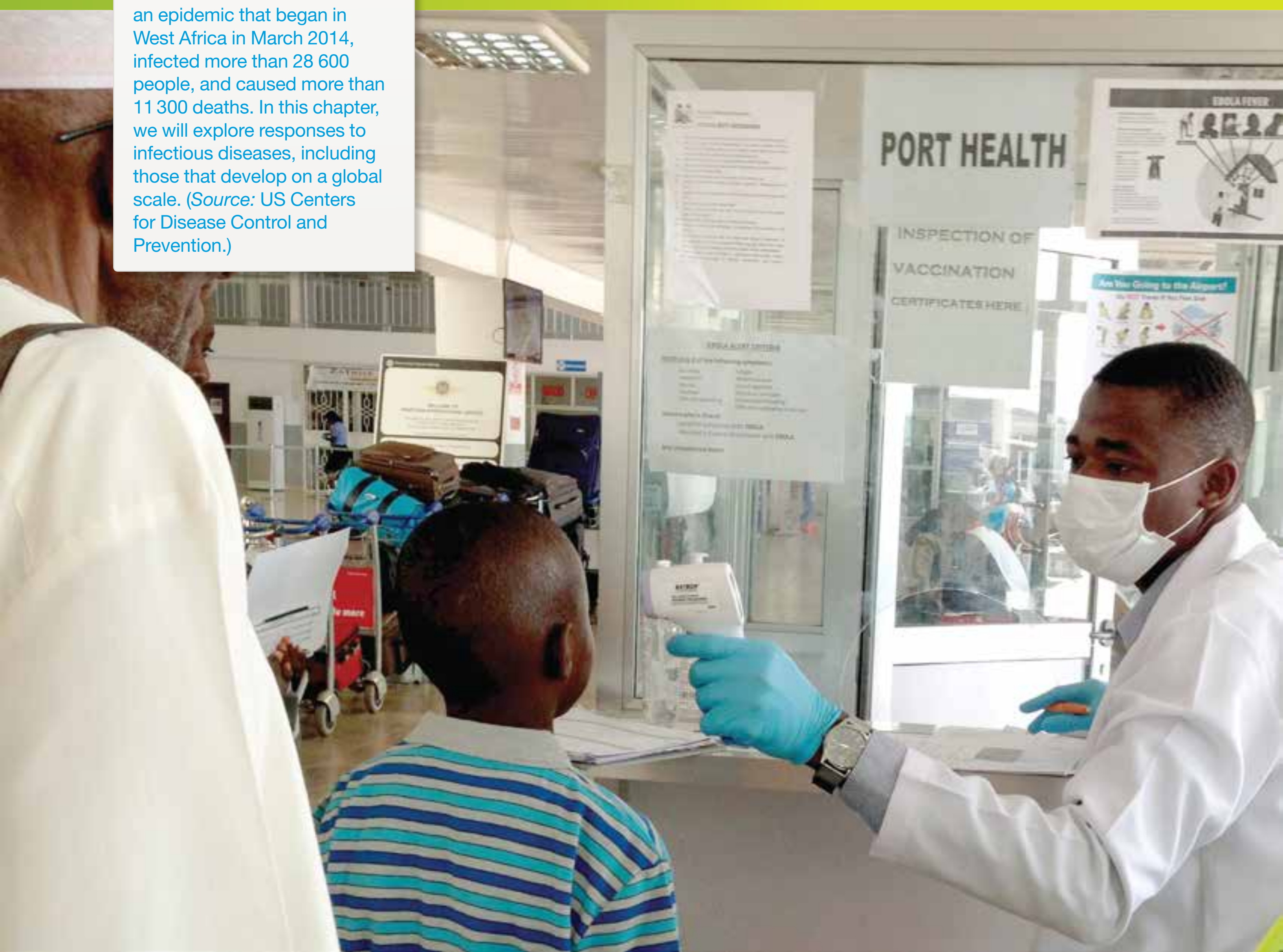
# 16 Biological knowledge in response to disease

FIGURE 16.1 A doctor using a non-contact thermometer to screen incoming passengers at the main airport of Sierra Leone in January 2015. Note the PPE (personal protective equipment) — face mask, lab coat and gloves — worn by the doctor. This screening program was one of many measures introduced to limit transmission of the Ebola virus disease. This disease became an epidemic that began in West Africa in March 2014, infected more than 28 600 people, and caused more than 11 300 deaths. In this chapter, we will explore responses to infectious diseases, including those that develop on a global scale. (*Source:* US Centers for Disease Control and Prevention.)

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- distinguish between pandemics and epidemics
- become familiar with techniques for the identification of pathogens
- develop knowledge of how viruses can change through mutation and re-assortment
- develop an awareness of strategies to deal with the emergence of new diseases
- gain knowledge and understanding of the concept of rational drug design
- understand the mode of action of chemical agents against pathogens
- identify the differences between antibiotics and antiviral drugs.



# Death in war and peace

The year 1918 marked the end of World War I (WW I). Peace came with the formal surrender of Germany on 11 November 1918 and the signing of an armistice — an agreement to stop fighting. The signing took place in a railroad car parked in a forest in France near the front lines. This global war resulted in the deaths of almost nine million people.

In 1918, during the final stages of WW I, another battle began. This time it was not between troops from warring nations. The enemy on this occasion was an influenza virus that caused a disease known as the Spanish flu. This virus raged during the cold months of 1918 and into 1919, spreading through America, Europe, Russia and even reaching Australia in early 1919, probably brought here by soldiers returning home from the war.

An estimated 500 million people worldwide were infected by this influenza virus, with the death toll estimated at up to 40 million, a total far greater than all the deaths resulting from World War I. This was the deadliest pandemic in human history to that date.

Although the disease is called Spanish flu, the origin of the influenza virus that caused this pandemic is not known. It may have originated in a military camp in the United States with an outbreak of flu among soldiers who were soon to be deployed to Europe. The virus involved was a new mutation of the influenza virus against which people had little or no immunity. **Virulence** refers to the degree to which a pathogen, such as a virus, can overcome the body's immune defences and cause disease. Virulence of a virus may be measured by the fatality rates in identified cases of the particular viral disease. The 1918 influenza virus was highly virulent, with case fatality rates much greater than those in later influenza pandemics. One estimate identifies case fatality rates for Spanish flu as more than 2.5 per cent, compared to less than 0.1 per cent in other influenza pandemics.

**ODD FACT**

The name 'Spanish flu' came from the early and extensive deaths from this flu in Spain where it allegedly killed eight million people.

Two features of this influenza virus were unexpected:

- One unusual feature was the rapidity of onset of the disease. People healthy at the start of the day were seriously ill by that evening, and, in many cases, death occurred days or even hours after the appearance of the first flu symptoms. At its worst, the virus caused uncontrollable haemorrhaging that filled the lungs of infected people, causing them to drown in their own body fluids. Those who did not die within the first few days of contracting the disease often died of complications, such as pneumonia, that were caused by bacteria.
- Another feature was the fact that the majority of people who died were young healthy adults, rather than the young children, old adults and sick people with weakened immune systems who were the usual fatalities resulting from influenza infections. Figure 16.2 shows the death rates from influenza for five-year age cohorts in New South Wales for the two waves of influenza outbreaks in 1919.

**ODD FACT**

The case mortality rate for the Spanish flu varied widely. In Unites States Army camps where statistics were kept, case mortality often exceeded 5 per cent and sometimes exceeded 10 per cent. In the British Army in India, case mortality for white troops was 9.6 per cent, and for Indian troops it was 21.9 per cent.

## Spanish flu in Australia

The Spanish influenza virus may have reached Australia with soldiers returning home from Europe after the end of World War I. Total deaths from this virus in Australia exceeded 12 000, and it appears that Australia experienced a milder strain of the virus compared to the one that devastated the population in Europe and America.

When the first influenza outbreaks occurred in Sydney and Melbourne, containment strategies were introduced: cinemas, hotels, churches and schools were closed, public gatherings were banned, travel by public transport was restricted, streets were sprayed, people were required to wear masks in public, and state borders were closed. Normal life was disrupted. The numbers of infected people requiring admission to hospitals far exceeded the number of hospital beds available, and demands on medical personnel were overwhelming. Large numbers of temporary hospitals were set up in whatever spaces were available, such as at schools, courthouses, showgrounds and churches, often staffed by volunteers. In Melbourne, the Royal Exhibition Buildings in Carlton Gardens became a temporary hospital between February and August 1919 (see figure 16.3).

**ODD FACT**

In 1919, almost 40 per cent of Sydney's population had influenza, and more than 4000 people died. In some regions of Sydney, up to 50 per cent of all deaths were due to influenza.

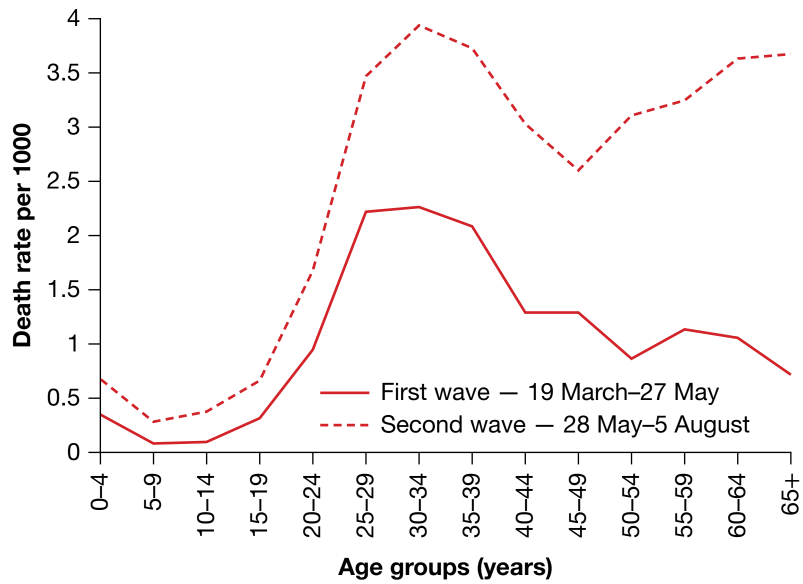



**FIGURE 16.2** Death rates for five-year age groups for the 1919 influenza pandemic in New South Wales. Note that the second wave of the pandemic was more deadly than the first wave. (Reproduced by permission, NSW Ministry of Health, 2016 — Influenza Report 1920.)

**FIGURE 16.3** Photograph showing part of the men's ward in the temporary influenza hospital set up in the Royal Exhibition Building in Melbourne in 1919. (*Source:* Museum Victoria, item 1602148.)

## Resurrecting the 1918 virus

The virus that was responsible for the 1918–1919 influenza pandemic naturally disappeared. Over the period from 1995–2005, scientists set about the task of reconstructing the virus under high-level conditions of biosecurity (see figure 16.4). The primary purpose of this research was stated as follows: 'to provide information critical to protect public health and to develop measures effective against future influenza pandemics'.

**FIGURE 16.4** A microbiologist in a Biosafety Level 3 -enhanced laboratory setting at the Centers for Disease Control and Prevention (CDC). He is using an electronic pipetter to extract reconstructed 1918 Pandemic Influenza virus from a vial. Note the use of an isolation chamber, gown, gloves, filtered respirator, mask and the filtration of exhaust air to prevent any possibility of accidental exposure of the laboratory workers or the public to the virus. (*Source:* CDC.)

To reconstruct the virus, scientists started by isolating fragments of viral RNA from preserved autopsy tissue of US soldiers who had died from Spanish flu during the 1918 pandemic. The scientists also succeeded in isolating further viral RNA fragments from an Inuit woman who was buried in the Arctic permafrost. From these various samples, the scientists were able to sequence the entire Spanish flu virus genome. Using the sequence data and a technique known as reverse genetics, the scientists reconstructed the virus. Among their findings was that the surface proteins of the Spanish flu virus were derived from a viral strain that normally infects birds, and that this virus infected the lower respiratory tract, producing acute respiratory distress.

This research was criticised by some scientists as 'dangerous' on the grounds that the benefits do not justify the risk to public health if this reconstructed virus was to escape from the laboratory, either accidentally or deliberately (see discussion question, page 763).

# Pandemic or epidemic?

WHO member states are grouped into six WHO regions: Africa, the Americas, South-East Asia, Europe, the Eastern Mediterranean, and the Western Pacific.

The terms 'pandemic' and 'epidemic' both relate to the uncontrolled spread of infectious diseases, but they differ in geographic spread. However, there is no sharp boundary between an epidemic and a pandemic, but an epidemic can develop into a pandemic, not vice versa. For the World Health Organization (WHO) to define an event as a pandemic, the infection must spread easily and sustainably among human populations in at least three countries in at least two different WHO regions.

WHO has a Department of Pandemic and Epidemic Diseases (PED) that '*develops strategies, initiatives, and mechanisms to address priority emerging and re-emerging epidemic diseases, thereby reducing their impact on affected populations and limiting their international spread*.'

## What is a pandemic?

A **pandemic** (from the Greek: *pan* = all, *demos* = people) refers to the global outbreak of a disease. Other definitions include:

- pandemic = an outbreak of a disease that occurs over a wide geographic area and affects an exceptionally high proportion of the population (Merriam-Webster)
- pandemic = the worldwide spread of a new disease (WHO).

Note that the term ‘pandemic’ refers to the spread of a disease, not to the severity of the disease — a pandemic could involve the global spread of a pathogen that causes a mild disease. However, the pandemics that attract greatest public interest and greatest media coverage are those in which the disease involved is severe and has a high death rate. The outbreaks between 1347 and 1400 of the bubonic plague or black death that swept across Europe, killing an estimated one-third of the population, were part of what is probably the best known pandemic (refer back to chapter 6, page 218). However, pandemics are not just historical events, as they occur frequently, and epidemics are even more common.

Examples of pandemics since the start of the twentieth century include:

‘Shifts’ and ‘drifts’ in influenza A viruses are discussed on page 740.

**The influenza pandemics** — 1918 (the Spanish flu), 1951 (the Asian flu), 1968 (the Hong Kong Flu) and the 2009 pandemic (11 June 2009 – 10 August 2010). Each of these pandemics resulted from a major change in the influenza A virus, known as a ‘shift’, that created a new or novel influenza A virus to which the population had no immunity.

**The HIV/AIDS pandemic** — AIDS was recognised as a new disease in 1981. Its origin traces back to a cross-species transmission to humans of a virus (SIV) that naturally infects non-human African primates. Since the beginning of the pandemic, almost 71 million people have been infected with the HIV virus and about 34 million people have died of HIV.

**Zika virus infections** — while not defined as a pandemic by WHO, the spread of Zika virus infections to almost 50 countries by mid-2016 signals the risk of a pandemic. In 2016, WHO declared the Zika virus to be a Public Health Emergency of International Concern. (See the box at the end of this section on the Zika virus.)

## Conditions for a pandemic

The conditions that favour the start of a pandemic include the following:

1. A new pathogen or a novel strain of an existing pathogen suddenly appears in geographic areas where the human populations have not previously come into contact with the pathogen.
   The suddenness of appearance means that:
   - most people will have little or no immunity to the pathogen
   - a vaccine is unlikely to exist or the vaccine is available in limited amounts that are not sufficient to prevent the spread of the pathogen.
2. The pathogen is the cause of an illness, often serious, in people. In many cases, a disease that can develop into a pandemic is caused by a pathogen that infects not only people but also other hosts, such as birds, pigs, bats and monkeys. These various non-human hosts act as reservoirs of the pathogen, moving from these to human hosts via a suitable vector.
3. The pathogen can be transmitted easily from person to person.
   This spread may occur:
   - by airborne particles coughed or sneezed by an infected person (as in influenza virus)
   - by contact with blood or other body fluids of an infected person (as in AIDS and Ebola virus fever)
   - via vectors, such as mosquitoes (as in Zika virus) or fleas (as in the plague), that transmit the pathogen from an infected person to an uninfected person.
4. Uncontrolled spread of the pathogen occurs across a wide geographic area. This spread may be facilitated by the movement of infected individuals from their home range and/or by the migration of infected vectors that transmit the pathogen from infected to non-infected individuals.

In 1918, the spread of the influenza virus was facilitated by the deployment of large numbers of soldiers from different countries to the battlefields of Europe from where it spread to Russia and even Greenland. By 1919, the spread of the virus was further facilitated by the demobilisation of soldiers who returned to their home countries.

**ODD FACT**

In 2014, commercial aircraft carried a record 3.3 billion passengers.

The mobility of people in the twenty-first century is much greater, owing to the massive increase in the movement of people by air. This movement creates the opportunity for pathogens to be exported by infected passengers in a day or less from one country where an infection is present, to a distant country. International travel means that people infected in one region can easily move thousands of kilometres and export the pathogen from that region to other countries. If the vector for the disease exists in the home country, the possibility of subsequent active transmission of the pathogen exists.

studyon
Unit 4 | AOS 2 | Topic 2 | Concept 6
**Epidemics and pandemics**
Summary screen and practice questions

Documented cases exist of travellers who spent time in countries where Zika virus infection is present and who inadvertently brought the virus back to their home countries, for example:

- In 2012, an Australian woman who returned home from Indonesia became ill and tested positive for Zika virus.
- The first case of Zika virus infection imported into Europe occurred in 2013 when a German man returned home with the Zika virus infection.
- The first case of Zika virus being imported to Canada was in 2014 with the return of an infected person from Thailand.
- By February 2016, the US Centers for Disease Control and Prevention in Atlanta, Georgia, reported the existence of 107 travel-related cases of Zika virus across various American states.

## What is an epidemic?

An **epidemic** (from the Greek: *epidemios* = prevalent; from *epi* = upon + *demos* = people) refers to the widespread occurrence of an infectious disease in a community or in a restricted geographic area at a particular time. Other definitions include:

- epidemic = outbreak of a disease or illness that spreads rapidly among individuals in an area or population at the same time (Dictionary.com)
- epidemic = the occurrence of more cases of a disease than would be expected in a community or region during a given time period (Medicine Net).

With any epidemic, a concern is that the disease may spread more widely and become a pandemic.

studyon
Unit 4 | AOS 2 | Topic 2 | Concept 7
**Strategies to deal with new diseases**
Summary screen and practice questions

Examples of epidemics in the early twenty-first century include:

- the SARS (severe acute respiratory syndrome) epidemic in China (2002–2003) caused by a new coronavirus that jumped from another species to humans; transmission of SARS occurred by respiratory droplets from coughs or sneezes of infected individuals
- the cholera epidemic in Haiti (2010–present), caused by the bacterial species *Vibrio cholera*, which is transmitted through faecal-contaminated water or food; these bacteria produce a toxin that binds to cells of the intestinal wall and interferes with the normal flow of water, sodium and chloride ions
- the Ebola virus disease epidemic in West Africa (2013 – March 2016); the Ebola virus is transmitted by direct contact of broken skin or mucous membranes with blood or other body fluids
- the yellow fever epidemic in Angola (January 2016–); as at June 2016, the number of suspected cases is 3137, including 847 confirmed cases with 345 deaths, throughout all provinces of the country. *Yellow fever* virus is an RNA virus that belongs to the genus *Flavivirus*, the same genus as the Zika virus. Yellow fever disease is transmitted by the bite of infected female mosquitoes that acquire the virus when they feed on infected people or infected monkeys.

## THE ZIKA VIRUS STORY

The Zika virus is usually transmitted to people through the bite of virus-infected female mosquitoes, mainly of the species *Aedes aegypti* and *Aedes albopictus*. Of particular concern is the observation that babies born to women infected by Zika virus early in pregnancy are at risk of suffering from **microcephaly** (= tiny head) (see figure 16.5). The Zika virus is also the cause of Guillain-Barré syndrome in adults, a disorder in which the body's immune system attacks a segment of the peripheral nervous system. Severe cases can result in paralysis and even death.



**FIGURE 16.5** (**a**) Poster produced by the World Health Organization (WHO) warning about the risk that Zika virus infections of pregnant women may cause microcephaly (= tiny head) in their babies. (*Source* for image (a): WHO.) (**b**) Comparison of a typical head size for a baby with a baby with microcephaly. (*Source* for image (b): CDC.)

The Zika virus is not a new virus. The Zika virus is a member of the *Flavivirus* genus of viruses, which are positive, single-stranded, enveloped RNA viruses that cause other diseases, including mosquito-borne yellow fever, dengue fever and Japanese encephalitis, and tick-borne encephalitis. The virus particle is shown in a transmission electron microscope (TEM) image in figure 16.6.

Zika virus was first identified in monkeys in the Zika forest in Uganda, Africa, in 1947. The first reported case of Zika infection in a person was in 1952. In a period of almost 60 years, from 1952 to 2007, a total of only 14 naturally occurring cases of Zika virus infections of people were identified, all in Africa and Asia. The associated mild symptoms included fever, rash and red eyes, and, in all cases, full recovery occurred. This evidence indicated that Zika virus caused a rare infection with mild flu-like symptoms.

However in 2007, all that began to change when the virus suddenly emerged on the island of Yap in the Federated States of Micronesia. Over a period of three months, the Zika virus spread, infecting about 5000 people, more than 70 per cent of the population. However, no serious complications were reported.

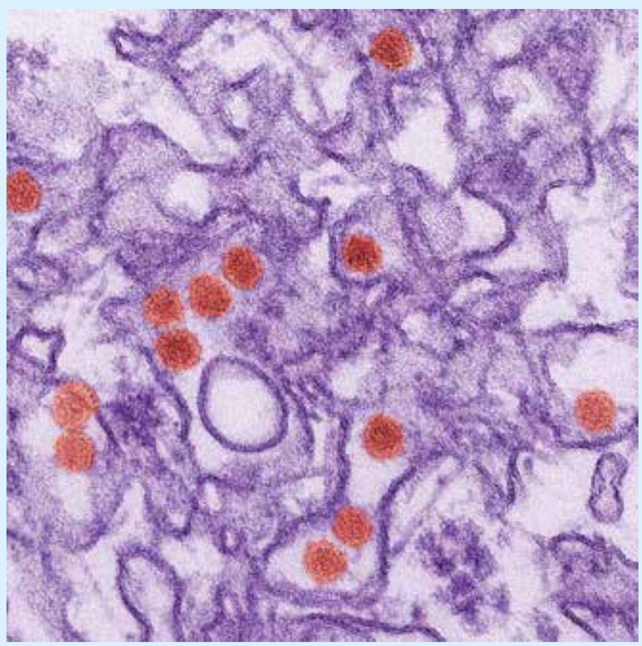


**FIGURE 16.6** Transmission electron microscope (TEM) image of Zika virus. These spherical viruses (shown in orange) are just 40 nm in diameter. (*Source*: CDC.)

*(continued)*

In 2013–2014, the Zika virus then spread across French Polynesia, infecting an estimated 30 000 people. On this occasion, Guillain-Barré syndrome, a severe neurological disorder, was first seen in infected individuals. The Zika virus was no longer just the cause of a mild flu-type illness. From there, the virus was exported by infected travellers to other Pacific islands, including Easter Island, the Cook Islands and the Solomon Islands.

Brazil then became the focus of Zika virus infections. In 2014, sporting events, including the FIFA World Cup and the World Sprint Championship canoe race, were held in Brazil, bringing tens of thousands of international visitors to that country. The strain of the Zika virus that spread throughout Brazil belongs to the Asian type found on Pacific islands, not the African type, suggesting that it may have been imported to Brazil by an infected traveller from one of the Pacific islands.

Because of the presence of large numbers of its mosquito vectors in densely populated urban areas, the Zika virus spread rapidly through Brazil and, from there, to other countries of South and Central America and the island countries of the Caribbean. The Zika virus has been imported to other countries by infected travellers returning from areas where the Zika virus is established. Figure 16.7 shows the time course of spread of the Zika virus.

The Zika virus epidemic in Brazil has been accompanied by an increase in the incidence of Guillain-Barré syndrome and, in October 2015, Brazil advised WHO of the occurrence of 54 cases of microcephaly in newborn babies in the preceding three-month period. The number of cases of microcephaly in Brazil increased to more than 4000 by 2016. In February 2016, WHO declared the current Zika outbreak to be a Public Health Emergency of International Concern.

As at May 2016, **active transmission of Zika virus by mosquito vectors** has been reported in nearly 50 countries, principally in South and Central America and the islands of the Caribbean, and also in Pacific islands, including Fiji, Tonga and Papua New Guinea. Other countries, such as Australia, experience imported cases of Zika virus through returning travellers but do not witness active transmission of the virus by vectors.

While transmission of the Zika virus occurs principally via its mosquito vectors, transmission has also been found to occur through the sperm of infected males. One case of transmission via a blood transfusion from an infected donor has also been confirmed.

Until 2007, the Zika virus infection was a mild (and probably often undetected) rare disease in Africa and Asia. Its movement into the Pacific and, from there, to other countries, is associated with a much higher incidence of infection and, in some cases, more severe symptoms. This signals that the Zika virus has mutated to a more virulent form as it has spread rapidly through populations having no pre-existing immunity.

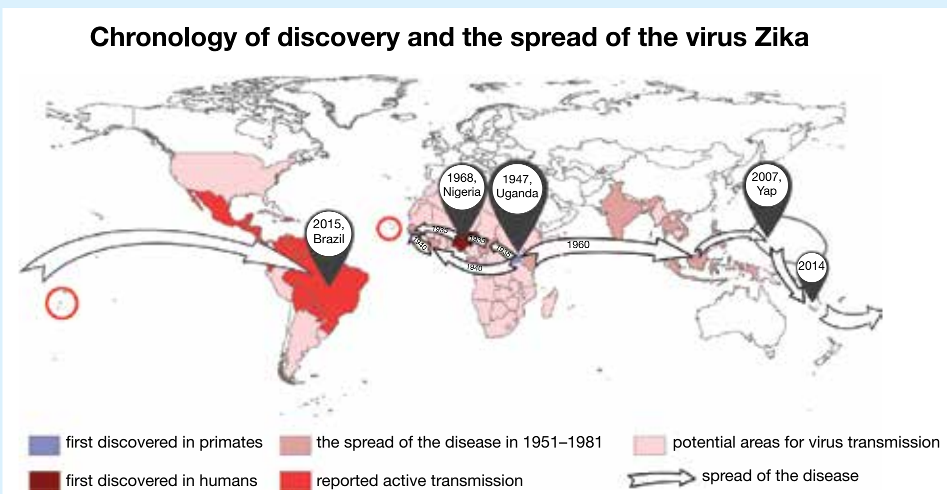



**FIGURE 16.7** Time course of spread of the Zika virus.

## KEY IDEAS

- A pandemic entails the uncontrolled worldwide spread of a new disease.
- An epidemic involves the uncontrolled spread of a disease over a restricted geographic area.
- Pandemics and epidemics involve new or novel pathogens to which people have not been previously exposed.
- The 1918–1919 Spanish flu pandemic caused the deaths of an estimated 40 million people.
- Several conditions must exist for a pandemic to develop.

## QUICK CHECK

1 Identify the following statements as true or false:
   a People exposed to new or novel pathogens would be expected to have little or no immunity to those pathogens.
   b The 1918 Spanish flu epidemic began in Spain.
   c Some pathogens have multiple hosts.
   d A pandemic could develop into an epidemic.
   e Deaths from the Spanish flu were higher in fit young adults than in older people.
2 Give an example of the following:
   a a virus that has a mosquito vector
   b an epidemic of the twenty-first century
   c a pandemic of the early twentieth century.
3 List two conditions that would be needed for a pandemic to develop.
4 Identify two ways in which the Zika virus can be transmitted.

# Meet the influenza A virus

In 1918, the cause of human influenza was unknown. Likewise, the links of human influenza virus to the viruses that cause influenza in birds (avian influenza) and in pigs (swine influenza) were not known. It was not until 1933 that the influenza A virus was isolated from ferrets by scientists. (The influenza B virus was isolated in 1938.)

Figure 16.8 is a model of the overall structure of an influenza virus. In this particular view, a portion of the outer envelope of the virion has been cut away, revealing its inner core of ribonucleoprotein (RNA plus protein).

Critical structures that determine the antigenic properties of the influenza A virus are the two kinds of protein embedded in the lipid bilayer membrane that forms the outer envelope of the virus.

These surface proteins are:

- **Hemagglutinin (HA)**, shown in figure 16.8 as blunt blue spikes — each H-surface molecule is a trimer, built of three monomers. To date, 19 different hemagglutinin types have been identified.
- **Neuraminidase (NA)**, shown in figure 16.8 as pink tacks — each N-surface molecule is a tetramer, built of four monomers. To date, 11 different subtypes of neuraminidase have been identified.

Figure 16.9 shows an electron micrograph of influenza virions.

Influenza viruses belong to the family Orthomyxoviridae, and their key features are as follows:

**Size:** Spherical particles with diameters in the range of 80 to 120 nanometres.

**Genome:** Single-stranded RNA, negative sense, composed of 13 588 bases that are organised into eight separate segments encoding eleven proteins (see table 16.1). Note that three of the segments encode sub-units of RNA polymerase, an enzyme that is needed when the infected host cell makes copies of the viral genome.

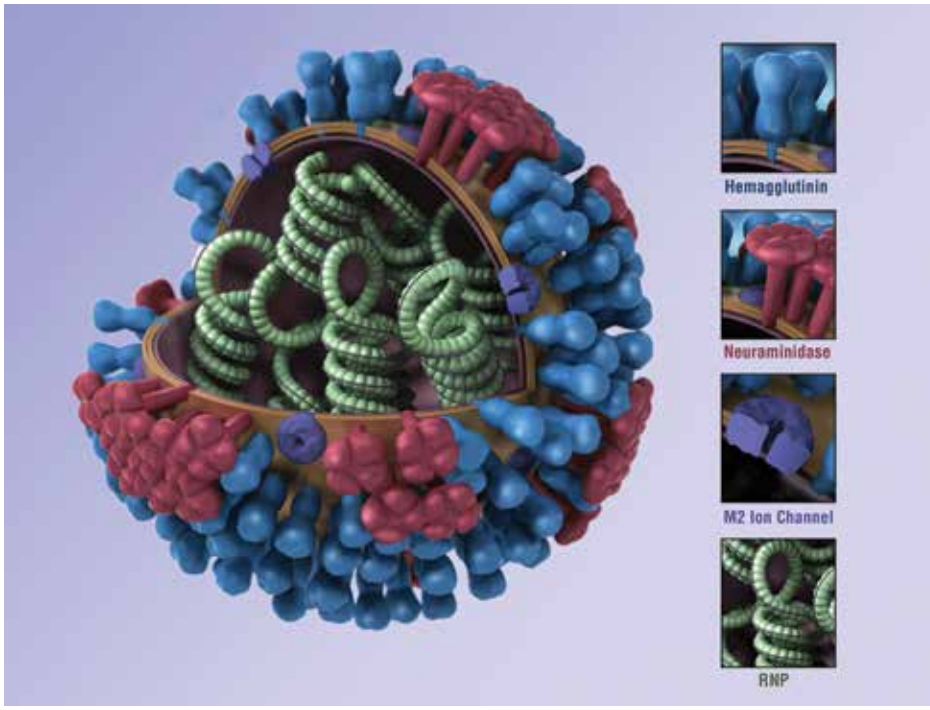



**FIGURE 16.8** Model showing a 3D representation of the ultrastructure of a typical influenza A virus, with part of its outer envelope cut away to show its internal ribonucleoprotein segments. The key at the right identifies the various components of this virus. (*Source*: CDC.)

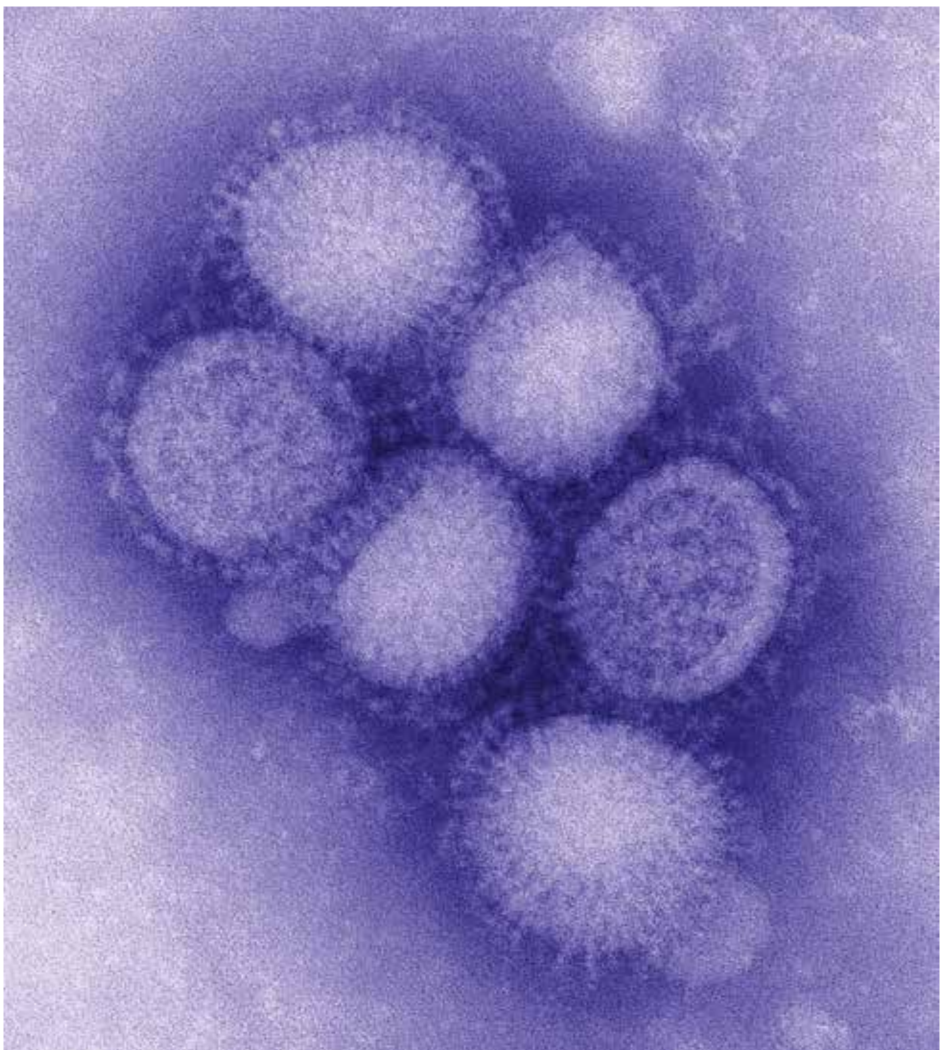


**FIGURE 16.9** Transmission electron micrograph of influenza A virus particles (virions). Note the rough outer margin that consists of the surface proteins, hemagglutinin (HA) and neuraminidase (NA), that determine the antigenic properties of the virus, such as whether or not it will be recognised by a person's immune system.

**TABLE 16.1** Functions of the eight RNA segments of the influenza A virus that comprise its genome. The surface proteins, HA (hemagglutinin) and NA (neuraminidase), determine the antigenic properties of a virus.

| RNA segment | Protein encoded by segment |
|---|---|
| **1** | **PB2**: RNA polymerase sub-unit |
| **2** | **PB1**: RNA polymerase sub-unit |
| **3** | **PA**: RNA polymerase |
| **4** | **HA**: hemagglutinin |
| **5** | **NP**: nucleoprotein |
| **6** | **NA**: neuraminidase |
| **7** | **M1, M2**: matrix proteins |
| **8** | **NS1, NS2**: non-structural proteins |

## Naming influenza viruses

The name or identifier of an influenza virus consists of:

(i) its type
(ii) a strain designation
(iii) for influenza A viruses only, a subtype based on the identity of their surface antigens, hemagglutinin (HA) and neuraminidase (NA).

**Type:** Based on their core proteins, influenza viruses are classified into three major types: A, B, and C. Types A and B cause human diseases of concern as they are responsible for seasonal outbreaks of influenza that occur during most winters and are the cause of epidemics and pandemics (see page 730).

Influenza A viruses infect not only people but also a range of other mammalian and avian species.

**Strain:** The strain of an influenza virus contains information including the host of origin (for strains isolated from non-human sources), the geographical origin of the virus, the strain number and the year of isolation.

HA and NA are used to denote the viral proteins hemagglutinin and neuraminidase. However, when they are part of a subtype, they are shortened to H and N, such as N1H1 and H3N2.

**Subtype**: Subtypes apply to influenza A only and they identify the antigenic subtypes of its surface proteins hemagglutinin, from H1 to H15, and neuraminidase, from N1 to N9. Examples of subtypes of influenza A viruses include H1N1, H5N1 and H3N2.

Figure 16.10 shows examples of the full identifier of a virus, which includes, in order, the type, the host (if not human), the geographical origin, the strain number and, in the case of influenza A, the subtype. Note that, as shown in figure 16.10b, when the virus is isolated from a person, the host is not specifically identified.

Subtypes of influenza A viruses are various combinations of the surface proteins, hemagglutinin (HA) and neuraminidase (NA), on the outer envelope of the virus. As indicated above, 18 different H subtypes and 11 different N subtypes have been detected in various species of mammals and birds.

- At present, only two subtypes of influenza A are in general circulation in human populations, namely H1N1 and H3N2.
- All possible subtype combinations, except for H17N10 and H18N11, have been detected in birds.

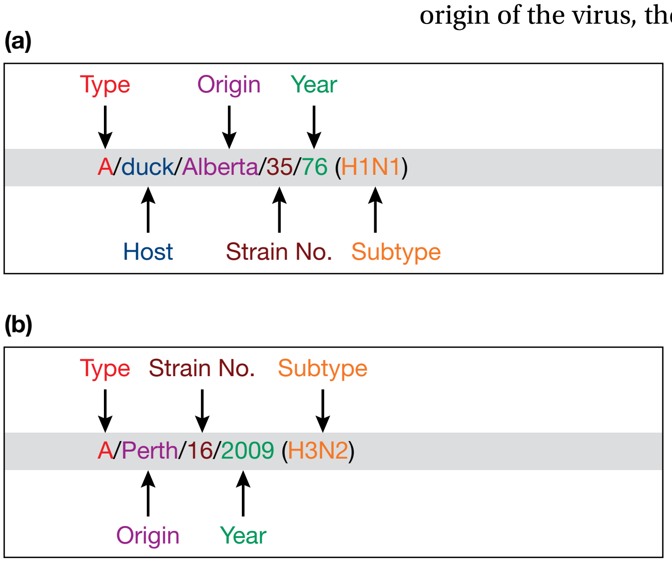



**FIGURE 16.10** Identifiers of two viruses. **(a)** Type A influenza virus found in a duck in Alberta, Canada, strain number 35, isolated in 1976. **(b)** Type A influenza virus (human), found in Perth, Western Australia, strain number 16, isolated in 2009.

- Pigs are naturally infected with H1N1, H3N2 and H1N2 viruses.
- Viruses of subtype H7N7 and H3N8 cause illness in horses.
- Dog flu is caused by H3N8 and H3N2.

### Influenza A virus can jump between species

Many examples exist of the 'jump' of an influenza A virus from one species to another.

**Bird to human:** Sporadic jumps to people of the H5N1 bird flu virus, which circulates among poultry in Asia and the Middle East, has occurred in 16 different countries. In particular, in Hong Kong in 1997 the direct transmission of bird (avian) H5N1 influenza virus from birds to people was detected. The cases concerned resulted in severe pneumonia with a death rate of about 60 per cent.

In 1999, another novel influenza A virus, subtype H9N2, present in birds, was found to be the cause of illness in several children. In 2013, the first cases were reported of the jump of the bird influenza virus H7N9 to humans. These cases resulted from close contact of people with infected poultry or exposure to poultry-contaminated environments. About one-third of these cases have resulted in death.

**Horse to dog**: The H3N8 virus, which causes influenza in dogs, originated in horses and has jumped from horses and become established in dogs. This virus was first detected in 2004 in a greyhound at a racing track that was used by both horses and greyhounds.

**Bird to dog**: The H3N2 virus, which causes influenza in dogs, is an avian virus that has apparently jumped directly from birds to dogs. Its presence in a dog was first detected in 2005 in South Korea and was then found in dogs in Thailand and China. In an outbreak of dog influenza in the United States in 2015, the causative agent was found to be this H3N2 strain of influenza A virus.

## Influenza pandemics

Of the many diseases that may become pandemic, influenza is one that occurs frequently. The viruses responsible for influenza pandemics are new subtypes of influenza A that arise as a result re-assortment of their genetic material such as can occur between human influenza viruses and those of birds and/or pigs. (Later in this chapter (see page 740), we will explore these changes in influenza viruses.)

Four influenza pandemics have occurred since the beginning of the twentieth century (see figure 16.11). In each case, a new or novel influenza A virus was the cause and the virus was easily spread from person to person.

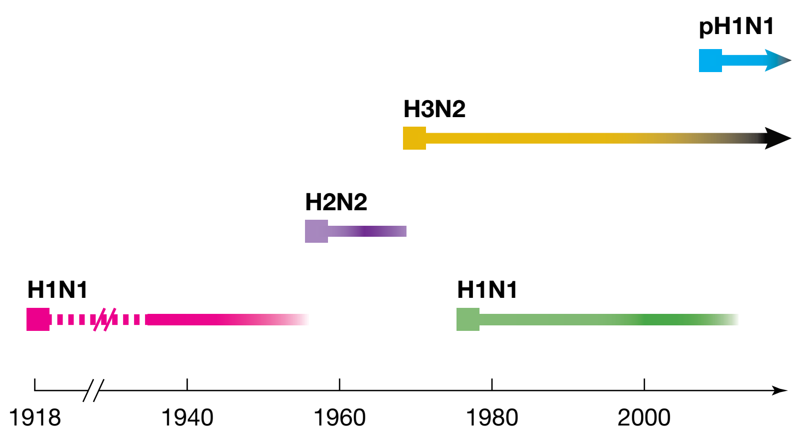



**FIGURE 16.11** Influenza A viruses in circulation since the start of the twentieth century. The notation pH1N1 denotes a new influenza virus in pigs that was first seen in 2009.

**1. 1918–1919: Spanish flu**

The Spanish influenza pandemic (refer back to page 728) was caused by an H1N1 subtype of influenza A virus. A pandemic with the severity of the Spanish flu, with global deaths estimated at up to 40 million people, has not been seen since.

**2. 1957: Asian flu**

In March 1957, the Asian influenza pandemic originated in Yunnan Province in China and was caused by an H2N2 subtype of influenza A. From there, the virus spread through South-East Asia and Japan and reached India, Indonesia and Australia in May 1957. By July, the viruses had spread worldwide, having reached Europe and the Americas. The major mortality was among the very young and the very old, and the total deaths from this H2N2 Asian flu pandemic were believed to have exceeded one million people.

**3. 1968: Hong Kong flu**

In early 1968, the Hong Kong influenza pandemic was first detected in Hong Kong, and was caused by an H3N2 subtype of influenza A. The death rate was greatest in people over 65 years old. The pandemic occurred in several waves, the first being from September 1968 to March 1969, and the same virus then returned in 1970 and 1972. Total deaths in this pandemic have been estimated between 500 000 and one million.

**4. 2009 flu**

The pandemic of 2009 was caused by a novel H1N1 virus of swine origin and was detected in humans in early April 2009. The first outbreak in Mexico was originally called swine flu, but it is now called 'pandemic (H1N1) 2009'. Although this virus was the same subtype (H1N1) as the virus that caused the Spanish flu pandemic of 1918–1919, this 2009 virus was a different strain, and it was less deadly. The global death toll from this pandemic has been estimated by the US Centers for Disease Control and Prevention (CDC) at 284 000. Unlike the pattern with seasonal flu, about 80 per cent of deaths in this pandemic were those of people under 65. This matches the finding for Spanish flu: disproportionately high death rates among younger adults.

**ODD FACT**

The pandemic caused by the H1N1 (2009) influenza virus was far less deadly than the 1918 Spanish flu virus. The former had a death rate of about 0.01 to 0.03 per cent of those infected. The Spanish flu of 1918–1919 was far more deadly, with a case fatality rate of 2 to 3 per cent.

## KEY IDEAS

- The influenza A virus is a negative single-stranded RNA virus with its genome organised as eight separate segments of RNA.
- Key proteins, hemagglutinin (HA) and neuraminidase (NA) located on the outer viral envelope, determine the antigenic properties of the virus.
- Influenza A virus infections occur in many bird species and several mammalian species.
- An international convention for the identification of influenza viruses exists.
- Influenza A viruses can jump from species to species.
- Influenza A pandemics are caused by the emergence of new or novel viruses that can be easily transmitted from person to person.

## QUICK CHECK

**5** Give an example of the following:
   **a** the jump of the influenza A virus from one species to another
   **b** the subtype of virus that is presently circulating in human populations
   **c** a protein that acts as an antigen and is located on the outer envelope of an influenza A virus particle
   **d** a segment of viral RNA that encodes an RNA polymerase sub-unit.

**6** Answer the following about an influenza virus that has the identifier A/California/7/2009 (H1N1):
   **a** What is its type?
   **b** What is its geographic origin?
   **c** What is its subtype?
   **d** When was it first isolated?
   **e** What is its host species?

# The changing influenza virus

Immunological memory means that once a person has been vaccinated against specific pathogens or has suffered a childhood infection, that person usually has immunity against those pathogens for a lifetime. This means that if people are later re-exposed to the pathogen, their immune systems have the ability to respond immediately by rapidly producing specific antibodies against that pathogen. As a result, the infection cannot develop into a disease state and no disease symptoms appear. However, this is not the case for influenza. Why then can a person have influenza several times over a period of many years? In the next sections, you will see that the influenza A virus is a constantly changing pathogen.

## Antigenic drift — gradual small changes

The RNA of the genetic material of the influenza virus continually undergoes frequent mutations as it replicates. This situation arises because there is no proofreading mechanism to edit any mistakes that occur in the replication of viral RNA. Each newly replicated virus can have a nucleotide change or point mutation in one of its genes. At first, these mutated viruses are closely related to one another and have similar antigenic properties. Under these circumstances, the memory cells of a person's immune system will recognise and respond to these viruses. These small changes continually occurring in the influenza virus are termed **antigenic drift**. All influenza virus types A, B and C are subject to the gradual change of antigenic drift.

Over time, however, the accumulation of these small changes means that antigenic properties of the mutated viruses have been altered to a major degree, and a new subtype is identified. When this happens, the virus will no longer be recognised by immune system memory cells but will be treated as a new pathogen. This means that a time delay will occur before antibodies are produced, with the result that the infected person comes down with the symptoms of influenza. A major change such as this can result in a pandemic.

## Antigenic shift — a sudden big change

Influenza A viruses can also undergo an occasional but major sudden change in their two surface proteins, hemagglutinin (HA) and neuraminidase (NA). These proteins are embedded in the outer envelope of the influenza A virus and give the virus its antigenic identity.

A major and sudden change of the HA protein or the NA protein or both proteins is termed an **antigenic shift**. Antigenic shift is restricted to influenza A viruses because these are the only influenza viruses that infect not only people but also other species, including birds and mammals. By being able to infect a variety of hosts — people, birds and pigs — strains of influenza A virus can become very different during their transmission within these animal groups.

When one host is infected with two different kinds of influenza A virus, a new combination of genetic material can be produced by **re-assortment**. As a result, the new virus cannot be identified by the immune system. This re-assortment is made possible because the influenza A genome is not one long single-stranded RNA chain but, instead, the genome consists of eight separate RNA segments that can be swapped.

Figure 16.12 is a simplified diagram showing how re-assortment of the genetic material of two different influenza A viruses, H3N2 and H5N1, can occur when one host is infected with both viruses. What is the product of this re-assortment?

Re-assortment of the viral genetic material most commonly occurs when a host becomes infected with two or more different influenza viruses at the same time. When this occurs, the influenza viruses can mix and exchange genetic information so that new influenza virus subtypes are formed.

**ODD FACT**

The HA protein plays a key role in the attachment of the virus to its host cell and its penetration into the cell. The NA protein plays a role in the detachment of new virions from the host cell.

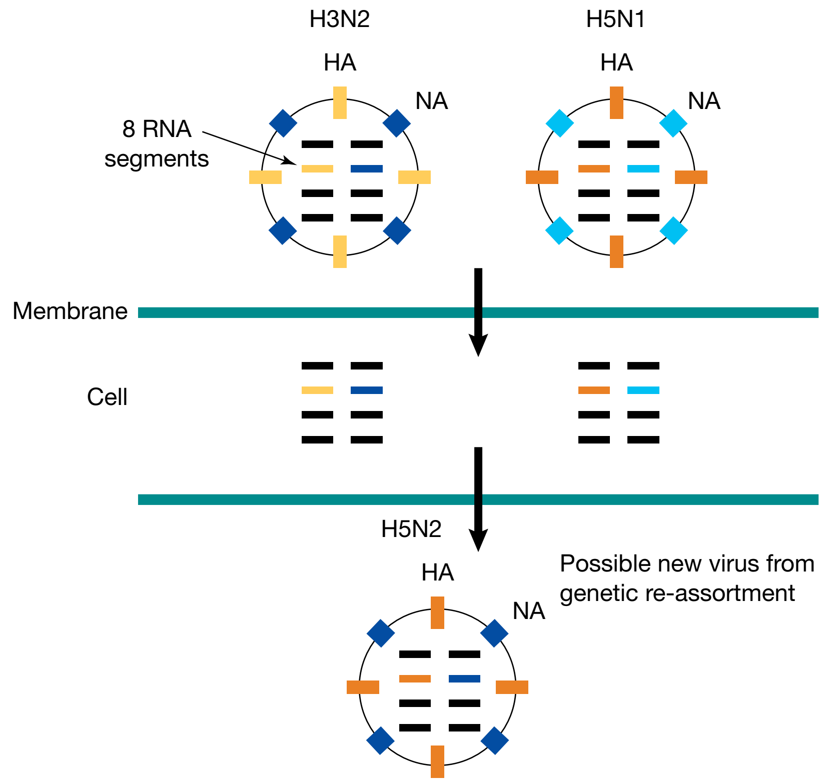



**FIGURE 16.12** Diagram showing the re-assortment of the genetic material of two influenza A viruses to form a new subtype of this virus. Re-assortment can occur when a host cell is infected by two (or more) different subtypes of the influenza A virus. What do the eight bars inside the virus represent?

The respiratory tract of pigs has receptors for both human and avian influenza viruses. As a result, pigs can be readily infected with influenza viruses from people and from birds, as well as from other pigs. Pigs represent a cellular environment in which re-assortment of influenza viruses can occur. Environments in which this is possible include farms and marketplaces where people, poultry and pigs come into close contact (see figure 16.13). If a newly re-assorted influenza A virus infects people, influenza will break out and spread rapidly because of the lack of immunological memory in the population. This situation has the potential to produce an epidemic, or even a pandemic, such as occurred in 2009, beginning in Mexico.

An antigenic shift produces a new subtype of influenza A that is markedly different from the subtype circulating in human populations. As a result, no one is likely to have immunity to this new viral subtype so that, if people are exposed to this new subtype of virus, they will contract influenza. This situation can develop into an epidemic, or even a pandemic.

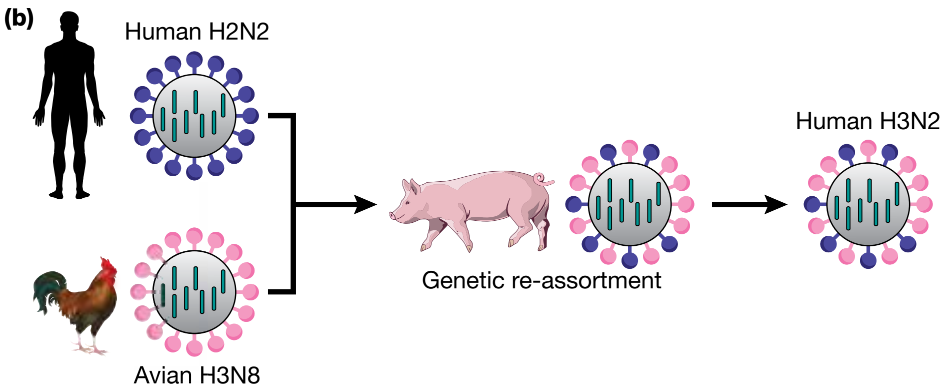
## Flu vaccines must be updated

New or novel subtypes of influenza virus that emerge suddenly through antigenic shift are not recognised by our immune systems. As a result, these new viruses are potentially very dangerous. Likewise, changes accumulated through antigen drift over long periods can also produce viruses that are no longer identified by our immune systems. Vaccines that protect against influenza are effective only against the particular viruses that are incorporated into the vaccine. This is why we can suffer from several bouts of influenza during our lifetimes and why flu vaccinations do not last a lifetime. This is also why the composition of a flu vaccine is reviewed each season and updated to adjust for any new subtypes of virus that have emerged and are circulating in the region (see figure 16.14).

**ODD FACT**

Because of the rapid rate of mutation of the influenza virus, the effectiveness of a given vaccine usually lasts for a year or so. Expert advice is that individuals at risk should be vaccinated yearly.

**FIGURE 16.13** **(a)** A livestock and poultry market provides an opportunity for cross infection of a host by influenza viruses from different species. **(b)** The genetic re-assortment produces an antigenic shift that creates a new subtype of influenza A virus that can infect people. What other re-assorted virus could possibly be produced in the pig?


(a)

(b)

| Antigen | Description |
|---|---|
| A (H1N1) | an A /California/7/2009 (H1N1)pdm09 like virus |
| A (H3N2) | an A/HongKong/4801/2014 (H3N2) like virus |
| B | a B/Brisbane/60/2008 like virus |
| B | a B/Phuket/3073/2013 like virus |

**FIGURE 16.14** **(a)** Person receives an annual influenza shot. **(b)** Formulation of the influenza vaccine for use in Australia in 2016. This is a quadrivalent (four-component) vaccine that gives protection against the two influenza A viruses and two influenza B viruses that are known to be present in the southern hemisphere.

## KEY IDEAS

- Influenza A viruses can change by the processes of antigenic drift and antigenic shift.
- Antigenic drift occurs gradually and by frequent single point mutations in the RNA segments that encode the surface proteins determining the subtype or antigen identity of the virus.
- The accumulated changes arising from antigenic drift over an extended period of time can generate new subtypes of influenza A viruses that are no longer recognised by a person's immune system.
- Antigenic shift is a process of re-assortment of RNA segments from two (or more) different virus subtypes present in one host that can result in the sudden appearance of new subtypes of the influenza A virus.
- Influenza epidemics and pandemics can result from the emergence of new subtypes of influenza A viruses.
- Because of constant changes in the influenza A virus, flu vaccines have a limited period of effectiveness and need to be updated, often on an annual basis.

## QUICK CHECK

7 Briefly explain the following terms:
  - a *antigenic drift*
  - b a *quadrivalent* vaccine
  - c an *avian* influenza virus.

8 Identify the following statements as true or false:
  - a Antigenic shift events occur more frequently than antigenic drift events.
  - b For an antigenic shift to occur, at least two different influenza A viruses must be present in the one host.
  - c A flu vaccine is effective for a person's lifetime.
  - d The expectation for a vaccine against the H3N2 virus subtype would be that it is effective against an H3N2 virus with a few point mutations.
  - e A vaccine against the H3N2 virus subtype would be effective against an H1N1 virus.

# Rational drug design

We have seen that influenza viruses are continually changing into new subtypes. The proteins embedded in their outer envelope change continually. Our bodies learn to defend against one influenza virus. However, the next time we are infected with an influenza virus, if it is a different subtype from the previous one, we will develop influenza again. An Australian development, an anti-flu drug called Relenza (generic name: zanamivir) has reduced this problem.

The scientists came up with a solution to the recurring cases of influenza by using a technique called **rational drug design**. The process essentially involves finding out how enzymes in the infective agent work against a cell, and using that information to design a drug that prevents the infective agent from being able to do what it does. This approach is particularly useful when a drug can be designed to inhibit the action of a key enzyme by blocking its active site.

Three scientists were awarded the 1996 Australia Prize for scientific research benefiting humanity for their work on a potential cure for influenza: Dr Peter Colman (see figure 16.15), now Head of the Structural Biology Division at the Walter and Eliza Hall Institute of Medical Research, Professor Mark von Itzstein, now Director of Griffith University's Institute for Glycomics, and Dr Graeme Laver (1929–2008), then at the Australian National University's John Curtin School of Medical Research. The Australia Prize is an international award given by the government of the Commonwealth of Australia in a selected area of science and technology promoting human welfare.

**FIGURE 16.15** Dr Peter Colman, Head of the Structural Biology Division at the Walter and Eliza Hall Institute of Medical Research in Melbourne. Using X-ray crystallography, Dr Colman identified the 3D structure of neuraminidase, including its active site. (Image courtesy of the Walter and Eliza Hall Institute.)

## Using a strategic approach

It is known that the two kinds of protein on the outer surface (envelope) of the influenza virus have different roles in the viral infection of a cell.

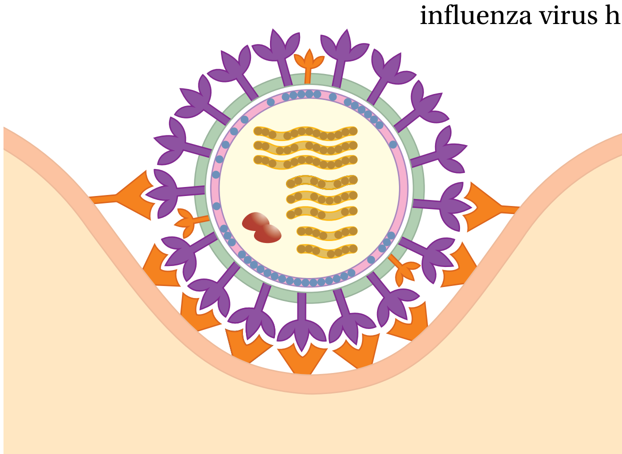


**FIGURE 16.16** Diagram showing the attachment of a virus particle (virion) to the plasma membrane of the host cell. The attachment involves the binding of the hemagglutinin proteins (shown in purple) to the sialic acid receptors (shown in orange) on the plasma membrane of the host cell.

- Hemagglutinin (HA) is active at the start of an infection by enabling the influenza A virus to gain entry to a host cell, where it will replicate. The virus first attaches to the host cell through its HA proteins that bind to sialic acid receptors on the plasma membrane of the host cell (see figure 16.16).
- Neuraminidase (NA) is active at the end of the infection of one cell by facilitating the exit of new virus particles from a cell, freeing them to infect other cells. Newly synthesised influenza viruses aggregate at the plasma membrane of the host cell and are only released from the cell through the action of the viral enzyme neuraminidase. So, NA is essential for the release of the next generation of virions from the surface of an infected cell and for the spread of the infection to other cells. The substrate for NA is sialic acid on the plasma membrane surface.

If the action of either of these viral proteins, HA and NA, could be interrupted, then treatment for influenza might be possible. This was the dream of scientists working on the influenza virus. Attempts to block entry to the flu virus failed, so attention turned to neuraminidase.

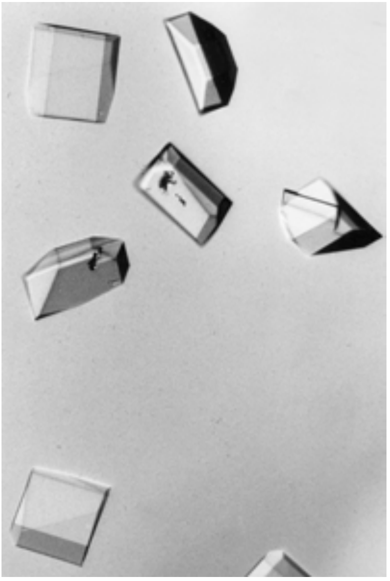


**FIGURE 16.17** The isolation of neuraminidase crystals meant that the 3D structure of the enzyme could be determined. This achievement resulted from utilising the power of a synchrotron. (Image courtesy of Bentham Science Publishers Ltd.)

Neuraminidase is an enzyme that varies in structure from subtype to subtype. The first breakthrough was made by Dr Peter Colman when he carried out **X-ray crystallography on crystals of neuraminidase** (see figure 16.17) and **determined its 3D structure**. Examination of subtypes of the virus from past years demonstrated that, although most of the NA enzyme changed dramatically, one small part remained constant. Fortunately, this non-variable part is the active site.

Having identified the active site of the neuraminidase enzyme, the next challenge was to design a drug to inhibit the active site of neuraminidase. This required that the shape and structure of the active site be worked out. This was done by computer modelling, which revealed the exact shape of the active site and identified the spatial arrangements of the atoms surrounding it. Figure 16.18 shows the key amino acids that form the active site of the neuraminidase enzyme and identifies how an inhibitor with a complementary shape could occupy the active site and exclude the normal substrate.

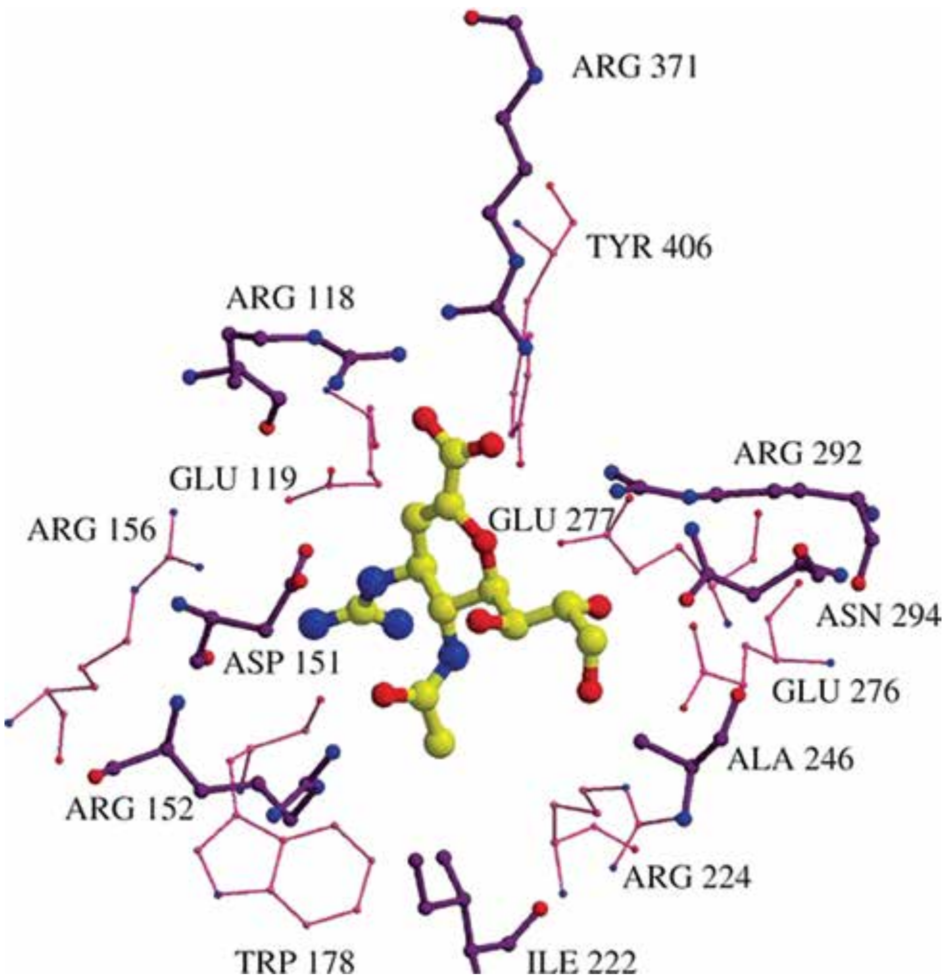



**FIGURE 16.18** Computer representation of the anti-flu drug Relenza, located within the active site of viral neuraminidase (NA). The anti-flu molecule is shown as the heavier structure and it is held in the active site by hydrogen bonds and ionic bonds. The side chains of the amino acids that form the active site of the neuraminidase enzyme are shown as lighter structures, and their shorthand symbols are shown. The number after each amino acid identifies its position in the protein; for example, the symbol ARG 224 denotes an arginine at the 224th position in the protein chain. Is the active site made of consecutive amino acids? (*Source:* Figure 6a, 'Antiviral adhesion molecular mechanisms for influenza: W.G. Laver's lifetime obsession', Garman, E.F., 2014.)

Once the detailed structure of the active site of the neuraminidase enzyme was known, the next challenge was to design a molecule to fit into and bind to the active site. To do this, (i) the molecule would need to have a complementary shape to that of the active site so that it would fit within the active site, and (ii) any charged groups on the molecule would need to be ones enhancing the binding of the molecule to the active site. The product of this design process was the creation of an anti-flu drug with the generic name zanamivir and trade name Relenza.

This technique, in which **the active site of a molecule, such as an enzyme, is determined and a second molecule (the drug) is constructed to fit into that active site to inhibit the activity of the first molecule, is called rational drug design**.

### Relenza in action

Normally, an influenza virus enters a cell and takes control. Many viral parts are rapidly synthesised using the machinery of the host cell. These viral parts are assembled and move to the surface of the cell and 'bud' out of the membrane prior to exiting from the cell. The newly synthesised viruses stay on the surface of the cell until neuraminidase cuts the sialic acid attachment between the cell and the viruses. The viruses then infect other cells. Many viruses are produced within a host cell so that, as infection proceeds, the host system is filled with millions of viruses, and influenza symptoms appear.

In the presence of the anti-influenza drug, the initial stages of infection of a cell by a virus are the same as described in the previous paragraph. However, **the anti-influenza drug interacts with the active site of neuraminidase so that the new viruses are trapped at the surface of the cell**. Because the viruses are prevented from infecting other cells, there is no massive increase in the numbers of viruses usually occurring, and virus influenza symptoms fail to develop. The immune system acts in the usual way to remove the bound virus particles.

The anti-influenza drug that inhibits the action of the enzyme neuraminidase developed by Dr Colman and his colleagues is available commercially in many countries. An Australian company, Biota, was involved in the final commercialisation of the drug, which is marketed under the name Relenza by GlaxoSmithKline (see figure 16.19). Four doses of the medicine are contained in a silver blister pack. This pack is placed in a wheel inside the diskhaler, which is then closed. A blister is broken and then the drug is inhaled through the mouthpiece of the diskhaler.

**studyon**

Unit 4 | AOS 2 | Topic 2 | Concept 8

**Rational drug design**
Summary screen and practice questions


**FIGURE 16.19** The anti-influenza drug developed by Dr Colman and his colleagues is now available in many countries around the world.

The drug is also effective in humans against the bird flu strain that became more active in 2005, particularly in some areas of South-East Asia. Biota continues research on improving influenza drugs in collaboration with an American company.

## Other designer drugs

The rational drug design approach has been used to produce several anti-influenza drugs. The principle is that these synthetic drugs are designed to have a molecular configuration and shape such that they can block or inhibit the action of molecules that carry out essential steps in the viral replication cycle in host cells.

The mode of operation of viruses varies depending on the species of virus, but all viruses operate by a general shared pattern that includes the following key processes:

- attachment to a host cell
- uncoating and release of viral genes and enzymes into the host cell
- replication of viral components using host cell machinery
- assembly of viral components into complete viral particles
- release of viral particles to infect new host cells.

Based on this general pattern of operation, possible designer drugs might include mechanisms that:

1. prevent viral attachment and/or entry to host cells
2. prevent replication of the viral genome
3. block synthesis of specific viral protein(s)
4. prevent assembly of new infectious virions
5. prevent release of new infectious virions.

You will recall that Relenza was designed to prevent process 5, the release of new infectious virions.

Examine figure 16.20, which identifies these key processes in a virus infection cycle, and the enzymes of the virus that are the targets of various designer drugs. The designer drugs are identified by their names, in bold blue, and the enzyme-catalysed steps that are the targets of these drugs are indicated by the black arrows (solid and dashed). Table 16.2 summarises the antiviral action of various designer drugs, both actual and developmental.

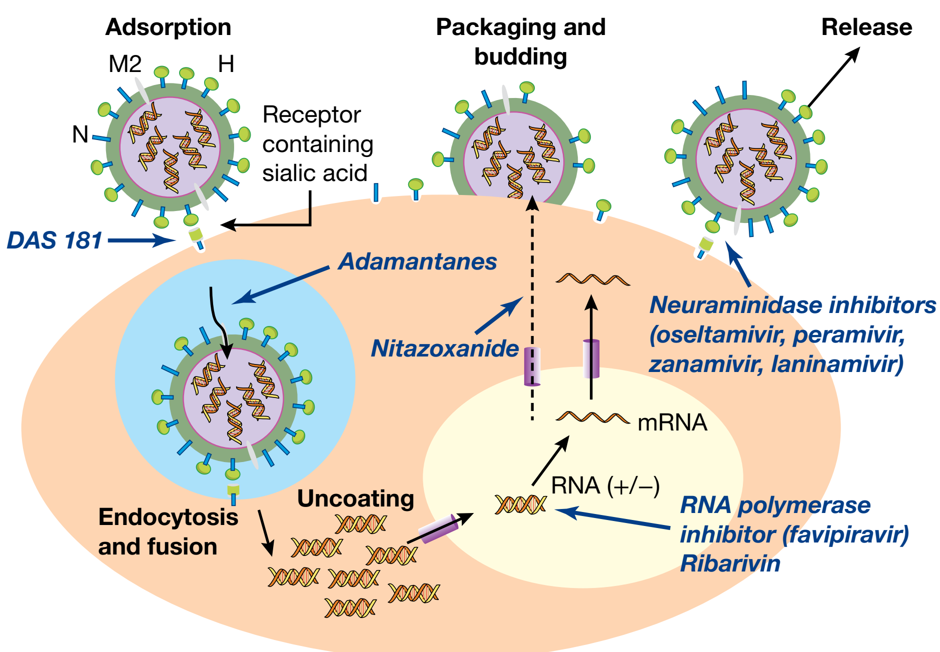



**FIGURE 16.20** Diagram showing the key processes in a viral infection cycle. Antiviral drug treatment designed to block specific molecular targets or processes is shown.

**TABLE 16.2** Summary of antiviral action of designer drugs for use against influenza A virus infections.

| Viral operation | Designer drug | Action of drug |
|---|---|---|
| viral attachment to host cell | DAS181 | cuts links of HA with sialic acid receptors |
| unpacking of virus and release of viral components into host cell | adamantane | inhibits ion channel |
| replication of genetic material | favipiravir<br>ribarivin | inhibits RNA polymerase |
| packaging of viruses | nitazoxanide | broad spectrum antiviral agent |
| release of new viruses | zanamivir<br>oseltamivir<br>peramivir | inhibits neuraminidase |

# Identifying pathogens

When an outbreak of a serious disease occurs and begins to spread rapidly through person-to-person contact, there is a risk of an epidemic. An epidemic, if not brought under control, has the potential to develop into a pandemic. In this situation, the 'boots-on-the-ground disease detectives' of the Epidemic Intelligence Service (EIS) are quick to move into the affected region to work with local officials. The EIS is part of the US Centers for Disease Control and Prevention.

EIS personnel seek rapid answers to key questions such as:

- What is the cause of the sickness?
- If a pathogen is identified as the cause, how can it be treated and how can the disease be prevented from spreading?
- What measures are needed to prevent another outbreak?

Figure 16.21 shows members of the EIS in West Africa during the largest outbreak of Ebola, which affected **Liberia, Sierra Leone, Guinea and Nigeria**. This Ebola virus disease epidemic, which began in March 2014, was caused by a virus of the family *Filoviridae*, genus *Ebolavirus*, and lasted until March 2016. This epidemic caused more than 11 000 deaths, with a death rate of two for every five infected people.


**FIGURE 16.21** Photograph taken in Liberia during the 2014 West African Ebola outbreak. A CDC staff member is being assisted by a Médecins Sans Frontières (MSF) staff member in an essential decontamination process before exiting an Ebola treatment unit. A hypochlorite spray is known to destroy Ebola virus. (*Source:* CDC.)

When an infectious disease breaks out, speedy identification of the causal pathogen is critical. Knowing the identity of the pathogen guides decisions about treatment of the disease and preventing it from spreading.

## Identifying viruses

Unlike most bacteria, viruses cannot be replicated in standard microbiological broths or on agar plates. Instead, because they are obligate intracellular parasites, viruses require living cells in order to replicate. Viruses must be cultured inside suitable host cells. Once a sample of the virus is available, identification can proceed.

A range of techniques can be used to identify viruses, and they include physical methods, immunological methods and molecular methods.

**1. Physical methods** can assist in identifying viruses based on size and shape. These methods include:

- X-ray crystallography, which has determined the structure of many viruses
- Electron microscopy, which has given us images that distinguish various kinds of virus. Figure 16.22 shows the contrasting shapes of Ebola virus and influenza A virus as revealed by transmission electron microscopy.

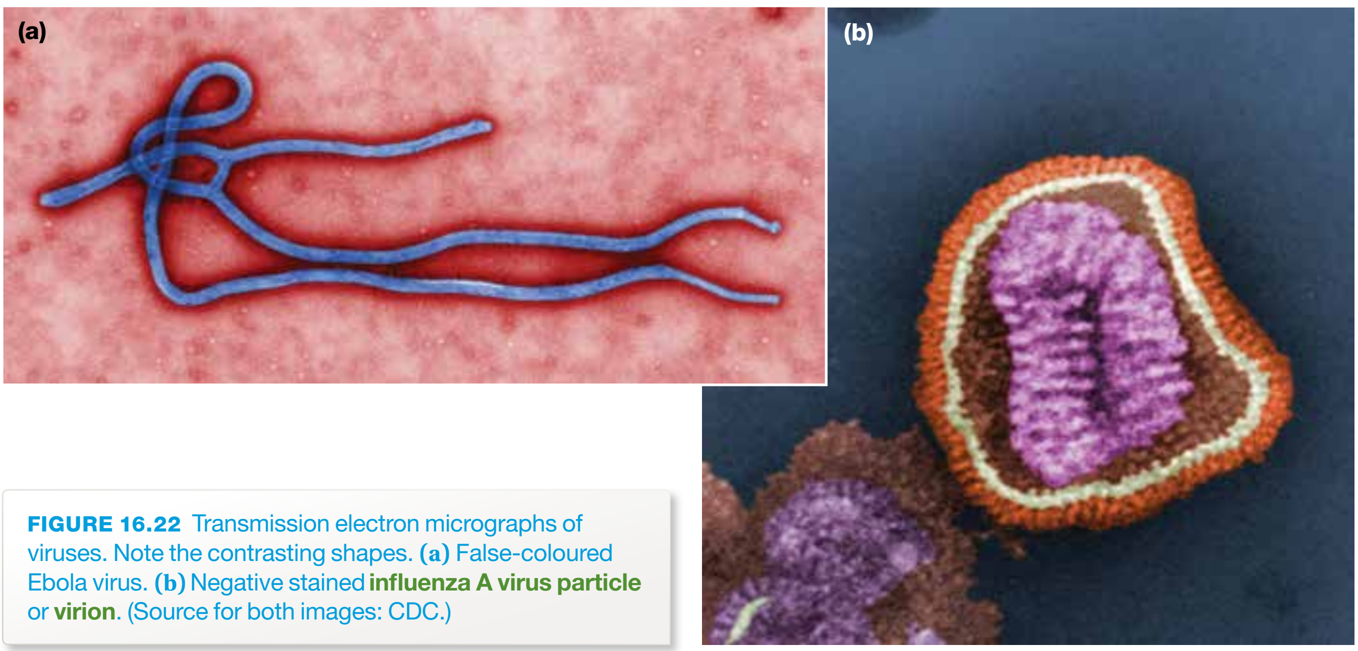



**FIGURE 16.22** Transmission electron micrographs of viruses. Note the contrasting shapes. **(a)** False-coloured Ebola virus. **(b)** Negative stained **influenza A virus particle** or **virion**. (Source for both images: CDC.)

**2. Immunological methods**, such as the Enzyme-Linked ImmunoSorbent Assay (ELISA) technique, enable the diagnosis of viral diseases. ELISA can be used, for example, to detect the presence of a specific viral antigen in a cell or in a sample of body fluid, such as serum. Examine figure 16.23a, which shows the use of a 'capture' antibody to identify the presence of a specific viral antigen through an antigen-antibody reaction. All unbound material is washed away. A second antibody with an enzyme indicator is then added, along with a substrate for this enzyme that enables production of a coloured complex. If colour appears, the specific viral antigen is present; if no colour appears, the specific viral antigen is absent.

The ELISA test can detect the presence of antibodies to specific viral disease in cells or body fluids (see figure 16.23b). This is exemplified in the screening test used to detect the presence of antibodies to human immunodeficiency virus (HIV) in a person's serum. In this ELISA screening test, a serum sample from the person to be tested is diluted 400-fold and applied to a plate on which HIV antigens are immobilised. If antibodies to HIV are present in the serum, they can bind to the HIV antigens. All unbound material is washed away. A second antibody linked to an enzyme is then added. Because of its high sensitivity and ease of application, ELISA is a powerful tool in the identification of viruses and is used routinely to test and diagnose many viral diseases, including HIV, adenovirus, West Nile virus and cytomegalovirus.

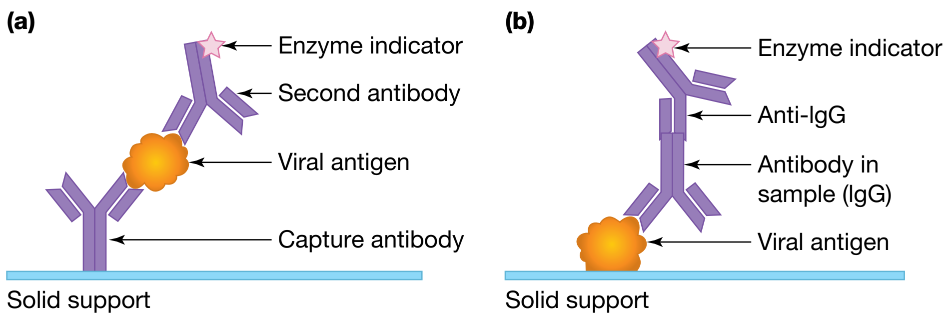



**FIGURE 16.23** The ELISA technique can be used to detect viral proteins acting as antigens and also detect viral antibodies in blood serum. **(a)** 'Capture' antibodies bound to a solid support, such as a plastic surface or plastic bead, can be used to detect viral proteins (antigens). **(b)** Viral antigen molecules bound to a solid support act as anchors to which specific viral antibodies can bind, forming aB-Ag complexes.

**3. Molecular techniques** can be used to identify viruses. Refer back to table 6.8 on page 250 to review the diversity of viral genomes such as ssDNA and ssRNA. Molecular techniques include the use of *in situ* hybridisation with DNA probes to detect and locate specific genetic sequences that are diagnostic of particular DNA viruses. Reverse techniques van be used to identify RNA viruses. In particular, restriction enzymes can be used to differentiate between different strains of the same DNA virus.

## BIOLOGIST AT WORK

### Professor Peter Timms — saving koalas from disease with smart science

'As a university lecturer and research professor at the University of the Sunshine Coast (USC), I have many interesting and varied jobs. One of my research interests is to develop a vaccine to help koalas fight infection and disease caused by the *Chlamydia* bacterium. The number of koalas in Australia is on the decline (unfortunately, quite a rapid decline in many areas) and one of the main reasons for this is the bacterial infection chlamydia, which causes eye infections leading to blindness and sexually transmitted disease leading to infertility. We are using what we have learned from developing vaccines for humans and applying this to koalas, and we are making progress.

**FIGURE 16.24** Professor Peter Timms with one of the koalas that are subjects in the trialling of a chlamydia vaccine. (*Source:* AAP Newswire.)

My original training was as a microbiologist when I studied for my science degree at the University of Queensland. While this is still a major part of my research and teaching interests, I also gained some additional qualification while doing my PhD in the area of molecular biology. In today's world of science, you need to have skills in lots of different areas to be able to solve the complex puzzles of nature. While my skills with molecular biology were initially used for my medical research interests, I am now able to apply the same skills to helping threatened wildlife, such as koalas.

So exactly how do you make a vaccine for koalas? First, you need to dissect the bacterium that causes the disease, chlamydia, into individual proteins that become the basis of the vaccine. Then you need to make large amounts of purified recombinant chlamydial protein. We combine this with an adjuvant and deliver it to the animal. There are several ways to give the vaccine, either via a needle shot under the skin or even as a spray up the nose. We are using all our molecular microbiology skills firstly inside the laboratory, but then we need to test the vaccine on actual animals, firstly under controlled conditions in zoos or sanctuaries, then with wild koalas out in the field. There are lots of small steps, but we are definitely making progress and have trialled our vaccine with a population of wild koalas recently. This is a good example of taking basic science for the laboratory out into the field, and hopefully we can make a real difference in making sure that Australia has healthy koalas into the future.'

## Identifying bacteria

Identification of bacterial species is important for several reasons, such as deciding on antibacterial therapy for patients or identifying the clinical significance of bacterial infections.

Various techniques can be used to identify bacteria and they fall into three categories, namely phenotypic, immunological and genotypic methods.

**1. Phenotypic methods** can assist in identifying bacteria, and include:

- use of microscopy to differentiate bacteria on the basis of differences in cell shape, size and response to Gram stain, and physical features such as the presence or absence of a capsule (refer back to figures 6.17 and 6.22 on pages 239 and 243, respectively)
- use of different media to differentiate bacteria on the basis of variation in growth patterns, which can distinguish aerobic bacteria from facultative anaerobes
- use of a range of biochemical tests eliciting different bacterial responses (see figure 16.25).

Because bacteria can grow in standard microbiological broths and on agar plates, it is possible to use differences in growth patterns under a range of conditions to identify bacteria, such as by growth on non-selective, selective and differential media. Non-selective media or standard methods agar can be used to detect and count the amount of bacteria in the sample. Selective media contain compounds, such as antibiotics or growth nutrients, that selectively inhibit or enhance the growth of specific bacteria. A differential medium contains a substrate that, under the action of an enzyme, produces a coloured or fluorescent product. Growth media of this kind can differentiate and directly identify bacteria according to various chemical reactions that are carried out during growth.

**2. Immunological methods** make use of techniques including monoclonal antibodies, ELISA and immunofluorescence to identify bacteria.

**3. Genotypic and molecular methods** involve the examination of the genetic material of bacteria and make use of techniques such as gene probes, sequence analyses and plasmid fingerprinting to identify bacteria. Because of their speed and accuracy, genotypic techniques are increasingly important in the identification of bacteria and identification of the various serotypes of a single bacterial species.

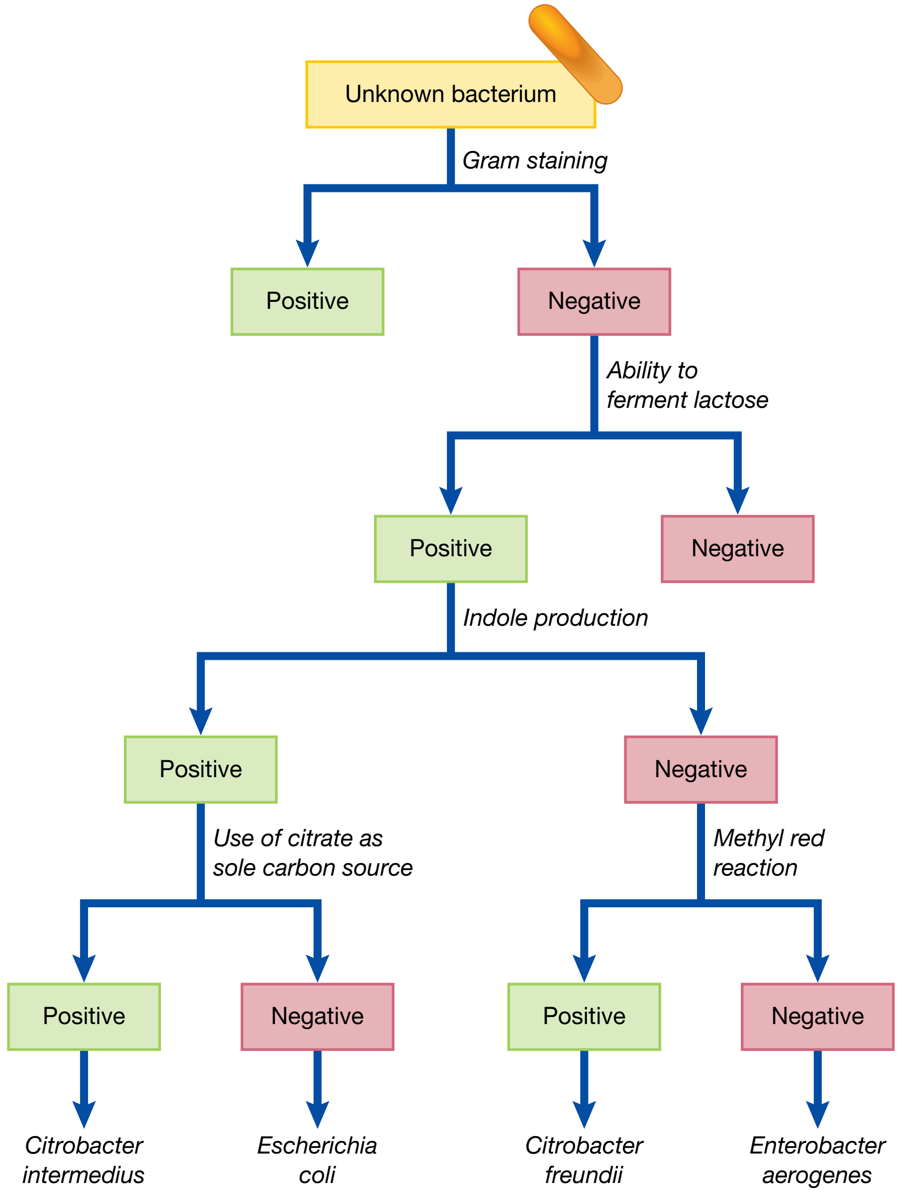



**FIGURE 16.25** Diagram showing the identification of four species of Gram-negative bacteria based on their reactions to several biochemical tests.

# Defence against infection

Treatments for bacterial and viral infections are different: **antibiotics** are used to treat bacterial infections, and antiviral agents are used for viral infections.

## Antibiotics

**Antibiotics** are a class of antimicrobial drug used in the treatment and prevention of bacterial infections. They act either by killing pathogenic bacteria or by inhibiting their growth. Some antibiotics are **narrow spectrum** and act against a limited variety of microorganisms; others are **broad spectrum** and act against many different kinds of pathogens.

One definition of an antibiotic is: A substance produced by a microorganism or a similar product produced wholly (**synthetic**) or partially (**semi-synthetic**)

**ODD FACT**

Ehrlich called Salvarsan 'compound 606' because he discovered it during an experiment with the 606th chemical he tested.

by chemical synthesis, that in low concentrations inhibits the growth of or kills microorganisms.

The first antibacterial substance that can be considered an **antibiotic** was Salvarsan, which was discovered in 1910 by the German physician and chemist Paul Ehrlich (1854–1915) and his team. Salvarsan was an arsenic compound that killed the bacteria *Treponema pallidum*, the cause of the sexually transmitted disease syphilis, which was prevalent at the time. The appearance of Salvarsan was an exciting advance because it marked the first compound that killed bacteria but was safe for the human patient. Salvarsan was the first of many antibiotic compounds that later appeared and markedly reduced the frightening mortality rates of infectious diseases, including post-surgical infections. It should be noted that other measures, in the late nineteenth century, also contributed to reducing the mortality rates of infectious diseases. These measures included the introduction of antiseptic procedures and improved sanitation and hygiene practices.

The next development occurred in 1932, when German pathologist Gerhard Domagk (1895–1964) discovered Prontosil, a sulfanilamide compound that had antibiotic properties. Domagk received the Nobel Prize in 1939 for this achievement.

studyon

Unit 4 | AOS 2 | Topic 2 | Concept 9

**Antibiotics**
Summary screen and practice questions

studyon

Unit 4 | AOS 2 | Topic 2 | Concept 9

**See more**
Antibiotics and natural selection

The mass production of penicillin during World War II and its success in treating a broad range of infections gave it the title of 'the miracle drug'. (See the box at the end of this section.) In 1943, streptomycin, isolated from the soil microbe *Streptomyces griseus*, was the first antibiotic to be effective in the treatment of tuberculosis. Then came the cephalosporins, new antimicrobial compounds with broad activity against Gram-positive and also some Gram-negative bacteria. Scientists continue to test new strains of bacteria and fungi in the search for new antibiotics.

## Actions of antibacterial drugs

The general action of an antibiotic is either by directly killing microorganisms (bactericidal) or by inhibiting their growth (bacteriostatic).

Different antibiotics act on different bacterial targets, as, for example:

- Penicillin inhibits cell wall synthesis in bacteria and so targets actively reproducing bacteria. Since human cells do not have peptidoglycan cell walls, penicillin does not attack human cells and so has low toxicity.
- Chloramphenicol, erythromycin, tetracyclines and streptomycin inhibit protein synthesis by acting on 70S ribosomes of prokaryotic cells.
- Sulfanilamide acts as an antimetabolite by competitively inhibiting enzyme activity in bacteria.
- Rifampin and quinolones inhibit nucleic acid synthesis.

Figure 16.26 shows the sites of action of various antibiotics.

**ODD FACT**

The primary concern in the use of penicillin is the risk of allergy that occurs in a low percentage of the population.

## Sensitivity tests

**Sensitivity tests are carried out to determine the antibiotic that would be most effective to treat a bacterial infection**. In one kind of sensitivity test, bacteria are spread across the surface of a solid nutrient plate and small discs with known concentrations of different antibiotics are put on the surface of the plate (see figure 16.27). If the bacteria are sensitive to a particular antibiotic, they will be killed by that antibiotic and so will not grow around the disc containing that drug. The lack of growth is shown by a clear zone around the disc of antibiotic. The larger the clear zone around a disc, the more effective is the antibiotic in that disc.

CIP = ciprofloxacin;
CN = gentamycin;
FD = nitrofurantoin;
P = penicillin;
VA = vancomycin

The important criterion for deciding which drug to use in chemotherapy is that it should be selectively toxic; that is, it should kill the infecting cells without destroying host cells. Unfortunately, some drugs that are effective against infective agents have some adverse effects on a host. These adverse effects are called **side effects** (see table 16.3). The side effects may not occur in all patients.

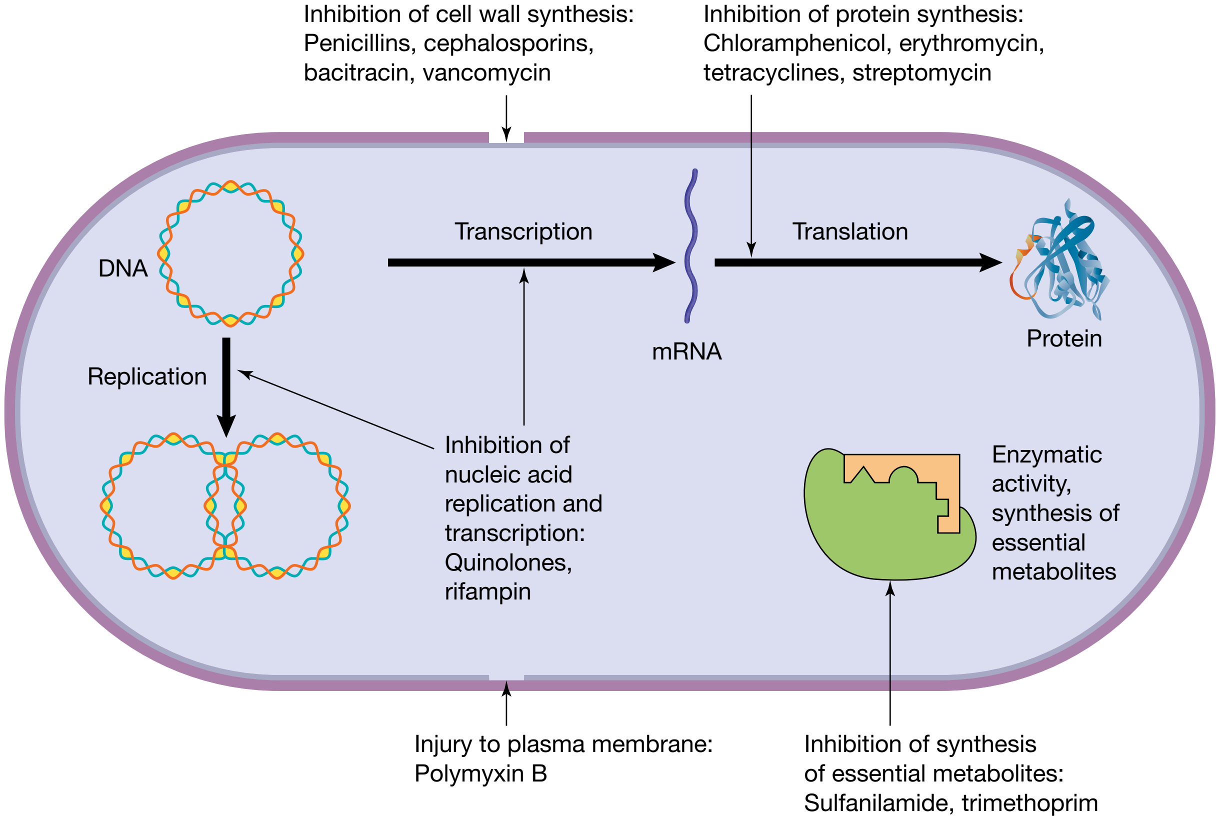



**FIGURE 16.26** Diagram showing the targets of various antibiotics on pathogenic bacteria.

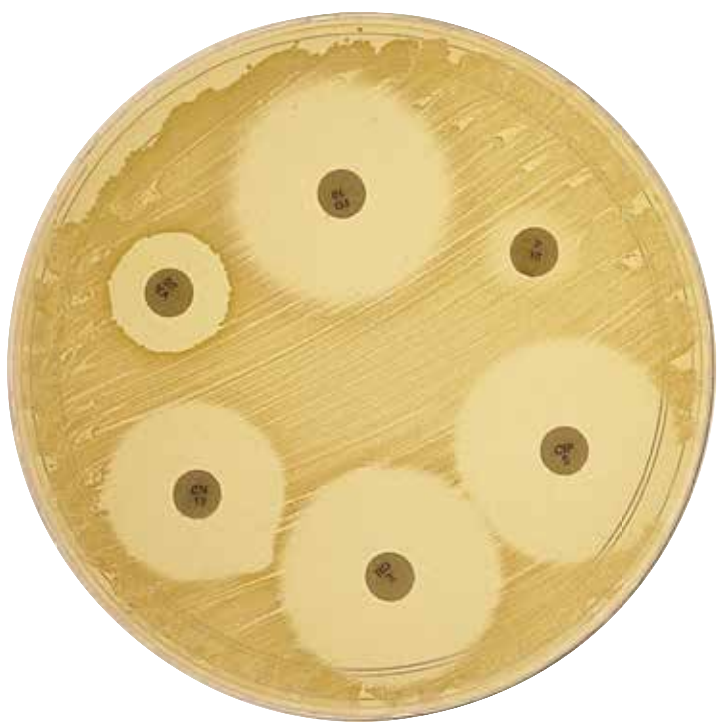


**FIGURE 16.27** Testing drug sensitivity of *Staphylococcus aureus*. From 12 o'clock, the drugs are FD10, P10, CIP5, RD2, CN10, VA30. Which drug is the most effective against the bacteria *Staphylococcus aureus*? Which drug is the least effective? (Image courtesy of Marjory Martin.)

**TABLE 16.3** Properties of some drugs. Some drugs are narrow spectrum and act on a limited variety of bacteria; others have a broad spectrum and act on many different kinds of bacteria. The side effects of a drug are taken into account when decisions are being made about which drug to use against a particular infection.

| Drug | Spectrum | Mode of action | Possible side effects |
|---|---|---|---|
| erythromycin | narrow (Gram +) | inhibits protein synthesis so bacteria unable to make essential compounds | gastrointestinal upset<br>liver damage |
| penicillin | narrow (Gram +) | inhibits cell wall synthesis so bacteria cannot reproduce | allergic responses |
| sulfonamides | broad (Gram + and −) | compete against and inhibit the bacteria | allergic responses<br>kidney and liver damage |
| tetracyclines | broad (Gram + and −, rickettsia and chlamydia) | inhibit protein synthesis so bacteria unable to make essential compounds | gastrointestinal upset<br>teeth discolouration<br>sometimes kidney and liver damage |

## Antiviral drugs

*Antiviral drugs* are a type of medication that is used specifically for treating viral infections. Typically, these drugs are effective only when the viruses are located within cells and are undergoing replication. Antibiotics are ineffective against viral infections.

The first types of partially effective antiviral drugs were nucleoside analogues that interfered with viral genome replication. The first specific inhibitors of viral protease activity were developed in the 1990s. This century has seen the development of a new drugs that inhibit viral enzymes. Because viruses use the host cells to replicate, designing safe and effective antiviral drugs is challenging. Antiviral drug targets must be ones that will not harm the host cells.

Most of the antiviral drugs presently available are designed to address the viruses that cause influenza A and B, hepatitis B and C (which can cause liver cancer), HIV (which causes AIDS), and various herpes viruses that cause many diseases ranging from cold sores to genital herpes infections. Advances in both our understanding of the details of the viral replication cycle in host cells and of the 3D structure of viral proteins has contributed to the development of many highly specific and effective antiviral drugs. (Refer back to figure 16.20 on page 747.)

Table 16.4 shows examples of antiviral drugs, their modes of action and the viruses they target.

Many millions of people living with HIV today now have access to treatment with **antiretroviral drugs**. These includes protease inhibitors, nucleoside reverse transcriptase inhibitors, integrase inhibitors and entry inhibitors (fusion inhibitors and CCR5). **For example, nucleoside reverse transcriptase inhibitors** (NRTIs) inhibit reverse transcriptase, the enzyme that transcribes a DNA copy of the RNA that is the genetic material of retroviruses, such as the human immunodeficiency virus (HIV). NRTIs compete with the substrates needed by the reverse transcriptase enzyme and block the replication of retroviruses. Examples of NRTI drugs include Epivir-HBV (generic name: lamivudine) used in the treatment of chronic hepatitis B infections and Retrovir (generic name: zidovudine) used in the treatment of HIV/AIDS.

studyon
Unit 4
AOS 2
Topic 2
Concept 10
**Antiviral drugs**
Summary screen and practice questions

**TABLE 16.4** Mode of action of various antiviral drugs and the viruses they target.

| Mode of action of antiviral drug | Example of drug | Target of drug |
|---|---|---|
| BLOCK entry of virus to host cell by<br>• binding to surface receptors on host cell | maraviroc | HIV/AIDS |
| BLOCK fusion of virus with host cell by<br>• binding to surface protein (gp41) on virus | enfuvirtide | HIV/AIDS |
| BLOCK uncoating of virus by inhibiting action of viral membrane protein M2 | amantadine<br>rimantadine | influenza A |
| BLOCK DNA replication through use of nucleoside analogues<br>• interfering with reverse transcriptase<br>• interfering with viral DNA polymerase | zidovudine (AZT)<br>rilpivirine<br>lamivudine<br>acyclovir | HIV/AIDS<br>HIV/AIDS<br>chronic hepatitis B<br>herpes simplex virus |
| BLOCK integration of viral DNA into host genome:<br>• use of integrase inhibitor | raltegravir | HIV/AIDS |
| BLOCK viral protease enzymes by<br>• use of protease inhibitors | saquinavir<br>boceprevir | HIV/AIDS<br>hepatitis C |
| INTERFERE with assembly of new infectious virions | mitoxantrone<br>daclatasvir | vaccinia virus<br>hepatitis C |
| BLOCK release of new virions from host cells<br>• inhibiting action of neuraminidase | zanamavir (Relenza)<br>oseltamavir (Tamiflu) | influenza A and B |
| STIMULATE immune system | interferon alpha | multiple viral infections, including hepatitis B |

Figure 16.28 shows the various stages of the HIV replication cycle in CD4 T-helper cells and the several targets of antiretroviral drugs.

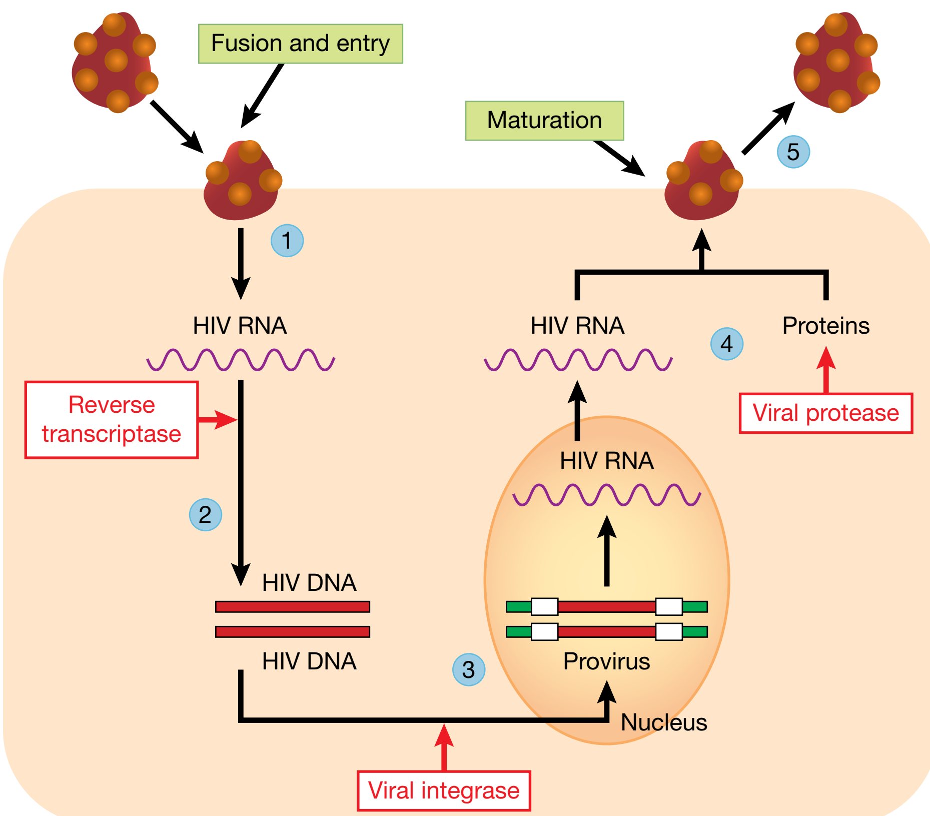



**FIGURE 16.28** Diagram showing the replication cycle of HIV in human CD4 T-helper cells. Three viral enzymes that are essential for this process are shown in red rectangles. These are logical targets for antiviral drugs.

**A new development in the clinical treatment of viral disease is the use of monoclonal antibodies (mAbs) as antiviral agents.** As at 2016, the only mAb approved for this purpose is **palivizumab**, which has been licensed for prevention of severe respiratory disease in high-risk infants and immunocompromised adults. Another mAb, **bavituximab**, has been in a Phase II clinical trial for hepatitis C since 2011, but no results have been reported as of 2016.

Refer back to chapter 8, page 363, for a description of monoclonal antibodies.

## Drug resistance is emerging

The effectiveness of almost all antibiotics and some antiviral drugs is declining. As pathogens mutate, drug resistance is becoming apparent in many pathogens, making them less susceptible to treatment with antibiotics.

Drug resistance has emerged as a result of the greatly increased, and in some cases unnecessary, use of antibiotics. The increase in the use of antibiotics is occurring not only for people but also for farm animals. This widespread use of antibiotics is acting as a positive selection factor for antibiotic resistance.

Figure 16.29 shows the increase in resistance of *Salmonella typhi,* the cause of typhoid fever, to the antibiotic ciprofloxacin, such that this antibiotic can no longer be used to treat the disease.

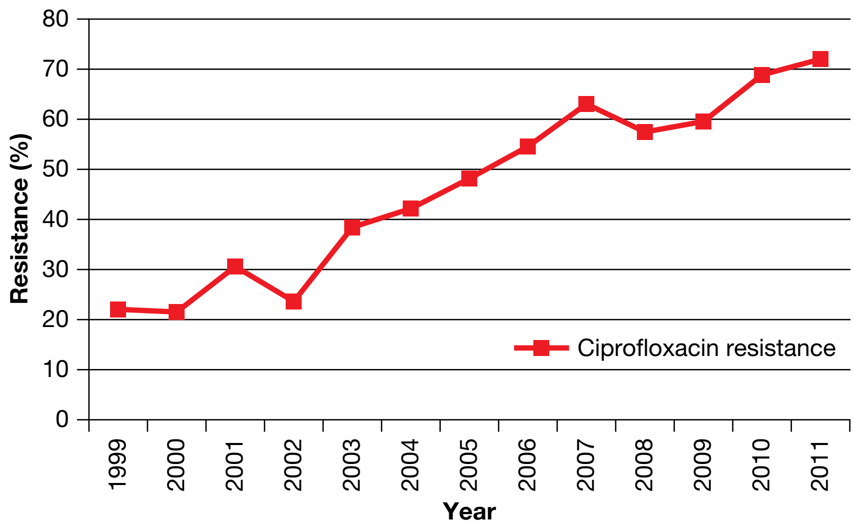



**FIGURE 16.29** Graph showing the increase in strains of *Salmonella typhi* resistant to the antibiotic ciprofloxacin in the period 1999 to 2011. Resistance has become so common that this drug can no longer be used to treat this disease. *S. typhi* is the cause of typhoid fever. (Data from US Centers for Disease Control and Prevention.)

Not only are more pathogens developing resistance to a particular drug, many pathogens now have multiple drug resistance. In May 2016, a woman with a persistent urinary tract infection was treated with **colistin**, an antibiotic that is regarded as ‘the antibiotic of last resort’, to be used only when all others have been ineffective. What was frightening was the finding that the bacteria responsible for this infection were not killed by the administration of colistin. Further, it was found that these bacteria had a total of 15 genes for resistance to other antibiotics. In many cases, the genes for resistance are carried on bacterial plasmids and so can be easily and rapidly spread from resistant to non-resistant bacteria.

**ODD FACT**

Colistin-resistant bacteria have been identified in a sample taken from the intestine of a pig on an American farm.

## PENICILLIN — THE MIRACLE DRUG

In 1928, a bacteriologist working at St Mary's Hospital in London (see figure 16.30a) noticed that a blue-green mould had contaminated one of his bacterial plates. Although bacteria were growing elsewhere on the plate, those near the mould had been killed. The bacteriologist, Alexander Fleming, correctly reasoned that the mould must have produced a substance that killed the bacteria. The mould was *Penicillium notatum* and the antibiotic it made was called **penicillin**. Fleming experimented with penicillin on bacterial cultures in glass containers but, because of poor results, he incorrectly concluded that penicillin would not work in living tissue and lost interest in the drug.

For ten years nothing more was done with penicillin until a South Australian medical graduate, Howard Florey, who was Professor of Pathology at Oxford University in England, worked on Fleming's mould. Florey and a German co-worker and chemist, Ernst Chain, extracted penicillin and carried out an important experiment with mice. They injected eight mice with a deadly strain of *Streptococcus* bacteria. Four mice were injected with penicillin, the other four were untreated. The mice receiving the penicillin all survived; the others died. Clinical trials with humans were carried out in 1941 when sufficient amounts of the drug could be extracted from the mould.

Penicillin was called the miracle drug because it could be used to treat many formerly lethal infections. Patients with septic infections filled whole hospital wards and here at last was a drug that ensured survival for virtually all these cases. Because of the war in Europe, Florey went to America in July 1941 to persuade American companies to produce penicillin on a large-scale basis. Eventually another strain, *Penicillium chrysogenum*, was used in cultures because it gave a much higher yield of penicillin. This higher yield species of *Penicillium* was originally found on a rotting cantaloupe.

Fleming was a brilliant observer, but it took someone with persistence and drive like Florey to carry through with the experiments that were necessary to prove the drug's effectiveness and potential for the treatment of infections. Sir Alexander Fleming (1881–1955), Lord Florey (1898–1968) and Ernst Chain (1906–1979) received many awards for their work, including the Nobel Prize in Physiology or Medicine, which they shared in 1945.

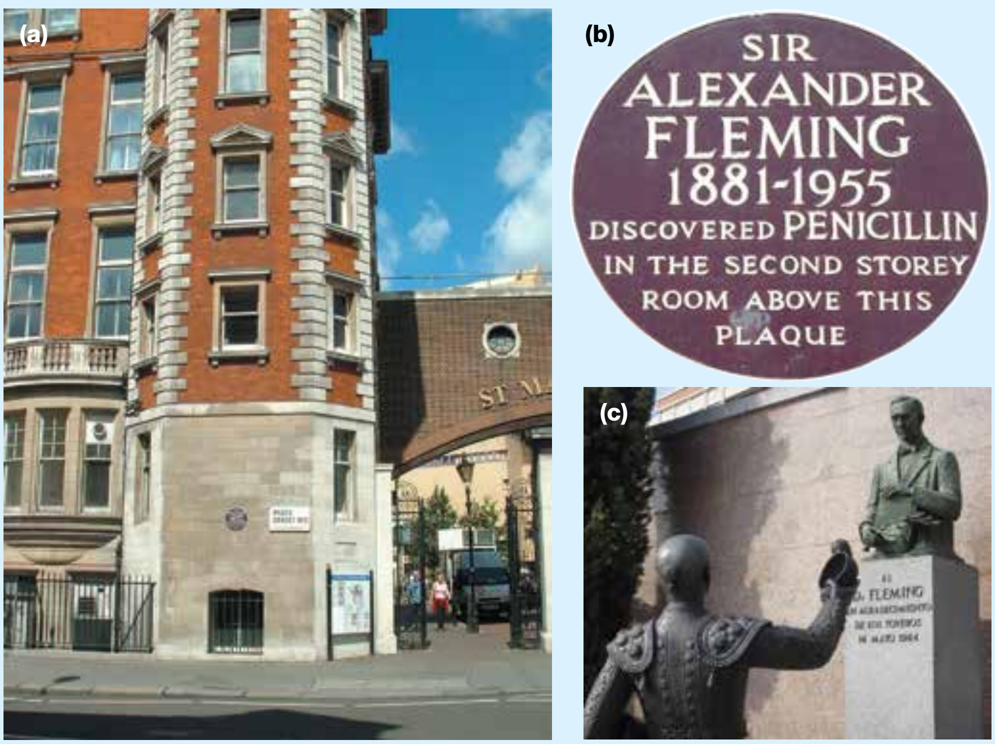



**FIGURE 16.30** **(a)** St Mary's Hospital, London, where Sir Alexander Fleming discovered penicillin. Note plaque near street sign. **(b)** The plaque commemorates Fleming's discovery and indicates the position of the laboratory. **(c)** A matador salutes Sir Alexander Fleming. The inscription reads 'To Dr Fleming with the gratitude of bullfighters'. Their lives were in less danger because of Fleming's discovery of penicillin. (Images courtesy of Marjory Martin.)

## ODD FACT

Alexander Fleming grew *Penicillium* mould on small discs of black paper, mounted them in glass frames and gave them as gifts to friends. One of these was auctioned by Christie's Auction House in July 1997. The Science Museum in London bought it for £14 950 (about A$34 100).

## KEY IDEAS

- **Identification of pathogens is important for both treatment and prevention.**
- **A range of different methods is available to identify both bacteria and viruses.**
- **Antibiotics are agents that kill or inhibit bacteria, using different bacterial targets.**
- **Antiviral drugs have been developed that target key viral enzymes involved in the viral replication cycle.**
- **Drug resistance among pathogens is increasing, driven by the greatly increased use of antibiotics for people and for farm animals.**

## QUICK CHECK

9 What is meant by 'multiple resistance' in regard to a bacterial pathogen?

10 Identify the following statements as true or false:

a Viruses are not destroyed by antibiotics.

b The presence of a particular viral protein in a person's serum could be identified using ELISA.

c Different viruses cannot be distinguished on the basis of their shape.

d Antibiotic resistance genes in bacteria can be carried on plasmids.

e Inhibiting the viral enzyme integrase is one strategy for controlling HIV.

# BIOCHALLENGE

The H7N9 subtype of influenza A virus is a subtype of bird influenza viruses that had been detected in various bird species from time to time. However, in March 2013 in China, this virus appeared for the first time in people and caused a severe form of influenza. Figure 16.31 shows the origin of this H7N9 subtype.

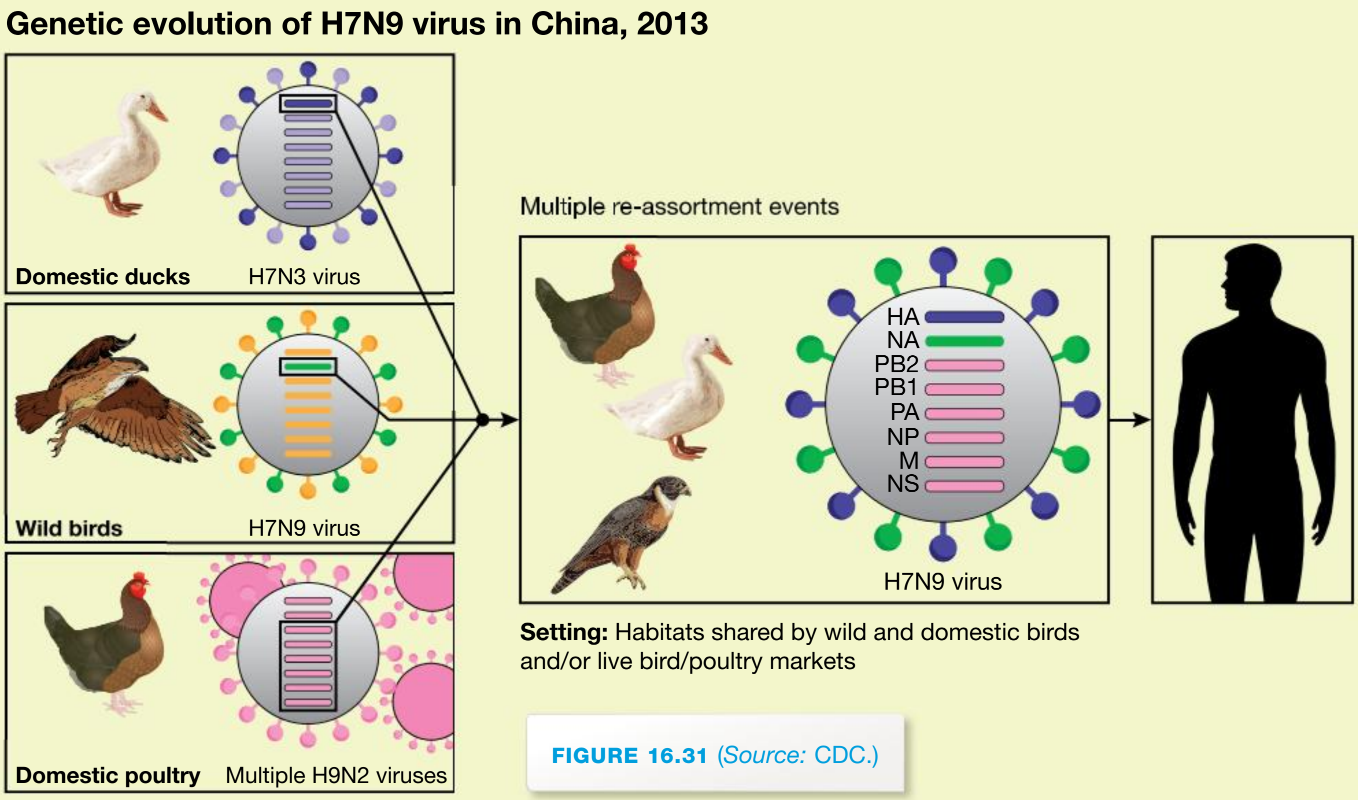



**FIGURE 16.31** (*Source:* CDC.)

1 Examine figure 16.31.
   a Outline the origin of the H7N9 influenza virus that infected people in China in 2013.
   b How does this new H7N9 subtype differ from the virus found in the past in wild birds?
   c Which group of birds was the source of the hemagglutinin in the new H7N9 subtype?
   d Would you predict that when this virus infected people, they would have been protected by their immune system? Explain.
   e What term is given to the process of producing new subtypes of influenza A viruses?
   f Two features of influenza A enable it to be involved in genetic re-assortments. Identify these two features.

2 Bacteria of the genus *Salmonella* are major foodborne pathogens that are a threat to public health. The main reservoir of *Salmonella* bacteria is the intestinal tract of many domestic animals whose products are part of the human food chain, including eggs, poultry and various meats. *Salmonella enterica* is a common cause of food poisoning. People infected with *S. enterica* from eating contaminated foods that are raw or undercooked can develop *Salmonella* poisoning or 'gastro'. When outbreaks of *Salmonella* infection occur, public health officials take rapid action to identify the source of the contamination. *S. enterica* bacteria are identical in microscopic appearance. However, samples from many different sources show slight variations in two antigens: a cell-surface antigen, O, and a flagellar antigen, H. These variants of one species of bacteria are termed serotypes or serovars. The many serotypes of *S. enterica* include serotype Typhimurium, serotype Montevideo and serotype Dublin.
   a Suggest how serotyping could help identify the cause of an outbreak of *S. enterica* gastroenteritis.
   b Which, if any, of the following foods might be a source of *S. enterica* that caused Salmonella poisoning among guests who dined at a hotel on the same afternoon: (i) raw egg mayonnaise, (ii) salami, or (iii) roast pork. Give a reason for your decisions.
   c Explain how *Salmonella* might move from chickens to people.
   d Common advice is to use a different chopping board for cutting up raw red meat and chicken from that used for cutting up salad vegetables. Does this advice have a solid foundation? Explain.

   In 2009, testing of chicken flocks in the UK revealed the presence of 36 different *S. enterica* serotypes.

   e Identify one technique that you would predict would enable you to distinguish between these different *Salmonella* serotypes.

# Chapter review

## Key words

antibiotics
antigenic drift
antigenic shift
antiretroviral drugs
bavituximab
broad spectrum
colistin
epidemic
hemagglutinin (HA)
influenza virus particle (virion)
microcephaly
narrow spectrum
neuraminidase (NA)
pandemic
palivizumab
penicillin
rational drug design
re-assortment
semi-synthetic
sensitivity tests
side effects
synthetic
virulence

## Questions

1 **Making connections** → Construct a concept map using as many of the key words above as possible.

2 **Developing logical explanations** → Formulate a biologically valid explanation for each of the following observations:
   a New influenza vaccines are typically developed on an annual basis.
   b Epidemics tend to begin without warning.
   c Antibiotics are not effective against virus infections.

3 **Interpreting graphical data and communicating ideas** → Examine figure 16.32, which shows the death rates from influenza for males in various age groups in New South Wales during two influenza pandemics.

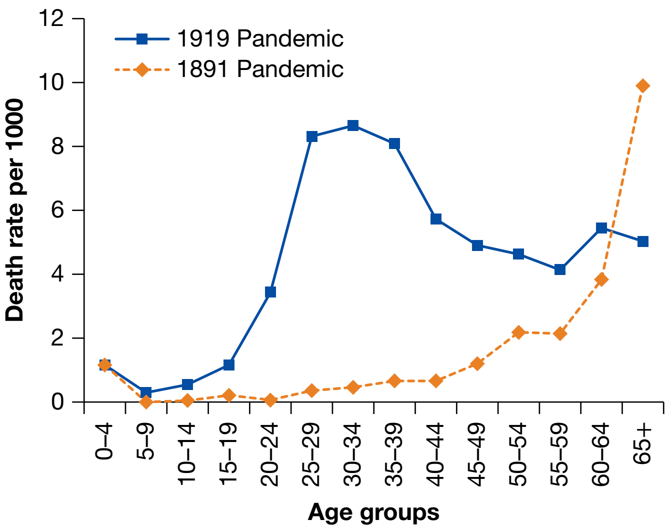



**FIGURE 16.32** Death rates from influenza for males in various age groups in New South Wales during the 1919 influenza pandemic and the 1891 influenza pandemic. (Reproduced by permission, NSW Ministry of Health, 2016.)

   a What pattern in the death rates is apparent in the 1919 influenza pandemic?
   b In 1919, what was the death rate for males in the 25 to 29 age group?
   c In 1919, what was the death rate for males in the 65+ age group?
   d What pattern in the death rates is apparent in the 1891 influenza pandemic?
   e In 1891, what was the death rate for males in the 25 to 29 age group?
   f In 1891, what was the death rate for males in the 65+ age group?
   g Do these data indicate that the same strain of influenza was responsible for these two pandemics? Explain.

   Overall in Australia, the death rate from influenza from the 1919 pandemic was about three deaths per 1000 population. However, death rates differed in various countries.

   h In the Pacific country of Western Samoa during the 1918–1919 pandemic, there were 8500 deaths in a total population of only 38 000. What was the overall death rate for that population?

4 **Demonstrating understanding and communication** → SARS or severe acute respiratory syndrome was first reported in China in November 2002. It took just a few months for this highly contagious and sometimes fatal respiratory illness to spread through 26 countries in Asia, North and South America, and Europe. This outbreak ceased in 2004. The causative agent was a coronavirus.
   a Does this suggest that SARS was an existing or a new virus? Explain.
   b Identify a factor that might have contributed to the speed of this spread.
   c Would this event be identified as a SARS epidemic or a SARS pandemic? Explain.

   In Singapore, there were 238 cases of SARS with 33 deaths. Across all countries affected by SARS, there were 8096 cases with 774 deaths.

   d What was the overall case mortality rate for SARS in Singapore?

**e** Which, if any, of the following agents might have been tried in the treatment of SARS: penicillin or interferon? Explain.

However, antiviral drugs were of little benefit in the treatment of SARS.

**f** In searching for an effective antiviral drug, what approach might be used?

Read the following, noting in particular the statement in italics:

'When SARS was first reported, its cause was not known. A key task was to find the causative agent and the top suspect was a coronavirus, now known as SARS coronavirus. *Scientists in the Netherlands demonstrated that the SARS coronavirus fulfilled Koch's postulates and so could confirm that the SARS coronavirus was indeed the causative agent of SARS.*'

**g** Explain why the use of Koch's postulates was a key to identifying the cause of SARS. (You may wish to review Koch's postulate (refer back to Chapter 6, page 222).

**h** Why is identifying the causative agent of a new disease important?

After SARS, changes to public health practices and procedures were introduced into the affected countries in order to minimise or prevent epidemics of SARS in the future. One change in Taiwan was to replace handshakes (for greeting people) with courtesy bows.

**i** Would this change be expected to reduce the risk of transmission of SARS?

**j** Suggest three other possible changes that might have been introduced. (Use your imagination, as there is no single correct answer.)

**5** **Demonstrating understanding** → Two different subtypes of influenza viruses that affect non-human mammals include viruses with the following identifiers:

i A/Equine/Prague/1/56 (H7N7)

ii A/Tern/South Africa/1/61(H5N3).

For each virus, outline the information that the identifier provides.

**6** **Demonstrating understanding and communication** → The virus responsible for the influenza pandemic of 2009 was examined, and the eight RNA segments that make up its genome were shown to have been derived from multiple sources over time. The purported history of this 2009 influenza A virus is shown in figure 16.33 below.

Examine this figure and answer the following questions:

**a** What is the subtype of this virus?

**b** Of what nucleic acid are these eight segments of the virus genome composed?

**c** What biological term is given to the process that produced the 2009 H1N1 pandemic virus?

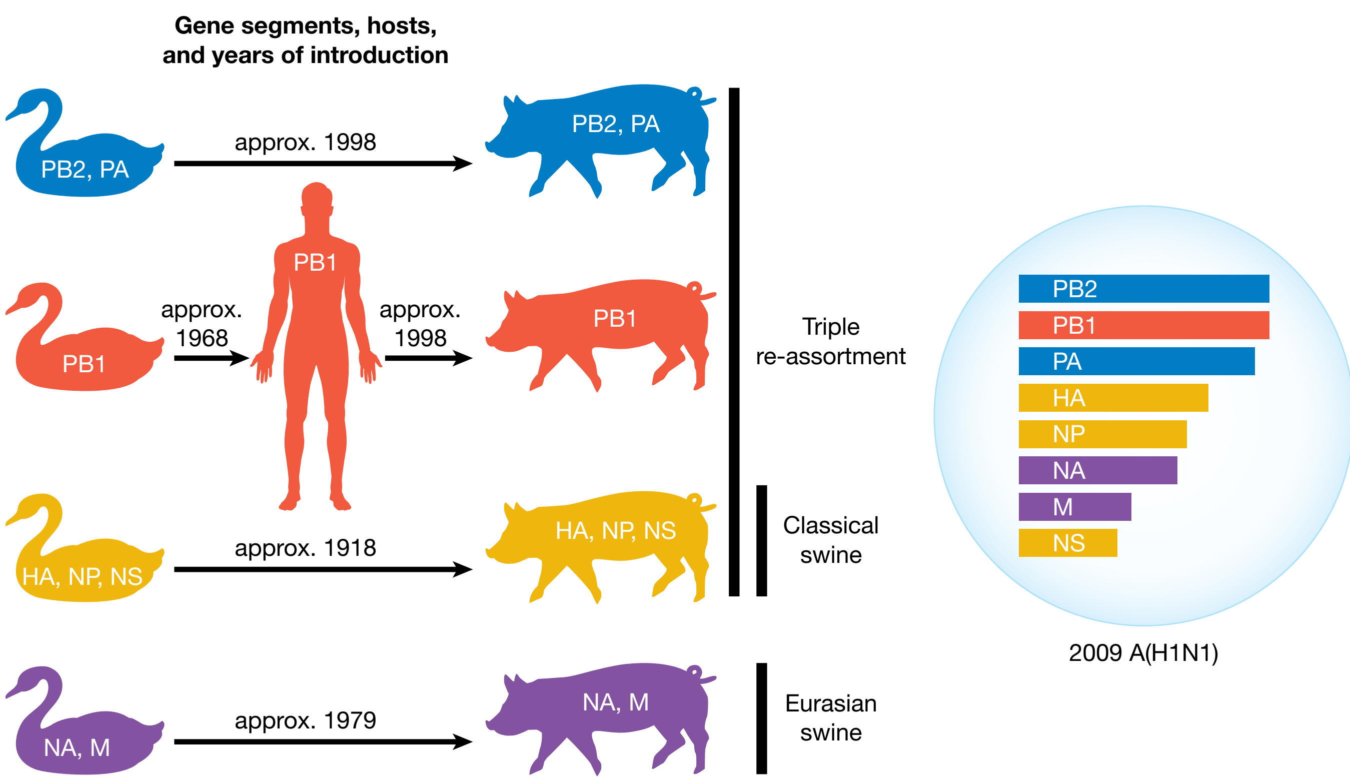



**FIGURE 16.33** Genetic origin of the 2009 pandemic H1N1 influenza A virus.

**d** What was the origin of the gene segment encoding the hemagglutinin protein?

**e** What was the origin of the gene segment that encoded the neuraminidase enzyme?

**f** Describe in your own words the exact history of the PB1 segment.

**7** **Discussion question ➜ Read the following statement from a commentary entitled:** *Ethics of reconstructing Spanish Flu: Is it wise to resurrect a deadly virus?* This commentary was **published in 2007 after the publication of the full genomic sequence of the Spanish flu virus in 2005. (Source: van Aken, J.** *Heredity* (2007) 98, 1–2, published online October 2006.)

'Recently, a team of US scientists resurrected a virus that has since been labelled "perhaps the most effective bioweapons agent now known" . . . . Now the full sequence of the Spanish Flu virus has been published . . . and the virus itself reconstructed. It proved to be as fatal as the original. When tested on mice, it killed the animals more quickly than any other flu virus ever tested.'

**a** The writer argues that the published information poses a biosecurity risk and that it should not have been published. Do you and your classmates share this view or not? For what reasons?

**b** The writer asserts that experiments of this nature — ones that produce sensitive information that could be put to damaging use by rogue scientists — should be stopped before they start. Do you agree or disagree? What are your reasons?

CHAPTER

# 17 Practical investigations

FIGURE 17.1 Some biological investigations involve the performance of laboratory-based experiments that are designed to test a particular hypothesis that has been generated by questions resulting from prior observations. Practical experimentation is part of a process of scientific research methodology that begins with observations and hypothesis testing, proceeding through several steps, including the publication of results.

## KEY KNOWLEDGE

This chapter is designed to enable students to:

- understand scientific research methodology and the typical steps in a scientific approach
- design and conduct a laboratory investigation that follows scientific process
- recognise that investigations must adhere to relevant safety, health and bioethical guidelines
- recognise the nature and characteristics of experimental variables
- distinguish between quantitative and qualitative datasets and to assess their validity, accuracy and precision
- identify that experimental results may either confirm or disprove a hypothesis, model or theory
- apply the conventions of scientific report writing and poster presentation.

# How does science operate?

In the medieval period (fifth to the fifteenth century), diseases were not understood. Their causes were variously identified as miasmas (bad smells), or an imbalance in the four bodily fluids (blood, black bile, yellow bile and phlegm), or the action of demons, or the result of sin, or even the alignment of planets in the night sky. The most commonly held view was the miasma theory of infectious disease.

In the twenty-first century, we have a very different view. The investigations of Louis Pasteur and Robert Koch in the 1870s led to the rejection of the miasma theory of disease, replacing it with the germ theory of infectious diseases. This total replacement of one theory by another reminds us that **biological knowledge does not only grow by adding new truths to old ones; instead, earlier views are often completely rejected and replaced by new theories or explanatory frameworks.**

Rejection of earlier theories and their replacement by new theories is not the end of the story. Over time, new theories may be refined or 'tweaked', such as occurs when new data or exceptions are identified. For example, the germ theory of disease, established through the research of Pasteur and Koch in the late nineteenth century, initially provided an adequate framework to explain infectious diseases. However, this theory recognised bacteria as the only cause of infectious diseases. The germ theory was later refined to accommodate non-cellular pathogenic agents — firstly, the inferred existence of viruses at the turn of the twentieth century, and then the discovery of prions (misfolded proteins) in the late twentieth century.

How are changes in our understanding of aspects of biology achieved? What might cause the rejection of an earlier theory? How might refinements of an existing theory come about? These changes typically begin with observations that raise questions that may be investigated in various ways, such as:

- carefully planned laboratory-based or field-based experiments designed to confirm or reject a particular hypothesis, such as the discovery that variant Creutzfeldt-Jakob disease (vCJD) in people was caused by the ingestion of beef from cattle with bovine spongiform encephalopathy (BSE) (refer back to chapter 6, in the section titled Other prion diseases)
- critical reinterpretation of previously accepted facts, producing a new framework, as, for example, the replacement of the theory of special creation of species with the Darwin–Wallace theory of evolution by natural selection (refer back to chapter 10, in the section titled Species: unchanging or not?)
- collection and analysis of new data, as, for example, the discovery and study of previously unknown fossilised remains, such as the fossilised remains of Denisovan humans and *Homo floresiensis*, which provide new insights to hominin evolution (refer back to chapter 12, in the sections titled The Denisovans, and The Hobbit: *Homo floresiensis*, respectively)
- identification and exploration of patterns or anomalies, as, for example, the research of Dr John Snow (1813–1858), which challenged the miasma theory of disease by providing evidence that cholera was spread by contaminated water. Snow mapped the occurrence of cases of cholera during an epidemic, which revealed a cluster of cholera cases around one water pump in one street in the London suburb of Soho.

Not all advances in biological knowledge are initiated from the results of carefully planned experiments. Some major advances in biology are initiated by unplanned chance events, such as a discovery in 1928 by Sir Alexander Fleming (1881–1955). Fleming was studying *Staphylococcus* bacteria, which he was growing on nutrient agar in petri dishes. Some of his plates became contaminated by the growth of airborne mould spores. Instead of discarding these plates as contaminated and useless, Fleming recognised that clear areas around the growing mould were a signal that the mould was producing a substance that inhibited bacterial growth — indeed it was, and this substance turned out to be the antibiotic penicillin.

## Why wash your hands?

One important advance in the prevention of infectious diseases was handwashing. This routine practice is important, not only in hospitals (see figure 17.2), but also in food-handling and food-preparation settings and as part of personal hygiene after going to the toilet. Let's explore how the value of this simple disease-prevention practice was first recognised. This exploration will introduce you to the key steps in an investigation carried out according to scientific research methodology.

**FIGURE 17.2** Today, washing hands or 'scrubbing up' is an essential pre-surgical routine; however, prior to the mid 1800s, this was seen as an odd and quite unnecessary practice. The perception began to change though as a result of investigations by one doctor, Ignaz Semmelweis.

### The scientific approach of Semmelweis

In 1846, the Hungarian doctor Ignaz Semmelweis (1818–1865) was appointed to the position of Assistant in Obstetrics at the Vienna General Hospital. The maternity section at the hospital was organised into two clinics, and each month hundreds of expectant mothers were admitted to these clinics. Depending on the day of the week of her admission, an expectant mother was admitted to either Clinic 1 or Clinic 2.

Semmelweis became acutely aware of the tragic deaths of many new mothers in the maternity clinics from a disease known as **puerperal fever**, commonly called 'childbed fever'. For example, in 1846 the average monthly death rate was almost 14 new mothers per 100 births (see figure 17.3). Translating this rate to the actual number of people reveals a horrifying death toll of 564 new mothers in that year: 459 in Clinic 1 and 105 in Clinic 2.

puerperal: pronounced poo-ER-per-al

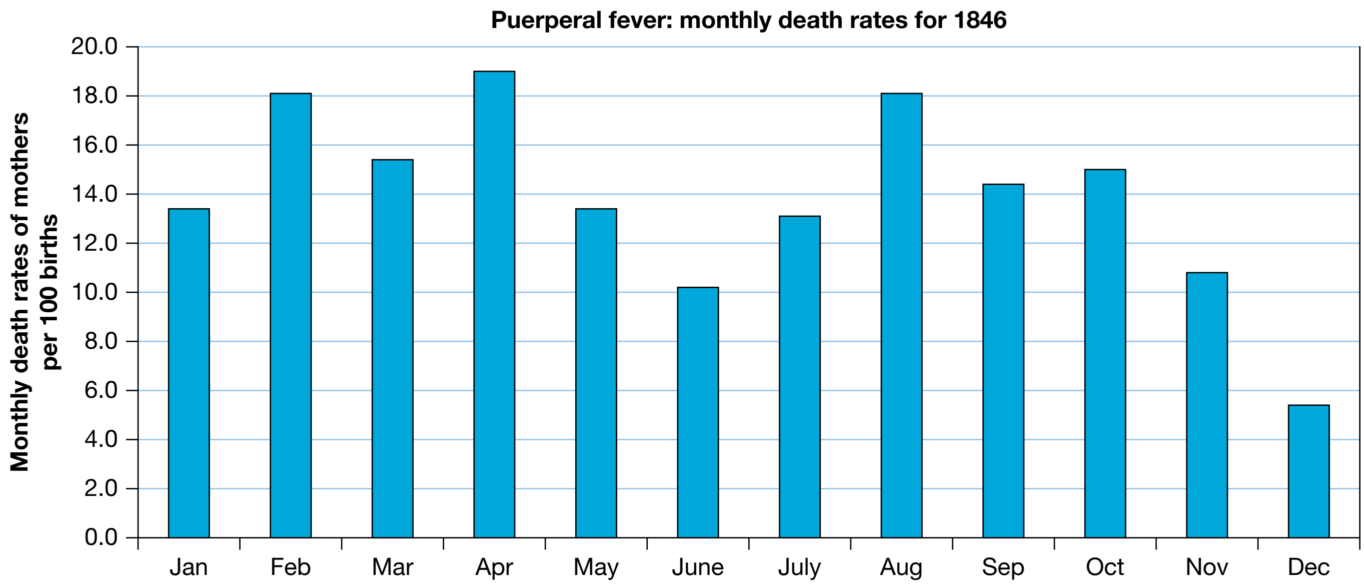



**FIGURE 17.3** Bar graph showing the monthly death rates of new mothers per 100 births during 1846 in Clinic 1 of the Vienna General Hospital. The average monthly death rate in this clinic over the year was 13.9 deaths per 100 births.

At that time, the cause of puerperal fever and its mode of transmission were not known. The prevailing view of the cause of disease was miasma or 'bad air'. Semmelweis was disturbed by the tragic loss of life of new mothers from this disease, so he decided to investigate. **The research question asked by Semmelweis was: *What is the cause of puerperal fever?*** He wanted to investigate the research question in order that he might identify ways of reducing the shocking death rate of new mothers from this disease.

In undertaking his investigation, Semmelweis used an approach that involved a logical sequence of steps that is termed **scientific research methodology**, or, simply, **scientific method**. Let's look at how Semmelweis approached this investigation into why so many new mothers were dying from puerperal fever.

## Making observations

To begin his investigation, Semmelweis made many observations relevant to puerperal fever. These observations provided Semmelweis with important background information to guide his investigation.

Semmelweis went to the hospital records and gathered data on death rates of new mothers from puerperal fever in previous years. He observed that the annual death rates of new mothers in the maternity section of the Vienna General Hospital had varied markedly in previous years. For example, in the period from 1813 to 1822, the average yearly death rate for new mothers was 1.9 per 100 births, but, in the period from 1823 to 1833, this death rate rose to a significantly higher level, averaging 5.4 deaths per 100 births, a nearly three-fold increase (see table 17.1).

It is not easy to see trends or patterns from data shown in a table. Check out figure 17.4, which shows the same dataset presented as a bar graph. Note that the change in the death rates from 1823 is far more apparent when the data is shown as a graph rather than as a table. (Keep this in mind when you make decisions about the presentation of your experimental results.)

**TABLE 17.1** Yearly death rates from puerperal fever in mothers giving birth in the Vienna General Hospital in the period 1813 to 1832. The average rate for the 10-year period 1813 to 1822 was 1.9 per cent, but it increased to 5.4 per cent in the 10-year period commencing 1823. This dataset is presented as a table. What other method might be used to show this data?

| Year | death rate |
|---|---|
| 1813 | 1.1 |
| 1814 | 3.2 |
| 1815 | 0.7 |
| 1816 | 0.5 |
| 1817 | 0.9 |
| 1818 | 2.2 |
| 1819 | 5.0 |
| 1820 | 2.5 |
| 1821 | 1.7 |
| 1822 | 0.8 |
| 1823 | 7.5 |
| 1824 | 4.9 |
| 1825 | 8.8 |
| 1826 | 8.1 |
| 1827 | 2.2 |
| 1828 | 3.6 |
| 1829 | 4.6 |
| 1830 | 4.0 |
| 1831 | 6.6 |
| 1832 | 3.2 |

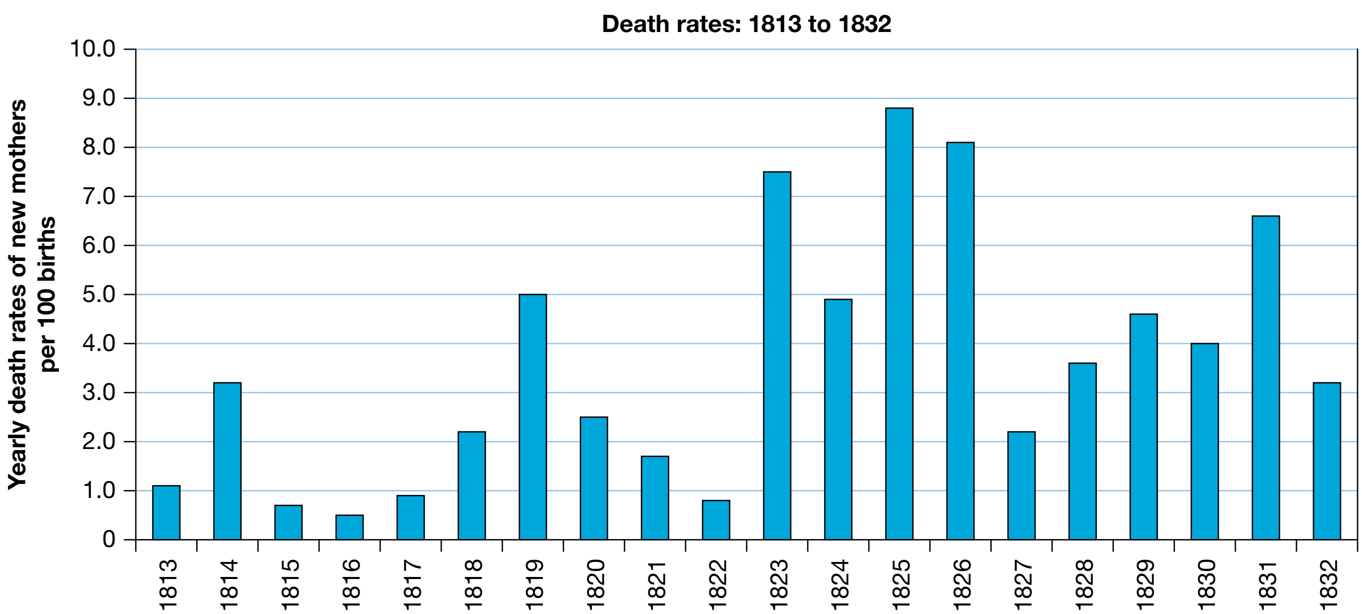



**FIGURE 17.4** Bar graph showing the yearly death rates from puerperal fever of new mothers per 100 births in the Vienna General Hospital in the period from 1813 to 1832.

Semmelweis also observed that, in 1833, the maternity section of the hospital was organised into two separate maternity clinics. He gathered more data on death rates in the two clinics and noted that:

- at first, from 1833 to 1839, the two clinics had similar death rates from puerperal fever, but
- from 1840 onwards, the death rates differed with the rate in Clinic 1 being higher than that in Clinic 2.

Refer to figure 17.5 to see a comparison of the yearly deaths rates from puerperal fever in the two clinics in the period 1833 to 1846.

Semmelweis also observed that, initially, expectant mothers in both clinics were treated and examined by doctors, midwives and medical students. He noted that, from 1840 onwards, staffing changes were made so that mothers in Clinic 1 were under the care of doctors and medical students only, while expectant mothers in Clinic 2 were cared for solely by midwives.

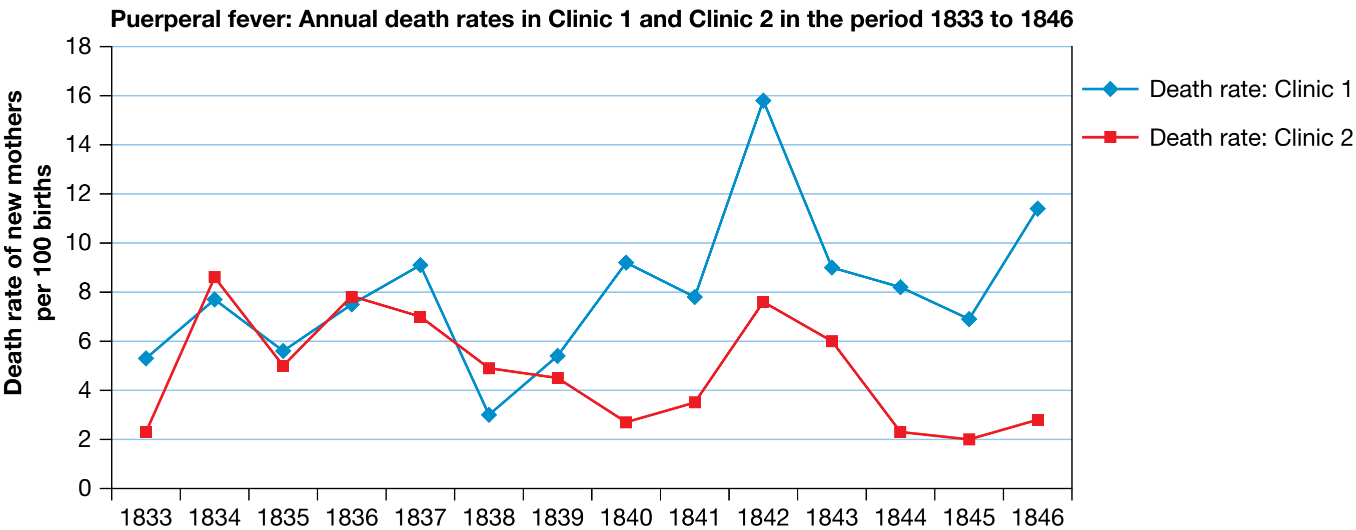



**FIGURE 17.5** Line graph showing the yearly death rates of mothers from puerperal fever after giving birth in the period from 1833 to 1846 in Clinic 1 and Clinic 2. Examine this figure carefully and observe that the death rates in the two clinics were initially similar. However, from 1840, the death rate in Clinic 1 was significantly greater than that in Clinic 2. What does this suggest?

Other data gathered by Semmelweis included the following facts:

A. In 1823, the Vienna General Hospital introduced the subject *Pathological Anatomy* for its medical students. This subject required students to dissect cadavers in the hospital morgue and study their diseased organs. These cadavers included the bodies of women who had died from puerperal fever in the maternity clinics. In addition, doctors routinely performed autopsies in the morgue on patients who had died in the hospital.

B. From 1823, doctors and medical students moved between the morgue, where they carried out autopsies and dissections, to the maternity clinic(s), where they performed genital examinations of women about to give birth and of women who had recently given birth.

C. The training of midwives at no time ever involved cadaver dissection.

Putting all his observations and data into order, Semmelweis established a picture of the variations in death rates from puerperal fever, which is summarised in figure 17.6. Note that death rates are given as yearly averages for stated periods. What is striking about these figures is the obvious variation in yearly death rates — this of course was noted by Semmelweis and raised questions for him (see following section).

| | | | | |
|---|---|---|---|---|
| **Average annual death rate (%)** | **1.9** | **5.4** | **Clinic 1: 6.6**<br>**Clinic 2: 5.4** | **Clinic 1: 9.9**<br>**Clinic 2: 4.0** |
| **Period** | **1813–1822** | **1823–1832** | **1833–1840** | **1841–1846** |
| | | ↑ *Pathological Anatomy* subject introduced for medical students | ↑ Two clinics: both staffed by doctors, students and midwives | ↑ Two clinics with different staff; only midwives in clinic 2 |

**FIGURE 17.6** Diagram summarising the average yearly death rates of new mothers per 100 births over four time periods at the Vienna General Hospital. The periods are separated by particular events in the history of the hospital. Arrows indicate changes at the beginning of each of the last three periods.

## Asking questions

Semmelweis reviewed the observations he had made and began to interpret the data he had gathered during the first step in his investigation into the cause of deaths of new mothers from puerperal fever. As he thought about this information, many questions arose in his mind:

- Were the changes shown in figure 17.6, above, related to the changes in the death rates from puerperal fever or were these changes incidental and unrelated?
- Was the introduction of the subject *Pathological Anatomy* in 1823 in some way related to the observed increase in death rates from puerperal fever from that time?
- Why did the two maternity clinics established in 1833 have similar death rates in the period 1833 to 1839 while, from 1840 onwards, the death rate from puerperal fever was higher in Clinic 1 than in Clinic 2?

After considering his observations and rejecting some possible causes, Semmelweis reached the following tentative conclusions:

- The dissections carried out by medical students of the bodies of women who died from puerperal fever contaminated the students' hands with an unknown agent of some sort, which Semmelweis called 'cadaveric particles'.
- Likewise, the hands of doctors who carried out demonstration dissections or who carried out autopsies on patients who died of puerperal fever could become contaminated by cadaveric particles.
- These cadaveric particles could be transferred from the hands of doctors and medical students to mothers-to-be and new mothers in the maternity clinic when they underwent genital examinations.
- This transfer of cadaveric particles to new mothers caused them to develop puerperal fever.

Can you suggest why Semmelweis called this unknown agent cadaveric particles? Why didn't he simply label this agent 'pathogenic bacteria' or 'germs'? Recall that this was the 1840s, and the germ theory formulated from the research of Pasteur and Koch was yet to come. So, although Semmelweis named the unknown cause of puerperal fever as cadaveric particles, he had no idea what these particles were.

Semmelweis realised that cadaveric particles and their transfer from the hands of doctors and students to the maternity clinic provided a tentative explanation supporting the observations that he had made, namely:

1. The initial increase in death rate from 1823 (refer back to figure 17.4) might be explained as a result of the introduction of *Pathological Anatomy,* a subject that exposed students to cadaveric particles through their dissections and handling of diseased organs of women who died of puerperal fever.

2. The higher death rate in Clinic 1 as compared to that in Clinic 2 after 1840 (refer back to figure 17.5) might be explained as a consequence of the fact that, from 1840, Clinic 1 was staffed by medical students and doctors who could introduce cadaveric particles to mothers in that clinic. In contrast, Clinic 2 was staffed by midwives who did not come into contact with cadavers.

### Formulating a testable hypothesis

Based on his research, Semmelweis realised that if these cadaveric particles could somehow be destroyed, the death rate from puerperal fever would fall.

**ODD FACT**

Chlorinated lime is a powerful oxidising agent that has various uses today: as a disinfectant, a water purifier, a sanitiser, and a bleach.

The question remained: How might these cadaveric particles be removed? Semmelweis noted that washing hands with soap did not remove the strong smell that remained after performing autopsies and dissections. Because of this, he assumed that the smell indicated that cadaveric particles were still present. Because chlorinated lime solution removed this smell, Semmelweis believed that washing hands in this solution might also remove the cadaveric particles. As it turned out, this was a good choice (see odd fact).

The working hypothesis that now shaped Semmelweis's investigation can be stated:

*HYPOTHESIS: 'IF cadaveric particles acquired by doctors and medical students from dissections of bodies of people who died from puerperal fever are the cause of puerperal fever, THEN their removal from the hands of doctors and medical students by washing in chlorinated lime solution on leaving the morgue will reduce the death rate of new mothers from puerperal fever who are later attended by those doctors and medical students.'*

This hypothesis tentatively identifies a relationship between handwashing in chlorinated lime solution by doctors and medical students and the death rate from puerperal fever of new mothers who are later examined by these personnel.

The next task for Semmelweis was to test a prediction from this hypothesis. A testable prediction from this hypothesis is: *Handwashing in chlorinated lime solution by doctors and medical students on leaving the morgue will reduce the death rates from puerperal fever of new mothers later examined by those doctors and students.*

The outcome of his test would provide the evidence either to support the hypothesis or to reject it.

### Testing the hypothesis by experimentation

Semmelweis tested his hypothesis by imposing a handwashing regime. He directed doctors and medical students to wash their hands in a solution of chlorinated lime after performing dissections or autopsies. Semmelweis predicted that this handwashing procedure would (1) remove cadaveric particles from their hands, (2) minimise the transfer of cadaveric particles from the hands of these doctors and medical students to the genital tracts of new mothers in the maternity clinic, and (3) reduce the death rates of new mothers from puerperal fever who were treated by the doctors and medical students.

Semmelweis carefully monitored the death rates of mothers from puerperal fever over the months following the introduction of this handwashing procedure and compared these rates against those in previous years.

Tests of a hypothesis often involve setting up a **control group** of subjects and an **experimental group**, also known as a test group. The control group receives no treatment, and the experimental group receives the treatment under investigation. This was not done by Semmelweis. Instead, he used the past death rates as the background values against which the new death rates could be compared to see if they increased, decreased or were unaltered. (See question 10 at the end of the chapter.)

## Analysing the results

Figure 17.7 shows the result obtained by Semmelweis. This figure shows the average death rates for puerperal fever in the hospital over two periods before the chlorinated lime handwashing procedure was introduced. This figure also shows the average death rates for two periods that followed.

Examine this figure and note that the death rates of mothers from puerperal fever decreased markedly after the handwashing in chlorinated lime solution was introduced. This evidence provided support for the working hypothesis of Semmelweis. If the death rates had remained unchanged after the new handwashing procedure was introduced, the working hypothesis of Semmelweis would have been disproven and rejected. Semmelweis would have needed to consider other possible causes of puerperal fever.

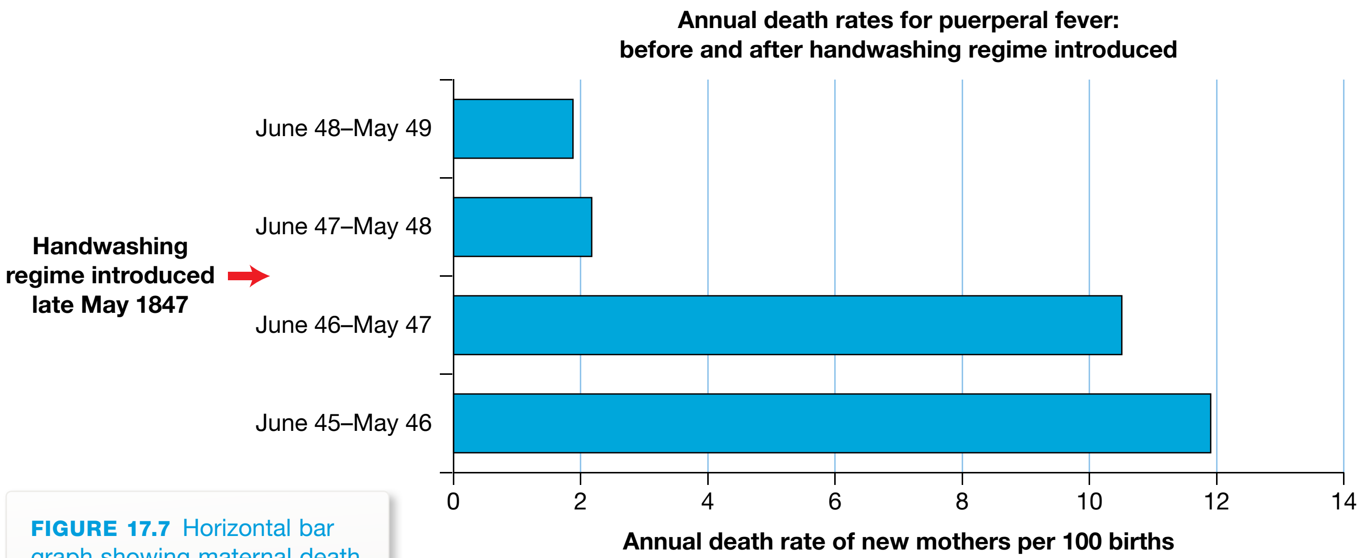



**FIGURE 17.7** Horizontal bar graph showing maternal death rates per 100 births from puerperal fever in the Vienna General Hospital for two 12-month periods immediately preceding the introduction of the chlorinated handwashing procedure, and two periods that followed. What pattern is obvious?

Semmelweis obtained additional supporting evidence for the involvement of some agent on the hands of medical students in the deaths of new mothers from puerperal fever. He was aware that large general hospitals had morgues where medical students were involved in cadaver dissection and that these medical students were also involved in the examination of new mothers. He was also aware that specialist maternity hospitals existed where medical students were involved in the examination of new mothers but not with dissecting cadavers.

In this case, the hypothesis formulated by Semmelweis was: '*IF cadaveric particles, acquired by doctors and medical students during cadaver dissections, are the cause of puerperal fever, THEN death rates of new mothers from puerperal fever are expected to be significantly higher in general hospitals where medical students engage in cadaver dissections than in specialist maternity hospitals where this type of anatomical teaching does not occur.*'

The prediction that Semmelweis tested was that the death rates from puerperal fever in new mothers in large general hospitals would be greater than those in smaller maternity hospitals where students were not involved in autopsies or dissections. To test this prediction, Semmelweis compared death rates from puerperal fever in the Vienna General Hospital, where cadaver dissection by students occurred from 1823 onwards, with death rates in the Dublin Maternity Hospital, where teaching of this type did not occur.

Figure 17.8 shows this comparison. Check out the graph and note that the annual death rate in the period in both hospitals up to 1823 was similar. Over the period from 1813 to 1822, the average Dublin death rate was 1.2 per cent, while that in Vienna was 1.6 per cent. Was this expected? From 1823, however, after medical students in Vienna became involved in cadaver dissections, the death rate from puerperal fever increased markedly in the Vienna General Hospital. In contrast, the death rate continued at the same low levels in Dublin Hospital, as predicted by Semmelweis. This provided support for his hypothesis.

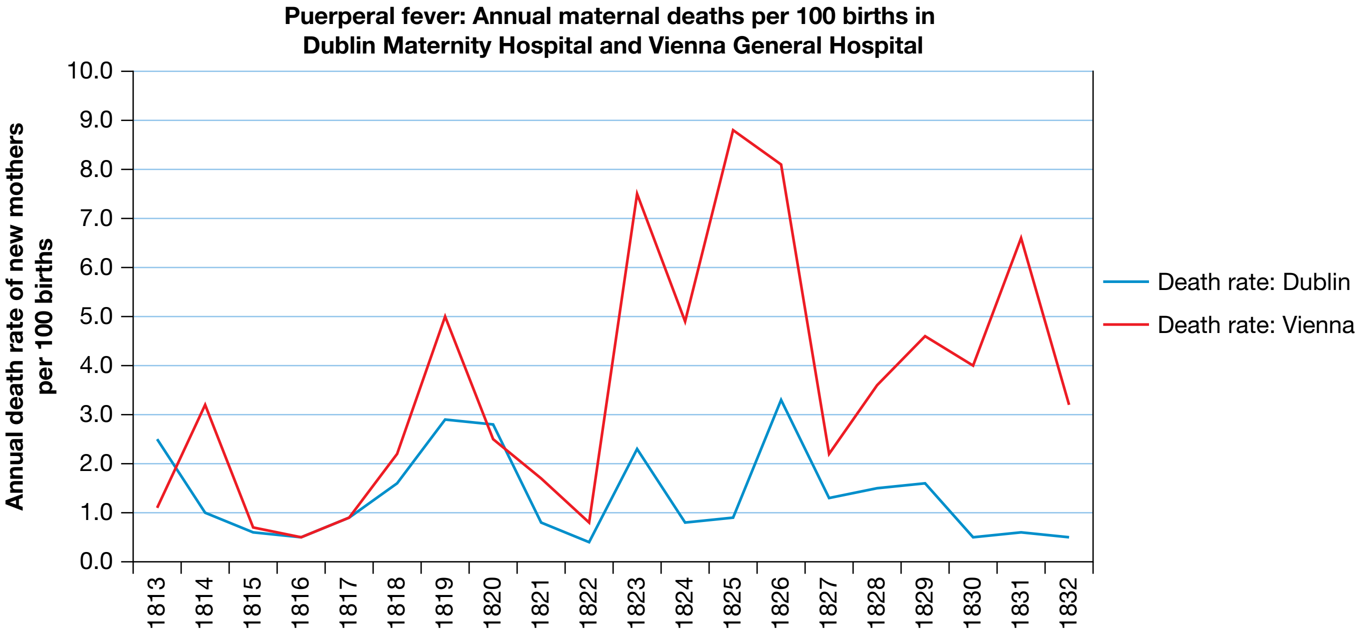



**FIGURE 17.8** Line graph showing the annual death rates for puerperal fever in the period from 1813 to 1832 in the maternity sections of two hospitals: Vienna General Hospital (red curve) and Dublin Hospital (blue curve). Note that in 1823, the subject *Pathological Anatomy* was introduced for medical students in Vienna, but teaching of this type was not carried out in Dublin. Prior to 1823, how do the death rates in the two hospitals compare? What general pattern is apparent after 1823?

## Communicating results

Having obtained support for his working hypothesis, the final step for Semmelweis was to communicate the findings of his investigation. He wrote to directors of maternity units in major hospitals outlining his results. Later, Semmelweis published a detailed account of his scientific investigations into the cause of puerperal fever in a book published in 1861.

Many of the professional colleagues of Semmelweis took offence at the suggestion that they were spreading puerperal fever and they rejected his findings. He was dismissed from his post at the Vienna General Hospital in March 1849. Semmelweis was unable to provide a scientific basis for his cadaveric particles and it was only after Pasteur and Koch developed the germ theory that the true nature of these 'particles' was identified as pathogenic bacteria, most commonly *Streptococcus pyogenes*. It was only then that Semmelweis was recognised as a pioneer in the development of antiseptic procedures in surgical and clinical settings and that he was 'ahead of his time'.

The hypotheses tested by Semmelweis were supported both by experimental and observational evidence. However, as noted above, his hypothesis that cadaveric particles caused puerperal fever was not correct. This is a reminder that **a scientific hypothesis can be disproven, but it can never be proven true; it can only be supported.**

The scientific investigation performed by Semmelweis followed a series of steps typically seen in rigorous scientific research and these are discussed in the following section.

**KEY IDEAS**

- Progress in science occurs when existing theories in science are modified or even rejected and replaced.
- Scientific investigations are multistep processes.
- Experiments are carried out only after observations relevant to the research question are made, questions asked about these observations and hypotheses formulated.
- Hypothesis testing is a process whereby experimental results provide evidence that either supports or disproves a hypothesis.

**QUICK CHECK**

1 Put the following actions of Semmelweis in the correct order over time:
  a performing an experiment
  b formulating a testable hypothesis
  c making observations
  d communicating results
  e asking questions
  f analysing results
2 True or false:
  a The hypothesis of Semmelweis was proved true by his experimental findings.
  b Puerperal fever is caused by cadaveric particles.
  c Semmelweis identified puerperal fever as a bacterial infection.
  d Advances in science may involve the rejection of an accepted theory.

# Scientific investigation

Figure 17.9 summarises a scientific approach to the investigation of a research question. The aim of the investigation may be to solve a problem, extend knowledge, answer questions or develop explanations. An example of this approach may be seen in the steps taken by Semmelweis when he set out to investigate the horrifying death rates of new mothers from puerperal fever in a maternity clinic in his hospital with the aim of trying to reduce it.

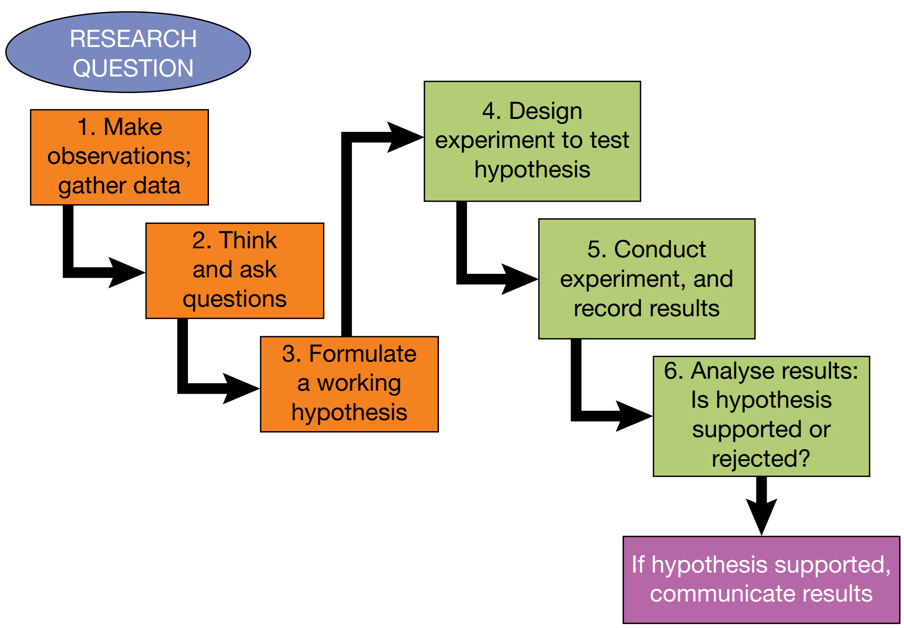



**FIGURE 17.9** Diagram outlining typical steps in a scientific approach to a research question. Note that an investigation does not start with an experiment. It starts by gaining background information about the research question by making observations and gathering data.

### Choosing a research question

As part of your studies, you are required to design an investigation or adapt an existing investigation that relates to (1) cellular processes, such as exploring transport across the plasma membrane or enzyme inhibition or (2) biological change and continuity over time, such as modeling genetic drift in a population. You are to carry out your investigation either in the laboratory or in the field.

In the case of Semmelweis, his initial research question was simply: 'What is the cause of puerperal fever?' Your research question may also be short and begin with: What . . .? How . . .? Why . . .? Once you have decided on the research question, you will use a scientific approach, as shown in figure 17.9 above and as discussed in the following sections. However, as you proceed through the early steps of your scientific investigation, your research question will become extended and will finally develop into a working hypothesis. Testing this hypothesis will be the focus of your laboratory experiment or your field investigation.

## 1. Observations produce data

**Observation is a fundamental skill used by scientists to gain data** — in other words, observations are a source of raw data. Observations provide the starting point of an investigation but, on their own, they do not provide explanations or answers.

**Making observations**: Observations can be made directly through your senses, such as by using your eyes to record the green and yellow colours of the variegated leaves of a flowering plant, your hands to identify the rough texture of its stem, your olfactory sense to note the smell of its flowers, and your tongue to detect the sweetness of the nectar from a nectary at the base of a flower.

As well as the unaided senses, many scientific observations are made using instruments that extend the power of the unaided senses and gather data that may otherwise be undetectable. Examples include thermometers to measure temperature, rulers to record length, scales to measure mass, microscopes to visualise cells and cell organelles, DNA sequencers to identify the base sequences of bacterial plasmids, magnetic resonance imaging (MRI) to make medical diagnoses (see figure 17.10). Not all observations are made using instruments in a laboratory. Other tools used to gather data include digital audio or video recorders, and digital cameras, questionnaires, surveys and interviews. Can you think of other tools for making observations and gathering data?

At the start of your investigation, you will make observations for the purpose of gathering data that will provide background information relevant to your research question. You may, for example, gather some data through a web search. You may also gather other background information from books, journal articles, and popular science magazines. However, later in your investigation, your observations and data gathering will focus on recording the results of the experiments that you perform to test your research hypothesis.

**FIGURE 17.10** Radiologists viewing an MRI result. Note the patient in the MRI scanner in the background. A typical MRI is qualitative in nature, and a radiologist assesses the image by eye, noting normal and abnormal areas. However, using specialised techniques, it is possible to quantify changes in tissues by measuring signal intensity.

In this chapter, 'data', when used as a mass noun, is singular, while 'data' used as a count noun is plural.

## Sources of data

Data is a set of facts that are collected, observed or generated. Data that you gather may come from primary sources or from secondary sources. Typically, data that you collect is raw data that must later be analysed and interpreted to produce useful information.

**Primary sources** of data provide direct or firsthand evidence about some phenomenon, as, for example, a research investigation, an event, an object or community attitudes. Sources of primary data include experimental results reported in scholarly journals by the researchers involved, video recordings of events, eyewitness accounts of events, historical documents, personal diaries, legal documents, completed questionnaires and surveys, and interview transcripts. Your completed logbook will be a primary source of data about pursuing an investigation into your research question. Your poster will be a primary source of data about the details of the experiment you designed to test your hypothesis.

**Secondary sources** of data are comments on or summaries and interpretations of primary data. Sources of secondary data include review articles in newspapers and popular science magazines written by one person who summarises and comments on the research of others. A book about Charles Darwin is a secondary source, but a book containing transcripts of Darwin's letters is a primary source. (Why?)

## Two kinds of data

Some data that you gather will be recorded in numbers, as, for example, an enzyme-denaturing temperature of 65 °C, a wavelength of 760 nm, a clear area of 13 mm diameter around an antibiotic disc on a bacterial culture plate, a C-G ratio of 40 per cent in a bacterial plasmid, or an allele frequency of $p = 0.64$. This type of data is termed quantitative. **Quantitative data can be both observed and measured and expressed numerically.** Can you think of other examples of quantitative data? Quantitative data can also be manipulated mathematically to produce useful statistics, such as means (averages).

Other data that you collect will be expressed in word descriptions. For example, in an experiment a student observed that treating the surface of a cut potato with iodine solution resulted in the appearance of a purple colour. 'Purple colour' is an example of qualitative data. Qualitative data is descriptive, not numerical, and can be easily observed, described but not measured. Other examples include descriptions of textures (rough, smooth, bristly, coarse, clammy, hairy, and so on) or smells (acrid, pungent, aromatic, sweet, musty, smoky, fragrant, and so on). Table 17.2 shows examples of how some attributes can be expressed both qualitatively and quantitatively.

**TABLE 17.2** Some attributes can be expressed either as qualitative or quantitative data. Which form provides more detailed information?

| Attribute | Qualitative | Quantitative |
|---|---|---|
| height (person) | tall | 183 cm |
| height (mountain) | very tall | 8848 metre |
| colour | green | 520 nm |
| sound | loud | 85 decibel |
| area (postage stamp) | small | 8.75 sq cm |
| speed | fast | 120 kph |
| temperature | hot | 100 °C |
| age (of rock) | ancient | 440 million years |

The results of laboratory experiments are most commonly quantitative data. In contrast, data gathered through the use of surveys, questionnaires and interviews are typically qualitative. What kind of data would you expect to gather if you used a survey-based investigation into community attitudes towards an issue, such as genetically modified organisms?

### Recording data

Scientists gather raw data or plain facts from their observations. Data must be recorded at the time in a suitable form, as, for example, text entries, sketches, tables and diagrams in logbooks (see figure 17.11), or in field notebooks. These may be supplemented by audio and video recordings.

**FIGURE 17.11** Record keeping in a logbook is an important part of the scientific method.

### KEEPING A LOGBOOK

A critical task for investigators is to keep a logbook. Your logbook may be a bound exercise book (*not* loose-leaf) or a digital document on your computer or iPad. It will contain all the details of your progress through the steps of your scientific investigation, starting from the process of deciding on your research question, and through each step of your investigation, which will follow a scientific approach. Your logbook is an important and valuable document. Note that even famous biologists kept notebooks; for example, Charles Darwin kept copious records of his observations in notebooks. During his voyage on the *Beagle* in 1831–36 (refer back to chapter 10, in the box titled The travels of Darwin and Wallace), Darwin filled 19 field notebooks with observations and thoughts that he made during shore excursions.

Make sure your name is prominently displayed on the front page of your logbook along with the title of your research topic. Your research question will be a topic of interest to you related to an aspect of cellular processes or to an aspect of biological change and continuity over time.

Keep a record of every session of your investigation. Your record might include (1) the date of the session, and (2) a heading that identifies the step in your investigation.

Possible headings you might use could include those shown below in italics:

- *Deciding on my Research Topic:* The starting point for your investigation will be to decide on your research topic. Briefly outline why you decided on a particular research question.
- *Background Data:* Through observations, you will gather background data (facts) on your research topic that can be recorded in your logbook as written notes, sketches, diagrams and tables; you might supplement these records with other media, such as photos, video recordings or audio recordings.

Be sure to record not only the data items themselves, but also their sources, so that, if necessary, you can easily locate and revisit them. Refer to the box titled Referencing data later in this chapter for further detail on data sources.

- *Questions arising*: In your logbook, record questions that arise and any thoughts that you have about your investigation. For each question, write your initial thoughts about possible answers. It does not matter if you later find that an answer is not correct. Simply record your amended answers — that's all part of a scientific approach. Identify any problems that arise during your investigation.
- *Developing my hypothesis*: Remember that a hypothesis is a tentative explanation and not simply a prediction. Refer to the section in this chapter titled Writing a hypothesis statement for more detail.
- *Planning my experiment:* Outline your plan, under the headings: Aim, Materials, Methods, Results, Analysis of Results, Conclusions.
- *Carrying out my experiment*: Observe and record your experimental results in an appropriate form, such as tables of values in the case of quantitative data, and word descriptions in the case of qualitative data. You can later convert tabular data into other presentations, such as bar graphs and line graphs that you will incorporate into a poster presentation.
- *Analysis and evaluation of my results*: Record your interpretation of the experimental results and decide whether they support or disprove your hypothesis. Record your conclusions.

## 2. Asking questions

The analysis and interpretation of raw data produces useful information. For Semmelweis, the data from his observations of death rates raised many questions. These included: Why did the death rates of new mothers suddenly increase, starting in 1823? Why did the death rates in Clinic 1 and Clinic 2 differ after 1841?

From a review of this information, questions will arise that relate to your research question.

Questions will arise in your mind from the observations you make and the data you gather from various sources about your research question. Record them in your logbook. Think about these questions and discuss them with your teacher and your classmates (see figure 17.12). This process will assist you to develop tentative conclusions and to then formulate a hypothesis that you will test by experimentation.

**FIGURE 17.12** Students discussing some of the data forming part of the background information to their research question.

## 3. Formulating a hypothesis

Formulating a hypothesis is an important step in the scientific process (see figure 17.13). What is a hypothesis?

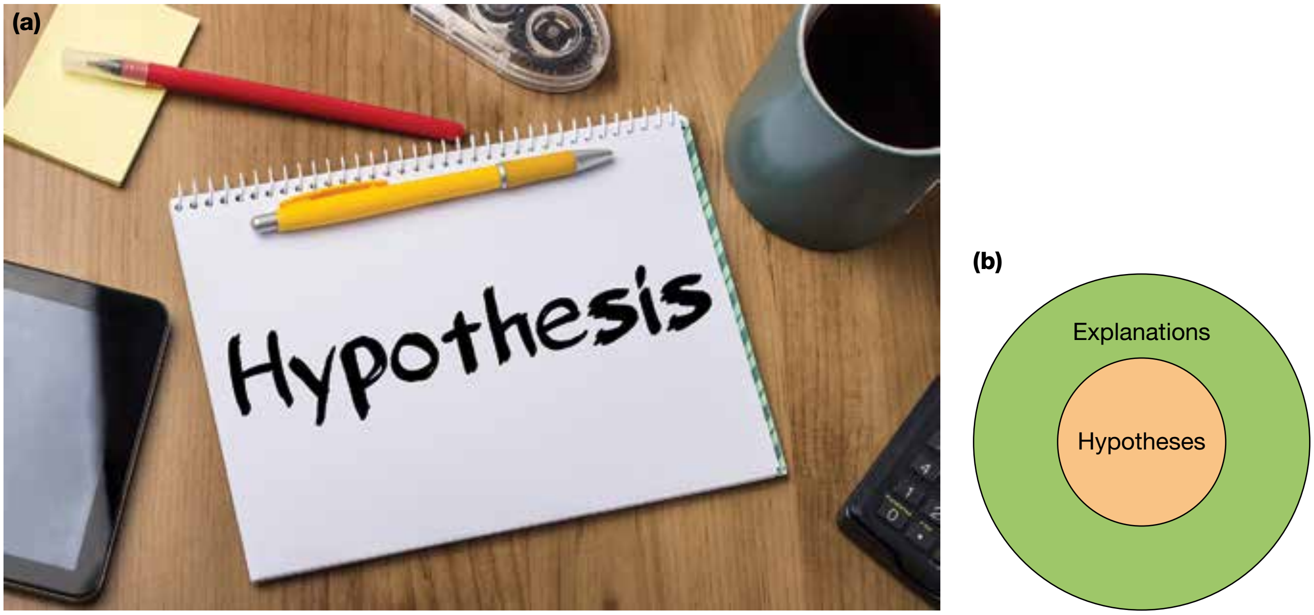



**FIGURE 17.13** **(a)** Formulating a hypothesis is an essential step in a research investigation carried out using a scientific approach. **(b)** Venn diagram illustrating that all hypotheses are explanations of observed phenomena, but not all explanations are hypotheses.

Some facts:

- A hypothesis is a tentative or provisional explanation for an observation — that is, it is an attempt to explain an observed phenomenon.
- A hypothesis is a statement, not a question.
- A hypothesis must be testable using the results of experiments or observations.
- A hypothesis must be falsifiable by a simple test or observation.

We can combine all these features into a single definition as follows:

**A hypothesis is a tentative, testable and falsifiable explanation for an observed phenomenon.**

If it is not an explanation, it is not a hypothesis, but . . . any explanation (correct or not) can be a hypothesis if it can be tested by observation or experimentation.

'**Testable**' means that a hypothesis can be easily tested by observations and/or experiments. Testing provides the evidence that allows a conclusion to be drawn about the hypothesis, namely that the hypothesis is either 'supported' or 'disproved (refuted).' **Scientific hypotheses are never stated as being 'proven correct'**.

A hypothesis can be any explanation for an observation on the condition that it is testable. Note that not every suggested explanation can be tested. One example of an untestable explanation comes from Carl Braun (1822–1891), who replaced Semmelweis after the latter was dismissed from the Vienna General Hospital. Braun's view was that puerperal fever was caused by 'unfavourable atmospheric-cosmic-terrestrial influences.' A testable hypothesis ***cannot*** be formulated from this. Try to complete the following statement: *'If unfavourable atmospheric-cosmic-terrestrial influences are the cause of puerperal fever, then . . .'*

'**Falsifiable**' means there must be a simple test, such as an observation, that could prove that a hypothesis is false. An example of a falsifiable statement is:

'*All cats are black*'. The testable prediction from this statement is '*If I see a cat, then it will be black*'. This hypothesis can be readily falsified by a simple test — it simply requires seeing just one cat that is not black. However, in reality, if a hypothesis is not testable, it is also not falsifiable.

### Writing a hypothesis statement

A hypothesis is a tentative, but testable, explanation for an observed phenomenon. In everyday life, you are often formulating simple hypotheses. You do this when you identify explanations for observations you make (see table 17.3).

**TABLE 17.3** Examples of simple but valid everyday hypotheses and how they can be tested.

| Observation | Hypothesis | Test of Hypothesis | Result | Conclusion |
|---|---|---|---|---|
| You turn on a torch and it does not work | Batteries are flat | Replace batteries and switch the torch on | Torch does not work | Hypothesis disproved |
| You are at the bus stop at the scheduled time, but no bus arrives | The bus is late | Wait and see if the bus arrives shortly | Bus arrives a few minutes later | Hypothesis supported |
| You turn on a torch and it does not work | The globe has blown | Replace the globe and switch the torch on | Torch works | Hypothesis supported |
| You become aware of the smell of burning | The toast is burning | Check toaster to see if the toast is burnt | Toast is black and smoking | Hypothesis supported |
| You hear the approach of blaring sirens | An ambulance is approaching | Look to identify the vehicle concerned | Police car with siren passes | Hypothesis disproved |

Typically, a scientific hypothesis is formulated as an 'IF . . . THEN . . . ' statement that starts with the tentative explanation and includes a prediction by which the hypothesis may be tested. In general, a good working hypothesis (regardless of whether or not it is later supported or is disproved) may have the form: 'IF hypothesis (explanation), THEN predicted outcome, as, for example:

**TABLE 17.4** Example of a good working hypothesis

| IF Hypothesis (explanation) | THEN Prediction |
|---|---|
| IF the rate of diffusion of molecules across a plasma membrane is directly related to their molecular size | THEN smaller molecules will cross the plasma membrane faster than larger molecules |
| IF skin cancers are caused by exposure to ultraviolet light | THEN people receiving higher UV exposures will develop more skin cancers |

Hypotheses may also include an action and these have the general form:

IF hypothesis (explanation) + action to be taken, THEN **predicted outcome**.

For example:

IF the torch batteries are flat and I replace them, THEN the torch will work again.

IF the rate of photosynthesis is affected by the wavelength of incident light and plants are illuminated with light of different wavelengths, THEN the rate of fixation of radioactive carbon by these plants will differ.

Now consider the following statements and decide if each is an example of a working hypothesis:

*Statement 1: 'Seedlings treated with organic fertiliser grow faster than similar untreated seedlings.'*

No. This is simply a testable prediction. It does not include a tentative explanation.

*Statement 2: 'IF seedlings are treated with organic fertiliser, THEN they will grow faster than untreated seedlings.'*

No. This does not identify a tentative explanation. The statement is a method followed by a predicted outcome.

*Statement 3: 'IF organic fertiliser provides essential growth factors for plants, THEN seedlings treated with fertiliser will grow faster than untreated seedlings.'*

Yes. This identifies a tentative hypothesis (explanation) and a predicted outcome by which the hypothesis can be tested.

A hypothesis usually predicts the relationship between two kinds of variables: an independent and a dependent variable. Let's look at the kinds of variables in an experiment to test a hypothesis.

## Variables in an experiment

In an experiment, a variable is any factor that the researcher can control or change or measure. Three kinds of variables are commonly recognised (see figure 17.14).

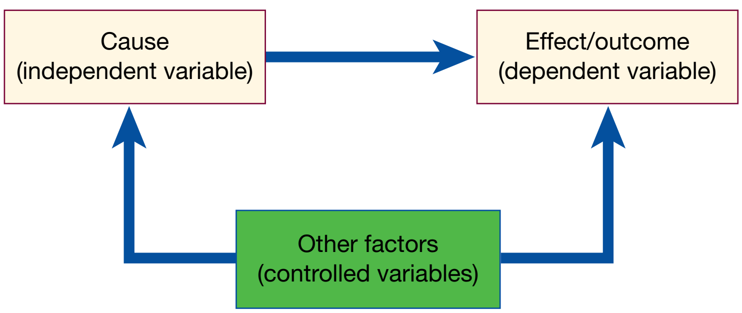



**FIGURE 17.14** Diagram showing the relationships between variables in an experiment. Which interactions should you aim to prevent?

**Variables: independent, dependent and controlled:**

- An **independent variable** is a factor that is under the control of and is deliberately manipulated by the investigator. The investigator then observes and measures the effect of the manipulation of the independent variable on a dependent variable. For example, in an experiment investigating the relationship between the wavelength of light and the growth rate of seedlings, the independent variable that the investigator changes is the wavelength of the light. The dependent variable that is observed and measured in response is the rate of photosynthesis.

  In some investigations, the investigator cannot manipulate an independent variable since it would be impractical to do so or it would be unethical to do so. For example, if a researcher wished to investigate the effect of an illegal drug on an aspect of human behaviour, it would clearly be unacceptable and unethical to ask one group of people to take such a drug. Research of this type might however be pursued through the use of a suitable questionnaire administered to both non-drug users and drug users.
- A **dependent variable** is also termed an outcome variable. This is the factor that the investigator observes to see how it changes in response to changes in the independent variable. The changes in the dependent variable are measured and recorded. So, in the above investigation, the dependent variable is the rate of the growth of the seedlings.
- **Controlled variables** are all the other factors that the investigator must maintain at constant values through the course of an experiment. In the above experiment, the plant species used, the age of the seedlings, the temperature, availability of water, and angle of incident light are all examples of controlled variables. The controlled variables are ones that the investigator keeps the same throughout the course of the experiment. If these factors are

not kept constant, they can confound the experimental results because they can cause changes in the dependent variable.

Typically, an experiment involves one independent variable that is systematically changed, and one dependent variable whose responses to these changes are observed and measured. By changing only one variable (the independent variable) at a time, an investigator can see and measure the effect on the dependent variable.

In summary:
- the independent variable is what an investigator changes
- the dependent variable is what an investigator observes and measures
- controlled variables are what the investigator keeps constant.

**Variables: categorical or numerical:**
The variables that are investigated in experimental research differ in their nature. They can be broadly divided into two classes — categorical variables and numerical variables. (These classes can be further subdivided, but this is not covered here.)

**Categorical variables** are qualitative and are expressed in words. A categorical variable can have a value that is one of a limited number of non-overlapping discrete values, and these values are expressed in words. Examples of categorical variables include the months of the year (12 values: January to December), the days of the week (7 values: Sunday to Saturday), ownership (or not) of a smart phone (2 values: Yes and No). Bar graphs are often used to display the frequencies of categorical variables in a study population.

**Numerical variables** are quantitative. Numerical variables can be precisely measured and they have values that are expressed in numbers. Examples of numerical variables include temperatures, masses, distances, rates of oxygen production in photosynthesis, wavelengths of light, salinity of water.

How would you classify these variables: breeds of dog? air temperatures? hair colours? rates of substrate production in an enzyme-catalysed reaction?

## 4. Designing an experiment to test a hypothesis

Designing an experiment means constructing a detailed experimental plan to test a hypothesis in advance of doing the experiment.

A typical plan might include details under the following headings, which you would record in your logbook:

*Title of experiment:*

*Aim*: Purpose of the experiment, which should include the hypothesis that you are testing

*Materials*: This would include the following items that should be checked with your teacher:
- the laboratory equipment and the consumables needed for your experiment, any personal safety equipment required, and their availability checked
- any safety issues associated with the conduct of your experiment, such as
  - potentially hazardous chemicals
  - equipment that can pose a potential safety hazard, as, for example, electrophoretic equipment or centrifuges.

For each safety issue identified, list in your logbook the safety controls and precautions to be taken.
- any ethical issues associated with the conduct of your experiment, as, for example, the need to seek informed consent if your experiment will involve human subjects.

*Methods*: Identify your independent variable and how you will change it during the experiment. Identify the dependent variable expected to respond to these changes and identify how you will measure the changes. Identify all the controlled variables to be kept constant throughout the experiment.

### Control and experimental groups

The design of many experiments includes a **control group** as well as one or more **experimental groups**. The experimental groups are exposed to the changing conditions determined by the independent variable, but the control group is independent and not affected by those conditions.

For example, in an experiment concerned with the effect of salinity on plant growth, a control group of seedlings would be a tray of seedlings allowed to grow under conditions of zero salinity, while the experimental groups would be exposed to conditions of different salinities.

The control group serves several purposes:

– it shows that the experiment is working
– it provides a baseline result against which the results of the experimental group can be compared.

You need to decide on the number of experimental groups required and the size of each group.

### Sample size

The size of the control and experimental groups is an important factor in experimental design. The size of each group must be sufficiently large so that (1) replicate results can be obtained and a meaningful comparison made between the outcomes in the control and the experimental groups, and (2) the results from the experimental group can be seen as applicable to the larger population. The upper limit on sample size is determined by cost and space considerations and other practicalities.

Living organisms, whether they be plants, algae or animals, are an uncontrolled source of genetic variation that can interact with the variables under investigation. Sample sizes that are sufficiently large produce replicate results or observations that (3) can reduce the effects of this uncontrolled variation.

Complicated formulae exist for calculating minimum sample sizes. However, one simple rule-of-thumb is that ten observations are required for each experimental variable (see question 14 in the end of chapter questions).

## 5. Conduct experiment and record results

Your teacher will supervise your laboratory investigation and will ensure that health and safety requirements are met and that your experiment is carried out according to any ethical principles that may apply.

Record your progress through the experiment and the results of your experiment in your logbook. Typically, if the variables in your experiment are numerical, you will set up a table and record the values of the independent variable and those of the corresponding dependent variable. If appropriate, calculate any average values.

### Displaying experimental results

Once you have your results recorded in a table, you must display this data in an appropriate form. Presenting results as a graph makes it easier to see relationships between the variables involved because graphs highlight trends and patterns. Time sequence graphs are important as they can display key turning points in a dataset.

Results can be displayed in various ways depending on the nature of the variables and what you want to show.

The most common representations include:

- scatter plots or scatter diagrams
- bar graphs
- line graphs
- frequency diagrams.

**Scatter plots**: Scatter plots are also termed X–Y graphs. When you have *paired* numerical data, a scatter plot can be used to see if a relationship exists between the two variables, as, for example, the weights of a sample of adults plotted against their heights. Each dot represents one observation, with the position of the dot determined along the x-axis and y-axis by the value of each of the variables concerned, namely weight and height (see figure 17.15). The independent variable is plotted on the x-axis and the dependent variable on the y-axis. Note that each axis has a label and measurement units. A scatter plot can show how much one variable is affected by the other. If data points fall along a line or a curve, the variables are linked or correlated, and the closer the data points are to the line, the higher the correlation. In the figure below, a strong positive correlation is displayed. What shape would you expect in a scatter plot showing ages of a sample of adults and their reaction times?

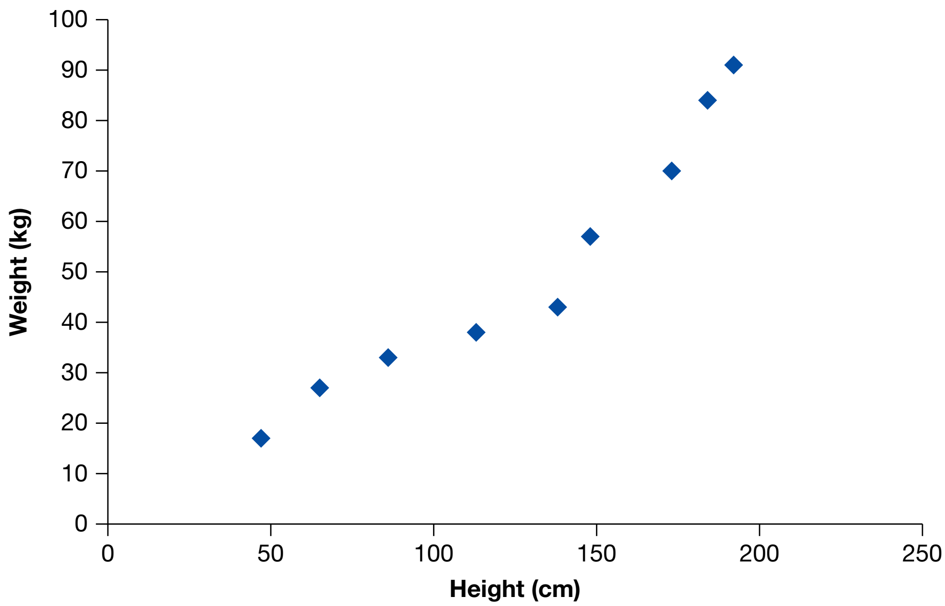


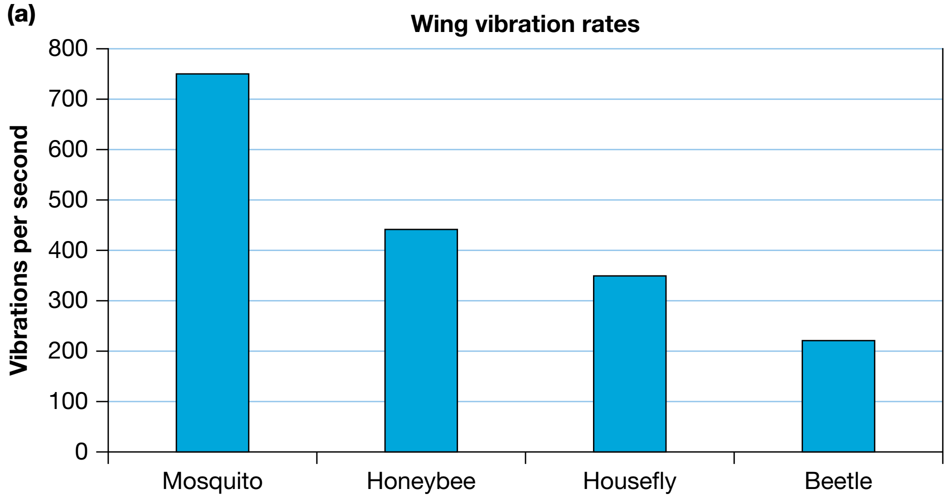

**FIGURE 17.15** Scatter plot showing the relationship between the height and weight of a sample of adults.

**Bar graphs**: Bar graphs can display numeric data, such as numbers or frequencies (percentages) or other measures of one or more discrete categories or groups. Figure 17.16a is a bar graph in which the vibration rate of the wings of various insects is represented by vertical bars. The horizontal or x-axis has no scale because it simply shows categories. The vertical or y-axis has a scale showing the units of measurements — in this case, vibrations per second. Note that the bars are separated by spaces.

Bar graphs can also be used to compare two sets of data by using side-by-side bars, as shown in figure 17.16b. This bar graph shows the percentage increase, relative to the control, in two algal species growing under conditions with added mineral nutrients. Which nutrient caused the greatest increase? Which species of alga (A or B) showed the greater increase?

Less commonly, bar graphs are drawn using horizontal bars. This orientation is used when the category label is long and complex and cannot be fitted along the x-axis.

**Histograms**: A histogram is a special kind of bar graph (see figure 17.17). In a histogram, the data represent continuous categories rather than discrete values. In contrast to the situation with a bar graph, a histogram has a scale on both the x-axis and the y-axis, and the bars are not separated by spaces.



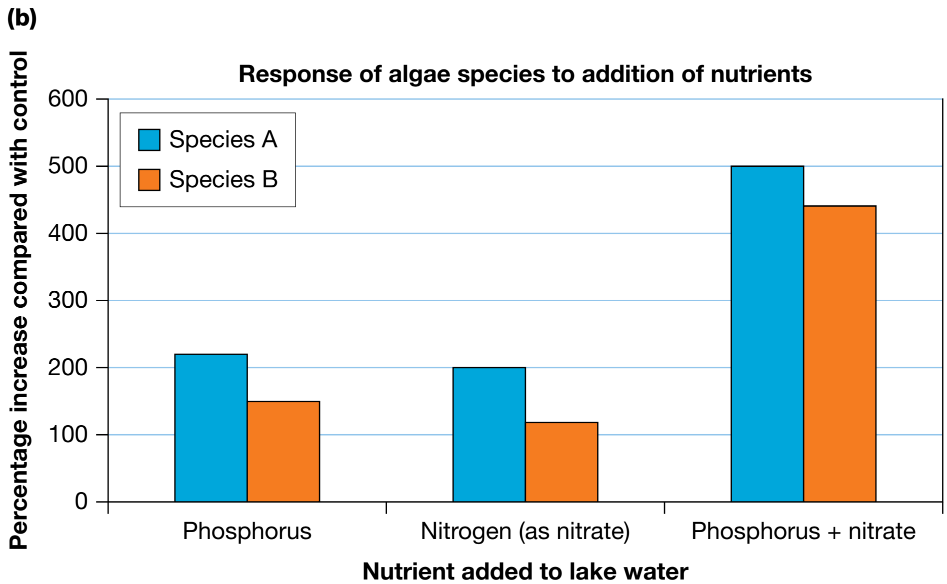




**FIGURE 17.16** **(a)** Bar graph showing wing vibration rates (vibrations per second) for four insects. **(b)** Percentage increases, compared to control, of two species of algae in response to the addition of nutrients to the water in which they were growing.

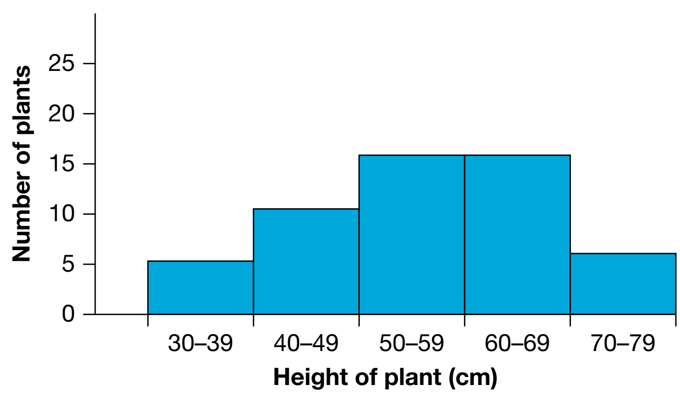



**FIGURE 17.17** A simple histogram showing the results of a survey of the heights of a particular plant species growing on a moorland.

**Line graphs**: In a line graph, a series of dots represents the values of a variable, and the dots are joined using a straight line. Note that the ends of a line graph do not have to meet the axes of the graph. Line graphs are often used to show changes over a continuous period of time, or over space. In particular, **line graphs can identify patterns, trends and turning points in a dataset**.

Data on two variables can be plotted on the same graphs to show the change in the two variables during the same time period. For example, refer back to figure 17.5, which shows the annual death rates in Clinic 1 and Clinic 2 over the period 1833 to 1846.

## 6. Analyse results and draw conclusions

Firstly, you should review your data and ask: Is the dataset complete? Do you need to collect more data?

The results of any experiment will produce data that you will need to analyse. Aspects of data that you may need to consider include the following: validity, precision, accuracy, reliability, minimisation of bias, limitations of data and sources of error.

### Analysing and evaluating results

1. **Validity of data**: The validity of data refers to the credibility of the research results from experiments or from observations. Validity refers to both the research findings (so-called internal validity) and to the applicability of the results to other situations (so-called external validity).
   - **Internal validity refers to the degree to which the experimental procedures measure what they are supposed to measure.** Testing internal validity asks questions such as: Can the results be believed? Could another unknown variable have influenced the results? Could any cause and effect relationships identified be explained by other factors? Internal validity is favoured when an experiment is carefully designed and a scientific approach is used.
   - **External validity refers to the extent to which research findings can be generalised to the greater population**. Testing external validity asks the questions: Is the sample of the population that was used in the research study representative of the greater population? Can we be reasonably sure that the results of the research are applicable to the greater population? External validity is favoured by an experimental design that includes the use of a control group, having control and experimental groups of sufficient sizes, and randomly assigning subjects to the control and test groups.
2. **Precision**: How close to one another are repeated measures of the same variable under the same conditions? If the measurements are close, as, for example, 19.5, 19.4, 19.6 and 19.5, they are precise. **Precision refers to the closeness of repeated measures or observations even if they are not accurate.** Precise values will differ from one another because of random observational error.
3. **Accuracy**: Accuracy is different from precision. **Accuracy refers to the closeness of an experimental measurement to an accepted or known value.** Accuracy in a research investigation relies on selecting the most appropriate instruments for any measuring tasks and using them correctly.

   Accuracy is an issue in a confirmatory experiment where the aim is to confirm a known value, as, for example, identifying the diploid number of the yellow fever mosquito, *Aedes aegypti*. If you concluded from your microscopic observations that the diploid number was six, then your result would be inaccurate. (The known number is $2n = 8$.) However, when an experiment is an exploratory investigation rather than a confirmatory exercise, matching an accepted value of a variable is not an issue.

4. **Reliability**: If your experiment were to be repeated a second and third time, would the results be the same? **If the same results are obtained in repeats of an experiment, the data can be said to be reliable.** Replication reduces variability in experimental results, increasing their significance and the confidence level with which a researcher can draw conclusions. Ideally, replications of experiments are carried out not only by the original investigators but, importantly, by third parties who use the same materials and the same methods as those of the original investigators.

**FIGURE 17.18** Repeating a biology experiment is a means of assessing the reliability of experimental findings.

Replication is an important means of verifying experimental results and showing that they are reliable. In January 2014, a young Japanese researcher, Haruko Obokata, published in the leading science journal *Nature* what was hailed as a major breakthrough in stem cell research. Obokata had apparently developed a very simple method of producing pluripotent stem cells from adult somatic cells — just soak them for 30 minutes in a weak solution of citric acid. However, in spite of the simplicity of her method, no one else was able to replicate her results. In reality, Obokata had falsified and fabricated her experimental data. The fact that her results were not reproducible was critical and resulted in her retracting her research findings.

5. **Minimisation of bias**: Bias is an intentional or unintentional influence on a research investigation as a result of systematic errors introduced by the researcher into the sampling or the testing procedures of an experiment. These biases will prejudice the research findings and raise questions about their validity and reliability.

   Consider a clinical study into the effects of a new drug that involves a control group of individuals who will receive a placebo and an experimental group of individuals who will receive the drug under test. Such a study could be subject to bias of various kinds, such as the following:

   - **Selection bias**: This type of bias can arise unless participants are equally and randomly assigned to the experimental and the control groups. All participants should have an equal chance of being placed in either group.

Selection bias can be minimised by randomly allocating participants to each group.

- **Sampling bias**: This type of bias can arise if the participants chosen for the study are not representative of the target population. If this occurs, the research results cannot be generalised to that population. Sampling bias can be minimised by ensuring that the participants in the study are a reasonable representation of the target population.
- **Response bias**: This type of bias arises when only certain members of the target population respond to an invitation to participate in the clinical trial, resulting in an unrepresentative sample of the larger population. Response bias can be minimised by ensuring that the participants in the study are a reasonable representation of the target population.
- **Measurement bias**: This bias can arise, even unintentionally, if the people measuring the test results are aware of whether they are measuring the results of individuals in the control group or in the experimental group. There is a possibility that the people measuring the results will, even unintentionally, focus on data that supports the drug treatment. Measurement bias can be minimised ideally by designing the clinical trials as a double-blind study in which neither the participants nor the researchers know which people are in which group.

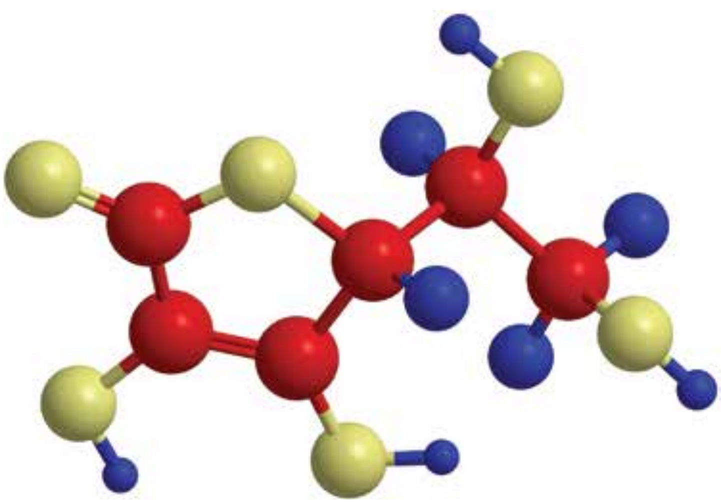
6. **Limitations of data**: The data that is gathered from the experimental results will have limitations. Limitations arise from several sources that can affect the quality of the data:
   - Experiments create artificial situations that do not necessarily represent real-life situations.
   - While every effort may be made to identify controlled variables and keep them constant throughout the course of an experiment, it is not always possible to identify and control every one of this type of variable.
   - The degree on which results obtained in the laboratory can be generalised to other situations and applied in the real world is limited.

As well as data limitations, **limitations exist in the models that we use in biology**. Models can be physical models, or mathematical models or conceptual models. Models can provide an explanatory framework that explains observed phenomena and helps organise myriad data into a related whole. Physical models can make an abstract concept tangible.

While they are very useful, models have their limitations:

- Because of the complexity of the processes that operate at every level of biological organisation, models cannot include all the details of the processes or the things that they represent.
- Models are necessarily approximations of the real world; for example, the Hardy–Weinberg equation is a mathematical model that allows us to predict changes in allele frequencies in the gene pool of a population. However, this model is constrained by certain conditions, and its predictions would not always match what happens in real-world populations.
- In making a concept or object concrete, a model must compromise accuracy; for example, a ball and stick model of vitamin C (see figure 17.19) is useful, but is a very highly simplified and stylised representation that reduces covalent bonds to sticks and atoms to solid balls. (Still, we wouldn't want the sub-atomic detail.)

**FIGURE 17.19** Ball and stick model of vitamin C (ascorbic acid). How do you think this differs from a 'real' vitamin C molecule?



7. **Sources of error**: Several sources of error can be identified in an investigation:
   - **Systematic errors** produce measurements that are consistently too high or too low, as, for example:
     - instrumental errors that arise because an instrument, such as a weighing balance for which the zero has not been reset before use
     - environmental errors that arise because of malfunctions, such as a power outage, that affect the conditions under which an experiment is being conducted.
   - **Random errors**: Random errors are chance variations in measurements that are equally likely to be high as low. An example of a random error is an error of judgement when reading the smallest division on the scale of a measuring instrument, such as a micrometer, a burette or a measuring balance. However, if multiple observations are made and an average calculated, these random errors are minimised.

It should be noted that so-called 'human errors' are not errors but are mistakes that result from carelessness. Examples include gross misreading of an instrument, and writing the wrong result in your logbook, such as 40 instead of 4.0.

### Hypothesis evaluation and other conclusions

Evaluating a hypothesis uses the evidence from the results of the experiment that was designed to test the prediction from that hypothesis. In Semmelweis's research, his prediction was that handwashing in chlorinated lime would reduce the incidence of deaths from puerperal fever.

If the prediction from your hypothesis was validated by your experimental results, you should evaluate your hypothesis as 'supported' and your tentative explanation of the scientific observation now stands as a reasonable explanation — for the time being at least. For Semmelweis, the observed decrease in deaths of new mothers from puerperal fever supported his hypothesis that this disease was caused by 'cadaveric particles'. This explanation remained accepted until it was later disproven by the germ theory of Pasteur and Koch.

If your predicted outcome was not supported by your experimental results, you should evaluate your hypothesis as 'rejected' or 'disproven'. In this case, you might modify your hypothesis.

Other conclusions might relate to the following questions:

- Were there any outliers in your dataset? An **outlier** apparently doesn't fit — it is a data point that is distant from all the other data points.

  If you had an outlier, what did you do about it? Rather than ignoring it, you should try to account for it. Most commonly, it will either be a systematic error or an outright mistake in measurement or in calculation, and so can be dropped. Occasionally, an outlier can be a legitimate observation that warrants further investigation:
- What future improvements can you suggest regarding your experimental design?
- In light of your experimental results, what follow-up experiments might you suggest?
- How do your research findings link to the key biological concepts relevant to your research question?

## 7. Communicating results

The results of a scientific investigation are typically communicated to others. This may be done through an oral presentation, a journal article, or a poster. In your case, the requirement is that you produce a poster, with a word count not exceeding 1000, either electronically or in hardcopy format.

**eBook*plus***

**Weblink**
Poster template: Biology Study Design (VCAA)

The template to be used for your poster is identified on page 12 of the Biology Study Design (see the website for the Victorian Curriculum and Assessment Authority (VCAA)).

## Producing a poster

Work through each of the headings in the format shown in the VCAA Biology Study Design and add the content specified. Your logbook will be an important resource that will assist you in filling out detail under each heading. Effective communication makes use of figures, such as photographs, drawings, or graphs, in addition to words and equations.

Experimental results are commonly presented as graphs. Why? Refer to the graphs shown earlier in this chapter to check on the various types and their formats.

Remember: If you use any material that is the work of another person, you *must* acknowledge its source. Do not claim it as your own work. Acknowledgments comes in two formats: a short version when it occurs in the body of your poster, and a longer version when it occurs in the *Reference and Acknowledgments* section at the end of your poster. These two formats are described in following box.

### REFERENCING DATA SOURCES

The examples below are based on the *Harvard Style Guide* published by Monash University in 2012; available at the library website of Monash University.

For each type of publication, the first style (A) gives full details and is used in reference lists at the end of manuscripts or posters. The second style (B) is short and is used in the body of the text of a manuscript or a poster.

- **Online authored articles**:
  **Style A**: Author name and initial(s). (year of publication). Title of document. Retrieved from http:// URL
  Example: **Geggel L (2016) Snakes used to have legs and arms . . . until these mutations happened. Retrieved on 4 Nov 2016 from www.livescience.com/56573-mutation-caused-snakes-to-lose-legs.html**
  If the online article is undated, put (n.d.) in place of the date
  If there is no author, use the first line of the text of the article in place of the author's name
  If there are up to three authors, list them all.
  If there are more than three authors, only use the name of the first author and follow it by the phrase 'et al' (meaning and others).
  **Style B**: **(Geggle 2016)**
- **Books**
  **Style A**: Author name and initials (year of publication). Book title. (Edition if applicable) Place of publication. Publisher. Page number(s) you accessed.
- Example: **Langley P & Jones M (2013) *The search for Richard III: The king's grave.* London, John Murray.**
  If there are more than three authors, only use the name of the first author and follow it by the phrase 'et al.'
  **Style B**: **(Langley & Jones 2013)**
- **Journal articles**:
  **Style A:** Author name (year of publication) *Title of article*. Title of journal. Volume number. Page numbers.
  Example: **Smith SL (2007) Stature estimation of 3–10 year children from long bone lengths. J Forensic Sci vol. 52 pp. 538–546.**
  If there are more than three authors, only use the name of the first author and follow it by the phrase 'et al.'
  **Style B:** **(Smith 2007)**
- **TV programs:**
  **Style A:** Title of program (date) (TV program) Channel identification.
  Example: **Gene editing made simple 2016, television program, ABC Sydney 30 August.**
  **Style B:** **(Gene editing 2016)**

# Chapter review

studyon

Sit exam

## Key words

accuracy
bar graphs
categorical variables
control group
control group
controlled variables
dependent variable
experimental group
falsifiable
histograms
independent variable
limitations of data
line graphs
making observations
measurement bias
numerical variables
outlier
precision
predicted outcome
primary sources
puerperal fever
random errors
reliability
response bias
sampling bias
scatter plots
scientific method
scientific research methodology
secondary sources
selection bias
sources of error
systematic errors
testable
validity of data

## Questions

1 **Consider some 'explanations' of the cause of the plague, such as divine punishment for sin or the alignment of planets (refer back to chapter 6 of the text).** ➔
   a Do you think that these 'explanations' could be easily tested?
   b Explain your answer.

2 **Examine the graph on the subject of Zebra mussel invasion data found on the Cary Institute website (eBook*plus* weblink) and answer the following questions:** ➔

   **eBook*plus***
   **Weblink**
   Zebra mussel invasion data

   a What kind of graph is this?
   b What variable is represented by the blue data points?
   c What variable is represented by the red data points?
   d Rotifers are microscopic aquatic animals that form part of the zooplankton in bodies of freshwater. What was the peak concentration of rotifers in 1991 in the Hudson River?
   e What trend is apparent in the rotifer population over the period shown?
   f Zebra mussels (*Dreissena polymorpha*) are filter feeders, and this non-native species reached the Hudson River in 1991. What happened to the mussel population in the period from 1991 to 1993?
   g Suggest a possible research question that might arise from the data in this graph.

3 **Use the eBook*plus* weblink for chlorophyll concentration in the Hudson River to answer these questions.** ➔ The decrease in chlorophyll a after 1992 shows the decrease in the phytoplankton population in the river water.

   **eBook*plus***
   **Weblink**
   Line graph: Chlorophyll concentration in the Hudson River

   a Write a hypothesis that might explain this sudden and persistent decline in the phytoplankton population.
   b Suggest how this hypothesis might be tested.

4 **Consider the hypothesis: *IF the growth rate of plants depends on the duration of sunlight they receive, and seedlings are exposed to varying periods of light, THEN seedlings exposed to longer period of light will grow faster than those exposed to shorter periods.*** ➔
   a Identify the independent variable.
   b Identify the dependent variable.
   c State the prediction that would be used to test this hypothesis.
   d Suggest a possible means by which the growth rates of seedlings might be measured.
   e Identify three controlled variables that would need to be taken into account when designing an experiment to test this hypothesis.

5 **A common explanation during the medieval period was that the plague was cause by miasma (bad smells).** ➔
   a Is this a hypothesis that could that be tested experimentally?
   b Is it possible to write a valid hypothesis statement in the form of an 'If . . . then . . . 'statement'?

6 **Which, if any, of the following statements is an appropriately crafted scientific hypothesis. In each case, give a reason for your decision.** ➔
   i Removing cadaveric particles from the hands of doctors and medical students will result in a reduction in the death rate from puerperal fever of new mothers.
   ii Do cadaveric particles that gain entry to the genital tracts of new mothers cause puerperal fever?
   iii If a plant receives fertiliser, then it will grow to be bigger than a plant that doesn't receive fertiliser.

iv The hypothesis is that the presence of cattle would have an adverse impact on the vegetation in the national park.

v If organic fertiliser is a source of essential plant nutrients, then seedlings treated with fertiliser will grow faster than untreated seedlings.

**7 Consider a simple experiment to investigate the relationship between the wavelength of light and the growth rate in which trays of seedlings are grown under conditions of different wavelengths. ➜**

**a** Identify the independent variable in this experiment.

**b** Identify the dependent variable.

**c** Identify at least four controlled variables that should be identified by the investigator.

**d** Which kind of variable is:

i kept constant throughout the experiment

ii manipulated by the investigator during the experiment

iii observed during the experiment and changes recorded?

**e** Would a sample size of two seedlings, for each selected wavelength be a good design? Explain.

**8 Consider the following data: ➜**

i labelling of S, M, L, XL, XXL on items of clothing

ii labelling of 300 mL, 600 mL, 1 litre, 2 litre on drink cartons

iii values of 73 °C, 86 °C, 54 °C, and 92 °C identified as maximum daily temperatures

iv classification of cats as Birman, Siamese, Tonkinese, Manx, Russian Blue.

**a** Which, if any, are examples of a categorical variable?

**b** Which, if any, are examples of a numerical variable?

**9 Identify a key difference between the members of the following pairs: ➜**

**a** independent and dependent variables

**b** qualitative and quantitative data

**c** control group and experimental group

**d** primary and secondary sources of data

**e** categorical variable and numerical variable.

**10 Semmelweis did not have a control group in his handwashing experiment. ➜**

**a** If he had had a control group, what would it have been?

**b** Can you identify any ethical issues that would be raised today by having such a control group?

**11 Consider an apple. ➜**

**a** Identify pieces of three quantitative data that you could record and measure.

**b** What instruments, if any, would you need to make these observations?

**c** Identify three pieces of qualitative data that you could observe.

**d** What instruments, if any, would you need to make these observations?

**12 In his investigations, Semmelweis noted that expectant mothers in Clinic 1 gave birth lying on their backs, while in Clinic 2, they gave birth lying on their sides. Briefly outline an experiment that would test whether this procedural difference was the cause of the difference in puerperal fever death rates in mothers in the two clinics. ➜**

**13 Consider table 17.5, below, which shows the results of an experiment that looked at the effects of salinity on plant growth: ➜**

**a** How many replicates were used in the control group? In each experimental group?

**b** Why not use just 2 or 3?

**c** What is the independent variable in this experiment?

**d** What is the dependent variable?

**e** Identify three possible controlled variables.

**14 You wish to see if a new drug is an effective treatment for asthma. One hundred volunteers are randomly divided into two groups: (A) and (B). Members of group B are given the drug and members of group A are given a placebo in the form of a sugar pill. ➜**

**a** Which is your test group, and which is your control group?

**b** A student said: 'Wouldn't it be better to have a test group of 100 volunteers rather than split the

**TABLE 17.5** The effects of salinity on plant growth

| Concentration of salt water (mg/L) | Replicates: seedling root length (mm) | | | | | | | | | | Average length |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| 0 (control) | 20 | 16 | 16 | 9 | 14 | 13 | 13 | 10 | 12 | 10 | 13.3 |
| 0.75 | 14 | 22 | 13 | 14 | 15 | 16 | 11 | 16 | 17 | 10 | 14.8 |
| 1.5 | 15 | 19 | 12 | 18 | 10 | 7 | 0 | 9 | 4 | 13 | 11.8 |
| 3 | 9 | 7 | 6 | 6 | 16 | 11 | 10 | 17 | 11 | 6 | 9.9 |
| 6 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.2 |
| 12 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.1 |

volunteers into a test group of 50 and a control group of 50?' How would you respond to that student?

**15 Consider figure 17.20, below, which shows the results of one experiment in which the effect of incubation temperature on the activity of a particular enzyme was examined. ➜**

- **a** Which variable is plotted along the horizontal axis?
- **b** Which variable is plotted along the vertical axis?
- **c** Suggest at least two controlled variables that the investigator would have kept constant during this experiment.
- **d** Interpret the graphical data; in other words, what story does this graph tell?

**16 Identify the following statement as true or false: ➜**

- **a** Interpretation follows data collection.
- **b** The dependent variable in an experiment must be kept constant throughout a course of an experimental constant.
- **c** If your experimental result causes you to reject the hypothesis under test, this means that your experiment was wrong.

**17 You conduct an experiment and record the data. You notice that some of the data doesn't fit with what you expected or doesn't agree with the other data. What of the following actions should you take? ➜**

- **a** Toss it out as it looks wrong.
- **b** Figure out what you want the data to show and then decide if you will keep it.
- **c** Try to identify why this occurred, then decide if it should be kept or discarded, and discuss it in your result analysis.

**18 Which of the following is a requirement for a good hypothesis? ➜**

- **a** You need to make the hypothesis broad so that the data can support it or reject it.
- **b** You need to be able to design and conduct an experiment that can test the hypothesis.
- **c** You need to know in advance whether or not the hypothesis is true.

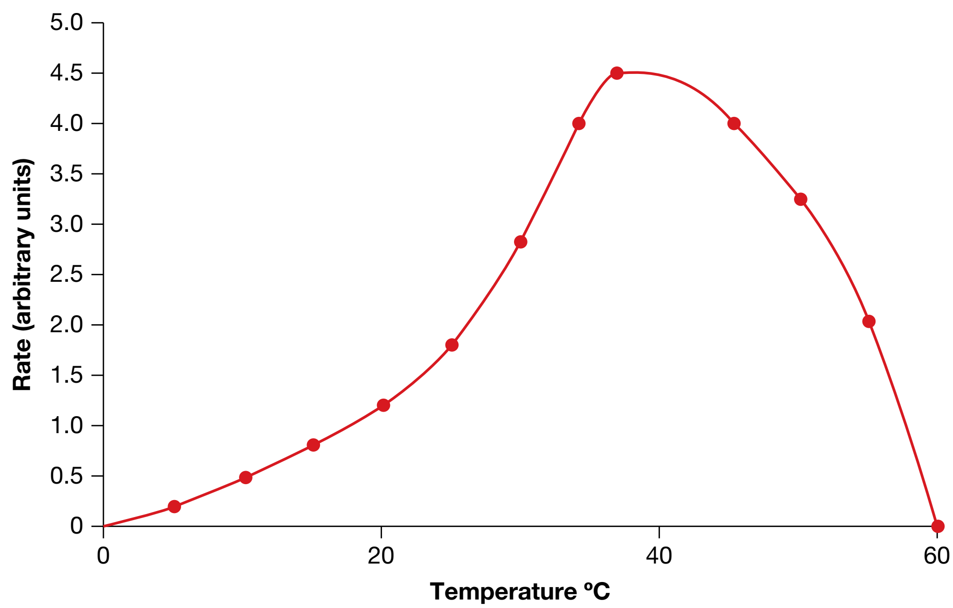



**FIGURE 17.20** The effect of incubation temperature on the activity of a particular enzyme

# Appendix

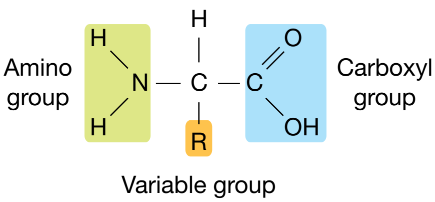



**FIGURE A1**

## AMINO ACIDS: building blocks of proteins

Amino acids are the basic building blocks of peptides, polypeptides and proteins. Figure A1 shows the general structure of amino acids. The common structure is a C atom to which are attached an amino group ($NH_2$), a carboxyl group (COOH) and an H atom. The fourth group, denoted R, is variable and distinguishes the different amino acids.

Table A1 shows the R group of each of the 20 common amino acids. The R groups are organised according to their chemical nature. The simplest R group is an H atom — as seen in glycine. This amino acid forms part of a group of amino acids with aliphatic and non-polar R groups.

**TABLE A1** The twenty common amino acids grouped according to the chemical properties of their R group.

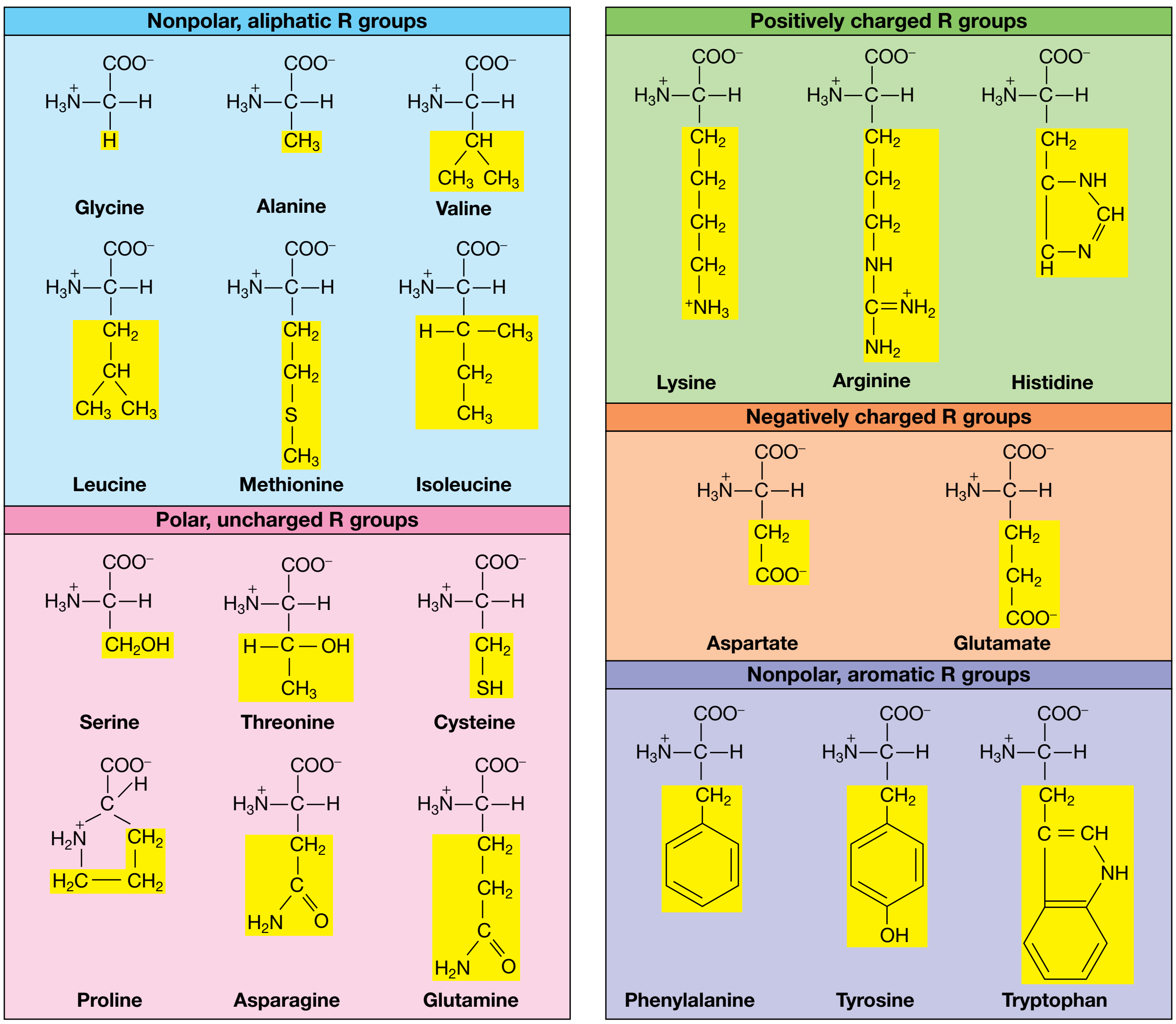


Table A2 shows abbreviations for the twenty common amino acids. Two schemes are in use, a three-letter abbreviation and a one-letter system.

**TABLE A2** Amino acids and their abbreviations.

| Amino acid | three-letter code | one-letter code |
|---|---|---|
| Alanine | Ala | A |
| Arginine | Arg | R |
| Asparagine | Asn | N |
| Aspartic acid | Asp | D |
| Cysteine | Cys | C |
| Glutamic acid | Glu | E |
| Glutamine | Gln | Q |
| Glycine | Gly | G |
| Histidine | His | H |
| Isoleucine | Ile | I |
| Leucine | Leu | L |
| Lysine | Lys | K |
| Methionine | Met | M |
| Phenylalanine | Phe | F |
| Proline | Pro | P |
| Serine | Ser | S |
| Threonine | Thr | T |
| Tryptophan | Trp | W |
| Tyrosine | Tyr | Y |
| Valine | Val | V |

# Glossary

## A

**abscission:** separation of a leaf or fruit from a plant

**absolute ages:** ages of objects, such as a rock or fossil, expressed in actual years

**acceptor molecules:** compounds such as NAD and FAD that can accept hydrogen atoms and electrons removed from compounds and transfer them to other compounds

**Acheulean technology:** production of stone tools, in particular multi-purpose hand axes, associated with *Homo erectus*, with the earliest tools dating from about 1.6 million years ago

**acquired immunity:** characterised by the presence in an individual of antibodies to a particular infective agent; in active acquired immunity, individuals produce their own antibodies as a result of exposure to an antigen; in passive acquired immunity, individuals receive antibodies from an outside source

**activation energy:** minimum amount of energy required to initiate a chemical reaction

**active site:** region of an enzyme that binds temporarily with the specific substrate of the enzyme

**active transport:** net movement of dissolved substances across a cell membrane by an energy-requiring process that moves substances against a concentration gradient from a region of lower to higher concentration

**adaptive convergence:** similarity in appearance of organisms that are not closely related but that have similar ways of life; also termed convergent evolution

**adaptive immunity:** form of specific immunity in which people produce their own antibodies to a particular antigen

**adaptive radiation:** outcome that occurs when one ancestral species gives rise to many new species that occupy different niches; a special case of divergent evolution

**aerobic respiration:** breakdown of glucose to simple inorganic compounds in the presence of oxygen and with release of energy that is transferred to ATP

**Afar Triangle:** region in present-day Ethiopia where many early hominin fossils have been found, the most famous being Lucy

**agent of selection:** see ‘selecting agent’

**alcoholic fermentation:** anaerobic respiration in yeasts

**allergens:** substances that elicit an allergic response

**allergic response:** (allergy) rapid immune responses to normally harmless antigens such as dust or pollen; involves action of mast cells

**allergy:** an abnormal immune response to a substance that is harmless for most people

**allopolyploid:** organism with more than two sets of chromosomes derived from the genomes of at least two different species

**allosteric site:** location on an enzyme molecule where a compound can bind and alter the shape of the enzyme, including its active site, so that the enzyme is inactivated

**alternative splicing:** production of different mRNA molecules from one kind of pre-mRNA through joining various combinations of exons from the pre-mRNA

**amino acids:** basic building blocks or sub-units of polypeptide chains and proteins

**anabolism:** chemical reactions in a cell in which complex molecules are built from simple molecules

**anaerobic respiration:** form of respiration occurring in the absence of oxygen in which glucose is broken down to smaller compounds, such as lactate, with release of energy that is transferred to ATP

**anaphylaxis:** acute and potentially lethal allergic reaction to an allergen to which a person has become hypersensitive

**aneuploid:** cell or organism in which the total number of chromosomes is not an exact multiple of the haploid number owing to a missing or an additional single chromosome

**antibiotics:** a class of antimicrobial drug used in the treatment and prevention of bacterial infections; they act either by killing pathogenic bacteria or by inhibiting their growth

**antibodies:** proteins produced by animals in response to antigens and which react specifically with the antigen that induced their formation

**anti-codon:** sequence of three bases in a transfer RNA molecule that can pair with the complementary codon of a messenger RNA molecule

**antigens:** compounds, usually proteins, that can trigger the immune system to respond in various ways, including antibody production

**antigen-binding sites:** region of an antibody molecule to which an antigen binds

**antigen-presenting cells (APCs):** cells of the immune system that display, on their surfaces, antigens complexed with MHC molecules

**antigenic drift:** in reference to viruses, small mutations of viral genes that do not change the antigenic properties of the virus so that it continues to be recognised and reacted to by the body's immune system

**antigenic shift:** in reference to viruses, accumulated mutations in viral genes that alter the antigenic properties of the virus so that it is no longer recognised and reacted to by the body's immune system

**antiretroviral drugs:** drugs that are active against the group of viruses termed retroviruses

**antisense mRNA:** a single-stranded mRNA molecule with a base sequence complementary to that of a target mRNA molecule transcribed by a particular gene

**apical dominance:** influence exerted by a terminal bud that suppresses growth of lateral buds

**apoptosis:** the natural death of cells, also called programmed cell death

**apoptotic bodies:** small membrane-bound vesicles produced during the programmed death of a cell by apoptosis

**aquaporins:** channel proteins that are specific for the facilitated diffusion of water molecules

**arboreal:** living in trees

**artificial insemination (AI):** process in which donor sperm is artificially introduced into the reproductive tract of a female

**artificial pollination:** process of cross pollination in which pollen collected from one plant is manually transferred to the stigma of a second plant

**asymptomatic carrier:** person with a microbial infection showing no clinical symptoms but able to infect others

**ATP:** compound containing adenosine and three molecules of phosphate, the major supplies of which are produced in the mitochondria; common source of chemical energy for cells

**attenuated:** refers to a microorganism that, while still living, has been treated so that it is no longer able to cause disease

**auto-antigen:** molecules normally present in the body that stimulate an abnormal immune response through antibody production

**autoantibodies:** antibodies produced in an abnormal response to one of the body's own molecules

**autoimmune diseases:** diseases in which the immune system fails to identify 'self' material and makes antibodies against the body's own tissues

**autopolyploidy:** organism with more than two sets of chromosomes derived from the genome of one species

**autotrophic:** describes an organism that, given a source of energy, can produce its own food from simple inorganic substances; also known as a producer

## B

**bacteria:** (singular = bacterium) microscopic, usually unicellular, organism, and member of Kingdom Monera

**bacterial transformation:** process in which bacterial cells take up segments of foreign DNA that become part of their genetic make-up

**bacteriochlorophyll:** light-capturing pigments present in photosynthetic bacteria

**bacteriophages:** types of virus that infect bacteria

**biological evolution:** process of change in members of a species under the influence of natural selection and requiring much longer time periods than so-called cultural evolution

**biosignatures:** chemical or physical traces that can be inferred to have resulted from the action of life forms; indirect evidence of life

**bipedal:** form of locomotion involving routine movement on two feet

**bone marrow:** fatty substance in the internal cavity of bones, the site of blood cell formation

**bottleneck effect:** chance effects on allele frequencies in a population as a result of a major reduction in population size

**bovine spongiform encephalopathy (BSE):** also known as 'mad cow disease'; a progressive and fatal neurodegenerative disease of cattle caused by abnormal prions

**brachiation:** mode of locomotion involving swinging from one handhold to another

**broad spectrum:** characterises an antibiotic that is effective against many different pathogens

**buboes:** swollen lymph nodes in the armpits and groin

**bubonic plague:** another term for black death

**budding:** in viruses, refers to the shedding of viral particles from the host cells infected with the virus

## C

**Calvin cycle:** cycle of reactions occurring in the stroma of chloroplasts in the light-independent stage of photosynthesis during which carbon dioxide is progressively reduced to sugar

**capsid:** protein shell enclosing the genetic material of a virus; in a naked virus, the capsid forms the outermost component of the viral particle

**capsule:** gelatinous layer surrounding the cell wall of some bacteria

**carbon fixation:** process by which atmospheric carbon dioxide is incorporated into organic molecules such as sugars

**carrier proteins:** trans-membrane proteins involved in the facilitated transport of hydrophilic substances across the plasma membrane in a process in which the shape of the carrier protein changes

**carrier testing:** testing carried out for the purpose of identifying if a person is a heterozygous carrier of an allele for a recessive disorder

**caspases:** key enzymes in the process of programmed cell death or apoptosis

**cast:** type of fossil formed when a buried structure decays leaving a cavity, which is later filled by different material that forms a model of the original structure

**catabolism:** chemical reactions in a cell in which complex molecules are broken down into simple molecules

**causal link:** relationship between an event and a result such that the event is the direct cause of the result

**cell membrane:** alternative term for plasma membrane

**cell surface markers:** proteins (or glycoproteins) present on the plasma membrane that distinguish various cell types and discriminate between self and non-self

**cell-surface receptors:** regions of a trans-membrane molecule exposed at the surface of a cell that act in cell signalling by receiving and binding to extracellular molecules

**channel proteins:** trans-membrane proteins involved in the transport of specific substances across a plasma membrane by facilitated diffusion

**chemotaxis:** movement of a cell or organism along the concentration gradient of a particular substance, such as a pheromone

**chloroplast:** chlorophyll-containing organelle that occurs in the cytosol of cells of specific plant tissues

**chromosome banding:** form of staining of chromosomes producing a reproducible and distinctive pattern of dark and light bands on each chromosome

**chromosome painting:** form of staining of chromosomes such that each chromosome fluoresces in a distinctive colour

**chronic inflammation:** condition that results if an acute inflammatory response is not resolved once an infection has been controlled

**cilia:** (singular = cilium) in eukaryote cells, fine hair-like outfoldings formed by extensions of the plasma membrane involved in synchronised movement

**clade:** group of organisms that, based on evidence, can be assumed to be the direct descendants of a common ancestor

**cladogram:** diagram, based on cladistic study, that shows the inferred relationship between different groups of organisms

**clonal expansion:** process of multiple cycles of cell division of one kind of B cell that carries receptors specific for a particular antigen, resulting in the production of large numbers of identical B cells

**clonal selection:** event occurring in lymph nodes in which those B cells with receptors that can recognise a new antigen come into contact with that particular antigen

**clonal selection theory:** theory first proposed by Sir Macfarlane Burnet to explain how the large numbers of antibodies specific to particular antigens are produced

**clone:** group of cells, organisms or genes with identical genetic make-up

**cloning:** process of making identical copies of a gene, a cell or an organism

**coding region:** part of a gene that contains the coded information for making a polypeptide chain

**codons:** sequences of three bases in a messenger RNA molecule that contain information either to bring amino acids into place in a polypeptide chain or to start or stop this process

**coenzyme:** organic compound that acts with an enzyme to alter the rate of a reaction

**cofactors:** non-protein, additional component that is essential for the normal functioning of some enzymes

**colistin:** a powerful antibiotic, derived from certain strains of a bacterial species, that is active against the endotoxins of Gram-negative bacteria

**colonoscopy:** procedure in which a thin flexible tube (colonoscope) is inserted via the anus for the purpose of examining the lining of the large intestine

**common ancestor:** an ancestral species from which later species evolved

**communicable disease:** an infectious disease that can be transmitted from person to person, either directly or indirectly

**comparative genomics:** comparative study of the genomes of several species

**complement:** a system of more than 30 plasma proteins that form part of the innate immune system, and, when activated by contact with the surface of a pathogen, set in train actions to destroy it

**complementary DNA (cDNA):** double-stranded DNA synthesised using single-stranded RNA, such as mRNA, as the template

**complete selection:** applies when no member of a population with a particular phenotype makes any genetic contribution to the next generation as, for example, when a phenotype results in childhood death. When a phenotype is selected against completely, it has zero genetic fitness.

**conjugate:** describes a condition in which amino acids of a protein, particularly proteins in the nucleus, associate with other groups; for example, nucleoproteins in which proteins conjugate with nucleic acids

**conservative missense mutation:** type of point mutation in which a base substitution does not change the resulting amino acid translated

**continuous:** within a population, describes variation that forms a continuum of phenotypes, differing to a small degree from one another, such as that seen for example in the heights of adult human males

**convergent evolution:** independent appearance of similar features in response to similar environmental pressures in different species that are not closely related

**coupling:** linking of an energy-requiring (endergonic) reaction or process with an energy-releasing (exergonic) one, such that the latter provides the energy needed to drive the former

**covalent:** describes a strong chemical bond formed between atoms when they share electrons

**cranium:** vertebrate skull minus the lower jaw

**Creutzfeldt-Jakob disease (CJD):** CJD is a fatal neurodegenerative disease in people that is caused by abnormal prions; three types of classic CJD exist as well as a variant form (vCJD) that resulted from people eating beef from cattle infected with bovine spongiform encephalopathy (mad cow disease)

**Cro-Magnons:** label given to humans who produced artifacts including finely crafted tools and other objects of art from stone, wood and bone and wall paintings, and who lived 40 000 to 10 000 years ago

**cultural change:** development over time of human societies from simple to more complex; also known as cultural evolution

**cultural evolution:** rapid changes in a population as a result of transmission of knowledge, in contrast to biological evolution, which requires a long time

**culture:** any form of learned behaviour acquired through formal teaching or by imitation, such as learning to write or learning a social rule

**cyanide:** a poison that is an irreversible inhibitor of a key enzyme in aerobic cellular respiration

**cytokines:** peptide, protein or glycoprotein molecules that act as messengers between cells and elicit a response in the receiving cell; regulate many parts of the immune system

**cytokinins:** plant hormones that regulate cell reproduction in shoots, roots and fruits

**cytotoxic T cells:** kind of T cell that specifically recognises and kills infected body cells

## D

**death receptor pathway:** one of the signal pathways involved in programmed cell death or apoptosis that is initiated in healthy cells when a signal from outside these cells activates death receptors on these cells

**decoded:** or translated; refers to the translation of genetic information held in DNA into amino acids

**degranulation:** process by which immune cells such as NK cells and cytotoxic T cells eliminate infected and cancerous body cells by releasing anti-microbial and toxic granules into those abnormal cells

**Denisovans:** members of an extinct hominin species in the genus *Homo*

**deoxyribonucleic acid (DNA):** nucleic acid containing the four bases — adenine, guanine, cytosine and thymine — which forms the major component of chromosomes and contains coded genetic instructions

**depolarisation:** process occurring during the transmission of a nerve impulse along an axon when the charge on a localised region of the axon changes in response to the sudden inflow of sodium ions

**derived character:** in cladistics, refers to a novel feature or character that is assumed to have evolved from a feature originally present in an ancestral species

**differential reproduction:** occurs when different inherited varieties in a population vary in their rates of production of viable offspring

**direct transmission:** mechanism of transmission of pathogenic agents that involves direct person-to-person contact, such as by kissing or sexual contact

**discontinuous:** describes variation within a population that consists of a few discrete non-overlapping phenotypes

**divergent evolution:** outcome that results when, over time, one ancestral species changes to give rise to several new species each occupying a different niche

**DNA arrays:** alternative term for microarrays or gene chips

**DNA editing:** a process by which changes to the DNA sequence of genes can be achieved; also termed gene editing

**DNA fingerprinting:** technique for identifying DNA from different individuals based on variable numbers of tandem repeats of short DNA segments near the ends of chromosomes. These regions are known as minisatellites or variable number of tandem repeats (VNTRs), and their detection uses a multi-locus probe.

**DNA profiling:** technique for identifying DNA from different individuals based on variable regions known as short tandem repeats (STRs) or microsatellites. Detection of variation uses a combination of single-locus probes, each specific to a different STR locus.

**DNA sequencers:** instruments that automate the identification of the order or sequence of bases along a DNA strand

**DNA-binding proteins:** proteins that bind to regions of nuclear DNA near genes and directly switch these genes on or off

**DNA–DNA hybridisation:** technique used to compare the similarity of the genomes of different species and so infer their degree of evolutionary relatedness

## E

**ecosystem:** a biological community involving interactions, both between the organisms of that community and between organisms and the physical environment of the ecosystem

**Ediacaran period:** recently defined period in the geological time scale that precedes the Cambrian period

**electrophoresis:** technique for sorting through an electric field a mixture of DNA fragments (and other molecules with a net charge) on the basis of different fragment lengths

**electroporation:** a technique that uses brief exposure of host cells to an electric field to enable the entry of segments of foreign DNA into the cells

**embryo transfer:** process of removing early embryos from one female and transferring them into the reproductive tracts of other females of the same species

**encoded:** refers to the holding of genetic information in DNA in coded form as a base sequence

**end-product inhibition:** inhibition of the early stage of a multi-step pathway by the final product of that pathway

**endergonic:** refers to a chemical reaction that is energy requiring

**endocrine-disrupting chemicals (EDCs):** contaminants in the environment with chemical structures that produce disruptions of the normal operation of the endocrine system resulting in developmental, reproductive and other disorders

**endocytosis:** bulk movement of solids or liquids into a cell by engulfment

**endoplasmic reticulum (ER):** cell organelle consisting of a system of membrane-bound channels that transport substances within the cell

**endosymbiosis:** a special case of symbiosis where one of the organisms lives inside the other

**enveloped virus:** those viruses with an outer envelope composed of part of the plasma membrane of the host cell when the viral particles bud from the cell

**environmental disease:** any disease resulting from exposure to toxic factors in the environment such as chemicals or radiation

**enzyme inhibitor:** a substance that can inactivate an enzyme, either temporarily or permanently, typically by binding to the enzyme, reducing its activity or by interfering with the enzyme in some way

**enzyme-substrate (E-S) complex:** transient compound produced by the bonding of an enzyme with its specific substrate, at the active site of the enzyme

**epithelial tissues:** include the external layers of the skin, and the internal linings or mucous membranes of the cavities of the airways, the gut and the urogenital tract

**eukaryote:** cell or organism with a membrane-bound nucleus

**eukaryotic:** describing cells that have a membrane-bound nucleus

**exergonic:** refers to a chemical reaction that is energy releasing

**exocytosis:** movement of material out of cells via vesicles in the cytoplasm

**exoenzymes:** enzymes that function outside the cells that produce them; that is, enzymes in an extracellular setting

**exon:** part of the coding region of a gene that is both transcribed and translated

**exon juggling:** production of different combinations of the exons in a gene transcript leading to different gene products from the same gene

**exotoxin:** toxin secreted into the surrounding medium by a micro-organism as it grows

**extinction:** permanent loss of a species, ranging from local to global

**extracellular:** refers to locations within the body that are outside cells, such as blood plasma and extracellular fluid

## F

**facilitated diffusion:** form of diffusion involving a specific carrier molecule for the substance that diffuses
**facultative anaerobe:** organism that can live regardless of whether oxygen is present or not
**familial search:** a type of DNA search used when a direct match cannot be found between a crime scene profile and an entry in the national DNA database, in order to identify possible partial matches that may exist in individuals related to a national entry
**first line of defence:** part of the defence against pathogens provided by the mechanical, chemical and microbiological barriers of the innate immune system that prevent entry of pathogens into the body
**flanking regions:** regions located either downstream or upstream of the coding region of a gene
**flehmen:** a behaviour seen in certain mammals in which an airborne pheromone signal is drawn into their vomeronasal organs
**flood basalts:** very thick layers of basalt covering large areas of land or ocean floor that are the result of a series of massive volcanic eruptions
**fluid mosaic model:** a model which proposes that the plasma membrane and other intracellular membranes should be considered as two-dimensional fluids in which proteins are embedded
**fossilisation:** process of preserving parts of organisms that lived in the geological past
**founder effect:** chance effects on allele frequencies in a population that is formed from a small unrepresentative sample of a larger population
**founder population:** refers to a small group of organisms, as small as a mating pair or individual inseminated female, that starts a new population
**frameshift:** type of mutation in which, as a result of insertion or deletion of a base, all codons from that point are affected.
**fusion:** the joining of two smaller chromosomes to form one larger chromosome

## G

**gene action:** processes of transcription and translation of a gene to produce a gene product, typically a polypeptide chain
**gene chips:** alternative name for microarrays
**gene cloning:** process of making multiple identical copies of a specific gene or segment of DNA
**gene conservation:** preservation of a DNA sequence in different species related by evolution
**gene editing:** see 'DNA editing'
**gene pool:** sum total of genetic information present in a population
**gene sequence:** order of bases in a DNA segment
**gene sequencing:** identification of the order or sequence of bases along the DNA of a specific gene
**genetic code:** representation of genetic information through a non-overlapping series of groups of three bases (triplets) in a DNA template chain
**genetic drift:** changes, unpredictable in direction, in allele frequencies from one generation to the next owing to the action of chance events
**genetic screening:** testing of individuals to detect those with the allele responsible for a particular genetic disorder
**genetic testing:** testing of individuals to detect those with the allele responsible for a particular genetic disorder
**genetic variation:** variation exhibited among members of a population owing to the action of genes
**genetically modified organisms (GMOs):** organisms whose genomes contain heritable changes created through the use of genetic engineering technology
**genome editing:** See 'DNA editing'
**germ theory of disease:** theory derived from the research of Pasteur and Koch, who demonstrated the causal link between microbes and infectious disease
**gestation period:** time from fertilisation to birth
**gibberellic acid (GA):** plant hormone responsible for leaf abscission
**gibberellin:** group of plant hormones that regulate aspects of cell growth and development
**glycoprotein:** combination formed when a carbohydrate group becomes attached to the exposed part of a trans-membrane protein
**Golgi complex:** organelle that packages material into vesicles for export from a cell (also known as Golgi apparatus or Golgi body)
**Gram-negative bacteria:** one major group of bacteria characterised by cell walls with an outer layer of lipopolysaccharide (LPS)
**Gram-positive bacteria:** bacteria containing a thick layer of peptidoglycan, which retains the crystal violet dye used in the Gram staining technique; Gram-positive bacteria are normally susceptible to antibiotics
**granum:** (plural = grana) a stack of flattened, fluid-filled membranous sacs, known as thylakoids, that are located within plant chloroplasts
**granzymes:** active protease enzymes present in granules that form part of the immune defences of NK cells and cytotoxic T cells
**guide RNA:** a component of the CRISPR gene-editing system
**gulono-lactone oxidase (GULO):** the enzyme catalysing the last step of the multi-step pathway that produces vitamin C from glucose

## H

**haplogroup:** a group of people with similar haplotypes who share a common ancestor as may be identified, for example, by their mitochondrial DNA sequences

**Hardy–Weinberg principle:** concept that allele frequencies in a population remain constant from one generation to the next unless an agent of change acts on the population

**helper T cells:** class of T cell whose actions include stimulating B cells to produce specific antibodies

**hemagglutinin (HA):** one of the surface proteins present on the outer envelope of influenza viruses

**herbicide:** chemical that kills plants

**herbivorous:** describes a plant-eating organism

**herd immunity:** indirect protection, at the population level, against an infectious disease; the protection is created by the presence in the population of a high proportion of individuals who are vaccinated against the particular disease

**heterotrophic:** describes an organism that cannot make its own food and must ingest or absorb organic material from its environment; also known as a consumer

**HLA typing:** the process by which the human leucocyte antigens (HLA) present on a person's cells are identified, typically for the purpose of matching a donor tissue or organ to the most compatible recipient

**homeotic genes:** master genes that are active during embryonic development and control other structural genes ensuring that various body parts are built in the correct locations

**hominoids:** collective term for apes including humans; members of the family Hominoidea

***Homo erectus*:** extinct human species appearing about two million years ago that was the first to migrate from Africa

**homologous:** in genetics, refers to members of a matching pair of chromosomes that carry the same gene loci

**human immunodeficiency virus (HIV):** the virus that attacks the helper T cells of the adaptive immune system and is the cause of acquired immunodeficiency syndrome (AIDS)

**human leucocyte antigens (HLA):** antigens present on human cell surfaces that determine the 'self' status of a person's cells, with all nucleated cells carrying the Class I HLA markers (A, B and C) and with the T and B immune cells also carrying the Class II HLA markers (DP, DQ and DR)

**Huntington's disease (HD):** an inherited neurodegenerative disease with a typical onset in middle age that can be identified before symptoms appear using a pre-symptomatic test that detects the defective allele

**hybridisation:** pairing between single-stranded complementary DNA segments from organisms from the same or even different species

**hybridoma technique:** procedure for producing pure monoclonal antibodies in unlimited amounts for use in the treatment of various cancers in both trial and clinical settings

**hydrophilic:** refers to compound that dissolves easily in water; also termed polar

**hydrophobic:** refers to substances that tend to be insoluble in water; also termed non-polar

**hypertonic:** having a higher concentration of dissolved substances than the solution to which it is compared

**hypotonic:** having a lower concentration of dissolved substances than the solution to which it is compared

## I

**immigration:** one of the agents that can result in changes in allele frequencies in the gene pool of a population

**immune system:** group of lymphoid tissues and organs and lymphatic vessels that assist the body to resist infection and disease through specialised cells such as macrophages, neutrophils, B cells and T cells

**immunity:** resistance to infectious disease

**immunoglobulins (Ig):** proteins that function as antibodies and as receptors that are produced by specialised B cells; five classes exist: IgA, IgD, IgG, IgE and IgM

**immunological memory:** See 'memory cells'

**immunosuppressive drugs:** drugs that suppress the normal immune response of recognising and attacking foreign antigens; such drugs are administered to individuals receiving organ or tissue transplants that are not perfectly matched to their own HLA genotype

***in situ* hybridisation:** pairing between a single-stranded, labelled probe and a specific complementary, single-stranded DNA segment that is not isolated but is in place within a chromosome

**incubation period:** the time period between infection and the first appearance of the symptoms of a disease

**index fossils:** fossils of geologically short-lived species that are widely distributed but are found in a restricted depth of rock strata

**indirect transmission:** mechanism of transmission of pathogenic agents that does not involve direct person-to-person contact, such as by airborne droplets or by ingestion of contaminated food

**induced fit model:** a model to account for the substrate specificity of enzymes but which entails a change in the active site of the enzyme to fit the substrate

**infectious disease:** disease caused by an organism, typically a microorganism; can be transmitted from person to person

**inflammation:** reaction to an infection, typically associated with reddening of the skin owing to an increased blood supply to that region

**inherited disease:** a disease that is transmitted from a parent to offspring via sperm or eggs carrying a defective allele or chromosomal error

**innate immunity:** the type of immunity that is present from birth, is fast acting, but not long lasting, acts at the site of infection, and produces non-specific (generic) responses against classes of pathogens

**inorganic cofactors:** most commonly metal ions or clusters of several ions that are required for some enzymes to function

**integral proteins:** fundamental components of the plasma membrane that are embedded in the phospholipid bilayer

**interferons (INFs):** group of proteins secreted by some cells, in response to a virus infection, that help uninfected cells resist infection by that virus

**intra-specific evolution:** evolutionary changes in allele frequencies within a population over time that does not result in new species

**intracellular receptors:** receptors for signalling molecules that are located within the cytosol or the nucleus of target cells

**introns:** parts of the coding region of a gene that are transcribed but not translated

**irreversible inhibition:** the permanent inactivation of an enzyme typically occurring when a compound forms strong covalent linkages with amino acids at the active site of an enzyme

**isotonic:** having the same concentration of dissolved substances as the solution to which it is compared

## J

**jasmonates (JAs):** a group of plant hormones that function by defending plants from attack by herbivores

## K

**Krebs cycle:** second stage of aerobic respiration, occurring mainly in mitochondria, in which pyruvate is broken down to carbon dioxide

## L

**lac operon:** a collection of adjacent genes in bacteria that encode the enzymes needed for the metabolism of lactose, and other sequences that regulate these genes including a promoter and a repressor

**lactase nonpersistent:** refers to individuals who, after infancy, no longer produce the lactase enzyme needed for digestion of the milk sugar lactose

**lactase persistent:** refers to individuals who, after infancy, continue to produce the lactase enzyme needed for digestion of the milk sugar lactose

**lactic acid fermentation:** the form of anaerobic cellular respiration that occurs in yeast cells

**last common ancestor (LCA):** the most recent organism from which a set of related organisms is directly descended

**law of fossil succession:** the principle, based on the changes over time in the kinds of fossils that exist, that rocks from different locations with similar fossils can be assumed to be of similar age

**leukocidins:** exoenzymes produced by some pathogenic bacteria that target and destroy the white blood cells of the adaptive immune system

**ligand:** any molecule that binds to a specific target to form an active complex, as, for example, signalling molecules that bind to cell surface receptors

**ligase:** enzyme that catalyses the joining of two double-stranded DNA fragments

**limiting factor:** environmental condition that restricts the types of organism that can survive in a given habitat

**loaded form:** refers to the form of co-enzymes, such as NAD and NADP, that can act as electron donors

**local mediators:** signalling molecules that diffuse over short distances from their cells of origin to their target cells

**lock-and-key model:** one model to account for the substrate specificity of enzymes that proposes that the active site of an enzyme has a shape complementary to that of its substrate

**Lucy:** the first and most famous fossil specimen of *Australopithecus afarensis*, which was found in 1974 in present-day Ethiopia

**lymph nodes:** organs of the lymphatic system where B cells and T cells are activated and adaptive immune responses occur

**lymphocytes:** class of white blood cells found in all tissues including blood, lymph nodes and spleen, and which play a role in specific immunity

**lysis:** destruction of cells

**lysosomes:** membrane-bound vesicles containing digestive enzymes

## M

**macrophages:** cells derived from monocytes that may be found in various tissues throughout the body and can engulf foreign material

**macroscopic:** of a size that is visible to the unaided eye

**mass extinctions:** periods of the Earth's geologic history during which extinctions occurred globally and on a massive scale

**mast cells:** non-motile cells containing histamine, which is involved in allergic responses

**mediator:** intracellular protein, such as a cytokine, that enhances and activates the functions of other proteins

**melting temperature ($T_m$):** temperature at which half of a sample of double-stranded DNA becomes single stranded

**membrane-attack complex (MAC):** one of the defence mechanisms resulting from activation of complement proteins that destroys pathogen cells by osmotic shock

**memory cells:** long-lived B cells that persist in lymph nodes after a specific disease has been resolved and that are capable of responding if the same disease re-occurs

**messenger RNA (mRNA):** single-stranded RNA formed by transcription of a DNA template strand in the nucleus; mRNA carries a copy of the genetic information into the cytoplasm

**metabolism:** total of all chemical reactions occurring in an organism

**miasma theory of disease:** a discredited theory in which 'bad air' was accepted as the cause of diseases

**microarrays:** surfaces, often glass, on which thousands of spots of single-stranded DNA are positioned in a precise arrangement and where each spot corresponds to a part of one particular gene

**microbe:** a label that typically includes microscopically small cellular organisms such as bacteria and archaea, and non-cellular agents such as viruses

**microbiota:** the community of various microbes found in a particular location, including the human gut

**microcephaly:** a congenital condition in which a baby is born with a head that is significantly smaller than normal

**microfossils:** very small fossils of a size that can only be studied using a microscope

**modulator:** any molecule that directly influences the biological effects of other molecules

**molecular clock:** concept that changes in the amino acid sub-units of a given protein occur at an approximately constant rate over geological time

**molecular clock principle:** a method for inferring the time since two species diverged from their most recent common ancestor based on the assumption that the longer the time since divergence, the greater the number of changes expected in the amino acid sequences of proteins common to both species

**molecular pattern:** microbe-specific (pathogen-associated) molecules present on the surface of pathogens that are recognised by phagocytic cells of the innate immune system

**monomeric enzymes:** enzymes that consist of a single polypeptide chain

**monomers:** small sub-units of the larger units of cells (polymers); also known as the building blocks of cells and include sugars, fatty acids, amino acids and nucleotides

**monomorphic:** refers to a population in which all members are identical with regard to a particular phenotypic trait

**motor end plate:** terminal of an axon that is located at the junction of a nerve and a muscle and that releases a neurotransmitter that binds to receptors on the muscle cells causing their contraction

**mould:** type of fossil that results when a buried organism decays, leaving an impression of the original organism

**mucous membranes:** cellular linings of the inner spaces within the airways, the gut and the urogenital tract

**mucus:** gelatinous fluid secreted by cells of the mucous membranes

**multiple cloning site (MCS):** also known as a polylinker, an MCS is a short DNA segment within a plasmid that contains recognition sites for a number of different restriction enzymes

**multiple ovulation:** process of stimulating a female through hormone treatment to release an increased number of eggs

**multiple ovulation and embryo transfer (MOET):** process of stimulating multiple egg release in one female and later transferring her early embryos to other females of the same species, where they develop to term

**mutagenic agents:** chemical or physical agents that can cause mutation in DNA

## N

**naked:** refers to viruses that lack an outer envelope

**natural immunity:** a form of specific immunity in which antibodies are produced naturally in response to a pathogen that gains entry to the body through natural means

**natural selection:** the process by which new heritable traits, whether morphological, physiological or behavioural, evolve and persist in a population in the wild

**necrosis:** a form of cell death resulting from the action of external factors, such as physical trauma, oxygen deprivation, and infection, that cause cells to disrupt with damage to surrounding tissue

**neuraminidase (NA):** one of the surface proteins present on the outer envelope of influenza viruses

**neurotransmitter:** chemical released by a neuron axon into the synaptic cleft between it and the target cell; stimulates or inhibits the target cell

**neutrophils:** most common type of white blood cell; one kind of phagocyte

**non-communicable disease:** any disease that cannot be transmitted from person to person
**non-conservative missense mutation:** a type of point mutation in which a base substitution changes the resulting amino acid that is translated
**non-GMO:** any organism that is not genetically modified
**non-infectious disease:** any disease resulting from a cause other than a pathogenic infection, such as a nutritional disease or an inherited disease
**non-specific immunity:** also called innate immunity; group of defences including physical barriers, action of different kinds of blood cells, chemical reactions, and the inflammation reaction that the body makes to all infections
**nuclear envelope:** membrane surrounding the nucleus of a eukaryote cell
**nucleic acids:** various types of DNA and RNA
**nucleocapsid:** the structure formed by a viral capsid in combination with the viral genetic material
**nucleoproteins:** proteins associated structurally with nucleic acids, either DNA or RNA
**nucleotides:** basic building blocks or sub-units of DNA and RNA and consisting of a phosphate group, a base and a sugar; the sugar in DNA is deoxyribose and that in RNA is ribose
**nucleotide sequence:** in DNA, refers to the specific order of nucleotides along part or all of a DNA molecule; also known as base sequence
**nutritional deficiency disease:** any disease that is the result of a deficiency in a person's diet

## O

**obligate aerobes:** organisms that can only carry out cellular respiration when oxygen is present
**obligate anaerobes:** organisms that can grow only in environments where oxygen is absent
**oestrus:** period of sexual receptivity in a nonhuman female mammal; also termed heat
**oestrus synchronisation:** process of making all mature females in a flock or herd sexually receptive within a predictable and narrow time frame
**Oldowan technology:** techniques used to fashion the oldest known stone tools, which date from 2.5 million years ago and are associated with *Homo habilis*
**oligomers:** a molecule that is built of a few monomer units
**operon:** in bacteria, a cluster of adjacent structural genes controlled by a single promoter and operating as a coordinated unit
**opposable:** describes a digit, such as the human thumb, whose ventral surface can be brought into contact with those of the fingers
**opsonisation:** the coating of the surface of pathogen cells by complement proteins, making the pathogens more susceptible to phagocytosis
**organic cofactors:** organic compounds needed for the functioning of particular enzymes that fall into two groups — prosthetic groups and coenzymes
**origin of replication:** the segment of DNA in a plasmid at which DNA replication begins
**original character:** in cladistics, refers to a feature or character that is present in an ancestral group from which other groups have descended
**osmosis:** net movement of water molecules across a partially permeable membrane and down a concentration gradient

## P

**pandemic:** refers to a situation when, over a relatively short time, many people worldwide contract a specific disease as it spreads from a region of origin
**partial selection:** applies when members of a population with a particular phenotype make less genetic contribution to the next generation than other phenotypes controlled by the same gene. Such a phenotype is said to have reduced genetic fitness.
**pasteurisation:** process of exposure of milk, other fluids, and canned foods to elevated temperatures (62 to 100 °C) that are sufficiently high to kill bacteria but not to chemically alter the substance
**pathogenic:** term used to describe a disease-causing agent
**pattern recognition receptor (PRR):** protein receptors present on phagocytic cells of the innate immune system that enable these cells to recognise and bind to pathogens, with recognition being at a generic level
**penicillin:** an antibiotic derived from a fungus that is typically effective against Gram-positive bacteria
**peptidoglycan:** a polymer consisting of sugars and amino acids that forms a major part of the cell wall of Gram-positive bacteria
**perforin:** a protein, released by some immune cells, that produces a pore in the membrane of pathogen cells undergoing an immune attack
**peripheral blood stem cell transplantation (PBSCT):** process of transplantation of peripheral blood stem cells from donor to recipient in which the stem cells have been mobilised from the donor's bone marrow by treatment with colony stimulating factors (CSFs)
**peripheral proteins:** are either anchored to the exterior of the plasma membrane through bonding with lipids or are indirectly associated with the plasma membrane through interactions with integral proteins in the membrane.
**phagocytes:** types of white blood cell, including neutrophils and monocytes, that can engulf and destroy foreign material, such as microorganisms, that enter the body
**phagocytic:** refers to immune cells that can engulf pathogens

**phagocytosis:** bulk movement of solid material into cells
**phagosome:** membrane-bound vesicle formed within a phagocytic cell that encloses the engulfed pathogen
**pheromones:** chemicals that, when released by one animal, elicit a response in another animal of the same species; in particular, act as a sex attractant for mating in many insect species
**phospholipids:** major type of lipid found in plasma membranes and the main structural component of plasma membranes
**photoautotroph:** any organisms, such as plants and algae, that trap the energy of sunlight to build complex organic molecules from simple inorganic substances
**photosystem:** refers to one of two kinds of structure (termed PS-I and PS-II) embedded in the thylakoid membranes of chloroplasts; each is composed of a light-harvesting complex of pigments (mainly chlorophylls) and various proteins, and a reaction centre that converts light energy to the chemical energy of energised electrons
***phyloGenerator*:** a website that can be used to generate phylogenetic trees
**phylogenetic tree:** a branching diagram showing inferred evolutionary relationships between life forms based on their observed physical and genetic similarities and differences
***Phylogeny.fr*:** a website that can be used to generate phylogenetic trees
**pinocytosis:** bulk movement of material that is in solution being transported into cells
**plasma membrane:** partially permeable boundary of a cell separating it from its physical surroundings; boundary controlling entry to and exit of substances from a cell
**plasmid:** small, circular fragment of double-stranded DNA, separate from the main chromosome, that is present in many copies in bacterial cells
**polar:** refers to a compound that dissolves easily in water; also termed hydrophilic
**polygenes:** a group of several genes in which each acts in a small but cumulative manner on a trait
**polylinker:** see 'multiple cloning site (MCS)'
**polymers:** large molecules made of identical or similar single units, for example, starch, which is made of many glucose units
**polymorphic:** refers to a population whose members show several variants of a particular trait
**polymorphism:** the presence in a population of different inherited discontinuous phenotypes, each being present in a significant proportion
**polypeptide:** macro-molecule built of amino acid sub-units and linked by peptide bonds to form a single chain
**population:** members of one species living in one region at a particular time
**position effects:** alteration of the activity or the expression of a gene, caused by the relocation of some genes on a chromosome
**post-mating isolation mechanisms:** mechanisms that operate after members of two different species have mated that prevent the production of viable fertile offspring
**post-transcription modification:** process occurring after transcription in which pre-mRNA is altered to become mRNA
**predictive testing:** DNA-based technique for distinguishing whether a person who is 'at risk', but shows no clinical signs, has inherited the allele(s) responsible for a particular genetic disease; also known as pre-symptomatic testing
**pre-implantation genetic diagnosis (PGD):** a test for the presence of alleles causing genetic disorders that uses one cell taken from an early embryo produced by in vitro fertilisation and occurs before the embryo is transferred to the mother's uterus (provided the test result is negative)
**pre-mating isolation mechanisms:** mechanisms that operate in nature to prevent the mating of members of two different species
**prenatal:** before birth
**presymptomatic testing:** DNA-based technique for distinguishing whether a person who is 'at risk' but who shows no clinical signs has inherited the allele(s) that is the cause of a particular genetic disease; also known as predictive testing
**primary antibody response:** production of antibodies induced in an individual by the first injection of antigen during vaccination
**primary lymphoid organs:** part of the lymphatic system that comprises the bone marrow and thymus
**probe:** single-stranded segment of DNA (or RNA) with a base sequence complementary to that in a target strand of DNA and that carries a radioactive or fluorescent label to detect it
**producer:** photosynthetic organisms and chemosynthetic bacteria that can build organic matter from simple inorganic substances given a source of energy; also known as autotrophs
**product:** the end point of a reaction that starts with a substrate
**proenzymes:** inactive form of enzymes; enzymes must be activated to perform their catalytic function
**programmed cell death:** alternative term for the process of apoptosis
**prokaryotes:** any cells or organisms without a membrane-bound nucleus
**prokaryotic:** refers to unicellular organisms that lack a membrane-bound nucleus

**promoters:** part of the upstream flanking region of a gene containing base sequences that control the activity of that gene

**prosthetic group:** an organic cofactor that is tightly bound to a particular enzyme

**proteins:** macromolecules built of amino acid sub-units and linked by peptide bonds to form a chain, sometimes termed a polypeptide; usual product of gene translation; some proteins consist of a single polypeptide while other proteins consist of two or more polypeptides

**proteome:** the complete array of proteins produced by a single cell or an organism in a particular environment

**proteomics:** the study of the proteome, the complete array of proteins produced by an organism

**proton:** alternative term for a hydrogen ion, that is, a hydrogen nucleus, which carries a unit positive charge

**protoplast:** a plant or bacterial cell whose outer cell wall has been removed

**pumps:** special transport proteins embedded across the plasma membrane that carry out the process of active transport

## R

**radiometric dating:** technique for obtaining the age of objects that depends on the known rate of decay of a radioactive parent isotope to a stable daughter product, such as potassium–argon dating

**radiometric dating technique:** a technique for estimating the absolute ages of rocks, using the rate of decay of certain radioactive elements present in the rocks

**random mating:** situation that applies in a population when all possible matings are equally likely to occur; this is in contrast to a situation in which non-random or assortative mating occurs

**rational drug design:** construction of a drug to fit the active site of a molecule so that the natural action of the molecule cannot occur

**receptors:** chemical structures, often on the surface of cells, that receive signals from hormones, neurons or cytokines

**recombinant plasmids:** plasmids that have been engineered to contain a segment of foreign DNA

**regulator genes:** class of gene that controls the activity of other genes

**relative ages:** age of objects expressed in relative terms so that they are identified as younger or older than other objects, but the actual ages of the objects concerned are unknown

**resting membrane potential:** the charge difference across the membrane of a nerve cell at rest with a net excess of positive charges outside the cell that is maintained by the sodium-potassium pump

**restriction enzymes:** enzymes, also known as endonucleases, that cut at specific sites within DNA molecules

**reverse transcriptase:** enzyme that directs the formation of copy DNA from a messenger RNA template

**reversible inhibition:** the temporary inactivation of an enzyme typically occurring when a compound forms weak non-covalent linkages with amino acids at the active site of an enzyme; the inhibitor can be dissociated from the enzyme

**ribonucleic acid (RNA):** type of nucleic acid consisting of a single chain of nucleotide sub-units that contain the sugar ribose, and the bases A, U, C and G; RNA includes messenger RNA (mRNA), transfer RNA (tRNA) and ribosomal RNA (rRNA)

**ribosomes:** organelles containing RNA that are major sites of protein production in cells in both eukaryotes and prokaryotes

**RNA interference:** process by which specific genes can be silenced, starting with double-stranded RNA

**RNA polymerase:** an enzyme that controls the synthesis of an RNA strand from a DNA template during transcription

**rough endoplasmic reticulum:** endoplasmic reticulum with ribosomes attached

## S

**sagittal crest:** a bony crest on the top of the skull of the robusts, a group of prehuman hominins belonging to the extinct genus *Paranthropus*

**scurvy:** a nutritional deficiency disease resulting from a vitamin C deficiency

**sebum:** the oily secretion produced by sebaceous glands of the skin

**second line of defence:** part of the defence provided by the immune cells and soluble proteins of the innate immune system against attacking pathogens that gain entry to the body; defences include phagocytosis, degranulation and inflammatory responses

**second messenger:** an intracellular signalling molecule produced during the process of signal transduction

**secondary antibody response:** the rapid production of high levels of specific antibodies to a foreign antigen that occurs in a person who was previously exposed to the same antigen

**secondary lymphoid organs:** part of the lymphatic system that comprises the lymph nodes and the spleen

**selectable marker gene:** genes carried by plasmids for certain traits, often for antibiotic resistance

**selecting agents:** any factors that affect the survival or fertility of members of a population such that variation is produced in different phenotypes

**selective advantage:** relative higher genetic fitness of a phenotype compared with other phenotypes controlled by the same gene

**selective breeding:** a process of mating that is not random, but uses parents chosen by the breeder on the basis of particular phenotypic characteristics that they display

**selectively permeable:** see 'semipermeable'

**self antigens:** antigens on cells that are recognised by self receptors as being part of the same body

**semen:** nutrient-rich fluid containing sperm

**semi-synthetic:** regarding an antibiotic, a substance produced partially by chemical synthesis

**semipermeable:** describes the plasma membrane as selectively permeable or partially permeable in that it allows only certain molecules to cross it by diffusion

**sensitivity test:** test to establish the most effective drug to use for treatment against a particular bacterial infection

**septicemia:** presence in the bloodstream of pathogenic microbes or their toxins, causing the serious condition of septicaemia, commonly called blood-poisoning

**severe combined immunodeficiency (SCID):** an inherited X-linked disorder of which there are several forms; the disease is 'severe' because the immune system is greatly weakened, and the disease is 'combined' because the disorder affects two types of white blood cells that are key components of the adaptive immune system

**sex selection:** process of producing offspring, predominantly of one sex, by artificially separating sperm from a semen sample into those with X chromosomes and those with Y chromosomes

**sexual dimorphism:** condition in some species, including some mammals and birds, in which adult males and females of the species can be readily distinguished because of marked differences in body size or in the colouring or patterning of hair/feathers

**short tandem repeats (STRs):** chromosomal sites where many copies of a short DNA sequence are joined end-to-end; also termed microsatellites; sequences are normally 2 to 4 base pairs, and the number of repeats are variable between unrelated people

**Siberian Traps:** massive deposits of 'flood basalts' seen today in Siberia that are the remnants of widespread volcanic eruptions occurring about 250 million years ago; the eruptions are associated with one of the major extinction events in Earth's history

**side effects:** adverse effects on a patient occurring in association with use of a particular drug

**signal transduction:** cascade of events linking an external signal (such as from a hormone) to a specific cellular response

**signal-binding domain:** the region of a trans-membrane receptor that is exposed to the outside of the cell in which it is embedded

**signalling proteins:** proteins involved in the transmission of signals across cell membranes or in signal transduction within cells

**simple diffusion:** the movement of substances across the phospholipid bilayer from a region of higher concentration to one of lower concentration of that substance; that is, *down* its concentration gradient

**small interfering RNA:** short fragments of RNA that play a role in the process of RNA interference

**sodium–potassium pump:** protein that transports sodium and potassium ions against their concentration gradients to maintain the differences in their concentrations inside and outside cells

**somatic mutation:** change in the genetic material (DNA) that occurs in a body cell and cannot be transmitted to the next generation

**special creation of species:** pre-Darwinian view that each species was separately and independently created

**speciation:** process of formation of new species

**spliceosomes:** complex molecules present in the nucleus that remove introns from the pre-mRNA transcript

**spores:** in bacteria, reproductive structure that is resistant to heat and desiccation; also formed by fungi and some plants

**stomata:** (singular = stoma) an opening, typically on a leaf surface, through which water vapour and carbon dioxide can move

**stroma:** in chloroplasts, the semi-fluid substance between the grana, which contains enzymes for some of the reactions of photosynthesis

**stromatolites:** layered structures built up over time in shallow water by the accretion of sediments trapped by microbial mats of cyanobacteria; fossil stromatolites provide indirect evidence of life

**structural genes:** genes that produce proteins that contribute to the structure or functioning of an organism

**substrate:** compound on which an enzyme acts

**substrate binding site:** the region of the active site of an enzyme where its specific substrate forms temporary bonds

**synapse:** in physiology, junction between two neurons; in genetics, the close pairing that takes place between homologous chromosomes during meiosis

**synaptic clefts:** the narrow space at the junction of two neurons where neurotransmitter molecules released from the end of the axon of the first neuron transmit the nerve impulse to the second neuron

## T

**TATA box:** short base sequence consistently found in the upstream flanking region of the coding region of genes of many different species

**taxa (singular, taxon):** a taxonomic group of any rank, such as a species, genus, class or phylum

**technological evolution:** progressive development over time of technologies giving greater human control over the environment

**template strand:** part of one strand of a DNA double helix that contains the coded information of a particular gene; sometimes called the sense strand

**third line of defence:** part of the defence provided by the immune cells of the adaptive immune system through the various actions of T cells and B cells when they attack extracellular and intracellular pathogens

**thylakoids:** flattened membranous sacs in chloroplasts that contain chlorophyll

**tissue culture:** in plants, technique in which several identical plants are produced from a single piece of plant tissue

**tissue typing:** see 'HLA typing'

**toxoids:** inactivated toxins used for active immunisation

**trans-membrane:** refers to proteins that are embedded within and span the plasma membrane but have parts exposed to the exterior and interior of the cells

**transcription:** process of copying the genetic instructions present in DNA to messenger RNA

**transcription factors:** proteins that bind to specific DNA sequences and control the rate of transcription of the DNA

**transfer RNA (tRNA):** one of a group of single-stranded RNA molecules that can attach to a specific amino acid and can also pair with an mRNA codon

**transgenic organisms:** organisms that carry in their genomes one or more genes artificially introduced from another species

**translation:** process of decoding the genetic instructions in mRNA into a protein (polypeptide chain) built of amino acids

**transmissible diseases:** contagious diseases with the particular pathogen concerned being transmittable from one host to another by either direct or indirect means

**transmutation of species:** view that species can change and give rise to new species

**triplet code:** identifies that the genetic code consists of triplets or three-base sequences

**tropism:** a directional growth response of a plant to an environmental stimulus

## U

**unconformity:** junction between two rock depositional sequences lying at different orientations

**unloaded form:** refers to the form of co-enzymes, such as NAD and NADP, that can act as electron acceptors

## V

**V(D)J recombination:** a mechanism of random assembly of gene segments that generates the great diversity of unique receptors on B cells and T cells during their early development in the bone marrow

**vaccines:** soluble antigens derived from the causative agents of diseases that are administered to individuals, providing them with protection against those diseases through a process of active artificial immunity

**vector:** in genetics, refers to an agent such as a plasmid or a virus that carries passenger DNA into a cell; in disease, an insect or other animal that carries a pathogenic organism from one host to another; also refers to an insect that carries pollen from one flower to another

**vesicles:** membrane-bound sacs found within a cell, typically fluid filled; one kind of vesicle containing digestive enzymes is known as a lysosome

**vestigial organs:** non-functional organs that are remnants of functional organs in ancestral species

**virion:** the extracellular form of a virus that, in the case of a naked virus, consists of the viral capsid and genetic material and, in the case of an enveloped virus, also includes an outer envelope

**virulence:** degree to which an organism can cause disease

**vomeronasal organ (VNO):** an auxiliary olfactory organ present in many vertebrates that is the site of detection of chemicals such as pheromones

## Z

**zoonotic disease:** a disease for which the pathogenic agent can be transmitted to people from its normal source in another vertebrate species; also termed a zoonosis

**zygomatic arches:** cheekbones comprising horizontal bony ridges on a mammalian skull

# Index

## A

## D

# E



# F

## G



## H

## I

## N



## O



## P

## S