INTRODUCTION TO GENETIC ANALYSIS

Anthony J. F. Griffiths

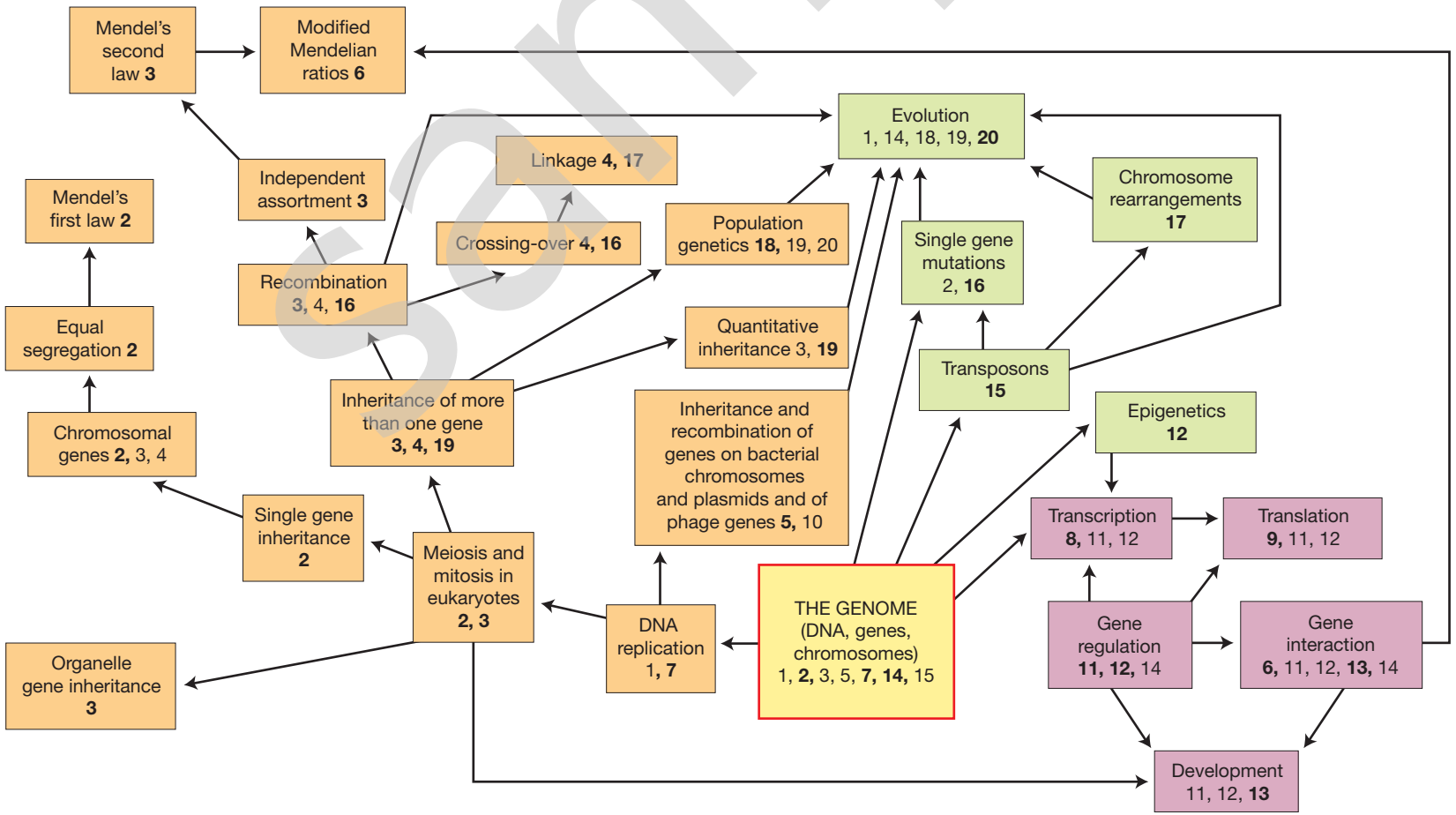
Susan R. Wessler

Sean B. Carroll

John Doebley

ELEVENTH EDITION

A Map of Genetics



The map displays the general divisions of genetics in boxes, with arrows showing the main connections between them covered in this book. Orange, broadly, is inheritance, purple is function, and green is change. Numbers are chapters covering the topic, with main discussions in bold.

INTRODUCTION TO

GENETIC ANALYSIS

About the Authors

[*Barbara Moon.*]

Anthony J. F. Griffiths is a Professor of Botany, Emeritus, at the University of British Columbia. His research focuses on developmental genetics using the model fungus *Neurospora crassa.* He has served as president of the Genetics Society of Canada and two terms as Secretary-General of the International Genetics Federation. He was recently awarded the Fellow Medal of the International Mycological Association.

[*Iqbal Pittawala.*]

Susan R. Wessler is a Distinguished Professor of Genetics in the Department of Botany and Plant Sciences at the University of California, Riverside. Her research focuses on plant transposable elements and their contribution to gene and genome evolution. Dr. Wessler was elected to the National Academy of Sciences in 1998. As a Howard Hughes Medical Institute Professor, she developed and teaches a series of dynamic genome courses in which undergraduates can experience the excitement of scientific discovery.

[*Sean Carroll.*]

Sean B. Carroll is Vice President for Science Education at the Howard Hughes Medical Institute and a Professor of Molecular Biology and Genetics at the University of Wisconsin–Madison. Dr. Carroll is a leader in the field of evolutionary developmental biology and was elected to the National Academy of Sciences in 2007. He is also the author of *Brave Genius, Endless Forms Most Beautiful: The Making of the Fittest,* and *Remarkable Creatures,* a finalist for the National Book Award in Nonfiction in 2009.

[*John Doebley.*]

John Doebley is a Professor of Genetics at the University of Wisconsin–Madison. He studies the genetics of crop domestication using the methods of population and quantitative genetics. He was elected to the National Academy of Sciences in 2003 and served as the president of the American Genetic Association in 2005. He teaches general genetics and evolutionary genetics at the University of Wisconsin.

INTRODUCTION TO GENETIC ANALYSIS

ELEVENTH EDITION

Anthony J. F. Griffiths
University of British Columbia

Susan R. Wessler
University of California, Riverside

Sean B. Carroll
Howard Hughes Medical Institute
University of Wisconsin–Madison

John Doebley
University of Wisconsin–Madison

W. H. Freeman & Company

A Macmillan Education Imprint

Publisher	*Kate Ahr Parker*
Senior Acquisitions Editor	*Lauren Schultz*
Executive Marketing Manager	*John Britch*
Marketing Assistant	*Bailey James*
Developmental Editor	*Erica Champion*
Editorial Assistant	*Alexandra Garrett*
Supplements Editor	*Erica Champion*
Executive Media Editor	*Amanda Dunning*
Media Editors	*Donna Broadman, Erica Champion, Tue Tran*
Art Director	*Diana Blume*
Cover and Interior Designer	*Vicki Tomaselli*
Senior Project Editor	*Jane O'Neill*
Permissions Manager	*Jennifer MacMillan*
Photo Editor	*Richard Fox*
Illustration Coordinator	*Janice Donnola*
Illustrations	*Dragonfly Media Group*
Production Supervisor	*Susan Wein*
Composition and Layout	*Sheridan Sellers*
Printing and Binding	*RR Donnelley*
Cover Photo	*Susan Schmitz/Shutterstock*

Library of Congress Preassigned Control Number: 2014957104

Hardcover: ISBN-13: 978-1-4641-0948-5
ISBN-10: 1-4641-0948-6

Loose-Leaf: ISBN-13: 978-1-4641-8804-6
ISBN-10: 1-4641-8804-1

Printed in the United States of America

First printing

W. H. Freeman and Company
A Macmillan Education Imprint
41 Madison Avenue
New York, NY 10010
Houndmills, Basingstoke RG21 6XS, England

www.macmillanhighered.com

Contents in Brief

Contents

This page intentionally left blank

Preface

Since its first edition in 1974, *Introduction to Genetic Analysis* has emphasized the power and incisiveness of the genetic approach in biological research and its applications. Over its many editions, the text has continuously expanded its coverage as the power of traditional genetic analysis has been extended with the introduction of recombinant DNA technology and then genomics. In the eleventh edition, we continue this tradition and show how the flowering of this powerful type of analysis has been used for insight into research in biology, agriculture, and human health.

Pedagogical Tools

One of the important new features in this edition is the inclusion of lists of **learning outcomes** at the beginning of each chapter. Learning outcomes are crucial components of understanding. One of the tenets of the constructivist theory of learning is that although understanding might be a series of new mental circuits, the learner can never be sure of what is in his or her brain until called upon for some type of performance. Indeed, understanding has even been defined by some as *flexible performance capacity.* The lists of goals show learners what precise performances are expected of them. The notes that follow show how the benefits of the learning outcomes in this book can be maximized for instructors who wish to use them.

LEARNING OUTCOMES

After completing this chapter, you will be able to

- Perform a quantitative analysis of the progeny of a dihybrid testcross to assess whether or not the two genes are linked on the same chromosome.
- Extend the same type of analysis to several loci to produce a map of the relative positions of loci on a chromosome.
- In ascomycete fungi, map the centromeres to other linked loci.
- In asci, predict allele ratios stemming from specific steps in the heteroduplex model of crossing over.

Classroom sessions large and small (for example, lectures and tutorials) should be structured as far as possible on learning outcomes closely paralleling those in these chapters. At various stages in the classes students should be asked to demonstrate their understanding of the material just covered by attaining one or more learning outcomes. In writing examination or test questions, the instructor should try to stick closely to learning outcomes. When reviewing test results, show in what ways the outcomes have been attained or not attained by the learner.

Students should read the list of learning outcomes before embarking on a chapter. Although it will not be possible to understand most of them before reading the chapter, their wording gives a good idea of the lay of the land, and shows the extent of what the instructor's expectations are. Ideally, after reading a section of the chapter, it is a good idea for a student to go back to the list and match the material covered to an outcome. This process should be repeated at the end of the chapter by scanning the sections and making a complete match with each outcome as far as possible. In solving the end-of-chapter problems, try to focus effort on the skills described in the learning outcomes. Students should use the learning outcomes for rapid review when studying for exams; they should try to imagine ways that they will be expected to demonstrate understanding through the application of the outcomes.

The general goal of a course in genetics is to learn how to think and work like a geneticist. The learning outcomes can fractionate this general goal into the many different skills required in this analytical subject.

In this edition we have replaced "Messages" with "**Key Concepts.**" Messages have been in the book since its first edition in 1974. In the 1960s and 1970s, perhaps due to the popularity of Marshall McLuhan's principle "The medium is the message," the word *message* was in common use, and teachers were often asked, "What is your message?" Although with the rise of electronic media it is perhaps time for a resurgence of McLuhan's principle, we felt that the word *message* no longer has the meaning it had in 1974.

New Coverage of Modern Genetic Analysis

One of our goals is to show how identifying genes and their interactions is a powerful tool for understanding biological properties. In the eleventh edition, we present a completely rewritten introductory Chapter 1, with a focus on modern applications of genetics. From there, the student follows the process of a traditional genetic dissection, starting with a step-by-step coverage of single-gene identification in Chapter 2, gene mapping in Chapter 4, and identifying pathways and networks by studying gene interactions in Chapter 6. New genomic approaches to identifying and locating genes are explored in Chapters 10, 14, and 19.

Flood-intolerant and flood-tolerant rice

FIGURE 1-20 An Indian farmer with rice variety *Swarna* that is not tolerant to flooding (*left*) compared to variety *Swarna-sub1* that is tolerant (*right*). This field was flooded for 10 days. The photo was taken 27 days after the flood waters receded. [Ismail et al., "The contribution of submergence-tolerant (Sub 1) rice varieties to food security in flood-prone rainfed lowland areas in Asia," Field Crops Research 152, 2013, 83–93, © Elsevier.]

SUB1 gene increases rice yield under flooding

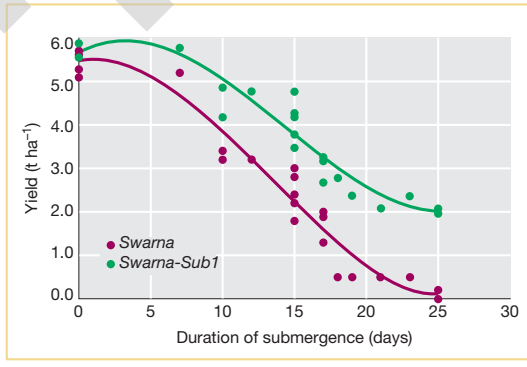


FIGURE 1-21 Yield comparison between variety *Swarna* that is not tolerant to flooding (purple circles) and variety *Swarna-Sub1* that is tolerant (green circles). Yield in tons per hectare (y-axis) versus duration of flooding in days (x-axis). [Data from Ismail et al., "The contribution of submergence-tolerant (Sub 1) rice varieties to food security in flood-prone rainfed lowland areas in Asia," Field Crops Research 152, 2013, 83–93.]

- A reconceptualized Chapter 1 now piques student interest in genetics by presenting a selection of modern applications in biology, evolution, medicine, and agriculture. After a brief history of the study of genetics and a review of some fundamentals, the chapter describes four stories of how genetics is used today.
- Classical genetic dissection is given a more gradual introduction in Chapters 2 and 4. Chapter 2 begins with a new introduction to forward genetics and the role of genetic analysis in identifying traits of single-gene inheritance. Crosses are depicted visually as well as mathematically. The concepts of dominance and recessiveness are explained in terms of haplosufficiency and haploinsufficiency. The use of chi-square analysis in Chapter 4 has been rewritten for clarity.
- The modern application of genetics introduced in Chapter 1 continues in Chapter 14 by applying new genomic techniques such as RNA-seq and exome sequencing, which are introduced to solve problems in medicine. The search for meaning in noncoding segments of the genome is an important frontier in genomics, and the ENCODE project has been added to this chapter to represent that search.

Focus on Key Advances in Genetics

We have enhanced coverage of several cutting-edge topics in the eleventh edition.

Chromatin remodeling and epigenetics: Previously spread among several chapters, the flourishing field of epigenetics is now consolidated and completely updated in Chapter 12. In section 12.3, "Dynamic Chromatin," we discuss the three major mechanisms of altering chromatin structure: chromatin remodeling, histone modification, and histone variants. Changes throughout this section provide more detail and clarity, based on recent advances in the field.

Genome surveillance: Cutting-edge research in transposable elements has uncovered genome surveillance systems in plants, animals, and bacteria similar to that previously identified in *C. elegans*. Chapter 15 now provides an overview of piRNAs in animals and crRNAs in bacteria, and allows students to compare and contrast those approaches to Tc1 elements in worms and MITEs in plants.

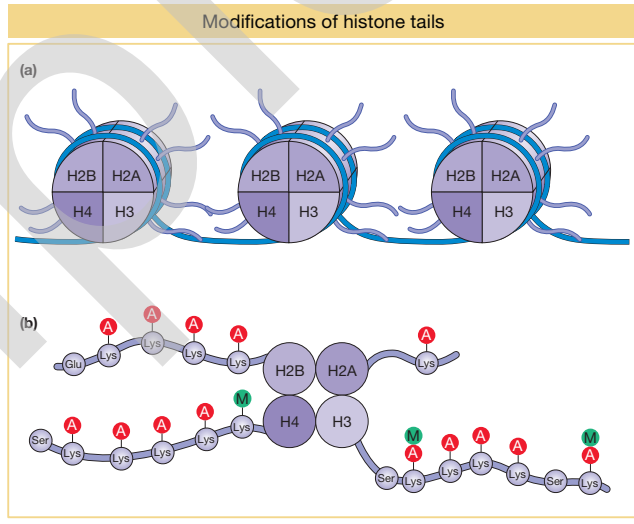


FIGURE 12-13 (a) Histone tails protrude from the nucleosome core (purple). (b) Examples of histone tail modifications are shown. Circles with A represent acetylation while circles with M represent methylation. See text for details.

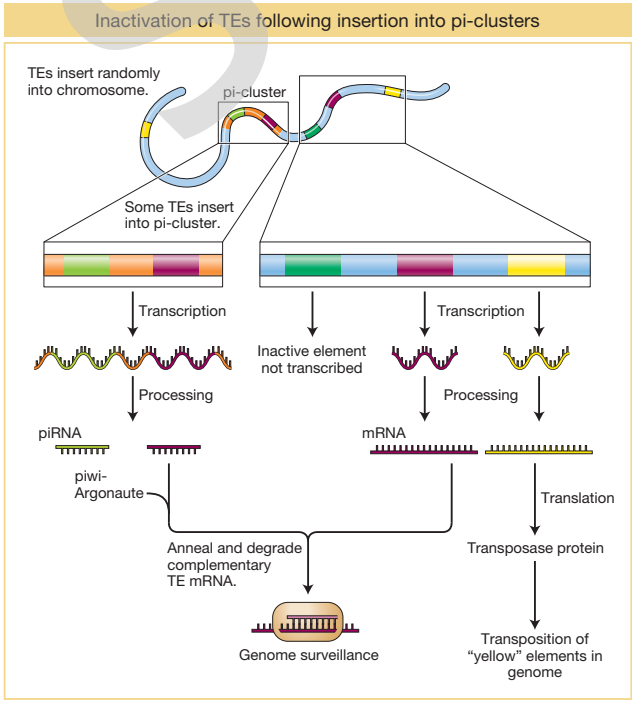


FIGURE 15-27 Insertion of the green and pink transposons into a pi-cluster in the genome results in the degradation of transcripts from these two transposons by the steps shown and described in the text. In contrast, the yellow transposon will remain active until copies insert by chance into a pi-cluster.

Enduring Features

Coverage of model organisms

The eleventh edition retains the enhanced coverage of model systems in formats that are practical and flexible for both students and instructors.

- Chapter 1 introduces some key genetic model organisms and highlights some of the successes achieved through their use.
- Model Organism boxes presented in context where appropriate provide additional information about the organism in nature and its use experimentally.
- A Brief Guide to Model Organisms, at the back of the book, provides quick access to essential, practical information about the uses of specific model organisms in research studies.
- An Index to Model Organisms, on the endpapers at the back of the book, provides chapter-by-chapter page references to discussions of specific organisms in the text, enabling instructors and students to easily find and assemble comparative information across organisms.

Problem sets

No matter how clear the exposition, deep understanding requires the student to personally engage with the material. Hence our efforts to encourage student problem solving. Building on its focus on genetic analysis, the eleventh edition provides students with opportunities to practice problem-solving skills—both in the text and online through the following features.

- **Versatile Problem Sets.** Problems span the full range of degrees of difficulty. They are categorized according to level of difficulty—basic or challenging.
- **Working with the Figures.** An innovative set of problems included at the back of each chapter asks students pointed questions about figures in the chapter. These questions encourage students to think about the figures and help them to assess their understanding of key concepts.
- **Solved Problems.** Found at the end of each chapter, these worked examples illustrate how geneticists apply principles to experimental data.
- **Unpacking the Problems.** A genetics problem draws on a complex matrix of concepts and information. "Unpacking the Problem" helps students learn to approach problem solving strategically, one step at a time, concept on concept.
- **NEW** Macmillan Education LaunchPad Multiple-choice versions of the end-of-chapter problems are available on our online LaunchPad for quick gradable quizzing and easily gradable homework assignments. The **Unpacking the Problem tutorials** from the text have been converted to in-depth online tutorials and expanded to help students learn to solve problems and think like a geneticist. New videos demonstrate how to solve selected difficult problems.

How genetics is practiced today

A feature called "What Geneticists Are Doing Today" suggests how genetic techniques are being used today to answer specific biological questions, such as "What is the link between telomere shortening and aging?" or "How can we find missing components in a specific biological pathway?"

Media and Supplements

LaunchPad

The *LaunchPad* is a dynamic, fully integrated learning environment that brings together all the teaching and learning resources in one place. It features the fully interactive e-Book, end-of-chapter practice problems now assignable as homework, animations, and tutorials to help students with difficult-to-visualize concepts.

This learning system also includes easy-to-use, powerful assessment tracking and grading tools, a personalized calendar, an announcement center, and communication tools all in one place to help you manage your course. Some examples:

- **Hundreds of self-graded end-of-chapter problems** allow students to practice their problem-solving skills. Most of the open-ended end-of-chapter questions have been carefully rewritten to create high-quality, analytical multiple-choice versions for assigning.
- **Animations** help students visualize genetics.
- **Unpacking the Problem tutorials** from the text have been converted and expanded to help students learn to solve problems and think like a geneticist. These in-depth online tutorials guide students toward the solution, offering guidance as needed via hints and detailed feedback.
- **NEW Problem-solving videos** walk students through solving difficult problems from the text.

Teaching resources for instructors

Electronic teaching resources are available online at the LaunchPad, at **http://www.whfreeman.com/launchpad/iga11e**

Includes all the electronic resources listed below for teachers. Contact your W. H. Freeman sales representative to learn how to log on as an instructor.

e-Book

The e-Book fully integrates the text and its interactive media in a format that features a variety of helpful study tools (full-text, Google-style searching; note taking; bookmarking; highlighting; and more). Available as a stand-alone item or on the LaunchPad.

Clicker Questions

Jump-start discussions, illuminate important points, and promote better conceptual understanding during lectures.

Layered PowerPoint Presentations

Illuminate challenging topics for students by deconstructing intricate genetic concepts, sequences, and processes step-by-step in a visual format.

All Images from the Text

More than 500 illustrations can be downloaded as JPEGs and PowerPoint slides. Use high-resolution images with enlarged labels to project clearly for lecture hall presentations. Additionally, these JPEG and PowerPoint files are available without labels for easy customization in PowerPoint.

67 Continuous-Play Animations

A comprehensive set of animations, updated and expanded for the eleventh edition, covers everything from basic molecular genetic events and lab techniques to analyzing crosses and genetic pathways. The complete list of animations appears on page xix.

Assessment Bank
This resource brings together a wide selection of genetics problems for use in testing, homework assignments, or in-class activities. Searchable by topic and provided in MS Word format, as well as in LaunchPad and Diploma, the assessment bank offers a high level of flexibility.

Student Solutions Manual
(ISBN: 1-4641-8794-0)

The Student Solutions Manual contains complete worked-out solutions to all the problems in the textbook, including the "Unpacking the Problem" exercises. Available on the LaunchPad and the Instructor's Web site as easy-to-print Word files.

Understanding Genetics: Strategies for Teachers and Learners in Universities and High Schools
(ISBN: 0-7167-5216-6)

Written by Anthony Griffiths and Jolie-Mayer Smith, this collection of articles focuses on problem solving and describes methods for helping students improve their ability to process and integrate new information.

Resources for students

at **http://www.whfreeman.com/launchpad/iga11e**

LaunchPad 6-month Access Card (ISBN: 1-4641-8793-2)

The LaunchPad contains the following resources for students:

- *Self-Graded End-of-Chapter Problems:* To allow students to practice their problem-solving skills, most of the open-ended end-of-chapter questions have been carefully rewritten to create high-quality, analytical multiple-choice versions for assigning.
- *Online Practice Tests:* Students can test their understanding and receive immediate feedback by answering online questions that cover the core concepts in each chapter. Questions are page referenced to the text for easy review of the material.
- *Animations:* A comprehensive set of animations, updated and expanded for the eleventh edition, covers everything from basic molecular genetic events and lab techniques to analyzing crosses and genetic pathways. The complete list of animations appears on the facing page.
- *Interactive "Unpacking the Problem":* An exercise from the problem set for many chapters is available online in interactive form. As with the text version, each Web-based "Unpacking the Problem" uses a series of questions to step students through the thought processes needed to solve a problem. The online version offers immediate feedback to students as they work through the problems as well as convenient tracking and grading functions. Authored by Craig Berezowsky, University of British Columbia.
- **NEW** *Problem-Solving Videos:* Twenty-five problem-solving videos walk students through solving difficult problems from the text.

Student Solutions Manual (ISBN: 1-4641-8794-0)
The Solutions Manual contains complete worked-out solutions to all the problems in the textbook, including the "Unpacking the Problem" exercises. Used in conjunction with the text, this manual is one of the best ways to develop a fuller appreciation of genetic principles.

Other genomic and bioinformatic resources for students:
Text Appendix A, *Genetic Nomenclature,* lists model organisms and their nomenclature.

Text Appendix B, *Bioinformatic Resources for Genetics and Genomics,* builds on the theme of introducing students to the latest genetic research tools by providing students with some valuable starting points for exploring the rapidly expanding universe of online resources for genetics and genomics.

Animations

Sixty-seven animations are fully integrated with the content and figures in the text chapters. These animations are available on the LaunchPad and the Book Companion site.

CHAPTER 1
A Basic Plant Cross (Figure 1-3)
The Central Dogma (Figure 1-10)

CHAPTER 2
Mitosis (Chapter Appendix 2-1)
Meiosis (Chapter Appendix 2-2)
X-Linked Inheritance in Flies (Figure 2-17)

CHAPTER 3
Punnett Square and Branch Diagram Methods for Predicting the Outcomes of Crosses (Figure 3-4)
Meiotic Recombination Between Unlinked Genes by Independent Assortment (Figures 3-8 and 3-13)
Analyzing a Cross: A Solved Problem (Solved Problem 2)

CHAPTER 4
Crossing Over Produces New Allelic Combinations (Figures 4-2 and 4-3)
Meiotic Recombination Between Linked Genes by Crossing Over (Figure 4-7)
A Molecular Model of Crossing Over (Figure 4-21)
A Mechanism of Crossing Over: A Heteroduplex Model (Figure 4-21)
A Mechanism of Crossing Over: Genetic Consequences of the Heteroduplex Model
Mapping a Three-Point Cross: A Solved Problem (Solved Problem 2)

CHAPTER 5
Bacterial Conjugation and Mapping by Recombination (Figures 5-11 and 5-17)

CHAPTER 6
Interactions Between Alleles at the Molecular Level, *RR:* Wild-Type
Interactions Between Alleles at the Molecular Level, *rr:* Homozygous Recessive, Null Mutation
Interactions Between Alleles at the Molecular Level, *r′r′*: Homozygous Recessive, Leaky Mutation
Interactions Between Alleles at the Molecular Level, *Rr:* Heterozygous, Complete Dominance
Screening and Selecting for Mutations
A Model for Synthetic Lethality (Figure 6-20)

CHAPTER 7
DNA Replication: The Nucleotide Polymerization Process (Figure 7-15)
DNA Replication: Coordination of Leading and Lagging Strand Synthesis (Figure 7-20)
DNA Replication: Replication of a Chromosome (Figure 7-23)

CHAPTER 8
Transcription in Prokaryotes (Figures 8-7 to 8-10)
Transcription in Eukaryotes (Figures 8-12 and 8-13)
Mechanism of RNA Splicing (Figures 8-16 and 8-17)

CHAPTER 9
Peptide-Bond Formation (Figure 9-2)
tRNA Charging (Figure 9-7)
Translation (Figure 9-14 to 9-16)
Nonsense Suppression at the Molecular Level: The *rod*^{ns} Nonsense Mutation (Figure 9-18)
Nonsense Suppression at the Molecular Level: The tRNA Nonsense Suppressor (Figure 9-18)
Nonsense Suppression at the Molecular Level: Nonsense Suppression of the *rod*^{ns} Allele (Figure 9-18)

CHAPTER 10
Polymerase Chain Reaction (Figure 10-3)
Plasmid Cloning (Figure 10-9)
Finding Specific Cloned Genes by Functional Complementation: Functional Complementation of the Gal$^-$ Yeast Strain and Recovery of the Wild-Type *GAL* gene
Finding Specific Cloned Genes by Functional Complementation: Making a Library of Wild-Type Yeast DNA
Finding Specific Cloned Genes by Functional Complementation: Using the Cloned *GAL* Gene as a Probe for *GAL* mRNA
SDS Gel Electrophoresis and Immunoblotting
Dideoxy Sequencing of DNA (Figure 10-17)
Creating a Transgenic Mouse (Figures 10-29 and 10-30)

CHAPTER 11
Regulation of the Lactose System in *E. coli:* Assaying Lactose Presence/Absence Through the Lac Repressor (Figure 11-6)
Regulation of the Lactose System in *E. coli:* O^C *lac* Operator Mutations (Figure 11-8)
Regulation of the Lactose System in *E. coli:* I$^-$ Lac Repressor Mutations (Figure 11-9)
Regulation of the Lactose System in *E. coli:* I^S Lac Superrepressor Mutations (Figure 11-10)

CHAPTER 12
Three-Dimensional Structure of Nuclear Chromosomes (Figure 12-11)
Gal4 Binding and Activation (Figures 12-6 through 12-9)
Chromatin Remodeling (Figures 12-13 and 12-14)

CHAPTER 13
Drosophila Embryonic Development
Sex Determination in Flies (Figure 13-23)

CHAPTER 14
DNA Microarrays: Using an Oligonucleotide Array to Analyze Patterns of Gene Expression (Figure 14-20)
DNA Microarrays: Synthesizing an Oligonucleotide Array
Yeast Two-Hybrid Systems (Figure 14-21)

CHAPTER 15
Replicative Transposition (Figure 15-9)
Life Cycle of a Retrovirus (Figure 15-11)
The *Ty1* Mechanism of Retrotransposition (Figures 15-13 and 15-14)

CHAPTER 16
Replication Slippage Creates Insertion or Deletion Mutations (Figure 16-8)
UV-Induced Photodimers and Excision Repair (Figure 16-19)
Base-Excision Repair, Nucleotide Excision Repair, and Mismatch Repair (Figures 16-20, 16-22, and 16-23)

CHAPTER 17
Autotetraploid Meiosis (Figure 17-6)
Meiotic Nondisjunction at Meiosis I (Figure 17-12)
Meiotic Nondisjunction at Meiosis II (Figure 17-12)
Chromosome Rearrangements: Paracentric Inversion, Formation of Paracentric Inversions (Figure 17-27)
Chromosome Rearrangements: Paracentric Inversion, Meiotic Behavior of Paracentric Inversions (Figure 17-28)
Chromosome Rearrangements: Reciprocal Translocation, Formation of Reciprocal Translocations (Figure 17-30)
Chromosome Rearrangements: Reciprocal Translocation, Meiotic Behavior of Reciprocal Translocations (Figure 17-30)
Chromosome Rearrangements: Reciprocal Translocation, Pseudolinkage of Genes by Reciprocal Translocations (Figure 17-32)

Acknowledgments

We extend our thanks and gratitude to our colleagues who reviewed this edition and whose insights and advice were most helpful:

Anna Allen, *Howard University*
Melissa Antonio, *California Baptist University*
Dave Bachoon, *Georgia College & State University*
Brianne Barker, *Drew University*
Lina Begdache, *Binghamton University*
Edward Berger, *Dartmouth College*
Aimee Bernard, *University of Colorado Denver*
Jaime Blair, *Franklin & Marshall College*
Jay Brewster, *Pepperdine University*
Doug Broadfield, *Florida Atlantic University*
Mirjana Brockett, *Georgia Institute of Technology*
Judy Brusslan, *California State University, Long Beach*
Gerald Buldak, *Loyola University Chicago*
Aaron Cassill, *University of Texas at San Antonio*
Helen Chamberlin, *Ohio State University*
Henry Chang, *Purdue University*
Randolph Christensen, *Coe College*
Mary Clancy, *University of New Orleans*
Craig Coleman, *Brigham Young University*
Matthew Collier, *Wittenberg University*
Shannon Compton, *University of Massachusetts-Amherst*
Diane Cook, *Louisburg College*
Victoria Corbin, *University of Kansas*
Claudette Davis, *George Mason University*
Ann Marie Davison, *Kwantlen Polytechnic University*
Elizabeth De Stasio, *Lawrence University*
Matt Dean, *University of Southern California*
Michael Dohm, *Chaminade University*
Robert Dotson, *Tulane University*
Chunguang Du, *Montclair State University*
Erastus Dudley, *Huntingdon College*
Edward Eivers, *California State University, Los Angeles*
Robert Farrell, *Penn State University*
David Foltz, *Louisiana State University*
Wayne Forrester, *Indiana University*
Rachael French, *San Jose State University*

Shirlean Goodwin, *University of Memphis*
Topher Gee, *UNC Charlotte*
John Graham, *Berry College*
Theresa Grana, *University of Mary Washington*
Janet Guedon, *Duquesne University*
Patrick Gulick, *Concordia University*
Richard Heineman, *Kutztown University*
Anna Hicks, *Memorial University*
Susan Hoffman, *Miami University*
Stanton Hoegerman, *College of William and Mary*
Margaret Hollingsworth, *University at Buffalo*
Nancy Huang, *Colorado College*
Jeffrey Hughes, *Millikin University*
Varuni Jamburuthugoda, *Fordham University*
Pablo Jenik, *Franklin & Marshall College*
Aaron Johnson, *University of Colorado School of Medicine*
Anil Kapoor, *University of La Verne*
Jim Karagiannis, *University of Western Ontario*
Kathleen Karrer, *Marquette University*
Jessica Kaufman, *Endicott College*
Darrell Killian, *Colorado College*
Dennis Kraichely, *Cabrini College*
Anuj Kumar, *University of Michigan*
Janice Lai, *Austin Community College*
Evan Lau, *West Liberty University*
Min-Ken Liao, *Furman University*
Sarah Lijegren, *University of Mississippi*
Renyi Liu, *University of California, Riverside*
Diego Loayza, *Hunter College*
James Lodolce, *Loyola University Chicago*
Joshua Loomis, *Nova Southeastern University*
Amy Lyndaker, *Ithaca College*
Jessica Malisch, *Claremont McKenna College*
Patrick Martin, *North Carolina A&T State University*
Presley Martin, *Hamline University*
Dmitri Maslov, *University of California, Riverside*
Maria Julia Massimelli, *Claremont McKenna College*
Endre Mathe, *Vasile Goldis Western University of Arad*
Herman Mays, *University of Cincinnati*
Thomas McGuire, *Penn State Abington*
Mark Meade, *Jacksonville State University*
Ulrich Melcher, *Oklahoma State University*
Philip Meneely, *Haverford College*
Ron Michaelis, *Rutgers University*
Chris Mignone, *Berry College*
Sarah Mordan-McCombs, *Franklin College of Indiana*
Ann Murkowski, *North Seattle Community College*
Saraswathy Nair, *University of Texas at Brownsville*
Sang-Chul Nam, *Texas A&M International University*
Scot Nelson, *University of Hawaii at Manoa*
Brian Nichols, *University of Illinois at Chicago*
Todd Nickle, *Mount Royal University*
Juliet Noor, *Duke University*
Mohamed Noor, *Duke University*
Daniel Odom, *California State University, Northridge*
Kirk Olsen, *East Los Angeles College*
Kavita Oommen, *Georgia State University*
Maria Orive, *University of Kansas*
Laurie Pacarynuk, *University of Lethbridge*
Patricia Phelps, *Austin Community College*
Martin Poenie, *University of Texas at Austin*
Jennifer Powell, *Gettysburg College*
Robyn Puffenbarger, *Bridgewater College*
Jason Rauceo, *John Jay College (CUNY)*
Eugenia Ribiero-Hurley, *Fordham University*
Ronda Rolfes, *Georgetown University*
Edmund Rucker, *University of Kentucky*
Jeffrey Sands, *Lehigh University*
Monica Sauer, *University of Toronto at Scarborough, UTSC*
Ken Saville, *Albion College*
Pratibha Saxena, *University of Texas at Austin*
Jon Schnorr, *Pacific University*
Malcolm Schug, *University of North Carolina at Greensboro*
Deborah Schulman, *Lake Erie College*
Allan Showalter, *Ohio University*
Elaine Sia, *University of Rochester*
Robert Smith, *Nova Southeastern University*
Joyce Stamm, *University of Evansville*
Tara Stoulig, *Southeastern Louisiana University*
Julie Torruellas Garcia, *Nova Southeastern University*
Virginia Vandergon, *California State University,* Northridge
Charles Vigue, *University of New Haven*
Susan Walsh, *Rollins College*
Michael Watters, *Valparaiso University*
Roger Wartell, *Georgia Institute of Technology*
Matthew White, *Ohio University*
Dwayne Wise, *Mississippi State University*
Andrew Wood, *Southern Illinois University*
Mary Alice Yund, *UC Berkeley Extension*
Malcom Zellars, *Georgia State University*
Deborah Zies, *University of Mary Washington*

Tony Griffiths would like to acknowledge the pedagogical insights of David Suzuki, who was a co-author of the early editions of this book, and whose teaching in the media is now an inspiration to the general public around the world. Great credit is also due to Jolie Mayer-Smith and Barbara Moon, who introduced Tony to the power of the constructivist approach applied to teaching genetics. Sean Carroll would like to thank Leanne Olds for help with the artwork for Chapters 11, 12, 13, 14, and 20. John Doebley would like to thank his University of Wisconsin colleagues Bill Engels, Carter Denniston, and Jim Crow, who shaped his approach to teaching genetics.

The authors also thank the team at W. H. Freeman for their hard work and patience. In particular we thank our developmental and supplements editor, Erica Champion; senior acquisitions editor Lauren Schultz; senior project editor Jane O'Neill; and copy editor Teresa Wilson. We also thank Susan Wein, production supervisor; Diana Blume, art director; Vicki Tomaselli, cover and text designer; Sheridan Sellers, page layout; Janice Donnola, illustration coordinator; Jennifer MacMillan, permissions manager; Amanda Dunning, executive media editor; and Alexandra Garrett, editorial assistant. Finally, we especially appreciate the marketing and sales efforts of John Britch, executive marketing manager, and the entire sales force.

This page intentionally left blank

CHAPTER 1

The Genetics Revolution

DNA (deoxyribonucleic acid) is the molecule that encodes genetic information. The strings of four different chemical bases in DNA store genetic information in much the same way that strings of 0's and 1's store information in computer code. [*Sergey Nivens/Shutterstock.*]

OUTLINE

LEARNING OUTCOMES

After completing this chapter, you will be able to

- Describe the way in which modern genetics developed.
- List the main cellular constituents involved in gene expression and action.
- Give some examples of how genetics has influenced modern medicine, agriculture, and evolution.

Genetics is a form of information science. Geneticists seek to understand the rules that govern the transmission of genetic information at three levels—from parent to offspring within families, from DNA to gene action within and between cells, and over many generations within populations of organisms. These three foci of genetics are known as transmission genetics, molecular-developmental genetics, and population-evolutionary genetics. The three parts of this text examine these three foci of genetics.

The science of genetics was born just over 100 years ago. Since that time, genetics has profoundly changed our understanding of life, from the level of the individual cell to that of a population of organisms evolving over millions of years. In 1900, William Bateson, a prominent British biologist, wrote presciently that an "exact determination of the laws of heredity will probably work more change in man's outlook on the world, and in his power over nature, than any other advance in natural knowledge that can be foreseen." Throughout this text, you will see the realization of Bateson's prediction. Genetics has driven a revolution in both the biological sciences and society in general.

In this first chapter, we will look back briefly at the history of genetics, and in doing so, we will review some of the basic concepts of genetics that were discovered over the last 100 years. After that, we will look at a few examples of how genetic analysis is being applied to critical problems in biology, agriculture, and human health today. You will see how contemporary research in genetics integrates concepts discovered decades ago with recent technological advances. You will see that genetics today is a dynamic field of investigation in which new discoveries are continually advancing our understanding of the biological world.

1.1 The Birth of Genetics

Like begets like

FIGURE 1-1 Family groups in the gray wolf show familial resemblances for coat colors and patterning. [*(Top) altrendo nature/Getty Images; (bottom) Bev McConnell/Getty Images.*]

Throughout recorded history, people around the world have understood that "like begets like." Children resemble their parents, the seed from a tree bearing flavorful fruit will in turn grow into a tree laden with flavorful fruit, and even members of wolf packs show familial resemblances (Figure 1-1). Although people were confident in these observations, they were left to wonder as to the underlying mechanism. The Native American Hopi tribe of the Southwestern United States understood that if they planted a red kernel of maize in their fields, it would grow into a plant that also gave red kernels. The same was true for blue, white, or yellow kernels. So they thought of the kernel as a message to the gods in the Earth about the type of maize the Hopi farmers hoped to harvest. Upon receiving this message, the gods would faithfully return them a plant that produced kernels of the desired color.

In the 1800s in Europe, horticulturalists, animal breeders, and biologists also sought to explain the resemblance between parents and offspring. A commonly held view at that time was the **blending theory** of inheritance, or the belief that inheritance worked like the mixing of fluids such as paints. Red and white paints, when mixed, give pink; and so a child of one tall parent and one short parent could be expected to grow to a middling height. While blending theory seemed to work at times, it was also clear that there were exceptions, such as tall children born to parents of average height. Blending theory also provided no mechanism by which the "heredity fluids" it imagined, once mixed, could be separated—the red and white paints cannot be reconstituted from the pink. Thus, the long-term expectation of blending theory over many generations of intermating among individuals is that all members of the population will come to express the same average value of a trait. Clearly, this is not how nature works. Human populations have people with a range of

heights, from short to tall, and we have not all narrowed in on a single average height despite the many generations that human populations have dwelled on Earth.

Gregor Mendel—A monk in the garden

While the merits and failings of blending theory were being debated, Gregor Mendel, an Austrian monk, was working to understand the rules that govern the transmission of traits from parent to offspring after hybridization among different varieties of pea plants (Figure 1-2). The setting for his work was the monastery garden in the town of Brünn, Austria (Brno, Czech Republic, today). From 1856 to 1863, Mendel cross-pollinated or intermated different varieties of the pea plant. One of his experiments involved crossing a pea variety with purple flowers to one with white flowers (Figure 1-3). Mendel recorded that the first hybrid generation

Gregor Mendel

FIGURE 1-2 Gregor Mendel was an Austrian monk who discovered the laws of inheritance. [*James King-Holmes/Science Source.*]

One of Mendel's experiments

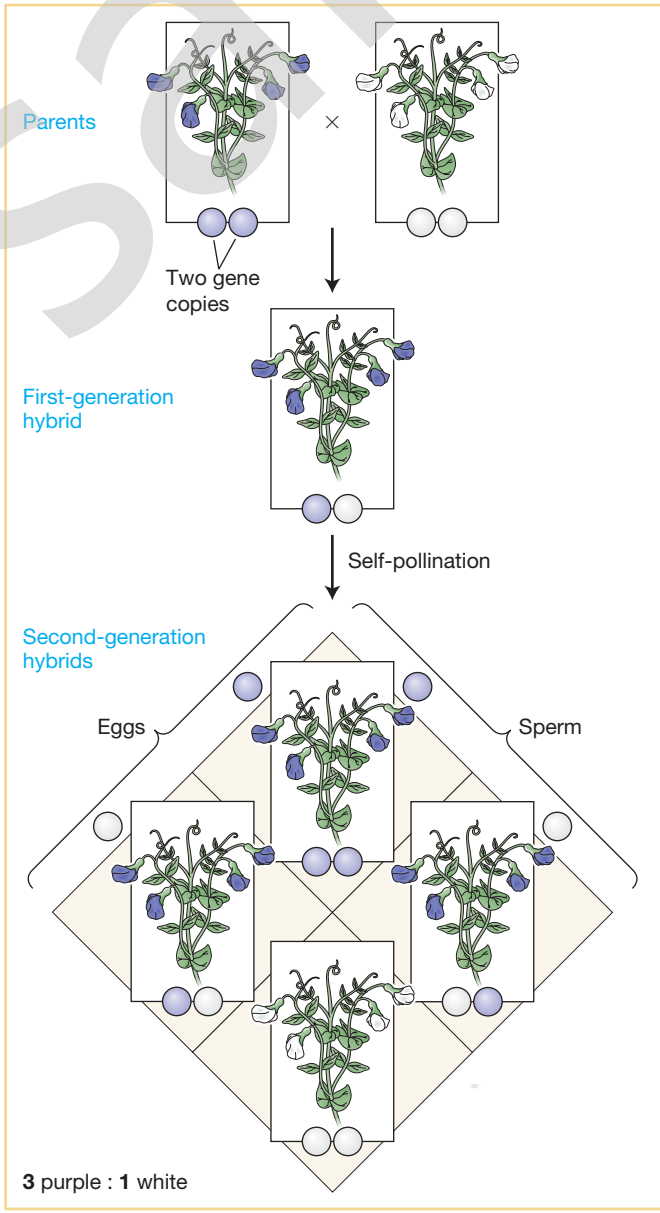


FIGURE 1-3 The mating scheme for Mendel's experiment involving the crossing of purple- and white-flowered varieties of pea plants. The purple and white circles signify the gene variants for purple vs. white flower color. Gametes carry one gene copy; the plants each carry two gene copies. The "×" signifies a cross-pollination between the purple- and white-flowered plants.

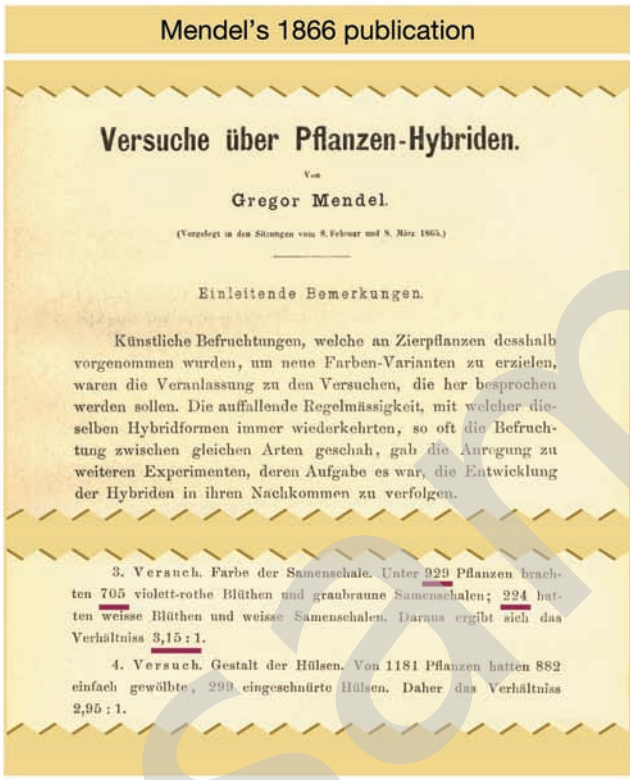


FIGURE 1-4 Excerpts from Mendel's 1866 publication, *Versuche über Pflanzen-Hybriden* (Experiments on plant hybrids). [*Augustinian Abbey in Old Brno, Courtesy of the Masaryk University, Mendel Museum.*]

of offspring from this cross all had purple flowers, just like one of the parents. There was no blending. Then, Mendel self-pollinated the first-generation hybrid plants and grew a second generation of offspring. Among the progeny, he saw plants with purple flowers as well as plants with white flowers. Of the 929 plants, he recorded 705 with purple flowers and 224 with white flowers (Figure 1-4). He observed that there were roughly 3 purple-flowered plants for every 1 white-flowered plant.

How did Mendel explain his results? Clearly, blending theory would not work since that theory predicts a uniform group of first-generation hybrid plants with light purple flowers. So Mendel proposed that the factors that control traits act like *particles* rather than fluids and that these particles do not blend together but are passed intact from one generation to the next. Today, Mendel's particles are known as **genes.**

Mendel proposed that each individual pea plant has two copies of the gene controlling flower color in each of the cells of the plant body (**somatic cells**). However, when the plant forms sex cells, or **gametes** (eggs and sperm), only one copy of the gene enters into these reproductive cells (see Figure 1-3). Then, when egg and sperm unite to start a new individual, once again there will be two copies of the flower color gene in each cell of the plant body.

Mendel had some further insights. He proposed that the gene for flower color comes in two gene variants, or **alleles**—one that conditions purple flowers and one that conditions white flowers. He proposed that the purple allele of the flower color gene is **dominant** to the white allele such that a plant with one purple allele and one white allele would have purple flowers. Only plants with two white alleles would have white flowers (see Figure 1-3). Mendel's two conclusions, (1) that genes behaved like particles that do not blend together and (2) that one allele is dominant to the other, enabled him to explain the lack of blending in the first-generation hybrids and the reappearance of white-flowered plants in the second-generation hybrids with a 3:1 ratio of purple- to white-flowered plants. This revolutionary advance in our understanding of inheritance will be fully discussed in Chapter 2.

How did Mendel get it right when so many others before him were wrong? Mendel chose a good organism and good traits to study. The traits he studied were all controlled by single genes. Traits that are controlled by several genes, as many traits are, would not have allowed him to discover the laws of inheritance so easily. Mendel was also a careful observer, and he kept detailed records of each of his experiments. Finally, Mendel was a creative thinker capable of reasoning well beyond the ideas of his times.

Mendel's particulate theory of inheritance was published in 1866 in the *Proceedings of the Natural History Society of Brünn* (see Figure 1-4). At that time, his work was noticed and read by some other biologists, but its implications and importance went unappreciated for over 30 years. Unlike Charles Darwin, whose discovery of the theory of evolution by natural selection made him world-renowned virtually overnight, when Mendel died in 1884, he was more or less unknown in the world of science. As biochemist Erwin Chargaff put it, "There are people who seem to be born in a vanishing cap. Mendel was one of them."

KEY CONCEPT Gregor Mendel demonstrated that genes behave like particles and not fluids.

Mendel rediscovered

As the legend goes, when the British biologist William Bateson (Figure 1-5) boarded a train bound for a conference in London in 1900, he had no idea how profoundly his world would change during the brief journey. Bateson carried with him a copy of Mendel's 1866 paper on the hybridization of plant varieties. Bateson had recently learned that biologists in Germany, the Netherlands, and Austria had each independently reproduced Mendel's 3:1 ratio, and they each cited Mendel's original work. This trio had rediscovered Mendel's laws of inheritance. Bateson needed to read Mendel's paper. By the time he stepped off the train, Bateson had a new mission in life. He understood that the mystery of inheritance had been solved. He soon became a relentless apostle of Mendel's laws of inheritance. A few years later in 1905, Bateson coined the term **genetics**—the study of inheritance. The genetics revolution had begun.

When Mendel's laws of inheritance were rediscovered in 1900, a flood of new thinking and ideas was unleashed. Mendelism became the organizing principle for much of biology. There were many new questions to be asked about inheritance. Table 1-1 summarizes the chronology of seminal discoveries made over the coming decades and the chapters of this text that cover each of these topics. Let's look briefly at a few of the questions and their answers that transformed the biological sciences.

Where in the cell are Mendel's genes? The answer came in 1910, when Thomas H. Morgan at Columbia University in New York demonstrated that Mendel's genes are located on chromosomes—he proved the **chromosome theory** of inheritance. The idea was not new. Walter Sutton, who was raised on a farm in Kansas and later served as a surgeon for the U.S. army during WWI had proposed the chromosome theory of inheritance in 1903. Theodor Boveri, a German biologist, independently proposed it at the same time. It was a compelling hypothesis, but there were no experimental data to support it. This changed in 1910, when Morgan proved the chromosome theory of inheritance using Mendelian genetics and the fruit fly as his experimental organism. In Chapter 4, you will retrace Morgan's experiments that proved genes are on chromosomes.

Can Mendelian genes explain the inheritance of continuously variable traits like human height? While 3:1 segregation ratios could be directly observed for simple traits like flower color, many traits show a continuous range of values in second-generation hybrids without simple ratios like 3:1. In 1918, Ronald Fisher, the British statistician and geneticist, resolved how Mendelian genes explained the inheritance of continuously variable traits like height in people (Figure 1-6). Fisher's core idea

William Bateson gave genetics its name

FIGURE 1-5 William Bateson, the British zoologist and evolutionist who introduced the term *genetics* for the study of inheritance and promoted Mendel's work. [*SPL/Science Source.*]

Continuous variation for height

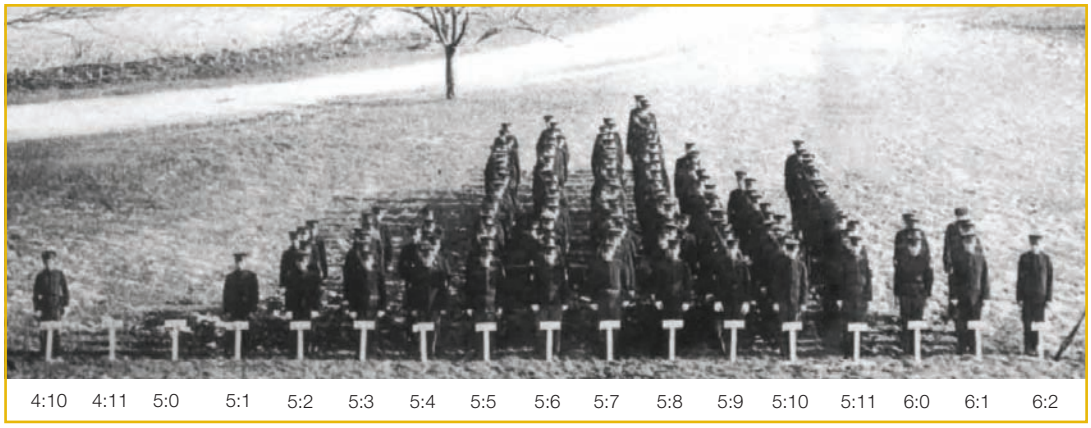


FIGURE 1-6 Students at the Connecticut Agriculture College in 1914 show a range of heights. Ronald Fisher proposed that continuously variable traits like human height are controlled by multiple Mendelian genes. [*A. F. Blakeslee, "Corn and Men," Journal of Heredity 5, 11, 1914, 511–518.*]

TABLE 1-1 Key Events in the History of Genetics

Year	Event	Chapters
1865	Gregor Mendel showed that traits are controlled by discrete factors now known as genes.	2, 3
1869	Friedrich Miescher isolated DNA from the nuclei of white blood cells.	7
1903	Walter Sutton and Theodor Boveri hypothesized that chromosomes are the hereditary elements.	4
1905	William Bateson introduced the term “genetics” for the study of inheritance.	2
1908	G. H. Hardy and Wilhelm Weinberg proposed the Hardy–Weinberg law, the foundation for population genetics.	18
1910	Thomas H. Morgan demonstrated that genes are located on chromosomes.	4
1913	Alfred Sturtevant made a genetic linkage map of the *Drosophila X* chromosome, the first genetic map.	4
1918	Ronald Fisher proposed that multiple Mendelian factors can explain continuous variation for traits, founding the field of quantitative genetics.	19
1931	Harriet Creighton and Barbara McClintock showed that crossing over is the cause of recombination.	4, 16
1941	Edward Tatum and George Beadle proposed the one-gene—one-polypeptide hypothesis.	6
1944	Oswald Avery, Colin MacLeod, and Maclyn McCarty provided compelling evidence that DNA is the genetic material in bacterial cells.	7
1946	Joshua Lederberg and Edward Tatum discovered bacterial conjugation.	5
1948	Barbara McClintock discovered mobile elements (transposons) that move from one place to another in the genome.	15
1950	Erwin Chargaff showed DNA composition follows some simple rules for the relative amounts of A, C, G, and T.	7
1952	Alfred Hershey and Martha Chase proved that DNA is the molecule that encodes genetic information.	7
1953	James Watson and Francis Crick determined that DNA forms a double helix.	7
1958	Matthew Meselson and Franklin Stahl demonstrated the semiconservative nature of DNA replication.	7
1958	Jérôme Lejeune discovered that Down syndrome resulted from an extra copy of the 21st chromosome.	17
1961	François Jacob and Jacques Monod proposed that enzyme levels in cells are controlled by feedback mechanisms.	11
1961–1967	Marshall Nirenberg, Har Gobind Khorana, Sydney Brenner, and Francis Crick "cracked" the genetic code.	9
1968	Motoo Kimura proposed the neutral theory of molecular evolution.	18, 20
1977	Fred Sanger, Walter Gilbert, and Allan Maxam invented methods for determining the nucleotide sequences of DNA molecules.	10
1980	Christiane Nüsslein-Volhard and Eric F. Wieschaus defined the complex of genes that regulate body plan development in *Drosophila.*	13
1989	Francis Collins and Lap-Chee Tsui discovered the gene causing cystic fibrosis.	4, 10
1993	Victor Ambrose and colleagues described the first microRNA.	13
1995	First genome sequence of a living organism (*Haemophilus influenzae*) published.	14
1996	First genome sequence of a eukaryote (*Saccharomyces cerevisiae*) published.	14
1998	First genome sequence of an animal (*Caenorhabditis elegans*) published.	14
2000	First genome sequence of a plant (*Arabidopsis thaliana*) published.	14
2001	The sequence of the human genome first published.	14
2006	Andrew Fire and Craig Mello win the Nobel prize for their discovery of gene silencing by double-stranded RNA.	8
2012	John Gurdon and Shinya Yamanaka win the Nobel prize for their discovery that just four regulatory genes can convert adult cells into stem cells.	8, 12

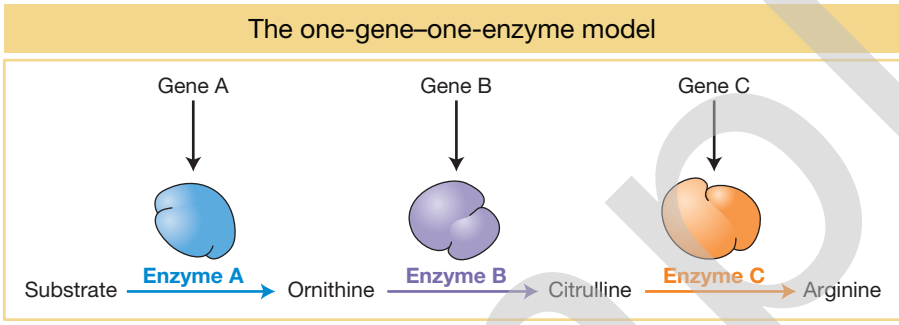


FIGURE 1-7 The one-gene–one-enzyme model proposed that genes encode enzymes that carry out biochemical functions within cells. Tatum and Beadle proposed this model based on the study of the synthesis of arginine (an amino acid) in the bread mold *Neurospora crassa*.

was that continuous traits are each controlled by multiple Mendelian genes. Fisher's insight is known as the **multifactorial hypothesis.** In Chapter 19, we will dissect the mathematical model and experimental evidence for Fisher's hypothesis.

How do genes function inside cells in a way that enables them to control different states for a trait like flower color? In 1941, Edward Tatum and George Beadle proposed that genes encode enzymes. Using bread mold (*Neurospora crassa*) as their experimental organism, they demonstrated that genes encode the enzymes that perform metabolic functions within cells (Figure 1-7). In the case of the pea plant, there is a gene that encodes an enzyme required to make the purple pigment in the cells of a flower. Tatum and Beadle's breakthrough became known as the **one-gene–one-enzyme hypothesis.** You'll see how they developed this hypothesis in Chapter 6.

What is the physical nature of the gene? Are genes composed of protein, nucleic acid, or some other substance? In 1944, Oswald Avery, Colin MacLeod, and Maclyn McCarty offered the first compelling experimental evidence that genes are made of deoxyribonucleic acid (DNA). They showed that DNA extracted from a virulent strain of bacteria carried the necessary genetic information to transform a nonvirulent strain into a virulent one. You'll learn exactly how they demonstrated this in Chapter 7.

How can DNA molecules store information? In the 1950s, there was something of a race among several groups of geneticists and chemists to answer this question. In 1953, James Watson and Francis Crick working at Cambridge University in England won that race. They determined that the molecular structure of DNA was in the form of a double helix—two strands of DNA wound side-by-side in a spiral. Their structure of the double helix is like a twisted ladder (Figure 1-8). The sides of the ladder are made of sugar and phosphate groups. The rungs of the ladder are made of four bases: **adenine (A), thymine (T), guanine (G),** and **cytosine (C).** The bases face the center, and each base is hydrogen bonded to the base facing it in the opposite strand. Adenine in one strand is always paired with thymine in the other by a *double hydrogen bond,* whereas guanine is always paired with cytosine by a *triple hydrogen bond.* The bonding specificity is based on the **complementary** shapes and charges of the bases. The sequence of A, T, G, and C represents the coded information carried by the DNA molecule. You will learn in Chapter 7 how this was all worked out.

How are genes regulated? Cells need mechanisms to turn genes on or off in specific cell and tissue types and at specific times during development. In 1961, François Jacob and Jacques Monod made a conceptual breakthrough on this question. Working on the genes necessary to metabolize the sugar lactose in the bacterium *Escherichia coli,* they demonstrated that genes have **regulatory elements** that regulate **gene expression**—that is, whether a gene is turned on or off (Figure 1-9). The regulatory elements are specific DNA sequences to which a regulatory protein binds and acts as either an activator or repressor of the expression of the gene. In Chapter 11, you will explore the logic behind the experiments of Jacob and Monod with *E. coli,* and in Chapter 12, you will explore the details of gene regulation in eukaryotes.

The structure of DNA

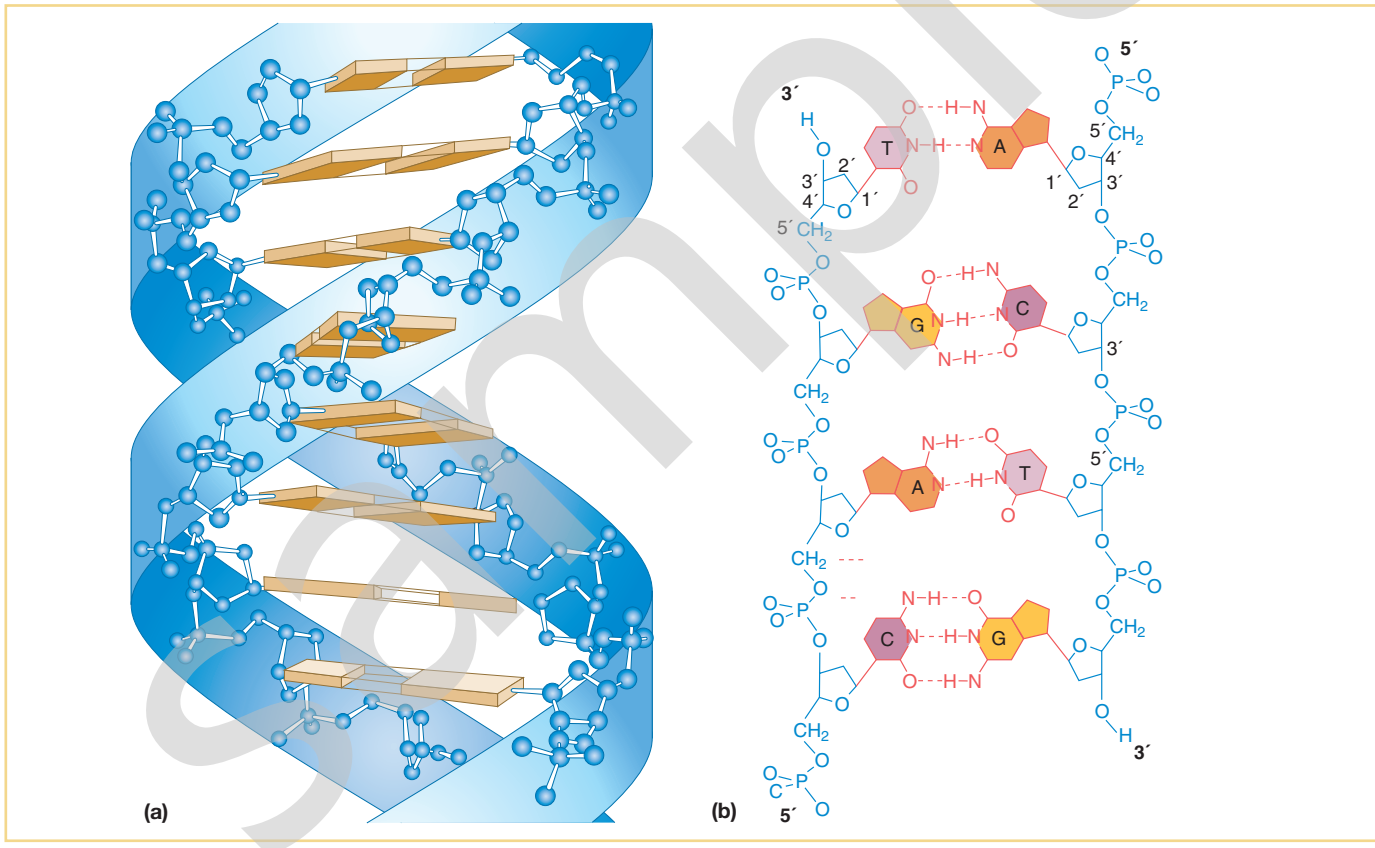


FIGURE 1-8 (a) The double-helical structure of DNA, showing the sugar–phosphate backbone in blue and paired bases in brown. (b) A flattened representation of DNA showing how A always pairs with T and G with C. Each row of dots between the bases represents a hydrogen bond.

How is the information stored in DNA decoded to synthesize proteins? While the discovery of the double-helical structure of DNA was a watershed for biology, many details were still unknown. Precisely how information was encoded into DNA and how it was decoded to form the enzymes that Tatum and Beadle had shown to be the workhorses of gene action remained unknown. Over the years 1961 through 1967, teams of molecular geneticists and chemists working in several countries answered these questions when they "cracked the genetic code." What this means is that they deduced how a string of DNA nucleotides, each with one of four different bases (A, T, C, or G), encodes the set of 20 different amino acids that are the building blocks of proteins. They also discovered that there is a messenger molecule made of ribonucleic acid (RNA) that carries information in the DNA in the nucleus to the cytoplasm where proteins are synthesized. By 1967, the basic flowchart for information transmission in cells was known. This flowchart is called the central dogma of molecular biology.

KEY CONCEPT The rediscovery of Mendel's laws launched a new era in which geneticists resolved many fundamental questions about the nature of the gene and the flow of genetic information within cells. During this era, geneticists learned that genes reside on chromosomes and are made of DNA. Genes encode proteins that conduct the basic enzymatic work within cells.

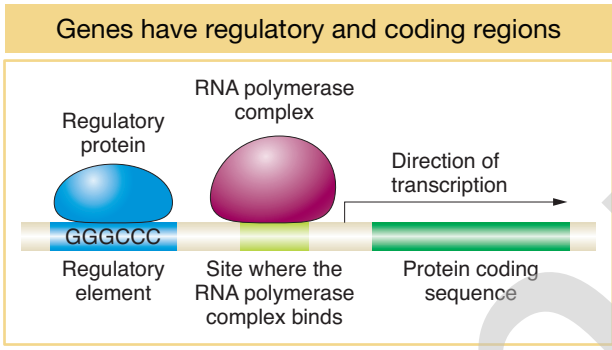


FIGURE 1-9 The structure of a protein-coding gene showing a regulatory DNA element (GGGCCC) to which a regulatory protein binds, the promoter region where the RNA polymerase complex binds to initiate transcription, and a protein-coding region

The central dogma of molecular biology

In 1958, Francis Crick introduced the phrase "central dogma" to represent the flow of genetic information within cells from DNA to RNA to protein, and he drew a simple diagram to summarize these relationships (Figure 1-10a). Curiously, Crick chose the word *dogma* thinking that it meant "hypothesis," which was his intention, unaware that its actual meaning is "a belief that is to be accepted without doubt." Despite this awkward beginning, the phrase had an undeniable power and it has survived.

Figure 1-10b captures much of what was learned about the biochemistry of inheritance from 1905 until 1967. Let's review the wealth of knowledge that this simple figure captures. At the left, you see DNA and a circular arrow representing **DNA replication,** the process by which a copy of the DNA is produced. This process enables each of the two daughter cells that result from cell division to have a

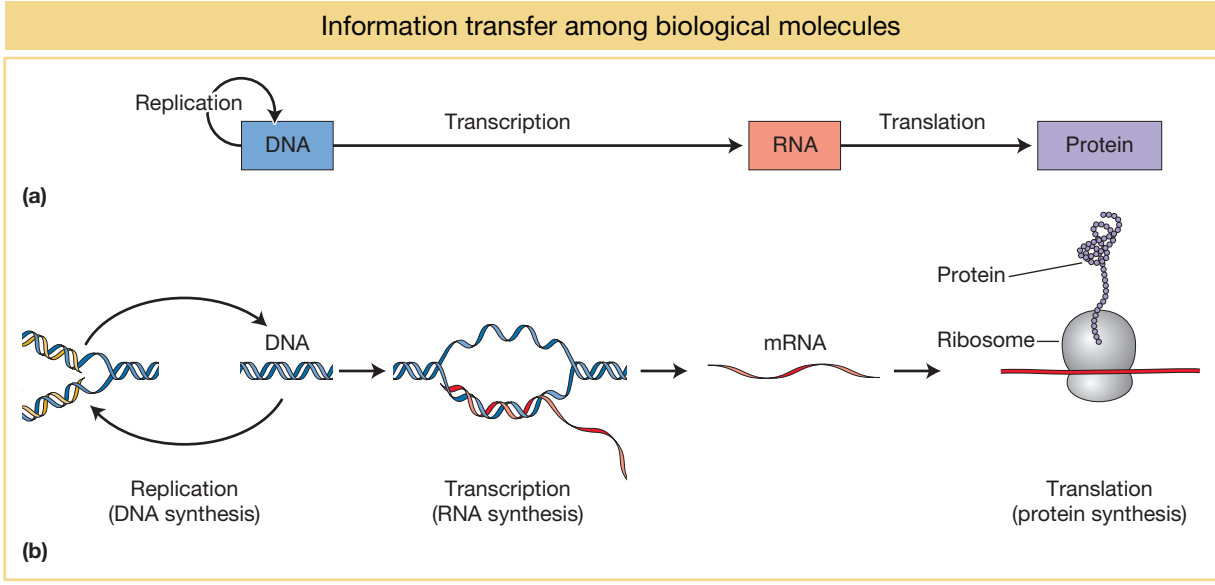


FIGURE 1-10 (a) One version of Francis Crick's sketch of the central dogma, showing information flow between biological molecules. The circular arrow represents DNA replication, the central straight arrow represents the transcription of DNA into RNA, and the right arrow the translation of RNA into protein. (b) More detailed sketch showing how the two strands of the DNA double helix are independently replicated, how the two strands are disassociated for transcription, and how the messenger RNA (mRNA) is translated into protein at the ribosome.

complete copy of all the DNA in the parent cell. In Chapter 7, you will explore the details of the structure of DNA and its replication.

Another arrow connects DNA to RNA, symbolizing how the sequence of base pairs in a gene (DNA) is copied to an RNA molecule. The process of RNA synthesis from a DNA template is called **transcription.** One class of RNA molecules made by transcription is **messenger RNA,** or **mRNA** for short. mRNA is the template for protein synthesis. In Chapter 8, you'll discover how transcription is accomplished.

The final arrow in Figure 1-10b connects mRNA and protein. This arrow symbolizes protein synthesis, or the **translation** of the information in the specific sequence of bases in the mRNA into the sequence of amino acids that compose a protein. Proteins are the workhorses of cells, comprising enzymes, structural components of the cell, and molecules for cell signaling. The process of translation takes place at the ribosomes in the cytoplasm of each cell. In Chapter 9, you will learn how the genetic code is written in three-letter words called **codons.** A codon is a set of three consecutive nucleotides in the mRNA that specifies an amino acid in a protein. CGC specifies the amino acid arginine, AGC specifies serine, and so forth.

Since Crick proposed the central dogma, additional pathways of genetic information flow have been discovered. We now know that there are classes of RNA that do not code for proteins, instances in which mRNA is edited after transcription, and cases in which the information in RNA is copied back to DNA (see Chapters 8, 9, and 15).

1.2 After Cracking the Code

With the basic laws of inheritance largely worked out by the end of the 1960s, a new era of applying genetic analysis to a broad spectrum of biological questions flourished. To this end, much effort has been and continues to be invested in developing the resources and tools to address these questions. Geneticists focused their research on a small number of species known as "model organisms" that are well suited for genetic analysis. They also developed an impressive array of tools for manipulating and analyzing DNA.

Model organisms

Geneticists make special use of a small set of model organisms for genetic analysis. A **model organism** is a species used in experimental biology with the presumption that what is learned from the analysis of that species will hold true for other species, especially other closely related species. The philosophy underlying the use of model organisms in biology was wryly expressed by Jacques Monod: "Anything found to be true of *E. coli* must also be true of elephants."[1]

As genetics matured and focused on model organisms, Mendel's pea plants fell to the wayside, but Morgan's fruit flies rose to prominence to become one of the most important model organisms for genetic research. New species were added to the list. An inconspicuous little plant that grows as a weed called *Arabidopsis thaliana* became the model plant species and a minute roundworm called *Caenorhabditis elegans* that lives in compost heaps became a star of genetic analysis in developmental biology (Figure 1-11).

What features make a species suitable as a model organism? (1) Small organisms that are easy and inexpensive to maintain are very convenient for research. So fruit flies are good, blue whales not so good. (2) A short generation time is imperative because geneticists, like Mendel, need to cross different strains and then study their

[1]F. Jacob and J. Monod, *Cold Spring Harbor Quant. Symp. Biol.* 26, 1963, 393.

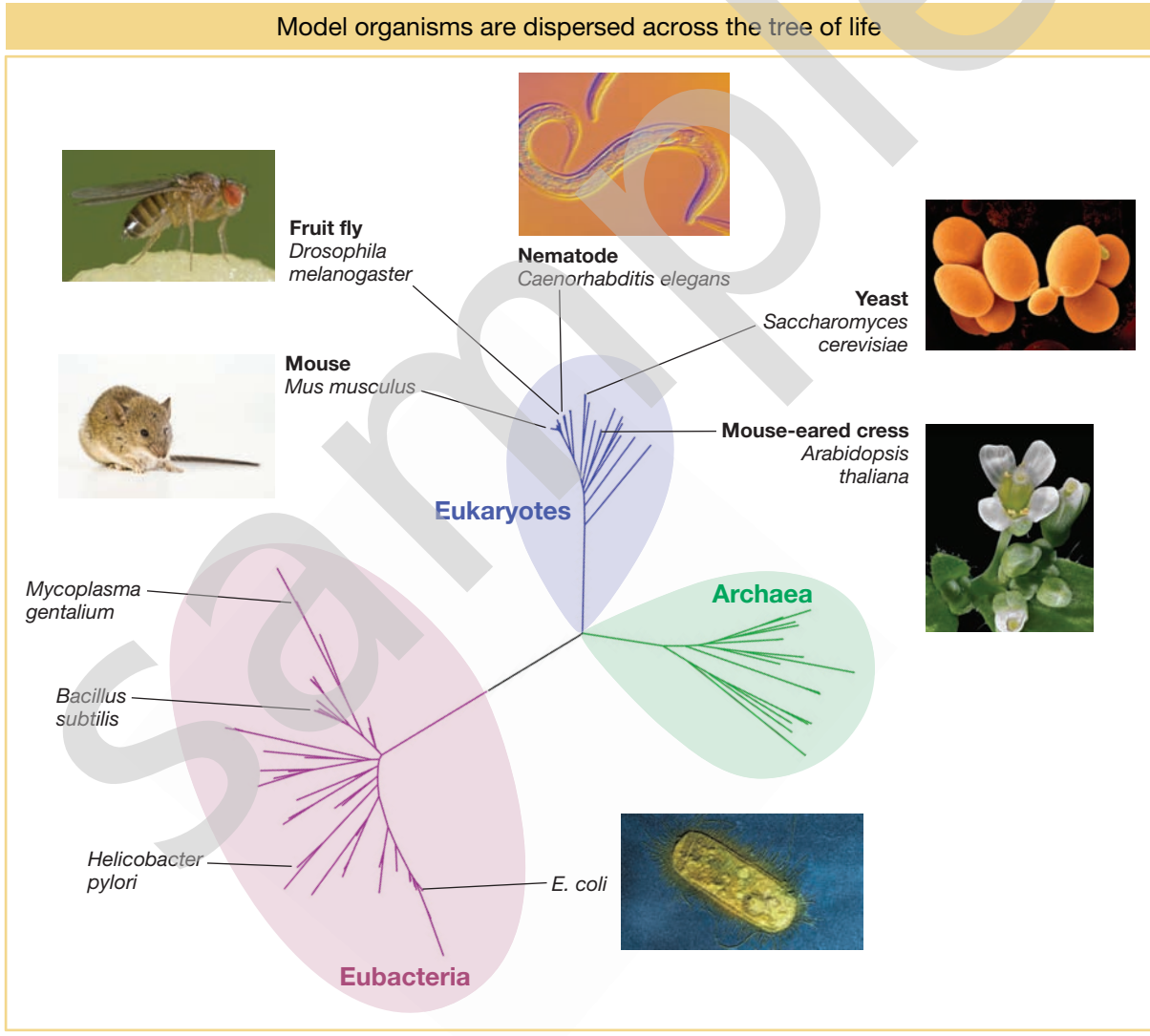


FIGURE 1-11 The tree shows evolutionary relationships among the major groups of organisms: Bacteria, Archaea, and Eukaryota (plants, fungi, and animals). [*(Clockwise, from top, center) Sinclair Stammers/Science Source; SciMAT/Science Source; Darwin Dale/Science Source; Biophoto Associates/Science Photo Library; Imagebroker.net/SuperStock; © blickwinkel/Alamy.*]

first- and second-generation hybrids. The shorter the generation time, the sooner the experiments can be completed. (3) A small genome is useful. As you will learn in Chapter 15, some species have large genomes and others small genomes in terms of the total number of DNA base pairs. Much of the extra size of large genome species is composed of repetitive DNA elements between the genes. If a geneticist is looking for genes, these can be more easily found in organisms with smaller genomes and fewer repetitive elements. (4) Organisms that are easy to cross or mate and that produce large numbers of offspring are best.

As you read this textbook, you will encounter certain organisms over and over. Organisms such as *Escherichia coli* (a bacterium), *Saccharomyces cerevisiae* (baker's yeast), *Caenorhabditis elegans* (nematode or roundworm), *Drosophila melanogaster* (fruit fly), and *Mus musculus* (mice) have been used repeatedly in experiments and revealed much of what we know about how inheritance works. Model organisms can be found on diverse branches of the tree of life (see Figure 1-11), representing bacteria, fungi, algae, plants, and invertebrate and vertebrate animals.

This diversity enables each geneticist to use a model best suited to a particular question. Each model organism has a community of scientists working on it who share information and resources, thereby facilitating each other's research.

Mendel's experiments were possible because he had several different varieties of pea plants, each of which carried a different genetic variant for traits such as purple versus white flowers, green versus yellow seeds, or tall versus dwarf stems. For each of the model species, geneticists have assembled large numbers of varieties (also called strains or stocks) with special genetic characters that make them useful in research. There are strains of fruit flies that have trait variants such as red versus white eyes. There are strains of mice that are prone to develop specific forms of cancer or other disease conditions such as diabetes. For baker's yeast, there is a collection of nearly 5000 *deletion stocks,* each of these having just one gene deleted from the genome. These stocks enable geneticists to study the function of each gene by examining how yeast is affected when the gene is removed. Since baker's yeast has about 6000 total genes, this collect of 5000 deletion stocks covers most of the genes in the genome.

The different strains of each model organism are available to researchers through stock centers that maintain and distribute the strains. Lists of available stocks are on the Internet (see Appendix B). To view an example for mouse stocks, go to the link http://jaxmice.jax.org/. Then, click the "Find JAX mice" button at the top of the page. Next, enter the word "black" in the search field and click the Search button. Now, click the "C57BL/6J" link. You will see an image and information on a commonly used C57-Black mouse strain. Other search terms such as "albino" or "obese" will link you with strains with other features.

KEY CONCEPT Most genetic studies are performed on one of a limited number of model organisms that have features that make them especially suited for genetic analysis.

Tools for genetic analysis

Geneticists and biochemists have also created an incredible array of tools for characterizing and manipulating DNA, RNA, and proteins. Many of these tools are described in Chapter 10 or in other chapters relevant to a specific tool. There are a few themes to mention here.

First, geneticists have harnessed the cell's own machinery for copying, pasting, cutting, and transcribing DNA, enabling researchers to perform these reactions inside test tubes. The enzymes that perform each of these functions in living cells have been purified and are available to researchers: **DNA polymerases** can make a copy of a single DNA strand by synthesizing a matching strand with the complementary sequence of A's, C's, G's, and T's. **Nucleases** can cut DNA molecules in specific locations or degrade an entire DNA molecule into single nucleotides. **Ligases** can join two DNA molecules together end-to-end. Using DNA polymerase or other enzymes, DNA can also be "labeled" or "tagged" with a fluorescent dye or radioactive element so that the DNA can be detected using a fluorescence or radiation detector.

Second, geneticists have developed methods to *clone DNA* and the genes it encodes. Here, cloning refers to making many copies (*clones*) of a DNA molecule. The common way of doing this involves isolating a relatively small DNA molecule (up to a few thousand base pairs in length) from an organism of interest. The DNA molecule might be an entire gene or a portion of a gene. The molecule is inserted into a host organism (often *E. coli*) where it is replicated many times by the host's DNA polymerase. Having many copies of a gene is important for a vast array of experiments used to characterize and manipulate it.

Third, geneticists have developed methods to insert foreign DNA molecules into the genomes of many species, including those of all the model organisms.

This process is called **transformation,** and it is possible, for instance, to transform genes from one species into the genome of another. The recipient species then becomes a **genetically modified organism (GMO).** Figure 1-12 shows a tobacco plant in which a gene from the firefly was inserted, enabling the tobacco plant to emit light or glow in the dark.

Fourth, geneticists have developed a large set of methods based on hybridizing DNA molecules to one another (or to RNA molecules). The two complementary strands of DNA in the double helix are bound together by hydrogen bonds, either G ≡ C or A = T. These bonds can be broken by heat (denatured) in an aqueous solution to give two single-stranded DNA molecules (Figure 1-13a). When the solution is cooled under controlled conditions, DNA molecules with complementary strands will preferentially hybridize with one another. DNA hybridization methods have enabled many discoveries. For example, the cloned DNA of a gene can be tagged with a fluorescent dye and then hybridized to chromosomes fixed on a microscope slide, revealing the chromosome on which the gene is located (Figure 1-13b).

Fifth, geneticists and biochemists have developed multiple methods for determining the exact sequence of all the A's, C's, G's, and T's in the genomes, chromosomes, or genes of an organism. The process used to decipher the exact sequence of A's, C's, G's, and T's in a DNA molecule is called **DNA sequencing,** and it has allowed geneticists to read the language of life.

Finally, over the last 20 years, researchers have created molecular and mathematical tools for analyzing the entire genome of an organism in a single experiment. These efforts gave birth to the field of **genomics**—the study of the structure and function of entire genomes (see Chapter 14). Genomic tools have enabled geneticists to assemble mind-boggling amounts of information on model organisms, including the complete DNA sequence of their genome, lists of all their genes, catalogs of variants in these genes, data on the cell and tissue types in which each gene is expressed, and much more. To get an idea of what is available, try browsing Fly Base (http://flybase.org/), the genomic data site for the fruit fly (see also Appendix B).

KEY CONCEPT Progress in genetics has both produced and been catalyzed by the development of molecular and mathematical tools for the analysis of single genes and whole genomes.

Genetically modified tobacco

FIGURE 1-12 This genetically modified tobacco plant has a gene from the firefly inserted into its genome, giving it the capability to emit light. [*D. W. Ow et al., "Transient and Stable Expression of the Firefly Luciferase Gene in Plant Cells and Transgenic Plants,"* Science *234, 4778, 1986, 856–859.*]

Strands of nucleic acids hybridize to complementary sequences

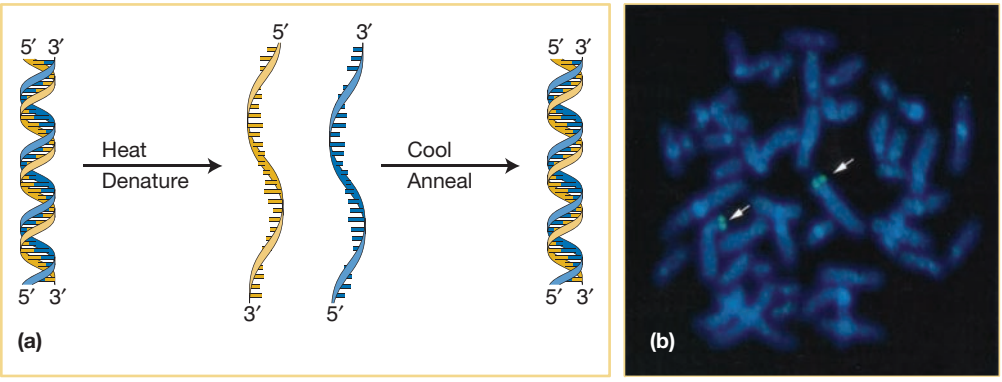


FIGURE 1-13 (a) The two strands of the DNA double helix can be dissociated by heat in aqueous solutions. Upon cooling under controlled conditions, strands reassociate, or *hybridize,* with their complement. (b) A cloned copy of the human *BAPX1* gene was tagged with a green fluorescent dye. The fluorescent-tagged DNA was then denatured and allowed to hybridize to the chromosomes in a single cell. The fluorescent-tagged clone hybridized to the location on chromosome 4 (green fluorescent regions) where the gene is located. [*(b) C. Tribioli and T. Lufkin, "Molecular cloning, chromosomal mapping and developmental expression of BAPX1, a novel human homeobox-containing gene homologous to Drosophila bagpipe,"* Gene, *203, 2, 1997, 225–233, Fig. 6, © Elsevier.*]

1.3 Genetics Today

In an interview in 2008, Princeton University geneticist Leonid Kruglyak remarked,

> "You have this clear, tangible phenomenon in which children resemble their parents. Despite what students get told in elementary-school science, we just don't know how that works."

Although Kruglyak's remark might seem disparaging to the progress made in the understanding of inheritance over the last 100 years, this was certainly not his intention. Rather, his remark highlights that despite the paradigm-shifting discoveries of the nineteenth and twentieth centuries, enigmas abound in genetics and the need for new thinking and new technologies has never been greater. Mendel, Morgan, Fisher, Watson, Crick, and many other others (see Table 1-1) delimited the foundation of the laws of inheritance, but the details that rest atop that foundation remain obscure in many ways. The six feet of DNA in the single cell of a human zygote encodes the information needed to transform that cell into an adult, but exactly how this works is understood only in the sparsest details.

In this section, we will review four recent advances in genetics—discoveries of enough importance and general interest that they were featured in the popular press. Reading about these discoveries will both reveal the power of genetics to answer critical questions about life and highlight how this knowledge can be applied to addressing problems in society. This textbook and the course of study in which you are engaged should convey a dual message—the science of genetics has profoundly changed our understanding of life, but it is also a youthful field in the midst of a dynamic phase of its development.

From classical genetics to medical genomics

Meet patient VI-1 (Figure 1-14a). Her name is Louise Benge, and as a young woman, she developed a crippling illness. Starting in her early 20s, she began to experience

Louise Benge has an undiagnosed disease

FIGURE 1-14 (a) Louise Benge developed an undiagnosed disease as a young woman. (b) An X ray revealed that Louise Benge's disease condition caused calcification of the arteries in her legs. [*(a) Jeannine Mjoseth, NHGRI/www.genome.gov; (b) National Human Genome Research Institute (NHGRI).*]

excruciating pain in her legs after walking as little as a city block. At first, she ignored the pain, then spoke with her primary care physician, and later visited a long line of specialists. She was given a battery of tests and X rays, and these revealed the problem—her arteries from her aorta on down to her legs were calcified, clogged with calcium phosphate deposits (Figure 1-14b). It was a disease for which her doctors had no name and no therapy. She had a disease, but not a diagnosis. There was only one thing left to do; her primary care physician referred Benge to the Undiagnosed Diseases Program (UDP) at the National Institutes of Health in Bethesda, Maryland.

The UDP is a group of MDs and scientists that has connections with specialists throughout the National Institutes of Health in every imaginable field of medicine. This is the team that is asked to tackle the most challenging cases. Working with Benge, the UDP team subjected her to nearly every test in their arsenal, and soon they found the underlying defect that caused her disease. Benge had a very low level of an enzyme called CD73. This enzyme is involved in signaling between cells, and specifically it sends a signal that blocks calcification. Now the UDP doctors could give Benge a diagnosis. They named her disease "arterial calcification due to deficiency of CD73," or ACDC.

Tracing a disease gene through a family tree

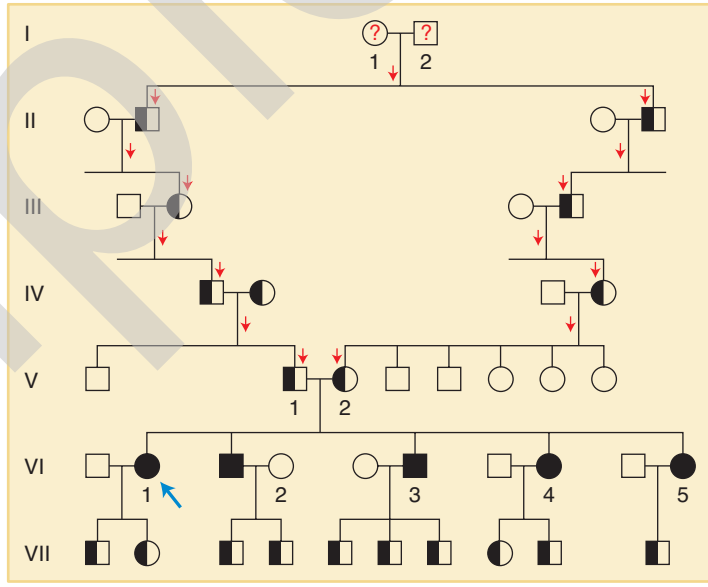


FIGURE 1-15 Family tree or pedigree showing the inheritance of the mutant gene causing arterial calcification due to deficiency of CD73 (ACDC). Squares are males, and circles are females. Horizontal lines connecting a male and female are matings. Vertical lines connect a mating pair to its offspring. Roman numerals designate generations; Arabic numerals designate individuals within generations. Half-filled squares or circles indicate an individual carrying one copy of the mutant gene. Filled squares or circles indicate an individual with two copies of the mutant gene and who have the ACDC disease. Either individual I-1 or I-2 must have carried the mutant gene, but which one carried it is uncertain as indicated by the "?". Blue arrow indicates Louise Benge. Red arrows show the path of the mutant gene through the generations. [*Data from C. St. Hilaire et al.,* New England Journal of Medicine *364, 2011, 432–442.*]

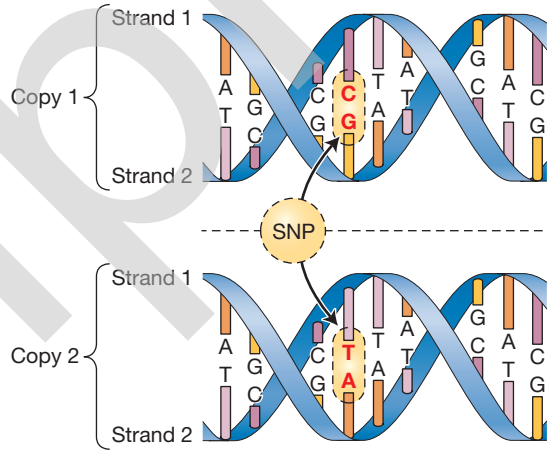
What intrigued the UDP team about Benge's case was that she was not alone in having this disease. Benge had two brothers and two sisters, and all of them had arterial calcification. Remarkably, however, Benge's parents were unaffected. Moreover, Benge and her siblings all had children and none of these children had arterial calcification. This pattern of inheritance suggested that the underlying cause might be genetic. Specifically, it suggested that Benge and all of her siblings inherited two defective copies of either CD73 or a gene that influences CD73 expression—one from their mother and one from their father. A person with one good copy and one defective copy can be normal, but if both of a person's copies are defective, then they lack the function that the gene provides. The situation is just like Mendel's white-flowered pea plants. Since the functional allele is dominant to the dysfunctional allele, ACDC, like white flowers, only appears if an individual carries two defective alleles.

The UDP team delved further into Benge's family history and learned that Benge's parents were third cousins (Figure 1-15). This revelation fit well with the idea that the cause was a defective gene. When a husband and wife are close relatives such as third cousins, there is an increased chance that they will both have inherited the same version of a defective gene from their common ancestor and that they will both pass on this defective gene to their children. Children with one copy of a defective gene are often normal, but a child who inherits a defective copy from both parents is likely to have a genetic disorder.

In Figure 1-15, we can see how this works. Benge's mother and father (individuals V-1 and V-2 in the figure) have the same great-great-grandparents (I-1 and I-2). If one of these great-great-grandparents had a mutant gene for CD73, then it could have been passed down over the generations to both Benge's mother and father (follow the red arrows). After that, if Benge received the mutant copy from both her mother and her father, then both of her copies would be defective. Each of Benge's siblings would also need to have inherited two mutant copies from their parents to explain the fact that they have ACDC. The chance of all of this happening is very small. If both of Benge's parents had one mutant copy, then the chance that Benge and all four of her siblings would receive a mutant copy from

BOX 1-1 Single Nucleotide Polymorphisms

Genetic variation is any difference between two copies of the same gene or DNA molecule. The simplest form of genetic variation one might observe at a single nucleotide site is a difference in the nucleotide base present, whether adenine, cytosine, guanine, or thymine. These types of variants are called **single nucleotide polymorphisms (SNPs),** and they are the most common type of variation in most, if not all, organisms. The figure shows two copies of a DNA molecule from the same region of a chromosome. Notice that the bases are the same in the two molecules except where one molecule has a CG pair and the other a TA pair. If we read strand 1 of the two molecules, then the top molecule has a "G" and the lower molecule an "A" at the SNP site.



both parents is only 1 in 1024. In Chapter 2, you'll learn how to calculate such probabilities.

With this hint from the family history, the UDP team now knew where to look in the genome for the mutant gene. They needed to look for a segment on one of the chromosomes for which the copy that Benge inherited from her mother is identical to the copy she inherited from her father. Moreover, each of Benge's siblings must also have two copies of this segment identical to Benge's. Such regions are very rare in people unless their parents are related, as in the case of Benge since her parents are third cousins. Generally, a segment of a chromosome that is just a few hundred base pairs long will have several differences in the sequence of A's, C's, G's, and T's between the copy we inherited from our mother and the one we inherited from our father. These differences are known as **single nucleotide polymorphisms,** or **SNPs** for short (see Box 1-1).

The UDP team used a new genomic technology, called a DNA microarray (see Chapter 18), that allowed them to study one million base-pair positions across the genome. At each of these base-pair positions along the chromosomes, the team could see where Benge's two chromosomal segments were identical, and whether all of Benge's siblings also carried two identical copies in this segment. For Benge, a portion of only 1/512 of her genome is expected to have two identical copies, and the chance that all four of her siblings will also have the same two identical copies is far smaller.

Looking over the genome-wide SNP data, the UDP team found exactly the type of chromosome segment for which they were looking. There was a small segment on one of Benge's chromosomes for which she and her siblings all had the same two identical copies. Furthermore, they discovered that the gene that encodes the CD73 enzyme is located in this segment. This result suggested that Benge and her siblings all had two identical copies of the same defective CD73-encoding gene. The team seemed to have found the needle in a haystack for which they were looking; however, there was one last experiment to perform.

The team needed to identify the specific defect in the defective CD73 gene that Benge and her siblings had inherited. After determining the DNA sequence for the CD73 gene from Benge and her siblings, the team found the defect in the gene—"the smoking gun." The defective gene encoded only a short, or truncated, protein—it did not encode the complete sequence of amino acids. One of the DNA

codons with letters TCG that encodes the amino acid serine was mutated to TAG, which signals the truncation of the protein. The protein made from Benge's version of the CD73 gene was truncated so it could not signal cells in the arteries to keep the calcification pathway turned off.

Louise Benge's journey from first experiencing pain in her legs to learning that she had a new disease called ACDC was a long one. The diagnosis of her disease was a triumph made possible by the integration of classic transmission genetics and genomics. Knowing the defect underlying the disease ACDC allowed the doctors to try a medication that they would never have considered before they knew that the cause was a defective CD73 enzyme. The medication in question is called etidronate, and it can substitute for CD73 in signaling cells to keep the calcification pathway turned off. Clinical trials with etidronate are currently underway for ACDC patients and are scheduled for completion in 2017.

KEY CONCEPT Classical transmission genetics provides the foundation for modern medical genetics. The integration of classical genetics and genomic technologies can allow the causes of inherited diseases to be readily identified.

Investigating mutation and disease risk

Shortly after the rediscovery of Mendel's work, the German physician Wilhelm Weinberg reported that there seems to be a higher incidence of short-limbed dwarfism (achondroplasia) among children born last in German families than among those born first. A few decades later, British geneticist J. B. S. Haldane observed another unusual pattern of inheritance. The genealogies of some British families suggested that new mutations for the blood-clotting disorder hemophilia tended to arise in men more frequently than in women. Taken together, these two observations suggested that the risk of an inherited disorder for a child is greater as the parents age and also that fathers are more likely than mothers to contribute new mutations to their children.

Over the ensuing decades, Weinberg's and Haldane's observations were supported by other studies, but the data were not conclusive. Tracing a new mutation in a child to the father versus the mother was fraught with uncertainty, and there was a scarcity of families well-suited for the study of the link between parental age and new disease mutations. These factors prevented definitive conclusions on the relationship between parental age and the occurrence of new mutations.

In 2012, advances in genomics and DNA sequencing technology (see Chapter 14) allowed new analyses proving that Weinberg's and Haldane's suspicions were correct and providing a very detailed picture of the origin of new mutations within families. Here is how it was accomplished. A team of geneticists in Iceland studied 78 "trios"—a family group of a mother, a father, and their child (Figure 1-16). For some families, they had data for three generations, including a child plus its parents and at least one set of grandparents. The researchers determined the complete genome sequence of each individual with DNA isolated from their blood cells, compiling genome sequences from a total of 219 individuals. Since each individual possesses two copies of every chromosome (i.e., two copies of the human genome), their data actually include the sequences of 438 genomes.

With these genome sequences in hand, the researchers could comb through the data for *new* or *de novo mutations*—unique DNA variants that exist in a child but neither of its parents. Their focus was on **point mutations,** or a change of one letter in the DNA code to another that can occur during DNA replication (see Chapter 16). For example, a change of an adenosine (A) to a guanine (G) (Figure 1-17).

The logic of the discovery process used by the Icelandic geneticists is outlined in Figure 1-17, which shows a segment of DNA for each member of a trio. Each

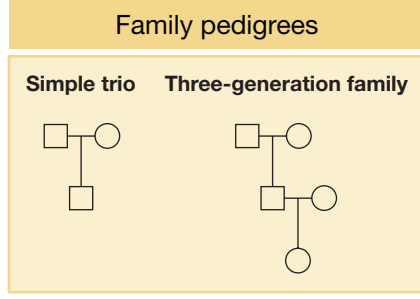


FIGURE 1-16 Squares are males, and circles are females. Horizontal lines indicate a mating. Vertical lines connect a mating pair to its offspring.

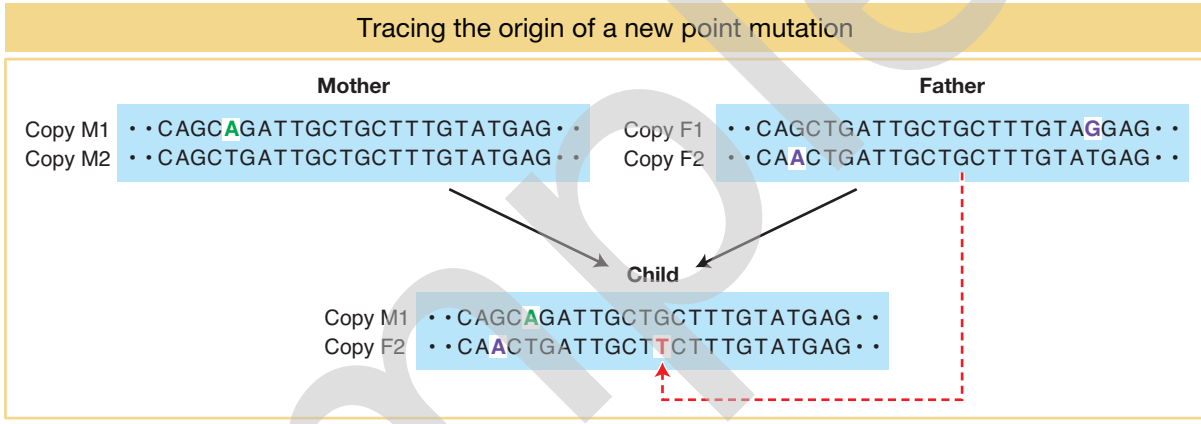


FIGURE 1-17 A short segment of DNA from one of the chromosomes is shown. Each individual has two copies of the segment. In the mother, these are labeled M1 and M2; in the father, F1 and F2. The child inherited copy M1 from its mother and F2 from its father. The version of F2 in the child carries a new point mutation (red). Single nucleotide polymorphisms (SNPs) that distinguish the different copies are shown in green (mother) and purple (father).

individual has two copies of the segment. Notice that copy M1 in the mother has a SNP (green letter) that distinguishes it from copy M2. Similarly, there are two SNPs (purple letters) that distinguish the father's two copies of this segment. Comparing the child to the parents, we see that the child inherited copy M1 from its mother and copy F2 from its father. Look closer at the child's two copies of the segment, and you'll notice something else. There is a unique variant (red letter) that occurs in the child but neither of its parents. This is a de novo point mutation. It this case, it is a mutation from a guanine (G) to a thymine (T). We can see that the mutation arose in the father since it is on the F2 copy of the segment.

Where and exactly when did the new mutation depicted in Figure 1-17 arise? Most of our bodies are composed of somatic cells that make up everything from our brain to our blood. However, we also have a special lineage of cells called the germline that divide to produce eggs in women and sperm in men. New mutations that arise in somatic cells as they divide during the growth and development of our bodies are not passed on to our offspring. However, a new mutation that occurs in the germline can be transmitted to the offspring. The mutation depicted in Figure 1-17 arose in the germline of the father.

With the genome sequence data for the trios, the Icelandic geneticists made some pretty startling discoveries. First, among the 78 children in the study, they observed a total of 4933 new point mutations. Each child carried about 63 unique mutations that did not exist in its parents. Most of these occurred in parts of the genome where they have only a small chance to pose a health risk, but 62 of the 4933 mutations caused potentially damaging changes to the genes such that they altered the amino acid sequence of the protein encoded. Second, among the mutations that could be assigned a parent of origin, there were on average 55 from the father for every 14 from the mother. The children were inheriting nearly four times as many new mutations from their fathers as their mothers. The Icelandic team had confirmed Haldane's prediction made 90 years earlier.

The genome sequences also allowed the team to test Weinberg's prediction that the frequency of mutation rises with the age of the parents. For each trio, the researchers knew the ages of the mother and the father at the time of conception. When they investigated whether the frequency of mutation rises with the mother's age when controlling for the age of the father, the team found no evidence that it did. Older mothers did not pass on more new point mutations to their offspring than younger ones. (Older mothers are known to produce more chromosomal aberrations than younger mothers, such as an extra copy of the 21st chromosome that causes Down syndrome; see Chapter 17.) Next, they examined the relationship between mutation and the age of the father when controlling for the age of the mother. Here, they found a powerful relationship. The older the father, the higher the frequency of new point mutations (Figure 1-18). In fact, for

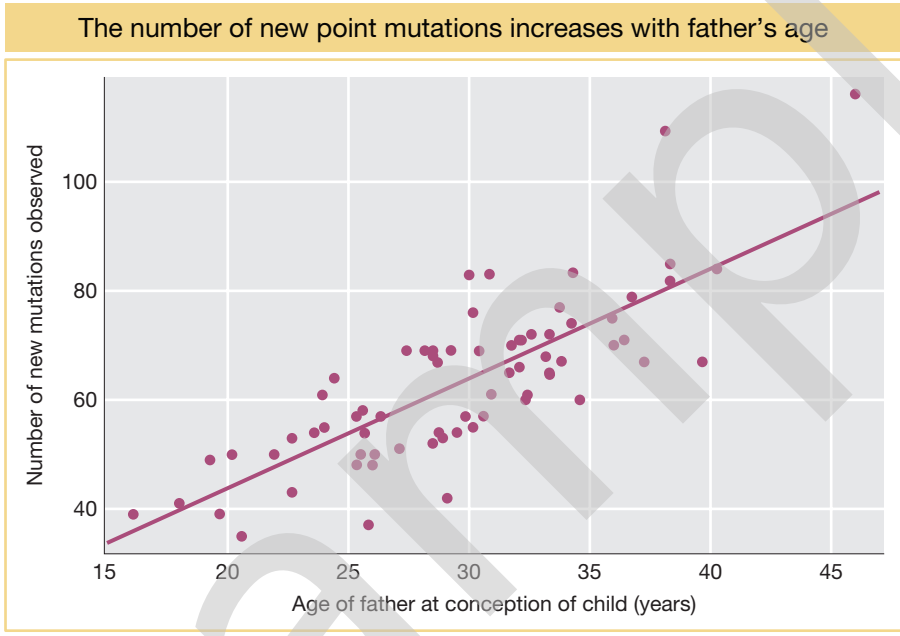


FIGURE 1-18 Plot of the number of new point mutations in each child (y-axis) by the age of the child's father (x-axis). Each dot represents one of the 78 children studied. The diagonal line indicates the rate of increase in new mutations with the father's age. [*Data from A. Kong et al.,* Nature *488, 2012, 471–475.*]

each year of increase in his age, a father will pass on two additional new mutations to his children. A 20-year-old father will pass on about 25 new mutations to each of his children, but a 40-year-old father will pass on about 65 new mutations. Weinberg's observation made 100 years earlier was confirmed.

Why does the age of the father matter, while that of the mother seems to have no effect on the frequency of new point mutations? The answer lies in the different ways by which men and women form gametes. In women, as in the females of other mammals, the process of making eggs takes place largely before a woman is born. Thus, when a woman is born she possesses in her ovaries a set of egg precursor cells that will mature into egg cells without further rounds of DNA replication. For a woman, from the point when she was conceived until the formation of the egg cells in her ovaries, there are about 24 rounds of cell division, 23 of which have a round of chromosome (DNA) replication and an opportunity for a copying error or mutation. All 23 of these rounds of chromosome replication occur before a woman is born, so there are no additional rounds after her birth and no chance for additional mutations as she ages. Thus, older mothers contribute no more new point mutations to their children than younger mothers.

Sperm production is altogether different. The cell divisions that produce sperm continue throughout a man's life, and there are many more rounds of cell division in sperm formation than in egg formation. Sperm produced by 20-year-old men will have experienced about 150 rounds of DNA replication from the time of the man's conception, almost seven times as many as for the eggs produced by 20-year-old women. By the time a man is age 40, his sperm will have a history that involves over 25 times as many rounds of DNA replication as for eggs in a woman of the same age. Thus, there is much more risk of new point mutations occurring during these extra rounds of cell division and DNA replication with the increase in the age of the father.

There is one final twist to the remarkable project performed by the Icelandic geneticists. The 78 trios that they studied were chosen because the children in most of the trios had inherited disorders. These included 44 children with autism spectrum disorder and 21 with schizophrenia. For all these children, there were no other cases of these disorders among their relatives, suggesting that their

condition was due to a new mutation. As anticipated, the researchers observed a correlation between the father's age and disease risk—older fathers were more likely to have children with autism and schizophrenia. In several cases, the DNA data for the child and parents also allowed the researchers to identify specific new mutations in genes that likely caused the disorder. For example, one child with autism inherited a new mutation in the *EPH receptor B2 (EPHB2)* gene that functions in the nervous system and in which a mutation had previously been found in an autistic child.

Studies such as this can have important implications for individuals and society. Some men who intend to delay parenting until later in life might choose to freeze samples of their sperm while still young. This study also informs us that changes in society can impact the number of new mutations that enter the human gene pool. If men choose to delay fatherhood for postsecondary education or establishing their careers, there will be an associated increase in the number of new mutations among their children. It is common knowledge that infertility rises with age for women—as is often stated, a woman's "biological clock" is ticking once she is past puberty. This work by the Icelandic geneticists informs us that a clock is ticking for men as well.

KEY CONCEPT Genome sequences of parents and their children clarify the factors that contribute to new point mutations. Fathers contribute four times as many new mutations to their offspring as do mothers. The number of new mutations passed on from a father to his children rises with the age of the father.

When rice gets its feet a little too wet

Among the cereal crops, rice is unique. Whereas wheat, barley, maize, and the other grain crops grow solely in dry fields, rice is commonly grown in flooded fields called paddies (Figure 1-19). The ability of rice to grow in flooded fields offers it an advantage: rice can survive modest flooding (up to 25 cm of standing water) in the paddies, but most weeds cannot. So rice farmers can use flooding to control the weeds in their field while their rice thrives.

The strategy works well where farmers have irrigation systems to control the water levels in their paddies and heavy rains do not exceed their capacity to

Rice growing in a flooded field or paddy

FIGURE 1-19 Rice is grown in fields with standing water called paddies. Rice is adapted to tolerate modest levels of standing water, but the water suppresses the growth of weeds that could compete with the rice. [© *Dinodia/AGE Fotostock.*]

control these levels. If the water in the paddies gets too deep (greater than 50 cm) for a prolonged period, then the rice plants, like the weeds, can suffer or even die.

Paddy agriculture, as practiced in the lowlands of India, Southeast Asia, and West Africa, relies on natural rainfall, rather than irrigation, to flood the fields. This circumstance poses a risk. When the rains are heavy, water depth in the paddies can exceed 50 cm and completely submerge the plants, causing rice plants to either suffer a loss in yield or simply die. Of the 60 million hectares of rain-fed lowland paddies, one-third experience damaging floods on a regular basis. The heavy rains and monsoons that flood the fields are estimated to cause a loss of rice worth more than US$1 billion each year. In India, Indonesia, and Bangladesh alone, 4 million tons of rice are lost to flooding each year, enough to feed 30 million people. Since this loss is mostly incurred by the poorest farmers, it can lead to malnourishment and even starvation.

In the early 1990s, David Mackill, a plant geneticist and breeder at the International Rice Research Institute, had an idea about how to improve rice so that it could tolerate being submerged in flood waters. He identified a remarkable variety of rice called FR13A that could survive submergence and even thrive after the plants remained fully submerged in deep water for up to two weeks. Unfortunately, FR13A had a low yield and the quality of its grain was marginal. So Mackill set out to transfer FR13A's genetic factor(s) for submergence tolerance into a rice variety with a higher yield and higher grain quality. He first crossed FR13A and a superior variety of rice and then for several generations crossed the hybrid plants back to the superior variety until he had created an improved form of rice that combined submergence tolerance and high yield.

Mackill had achieved his initial goal of transferring submergence tolerance into a superior variety, but the genetic basis for why FR13A was submergence tolerant remained obscure. Was FR13A's submergence tolerance controlled by many genes on multiple chromosomes, or might it be mostly controlled by just one gene? To delve into the genetic basis of submergence tolerance, Mackill and his team conducted a form of genetic analysis called **quantitative trait locus (QTL)** mapping (see Chapter 19). A QTL is a genetic locus that contributes incrementally or quantitatively to variation for a trait. Mendel's gene for flower color had two categorical alleles: one for purple flowers and the other for white flowers. QTL have alleles that usually engender only partial changes such as the difference between a pale purple and a medium purple. Using QTL mapping, Mackill learned that the secret to FR13A exceptionalism was mostly due to a single genetic locus or QTL on one of the rice chromosomes. He named this locus *SUB1* for "submergence tolerant."

With the chromosomal location of *SUB1* revealed, it was time to delve even deeper and identify the molecular nature of *SUB1.* What type of protein did it encode? How did the allele of *SUB1* found in FR13A allow the plant to cope with submergence? What is the physiological response that enables the plant to survive submergence?

To address these questions, molecular geneticists Pamela Ronald at the University of California, Davis, and Julia Bailey-Serres at the University of California, Riverside, joined the team. Working with Mackill, this expanded team zeroed in on the chromosome segment containing the *SUB1* QTL and determined that it encompasses a member of a class of genes called *ethylene response factors (ERFs).* ERF genes encode regulatory proteins that bind to regulatory elements in other genes and thereby regulate their expression. Thus, *SUB1* is a gene that regulates the expression of other genes. Moreover, they determined that the allele of *SUB1* in FR13A is switched on in response to submergence, while the allele of *SUB1* found in submergence-sensitive varieties is not switched on by submergence.

The next question was, how does switching on *SUB1* enable FR13A to survive complete submergence? To answer this question, let's review how ordinary rice plants respond to submergence. When a plant is completely submerged, oxygen levels in its cells drop to a low level, and the concentration of ethylene, a plant hormone, in the cells increases. Ethylene signals the plant to escape submergence by elongating its leaves and stems to keep its "head" above water. This *escape strategy* works fine as long as the water is not so deep that the plant fails to grow enough to position its stems and leaves above the flood waters. If the flood waters are too deep, then the plant cannot grow enough to escape. As a plant in such deeply flooded circumstances grows to escape the flood water, it uses up all its energy reserves (carbohydrates), becomes spindly and weak, and eventually dies.

How does the FR13A variety manage to survive submergence while many other types of rice cannot? FR13A has a different strategy that could be called *sit tight.* In response to complete submergence, rather than attempt rapid growth to escape the flood, an FR13A plant using the sit-tight strategy becomes quiescent. It stops the elongation growth response, thereby preventing itself from burning up all its reserve carbohydrates and becoming weak and spindly. With the sit-tight strategy, a plant can remain in a quiescent, submerged state for up to two weeks and then emerge healthy and resume normal growth when the flood waters recede.

The sit-tight strategy of FR13A is controlled by *SUB1,* which acts as the master switch or regulatory gene to activate this strategy. When the flood waters rise, the concentration of the plant hormone ethylene increases in plant cells. Because *SUB1* is an ERF, it is switched on in response to the elevated ethylene levels. Then, the protein that *SUB1* encodes orchestrates the plant's response by switching on (or off) a battery of genes involved in plant growth and metabolism. In FR13A plants that become submerged, genes involved in stem and leaf elongation as part of the escape strategy are switched off, as are genes involved in mobilizing the energy reserves (carbohydrates) needed to fuel the escape strategy. Using the tools of molecular genetics and genomics such as DNA microarrays (see Chapters 10 and 14), the rice team was able to decipher the extensive catalog of genes controlling

Flood-intolerant and flood-tolerant rice

FIGURE 1-20 An Indian farmer with rice variety *Swarna* that is not tolerant to flooding (*left*) compared to variety *Swarna-sub1* that is tolerant (*right*). This field was flooded for 10 days. The photo was taken 27 days after the flood waters receded. [*Ismail et al., "The contribution of submergence-tolerant (Sub 1) rice varieties to food security in flood-prone rainfed lowland areas in Asia,"* Field Crops Research *152, 2013, 83–93, © Elsevier.*]

organ elongation, carbon metabolism, flowering, and photosynthesis that are regulated by *SUB1* to achieve the sit-tight response.

With the basic genetics of *SUB1* elucidated, it was time to put this knowledge to work. The team repeated Mackill's early breeding work to transfer the flood tolerance into a superior variety. Now, however, since they knew the precise location of *SUB1* on one of the chromosomes, they could transfer it into a superior variety with surgical precision. This precision is important because it enabled the team to avoid transferring other undesirable genes at the same time. For this project, they worked with a submergence-intolerant, but superior, Indian variety, called *Swarna,* which is widely grown and favored by farmers. The new line they created is called *Swarna-Sub1,* and it has lived up to expectations. Field trials showed a striking difference in plant survival and yield between *Swarna* and *Swarna-Sub1* when there is complete submergence (Figure 1-20). As shown in Figure 1-21, *Swarna-Sub1* provides higher yield than the original *Swarna* under all different levels of flooding. In various trials, the *SUB1* improved yield between 1 to 3 tons of grain per hectare.

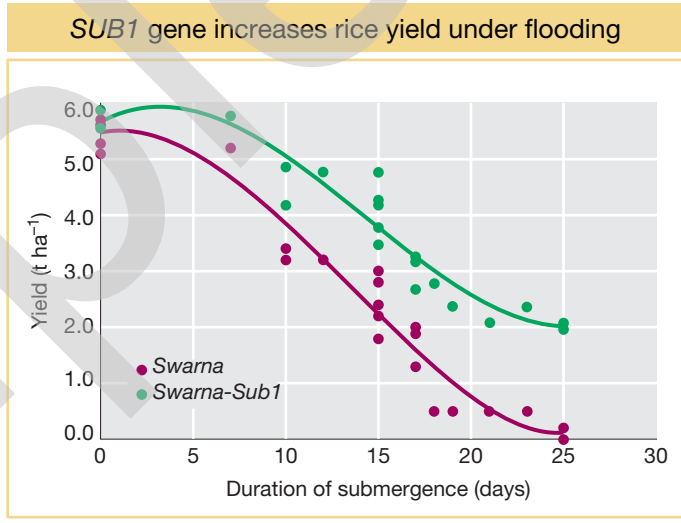


FIGURE 1-21 Yield comparison between variety *Swarna* that is not tolerant to flooding (purple circles) and variety *Swarna-Sub1* that is tolerant (green circles). Yield in tons per hectare (y-axis) versus duration of flooding in days (x-axis). [*Data from Ismail et al., "The contribution of submergence-tolerant (Sub 1) rice varieties to food security in flood-prone rainfed lowland areas in Asia,"* Field Crops Research *152, 2013, 83–93.*]

With the support and sponsorship of international research organizations, governmental agencies, and philanthropies, *Swarna-Sub1* and other superior varieties carrying the *SUB1* allele from FR13A have now been distributed to farmers. In 2008, only 700 farmers were growing *SUB1* enhanced rice, but by 2012, that number had grown to 3.8 million farmers. By 2014, the number of farmers growing rice with *SUB1* should climb to 5 million, adding considerably to food security among some of the world's poorest farmers.

In the long run, the impact of the *SUB1* research may not be limited to rice. Many crops are subjected to damaging floods that reduce yields or destroy the crop altogether. The genetic research on *SUB1* has provided a deep understanding of the molecular genetics of how plants respond to flooding. With this knowledge, it will be possible to manipulate the genomes of other crop plants so that they too can withstand getting their feet a little too wet.

KEY CONCEPT Genetics and genomics are playing a leading role in improving crop plants. The basic principles of genetics that you will learn during your genetics course are the foundation for these advances.

Recent evolution in humans

One goal of genetics is to understand the rules that govern how genes and the information they encode change over the generations within populations. The genes in populations change over time for several different reasons. For example, as we have seen, mutation in the germline can cause a new gene variant or allele to occur in the next generation that was not present in the current generation. Another factor is *natural selection,* which was first described by Charles Darwin. Briefly, if individuals with a certain gene variant contribute more offspring to the next generation than individuals who lack that variant, then the frequency of that variant will rise over time in the population. The last three chapters of the text focus on rules governing the transmission of genes from one generation to the next within populations.

Over the past decade, evolutionary geneticists have described in remarkable detail how genetic changes have enabled human populations to adapt to the

conditions of life on different parts of the globe. This work revealed that three factors have been particularly powerful in shaping the types of gene variants that occur in different human populations. These factors are (1) pathogens such as malaria or smallpox; (2) local climatic conditions including solar radiation, temperature, and altitude; and (3) diet, such as the relative amounts of meat, cereals, or dairy products eaten. In Chapter 20, you'll learn how a genetic variant in the hemoglobin gene has enabled people in Africa to adapt to the ravages of malaria. Let's look briefly at examples of genetic adaptations to climate and diet. We'll start with a case of human adaptation to life at high altitude.

Adaptation to high altitude In their effort to colonize the Andes mountains of South America, Spanish colonists established towns high up in the mountains near the settlements of the native peoples. Soon they realized something was wrong. Spanish parents were not producing children. At Potosi, Bolivia, which is situated 4000 meters above sea level, it was 53 years after the founding of the town before the first child was born to Spanish parents. As noted by the Spanish priest Father Cobo, "The Indians are healthiest and where they multiply the most prolifically is in these same cold air-tempers, which is quite the reverse of what happens to the children of the Spaniards, most of whom when born in such regions do not survive."[2] Unlike the Andean natives, the Spanish were experiencing chronic mountain sickness (CMS), a condition caused by their inability to obtain enough oxygen from the thin air of the mountains.

Since early observations like these, geneticists have invested much effort into the study of human adaptation to high altitude in South America, Tibet, and Ethiopia. What enables the natives of these regions to flourish while lowlanders who move to high elevations suffer the grave health consequences of CMS? Let's look at the case in Tibet, where the Tibetan highlanders live at altitudes up to 4000 meters above sea level (Figure 1-22). The high Tibetan Plateau was colonized by people about 3000 years ago, and the people who colonized Tibet are closely related to the modern Han Chinese. However, at high altitude, native Tibetans are far less likely than Han Chinese to experience CMS and conditions such as pulmonary hypertension and the associated formation of blood clots that underlie it.

To understand the genetics of how Tibetans adapted to life at high elevation, a research team led by Cynthia Beall of Case Western Reserve University compared Tibetans to Han Chinese at over 500,000 SNPs across the genome. Since Tibetans and Chinese are closely related, one expects each SNP variant to occur at about the same frequency in both groups. If the T variant of a SNP occurs at a frequency of 10 percent in Han Chinese, it should also be at about 10 percent in Tibetans. However, if the variant is associated with improved health at high elevation, its frequency would have risen among Tibetans over the many generations since they colonized the Tibetan Plateau, because Tibetans with this variant would have been healthier and have had more surviving children than those who lacked it. Charles Darwin's natural selection would be at work.

When the research team analyzed their SNP data, the SNPs in one gene stood out. The gene is called *EPAS1,* and some SNPs in it

Tibetans are genetically adapted to life at high elevation

FIGURE 1-22 A young Tibetan woman. Inset shows the location of Tibet in Asia. [*Stefan Auth/imagebroker/AGE Fotostock; (inset) Planet Observer/UIG/Getty Images.*]

[2] V. J. Vitzthum, "The home team advantage: Reproduction in women indigenous to high altitude," *Journal of Experimental Biology* 204, 2001, 3141–3150.

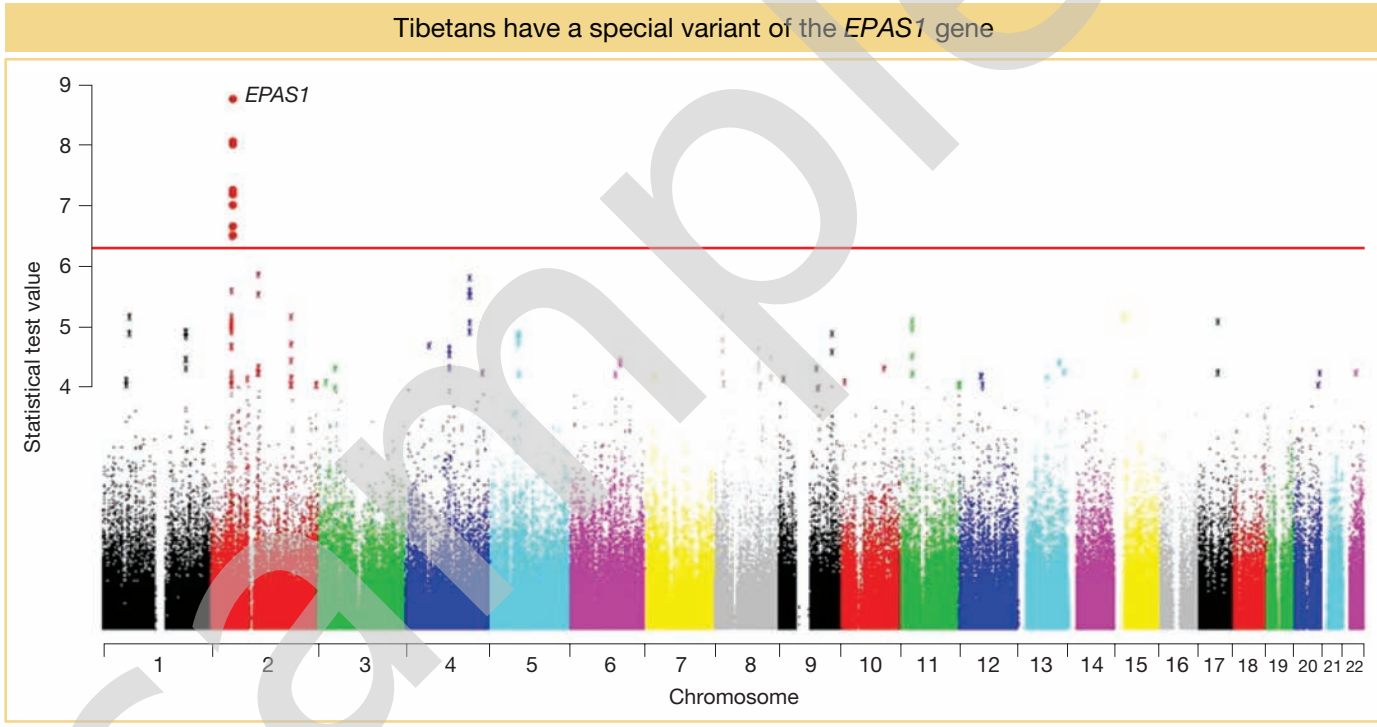


FIGURE 1-23 Twenty-two human chromosomes are arrayed from left to right. The y-axis shows results from a statistical test of whether there is a significant difference in SNP frequency between Tibetans and Han Chinese. Each small dot represents one of the SNPs that was tested. SNPs above the horizontal red line are significantly different. Only the SNPs in the *EPAS1* gene show a significant difference. [*C. Beall et al.* Proceedings of the National Academy of Sciences USA, *107, 25, 2010, 11459–11464, Fig. 1.*]

occur at very different frequencies in Tibetans (87 percent) and Han Chinese (9 percent). Their results are shown in Figure 1-23. In this figure, the human chromosomes, numbered 1 through 22, are along the x-axis, and a measure of the difference in SNP variant frequency between Tibetans and Chinese is on the y-axis. Each dot represents a SNP. SNPs that fall above the horizontal red line are those for which the frequency difference between Tibetans and Han Chinese is so large that the gene near these SNPs must have provided some advantage to people who colonized the Tibetan Plateau. The SNPs in *EPAS1* fall above this line.

These results suggest that Tibetans have a special variant of *EPAS1* that helps them adapt to life at high elevation. To understand this better, let's first review what is known about *EPAS1.* This gene regulates the number of red blood cells (RBCs) that our bodies produce. Moreover, it regulates the number of RBCs in response to the level of oxygen in our tissues. When oxygen levels in our tissues are low, *EPAS1* signals the body to produce more RBCs.

Why does *EPAS1* direct our bodies to produce more RBCs when the oxygen levels in our tissues are low? The *EPAS1* response to low oxygen may be how our bodies normally respond to anemia (too few red blood cells). People with low RBC counts get too little oxygen in their tissues, and so *EPAS1* could signal the body to make more RBCs to correct anemia. This mechanism could explain why people who live at low elevation need the *EPAS1* gene.

Now, let's think about how a person from low elevation would respond if they move to high elevation. Because of the thin air at high elevation, their tissues would get less oxygen. If their bodies interpreted low oxygen due to thin air as a sign of anemia, then *EPAS1* would try to correct the problem by signaling their

body to make more RBCs. However, since they are not anemic and already have enough RBCs, their blood would become overloaded with RBCs. Too many RBCs can cause pulmonary hypertension and the formation of blot clots, the conditions underlying CMS.

Finally, how could a new variant of *EPAS1* have helped Tibetans avoid CMS and adapt to high elevation? The answer to this question is not known, and it is now being actively investigated, but here is one hypothesis. Unlike lowlanders, Tibetans maintain relatively normal levels of RBCs at high elevation, and they have a lower risk of blot clot formation and pulmonary hypertension than lowlanders who move to high elevation. Thus, the Tibetan version of *EPAS1* may no longer cause the overproduction of RBCs at high elevation, while providing another mechanism to cope with the thin air. The Tibetan variant of *EPAS1* helps them live at high elevation without suffering from CMS.

Lactose tolerance Before the invention of agriculture about 10,000 to 12,000 years ago, human populations subsisted on foods harvested from nature by hunting wild animals and gathering wild fruits and vegetables. At that time, no human populations used dairy products. Cattle were yet to be domesticated, and methods for milking cows were not yet invented. Children nursed on mother's milk, but as they aged, the gene that encodes the enzyme *lactase,* which enables children to digest milk sugar (lactose), was switched off. Once weaned, a child in pre-agricultural societies no longer needed the lactase enzyme, and so the lactase gene had a "switch" or regulatory element that turned it off during late childhood.

With the origin of agriculture, cattle were domesticated from wild aurochs. The early farmers may have kept cattle as a source of meat at first. After milking was invented, milk offered another source of food. But there was a problem. Although children in these ancient societies could digest milk sugar, the adults could not. Adults could consume milk, but since they could not digest the lactose, they would experience bloating, cramps, and diarrhea. Adults who experience these symptoms from drinking milk are *lactose intolerant.* Importantly, because they could not digest milk sugar, they were not utilizing this source of nutrition.

In ancient societies, where food could be scarce at times, the difference between life and death could hinge on making the best use of all available food sources. Yet, because the lactase gene is switched off in adults, adults could not digest milk sugar.

Some human populations have lactase gene variants expressed in adults

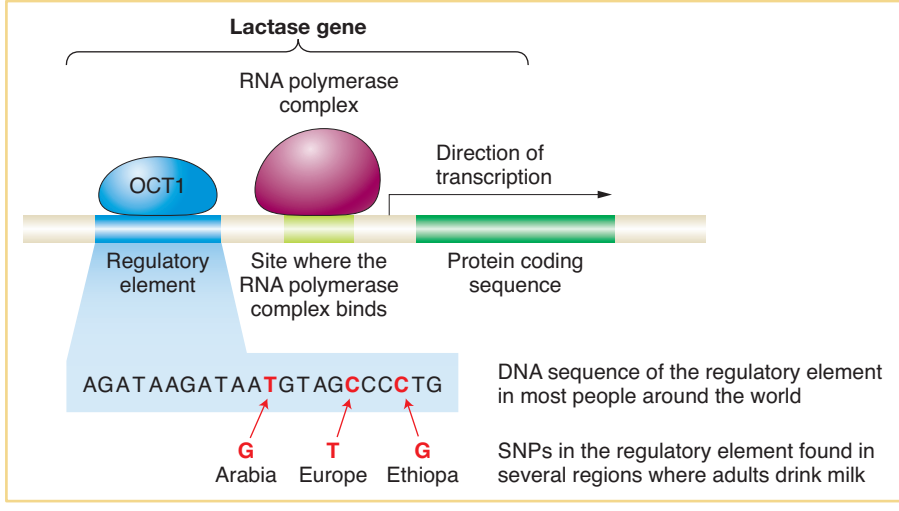


FIGURE 1-24 Simplified diagram of the lactase gene showing a regulatory element and protein coding region. OCT1 is a protein thought to regulate expression of the lactase gene. SNP variants in the regulatory element are found in some parts of the world. These SNPs are associated with OCT1 binding to the element and expression of the lactase gene in adults.

Now, suppose a new mutation entered the population and that this mutation allowed the lactase gene to be expressed in adults. Adults with this new mutation or variant could then benefit from drinking milk in a way that adults who lacked this variant could not. Such a benefit could increase their chances to survive and have children, and over time the variant that provides lactase persistence into adulthood would become more common in the population.

The scenario just described is what appears to have happened during human history in several areas of the world where people kept cattle (or camels) and used them for milk. It happened in Europe, the Middle East, and Africa. In Europe, some people have a variant of the lactase gene that has a "T" at a particular SNP, whereas people from other regions of the world have a "C" at this SNP. Recently, geneticists discovered that the "T" appears to be located in a regulatory element that controls when the lactase gene is turned on (Figure 1-24). People with the "T" variant have persistent expression of the lactase gene into adulthood, whereas people with the "C" variant have their lactase gene switched off after childhood. The "T" seems to enable a regulatory protein called OCT1 to bind near the lactase gene and thereby cause its expression in adults. Other variants that have the same effect appear to have arisen independently in the Middle East and Africa.

As shown in Figure 1-25, in northern Europe where cattle farming and dairy consumption are prominent, both lactase persistence and the "T" lactase variant that produce it are common, while these features are much less common in southern Europe. Geneticists infer that the early cattle farmers of northern Europe who had the "T" variant benefited from milk consumption, enabling them to survive and produce more offspring, and so this variant became more common in the population over time. Today, the "T" variant is at a frequency of 90 percent in northern Europe. Since milk was not as important a part of the diet in southern Europe, the T variant offered no special benefit and thus remained at a lower frequency (about 10 percent).

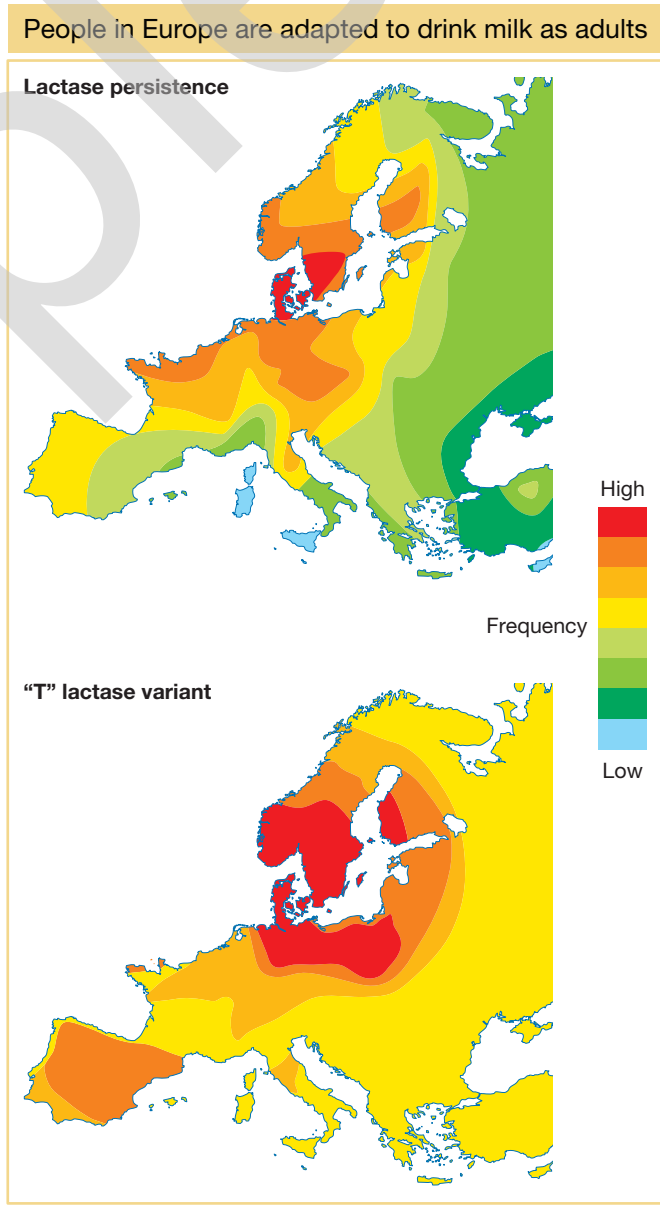


FIGURE 1-25 (a) Frequency in Europe of lactase persistence, the expression of the lactase enzyme in adults. (b) Frequency in Europe of the T variant in the lactase gene that appears to control lactase persistence. [*(a) Adapted from Y. Itan et al.,* BMC Evolutionary Biology *10, 2010, 36. (b) Adapted from A. Beja-Pereira et al.* Nature Genetics *35, 2003, 311–313.*]

These two examples highlight how human populations have evolved in recent times in response to the conditions of life such as the available food and climate. In the last three chapters of this text, you will learn the theory and methods used by geneticists to understand how populations evolve in response to their environment. You'll learn how SNP data are gathered, how frequencies of variants are calculated, and how comparisons are made to understand the forces that have influenced the types of gene variants that occur in different populations. Through this type of analysis, evolutionary geneticists have learned a vast amount about how different species of plants, animals, fungi, and microbes have evolved and continue to evolve in response to the conditions in which they live.

KEY CONCEPT Evolutionary genetics provides the tools to document how gene variants that provide a beneficial effect can rise in frequency in a population and make individuals in the population better adapted to the environment in which they live.

SUMMARY

As you begin your study of genetics, imagine yourself as a person at halftime on an amazing journey of discovery. The last 100 years have witnessed a remarkable revolution in human knowledge about how biological systems are put together and how they work. Genetics has been at the epicenter of that revolution. Genetic analysis has answered many fundamental questions about the transmission of genetic information within families, inside cells, and over the eons of evolutionary time. Yet, as you will learn, the discovery process in genetics has never been more dynamic and the pace of growth in knowledge never greater. Unanswered questions abound.

- How do all the genes in the genome work together to transform a fertilized egg into an adult organism?
- How do cells manage to seamlessly orchestrate the incredibly complex array of interacting genes and biochemical reactions that are found within them?
- How do genetic variants at hundreds or even thousands of genes control the yield of crop plants?
- How can genetics guide both the prevention and treatment of cancer, autism, and other diseases?
- How do genes give humans the capacity for language and consciousness?

Genetic analysis over the next 100 years promises to help answer many questions like these.

KEY TERMS

adenine (A) (p. 7)
alleles (p. 4)
blending theory (p. 2)
chromosome theory (p. 5)
codon (p. 10)
complementary (p. 7)
cytosine (C) (p. 7)
DNA polymerase (p. 12)
DNA sequencing (p. 13)
dominant (p. 4)
DNA replication (p. 9)
gametes (p. 4)
gene (p. 4)
gene expression (p. 7)
genetically modified organism (GMO) (p. 13)
genetics (p. 5)
genomics (p. 13)
guanine (G) (p. 7)
ligase (p. 12)
messenger RNA (mRNA) (p. 10)
model organism (p. 10)
multifactorial hypothesis (p. 7)
nuclease (p. 12)
one-gene–one-enzyme hypothesis (p. 7)
point mutation (p. 17)
quantitative trait locus (QTL) (p. 21)
regulatory element (p. 7)
single nucleotide polymorphism (SNP) (p. 16)
somatic cells (p. 4)
thymine (T) (p. 7)
transcription (p. 10)
transformation (p. 13)
translation (p. 10)

PROBLEMS

Most of the problems are also available for review/grading through LaunchPad http://www.whfreeman.com/launchpad/iga11e.

WORKING WITH THE FIGURES

1. If the white-flowered parental variety in Figure 1-3 were crossed to the first-generation hybrid plant in that figure, what types of progeny would you expect to see and in what proportions?

2. In Mendel's 1866 publication as shown in Figure 1-4, he reports 705 purple-flowered (violet) offspring and 224 white-flowered offspring. The ratio he obtained is 3.15:1 for purple:white. How do you think he explained the fact that the ratio is not exactly 3:1?

3. In Figure 1-6, the students have 1 of 15 different heights, plus there are two height classes (4′11″ and 5′0″) for which there are no observed students. That is a total of 17 height classes. If a single Mendelian gene can account for only two classes of a trait (such as purple or white flowers), how many Mendelian genes would be minimally required to explain the observation of 17 height classes?

4. Figure 1-7 shows a simplified pathway for arginine synthesis in *Neurospora*. Suppose you have a special strain of *Neurospora* that makes citrulline but not arginine. Which gene(s) are likely mutant or missing in your special strain? You have a second strain of *Neurospora* that makes neither citrulline nor arginine but does make ornithine. Which gene(s) are mutant or missing in this strain?

5. Consider Figure 1-8a.

a. What do the small, blue spheres represent?

b. What do the brown slabs represent?

c. Do you agree with the analogy that DNA is structured like a ladder?

6. In Figure 1-8b, can you tell if the number of hydrogen bonds between adenine and thymine is the same as that between cytosine and guanine? Do you think that a DNA molecule with a high content of A + T would be more stable than one with high content of G + C?

7. Which of three major groups (domains) of life in Figure 1-11 is not represented by a model organism?

8. Figure 1-13b shows the human chromosomes in a single cell. The green dots show the location of a gene called *BAPX1.* Is the cell in this figure a sex cell (gamete)? Explain your answer.

9. Figure 1-15 shows the family tree, or pedigree, for Louise Benge (Individual VI-1) who suffers from the disease ACDC because she has two mutant copies of the CD73 gene. She has four siblings (VI-2, VI-3, VI-4, and VI-5) who have this disease for the same reason. Do all of the 10 children of Louise and her siblings have the same number of mutant copies of the CD73 gene, or might this number be different for some of the 10 children?

BASIC PROBLEMS

10. Below is the sequence of a single strand of a short DNA molecule. On a piece of paper, rewrite this sequence and then write the sequence of the complementary strand below it.

GTTCGCGGCCGCGAAC

Comparing the top and bottom strands, what do you notice about the relationship between them?

11. Mendel studied a *tall* variety of pea plants with stems that are 20 cm long and a *dwarf* variety with stems that are only 12 cm long.

a. Under blending theory, how long would you expect the stems of first and second hybrids to be?

b. Under Mendelian rules and assuming stem length is controlled by a single gene, what would you expect to observe in the second-generation hybrids if all the first-generation hybrids were tall?

12. If a DNA double helix that is 100 base pairs in length has 32 adenines, how many cytosines, guanines, and thymines must it have?

13. The complementary strands of DNA in the double helix are held together by hydrogen bonds: G ≡ C or A = T. These bonds can be broken (denatured) in aqueous solutions by heating to yield two single strands of DNA (see Figure 1-13a). How would you expect the relative amounts of GC versus AT base pairs in a DNA double helix to affect the amount of heat required to denature it? How would you expect the length of a DNA double helix in base pairs to affect the amount of heat required to denature it?

14. The figure at the bottom of the page shows the DNA sequence of a portion of one of the chromosomes from a trio (mother, father, and child). Can you spot any new point mutations in the child that are not in either parent? In which parent did the mutation arise?

CHALLENGING PROBLEMS

15. **a.** There are three nucleotides in each codon, and each of these nucleotides can have one of four different bases. How many possible unique codons are there?

b. If DNA had only two types of bases instead of four, how long would codons need to be to specify all 20 amino acids?

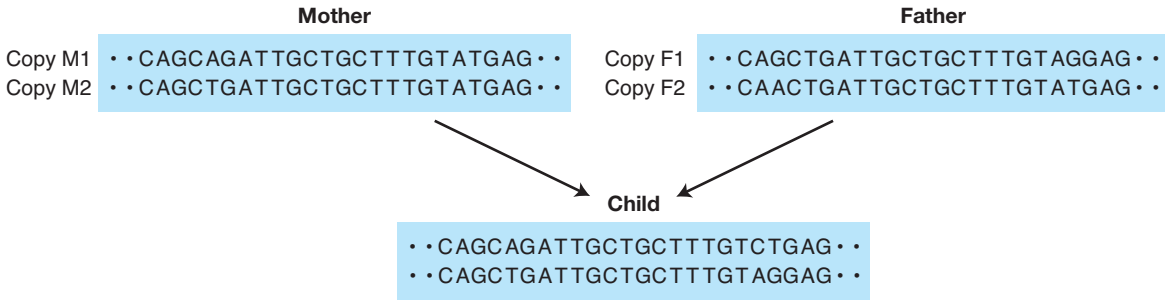
16. Fathers contribute more new point mutations to their children than mothers. You may know from general biology that people have sex chromosomes—two X chromosomes in females and an X plus a Y chromosome in males. Both sexes have the autosomes (A's).

a. On which type of chromosome (A, X, or Y) would you expect the genes to have the greatest number of new mutations per base pair over many generations in a population? Why?

b. On which type of chromosome would you expect the least number of new mutations per base pair? Why?

c. Can you calculate the expected number of new mutations per base pair for a gene on the X and Y chromosomes for every one new mutation in a gene on an autosome if the mutation rate in males is twice that in females?

17. For young men of age 20, there have been 150 rounds of DNA replication during sperm production as compared



to only 23 rounds for a woman of age 20. That is a 6.5-fold greater number of cell divisions and proportionately greater opportunity for new point mutations. Yet, on average, 20-year-old men contribute only about twice as many new point mutations to their offspring as do women. How can you explain this discrepancy?

18. In computer science, a bit stores one of two states, 0 or 1. A byte is a group of 8 bits that has $2^8 = 256$ possible states. Modern computer files are often megabytes (10^6 bytes) or even gigabytes (10^9 bytes) in size. The human genome is approximately 3 billion base pairs in size. How many nucleotides are needed to encode a single byte? How large of a computer file would it take to store the same amount of information as a single human genome?

19. The human genome is approximately 3 billion base pairs in size.

a. Using standard $8.5'' \times 11''$ paper with one-inch margins, a 12-point font size, and single-spaced lines, how many sheets of paper printed on one side would be required to print out the human genome?

b. A ream of 500 sheets of paper is about 5 cm thick. How tall would the stack of paper with the entire human genome be?

c. Would you want a backpack, shopping cart, or a semitrailer truck to haul around this stack?

CHAPTER 2

Single-Gene Inheritance

The monastery of the father of genetics, Gregor Mendel. A statue of Mendel is visible in the background. Today, this part of the monastery is a museum, and the curators have planted red and white begonias in a grid that graphically represents the type of inheritance patterns obtained by Mendel with peas. [*Anthony Griffiths.*]

OUTLINE

LEARNING OUTCOMES

After completing this chapter, you will be able to

- Discover a set of genes affecting a specific biological property of interest, by observing single-gene inheritance ratios of mutants affecting that property.
- In the progeny of controlled crosses, recognize phenotypic ratios diagnostic of single-gene inheritance (1 : 1 in haploids, and 3 : 1, 1 : 2 : 1, and 1 : 1 in diploids).
- Explain single-gene inheritance ratios in terms of chromosome behavior at meiosis.
- Predict phenotypic ratios among descendants from crosses of parents differing at a single gene.
- Propose reasonable hypotheses to explain dominance and recessiveness of specific alleles at the molecular level.
- Apply the rules of single-gene inheritance to pedigree analysis in humans, and recognize patterns diagnostic of autosomal dominant, autosomal recessive, X-linked dominant, and X-linked recessive conditions.
- Calculate risk of descendants inheriting a condition caused by a mutant allele in one or more specific ancestors.

What kinds of research do biologists do? One central area of research in the biology of all organisms is the attempt to understand how an organism develops from a fertilized egg into an adult—in other words, what makes an organism the way it is. Usually, this overall goal is broken down into the study of individual biological **properties** such as the development of plant flower color, or animal locomotion, or nutrient uptake, although biologists also study some general areas such as how a cell works. How do geneticists analyze biological properties? The genetic approach to understanding any biological property is to find the subset of genes in the genome that influence that property, a process sometimes referred to as **gene discovery.** After these genes have been identified, their cellular functions can be elucidated through further research.

There are several different types of analytical approaches to gene discovery, but one widely used method relies on the detection of *single-gene inheritance patterns,* and that is the topic of this chapter.

All of genetics, in one aspect or another, is based on heritable variants. The basic approach of genetics is to compare and contrast the properties of variants, and from these comparisons make deductions about genetic function. It is similar to the way in which you could make inferences about how an unfamiliar machine works by changing the composition or positions of the working parts, or even by removing parts one at a time. Each variant represents a "tweak" of the biological machine, from which its function can be deduced.

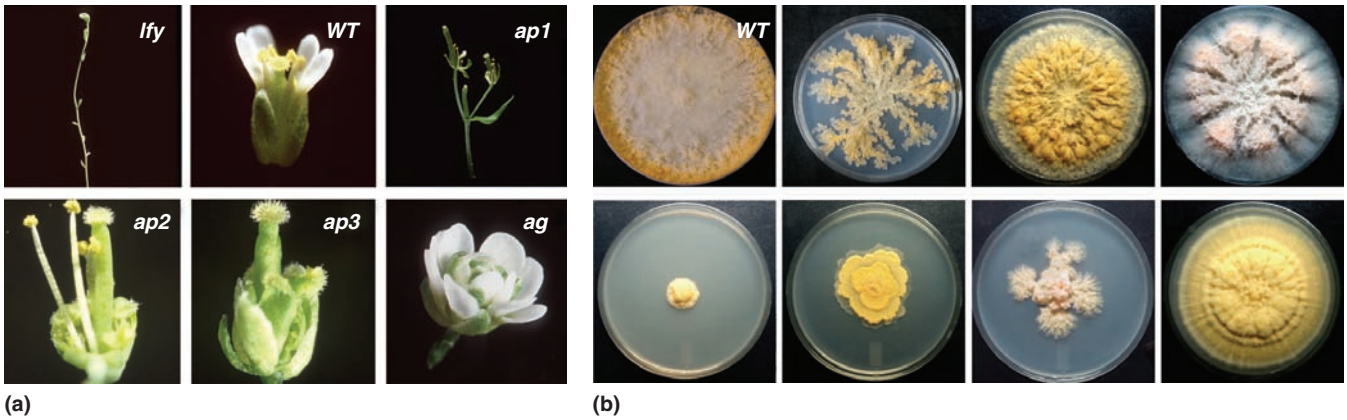
In genetics, the most common form of any property of an organism is called the **wild type,** that which is found "in the wild," or in nature. The heritable variants observed in an organism that differs from the wild type are **mutants,** individual organisms having some abnormal form of a property. As examples, the wild type and some mutants in two model organisms are shown in Figure 2-1. The alternative forms of the property are called **phenotypes.** In this analysis we distinguish a wild-type phenotype and a mutant phenotype.

Compared to wild type, mutants are rare. We know that they arise from wild types by a process called **mutation,** which results in a heritable change in the DNA of a gene. The changed form of the gene is also called a mutation. Mutations are not always detrimental to an organism; sometimes they can be advantageous, but most often they have no observable effect. A great deal is known about the mechanisms of mutation (see Chapter 16), but generally it can be said that they arise from mistakes in cellular processing of DNA.

Most natural populations also show **polymorphisms,** defined as the coexistence of two or more reasonably common phenotypes of a biological property,

FIGURE 2-1 These photographs show the range of mutant phenotypes typical of those obtained in the genetic dissection of biological properties. These cases are from the dissection of floral development in *Arabidopsis thaliana* (a) and hyphal growth in *Neurospora crassa,* a mold (b). *WT* = wild type. [*(a) George Haughn; (b) Anthony Griffiths/Olivera Gavric.*]

Genetic analysis begins with mutants



such as the occurrence of both red- and orange-fruited plants in a population of wild raspberries. Genetic analysis can (and does) use polymorphisms, but polymorphisms have the disadvantage that they generally do not involve the specific property of interest to the researcher. Mutants are much more useful because they allow the researcher to zero in on any property.

Simply stated, the general steps of functional analysis by gene discovery are as follows:

1. Amass mutants affecting the biological property of interest.

2. **Cross** (mate) the mutants to wild type to see if their descendants show ratios of wild to mutant that are characteristic of single-gene inheritance.

3. Deduce the functions of the gene at the molecular level.

4. Deduce how the gene interacts with other genes to produce the property in question.

Of these steps, only 1 and 2 will be covered in the present chapter.

Gene discovery starts with a "hunt" to amass mutants in which the biological function under investigation is altered or destroyed. Even though mutants are individually rare, there are ways of enhancing their recovery. One widely used method is to treat the organism with radiation or chemicals that increase the mutation rate. After treatment, the most direct way to identify mutants is to visually *screen* a very large number of individuals, looking for a chance occurrence of mutants in that population. Also, various selection methods can be devised to enrich for the types sought.

Armed with a set of mutants affecting the property of interest, one hopes that each mutant represents a lesion in one of a set of genes that control the property. Hence, the hope is that a reasonably complete gene pathway or network is represented. However, not all mutants are caused by lesions within one gene (some have far more complex determination), so first each mutant has to be tested to see if indeed it is caused by a single-gene mutation.

The test for single-gene inheritance is to mate individuals showing the mutant property with wild-type and then analyze the first and second generation of descendants. As an example, a mutant plant with white flowers would be crossed to the wild type showing red flowers. The progeny of this cross are analyzed, and then they themselves are interbred to produce a second generation of descendants. In each generation, the diagnostic ratios of plants with red flowers to those with white flowers will reveal whether a single gene controls flower color. If so, then by inference, the wild type would be encoded by the wild-type form of the gene and the mutant would be encoded by a form of the same gene in which a mutation event has altered the DNA sequence in some way. Other mutations affecting flower color (perhaps mauve, blotched, striped, and so on) would be analyzed in the same way, resulting overall in a set of defined "flower-color genes." The use of mutants in this way is sometimes called **genetic dissection,** because the biological property in question (flower color in this case) is picked apart to reveal its underlying genetic program, not with a scalpel but with mutants. Each mutant potentially identifies a separate gene affecting that property.

After a set of key genes has been defined in this way, several different molecular methods can be used to establish the functions of each of the genes. These methods will be covered in later chapters. Hence, genetics has been used to define the set of gene functions that interact to produce the property we call flower color (in this example).

This type of approach to gene discovery is sometimes called **forward genetics,** a strategy to understanding biological function starting with random single-gene mutants and ending with their DNA sequence and biochemical function. (We shall

see **reverse genetics** at work in later chapters. In brief, it starts with genomic analysis at the DNA level to identify a set of genes as candidates for encoding the biological property of interest, then induces mutants targeted specifically to those genes, and then examines the mutant phenotypes to see if they indeed affect the property under study.)

> **KEY CONCEPT** The genetic approach to understanding a biological property is to discover the genes that control it. One approach to gene discovery is to isolate mutants and check each one for single-gene inheritance patterns (specific ratios of normal and mutant expression of the property in descendants).

Gene discovery is important not only in experimental organisms but also in applied studies. One crucial area is in agriculture, where gene discovery can be used to understand a desirable commercial property of an organism, such as its protein content. Human genetics is another important area: to know which gene functions are involved in a specific disease or condition is useful information in finding therapies.

The rules for single-gene inheritance were originally elucidated in the 1860s by the monk Gregor Mendel, who worked in a monastery in the town of Brno, now part of the Czech Republic. Mendel's analysis is the prototype of the experimental approach to single-gene discovery still used today. Indeed, Mendel was the first person to discover any gene! Mendel did not know what genes were, how they influenced biological properties, or how they were inherited at the cellular level. Now we know that genes work through proteins, a topic that we shall return to in later chapters. We also know that single-gene inheritance patterns are produced because genes are parts of chromosomes, and chromosomes are partitioned very precisely down through the generations, as we shall see later in the chapter.

2.1 Single-Gene Inheritance Patterns

Recall that the first step in genetic dissection is to obtain variants that differ in the property under scrutiny. With the assumption that we have acquired a collection of relevant mutants, the next question is whether each of the mutations is inherited as a single gene.

Mendel's pioneering experiments

The first-ever analysis of single-gene inheritance as a pathway to gene discovery was carried out by Gregor Mendel. His is the analysis that we shall follow as an example. Mendel chose the garden pea, *Pisum sativum,* as his research organism. The choice of organism for any biological research is crucial, and Mendel's choice proved to be a good one because peas are easy to grow and breed. Note, however, that Mendel did not embark on a hunt for mutants of peas; instead, he made use of mutants that had been found by others and had been used in horticulture. Moreover, Mendel's work differs from most genetics research undertaken today in that it was not a genetic dissection; he was not interested in the properties of peas themselves, but rather in the way in which the hereditary units that influenced those properties were inherited from generation to generation. Nevertheless, the laws of inheritance deduced by Mendel are exactly those that we use today in modern genetics in identifying single-gene inheritance patterns.

Mendel chose to investigate the inheritance of seven properties of his chosen pea species: pea color, pea shape, pod color, pod shape, flower color, plant height, and position of the flowering shoot. In genetics, the terms **character** and **trait** are used more or less synonymously; they roughly mean "property." For each of these seven characters, he obtained from his horticultural supplier two lines that showed distinct and contrasting phenotypes. These contrasting phenotypes are illustrated

in Figure 2-2. His results were substantially the same for each character, and so we can use one character, pea seed color, as an illustration. All of the lines used by Mendel were **pure lines,** meaning that, for the phenotype in question, all offspring produced by matings within the members of that line were identical. For example, within the yellow-seeded line, all the progeny of any mating were yellow seeded.

Mendel's analysis of pea heredity made extensive use of crosses. To make a cross in plants such as the pea, pollen is simply transferred from the anthers of one plant to the stigmata of another. A special type of mating is a **self** (self-pollination), which is carried out by allowing pollen from a flower to fall on its own stigma. Crossing and selfing are illustrated in Figure 2-3. The first cross made by Mendel mated plants of the yellow-seeded lines with plants of the green-seeded lines. In his overall breeding program, these lines constituted the **parental generation,** abbreviated **P.** In *Pisum sativum,* the color of the seed (the pea) is determined by the seed's own genetic makeup; hence, the peas resulting from a cross are effectively progeny and can be conveniently classified for phenotype without the need to grow them into plants. The progeny peas from the cross between the different pure lines were found to be all yellow, no matter which parent (yellow or green) was used as male or female. This progeny generation is called the **first filial generation,** or **F_1.** The word *filial* comes from the Latin words *filia* (daughter) and *filius* (son). Hence, the results of these two reciprocal crosses were as follows, where × represents a cross:

female from yellow line × male from green line → F_1 peas all yellow

female from green line × male from yellow line → F_1 peas all yellow

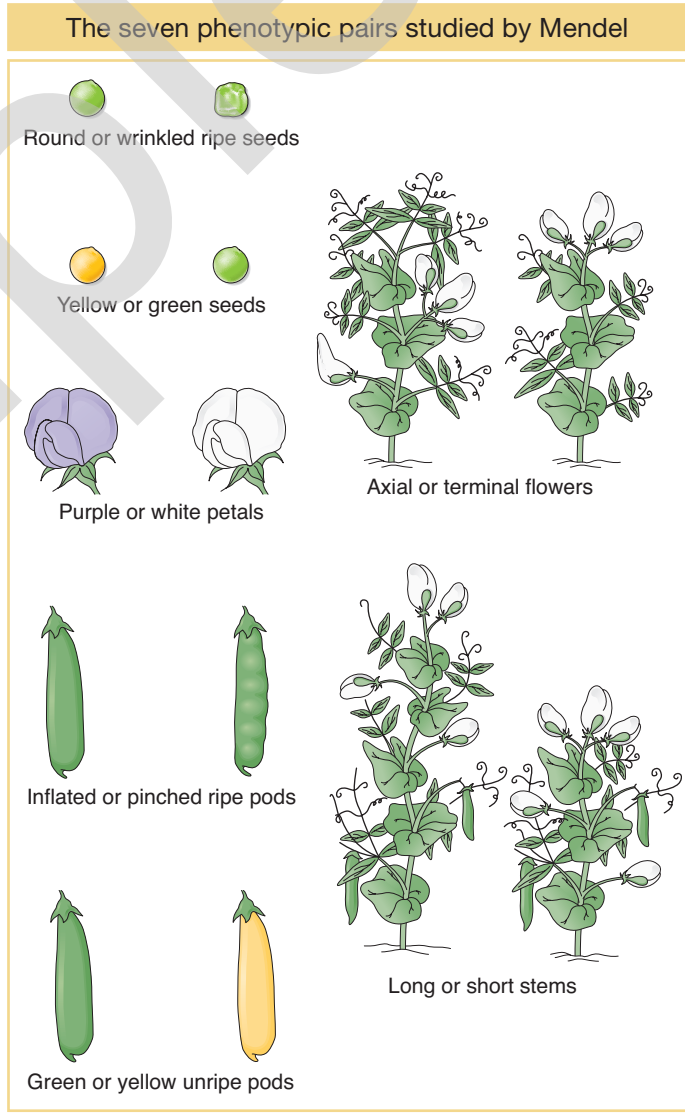


FIGURE 2-2 For each character, Mendel studied two contrasting phenotypes.

The results observed in the descendants of both reciprocal crosses were the same, and so we will treat them as one cross. Mendel grew F_1 peas into plants, and he selfed these plants to obtain the **second filial generation,** or **F_2.** The F_2 was composed of 6022 yellow peas and 2001 green peas. In summary,

yellow F_1 × yellow F_1 → F_2 comprised of	6022	yellow
	2001	green
Total	8023	

Mendel noted that this outcome was very close to a mathematical ratio of three-fourths (75%) yellow and one-fourth (25%) green. A simple calculation shows us that 6022/8023 = 0.751 or 75.1%, and 2001/8023 = 0.249 or 24.9%. Hence, there was a 3:1 ratio of yellow to green. Interestingly, the green phenotype, which had disappeared in the F_1, had reappeared in one-fourth of the F_2 individuals, showing that the genetic determinants for green must have been present in the yellow F_1, although unexpressed.

To further investigate the nature of the F_2 plants, Mendel selfed plants grown from the F_2 seeds. He found three different types of results. The plants grown from the F_2 green seeds, when selfed, were found to bear only green peas.

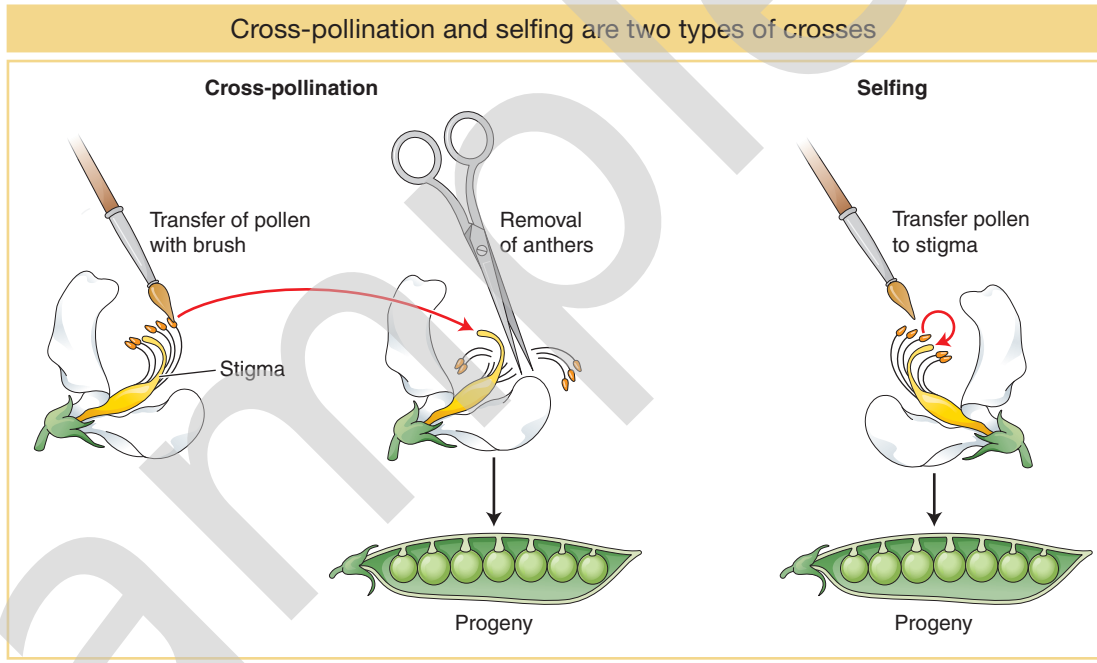


FIGURE 2-3 In a cross of a pea plant (*left*), pollen from the anthers of one plant is transferred to the stigma of another. In a self (*right*), pollen is transferred from the anthers to the stigmata of the same plant.

However, plants grown from the F_2 yellow seeds, when selfed, were found to be of two types: one-third of them were pure breeding for yellow seeds, but two-thirds of them gave mixed progeny: three-fourths yellow seeds and one-fourth green seeds, just as the F_1 plants had. In summary,

$\frac{1}{4}$ of the F_2 were green, which when selfed gave all greens

$\frac{3}{4}$ of the F_2 were yellow; of these $\frac{1}{3}$ when selfed gave all yellows

$\frac{2}{3}$ when selfed gave $\frac{3}{4}$ yellow and $\frac{1}{4}$ green

Hence, looked at another way, the F_2 was comprised of

$\frac{1}{4}$ pure-breeding greens

$\frac{1}{4}$ pure-breeding yellows

$\frac{1}{2}$ F_1-like yellows (mixed progeny)

Thus, the 3:1 ratio at a more fundamental level is a 1:2:1 ratio.

Mendel made another informative cross between the F_1 yellow-seeded plants and any green-seeded plant. In this cross, the progeny showed the proportions of one-half yellow and one-half green. In summary:

F_1 yellow × green $\rightarrow$ $\frac{1}{2}$ yellow

$\frac{1}{2}$ green

These two types of matings, the F_1 self and the cross of the F_1 with any green-seeded plant, both gave yellow and green progeny, but in different ratios. These two ratios are represented in Figure 2-4. Notice that the ratios are seen only when the peas in several pods are combined.

The 3:1 and 1:1 ratios found for pea color were also found for comparable crosses for the other six characters that Mendel studied. The actual numbers for the 3:1 ratios for those characters are shown in Table 2-1.

Mendel's law of equal segregation

Initially, the meaning of these precise and repeatable mathematical ratios must have been unclear to Mendel, but he was able to devise a brilliant model that not

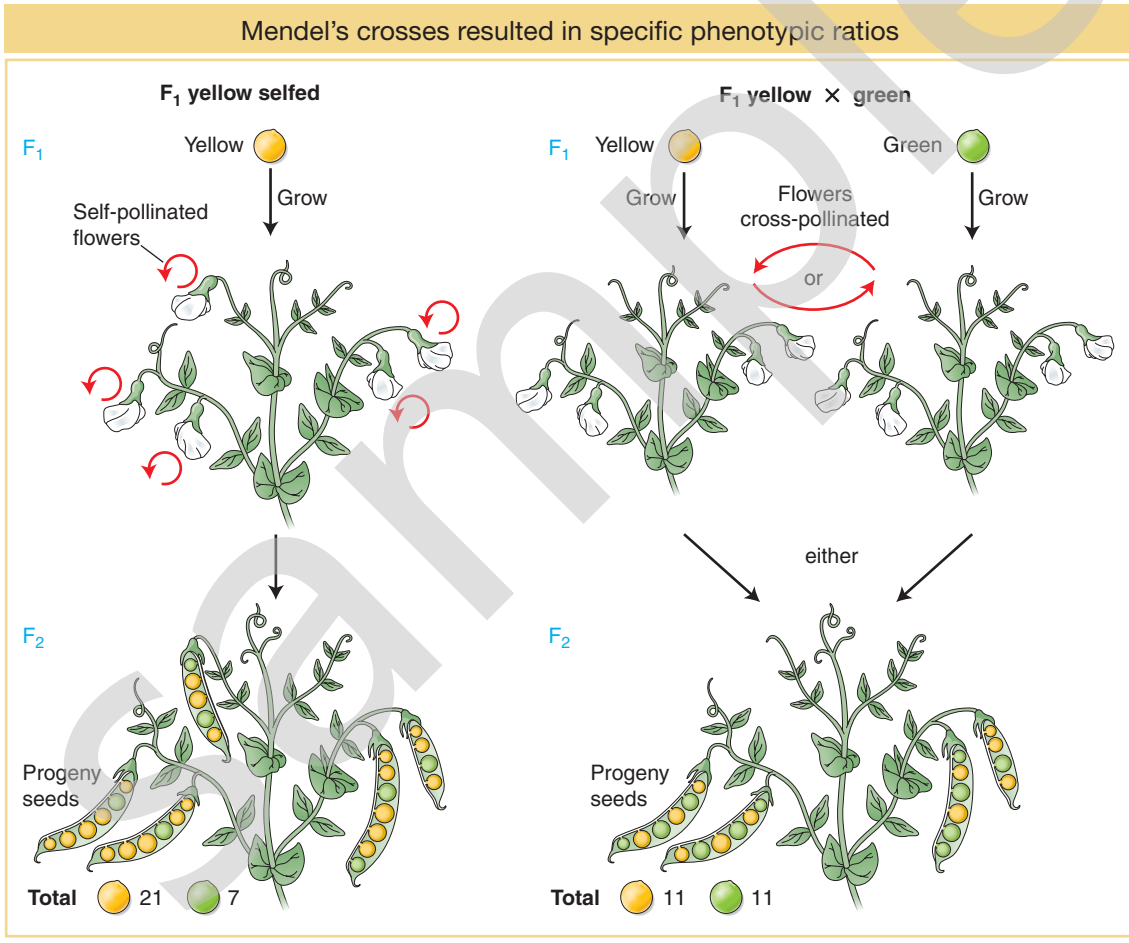


FIGURE 2-4 Mendel obtained a 3:1 phenotypic ratio in his self-pollination of the F_1 (*left*) and a 1:1 phenotypic ratio in his cross of F_1 yellow with green (*right*). Sample sizes are arbitrary.

only accounted for all the results, but also represented the historical birth of the science of genetics. Mendel's model for the pea-color example, translated into modern terms, was as follows:

1. A hereditary factor called a **gene** is necessary for producing pea color.

2. Each plant has a pair of this type of gene.

3. The gene comes in two forms called **alleles.** If the gene is phonetically called a "wye" gene, then the two alleles can be represented by *Y* (standing for the yellow phenotype) and *y* (standing for the green phenotype).

4. A plant can be either *Y/Y, y/y,* or *Y/y.* The slash shows that the alleles are a pair.

5. In the *Y/y* plant, the *Y* allele dominates, and so the phenotype will be yellow. Hence, the phenotype of the *Y/y* plant defines the *Y* allele as **dominant** and the *y* allele as **recessive.**

6. In meiosis, the members of a gene pair separate equally into the cells that become eggs and sperm, the *gametes.* This equal separation has become known as **Mendel's first law** or as the **law of equal segregation.** Hence, a single gamete contains only one member of the gene pair.

7. At fertilization, gametes fuse randomly, regardless of which of the alleles they bear.

Here, we introduce some terminology. A fertilized egg, the first cell that develops into a progeny individual, is called a **zygote.** A plant with a pair of identical

TABLE 2-1 Results of All Mendel's Crosses in Which Parents Differed in One Character

Parental phenotypes	F_1	F_2	F_2 ratio
1. round × wrinkled seeds	All round	5474 round; 1850 wrinkled	2.96 : 1
2. yellow × green seeds	All yellow	6022 yellow; 2001 green	3.01 : 1
3. purple × white petals	All purple	705 purple; 224 white	3.15 : 1
4. inflated × pinched pods	All inflated	882 inflated; 299 pinched	2.95 : 1
5. green × yellow pods	All green	428 green; 152 yellow	2.82 : 1
6. axial × terminal flowers	All axial	651 axial; 207 terminal	3.14 : 1
7. long × short stems	All long	787 long; 277 short	2.84 : 1

alleles is called a **homozygote** (adjective homozygous), and a plant in which the alleles of the pair differ is called a **heterozygote** (adjective heterozygous). Sometimes a heterozygote for one gene is called a **monohybrid.** An individual can be classified as either **homozygous dominant** (such as *Y/Y*), **heterozygous** (*Y/y*), or **homozygous recessive** (*y/y*). In genetics generally, allelic combinations underlying phenotypes are called **genotypes.** Hence, *Y/Y, Y/y,* and *y/y* are all genotypes.

Figure 2-5 shows how Mendel's postulates explain the progeny ratios illustrated in Figure 2-4. The pure-breeding lines are homozygous, either *Y/Y* or *y/y.* Hence, each line produces only *Y* gametes or only *y* gametes and thus can only breed true. When crossed with each other, the *Y/Y* and the *y/y* lines produce an F_1 generation composed of all heterozygous individuals (*Y/y*). Because *Y* is dominant, all F_1 individuals are yellow in phenotype. Selfing the F_1 individuals can be thought of as a cross of the type $Y/y \times Y/y$, which is sometimes called a **monohybrid cross.** Equal segregation of the *Y* and *y* alleles in the heterozygous F_1 results in gametes, both male and female, half of which are *Y* and half of which are *y.* Male and female gametes fuse randomly at fertilization, with the results shown in the grid in Figure 2-5. The composition of the F_2 is three-fourths yellow seeds and one-fourth green, a 3 : 1 ratio. The one-fourth of the F_2 seeds that are green breed true as expected of the genotype *y/y.* However, the yellow F_2 seeds (totaling three-fourths) are of two genotypes: two-thirds of them are clearly heterozygotes *Y/y,* and one-third are homozygous dominant *Y/Y.* Hence, we see that underlying the 3 : 1 phenotypic ratio in the F_2 is a 1 : 2 : 1 genotypic ratio:

$$\left.\begin{array}{l} \frac{1}{4}\ Y/Y \text{ yellow} \\ \frac{2}{4}\ Y/y \text{ yellow} \end{array}\right\} \ \frac{3}{4} \text{ yellow } (Y/-)$$
$$\frac{1}{4}\ y/y \text{ green}$$

The general depiction of an individual expressing the dominant allele is *Y/−*; the dash represents a slot that can be filled by either another *Y* or a *y.* Note that equal segregation is detectable only in the meiosis of a heterozygote. Hence, *Y/y* produces one-half *Y* gametes and one-half *y* gametes. Although equal segregation is taking place in homozygotes too, neither segregation $\frac{1}{2}\,Y : \frac{1}{2}\,Y$ nor segregation $\frac{1}{2}\,y : \frac{1}{2}\,y$ is meaningful or detectable at the genetic level.

We can now also explain results of the cross between the plants grown from F_1 yellow seeds (*Y/y*) and the plants grown from green seeds (*y/y*). In this case, equal segregation in the yellow heterozygous F_1 gives gametes with a $\frac{1}{2}\,Y : \frac{1}{2}\,y$ ratio. The *y/y* parent can make only *y* gametes, however; so the phenotype of the progeny depends only on which allele they inherit from the *Y/y* parent. Thus, the $\frac{1}{2}\,Y : \frac{1}{2}\,y$ *gametic* ratio from the heterozygote is converted into a $\frac{1}{2}\,Y/y : \frac{1}{2}\,y/y$

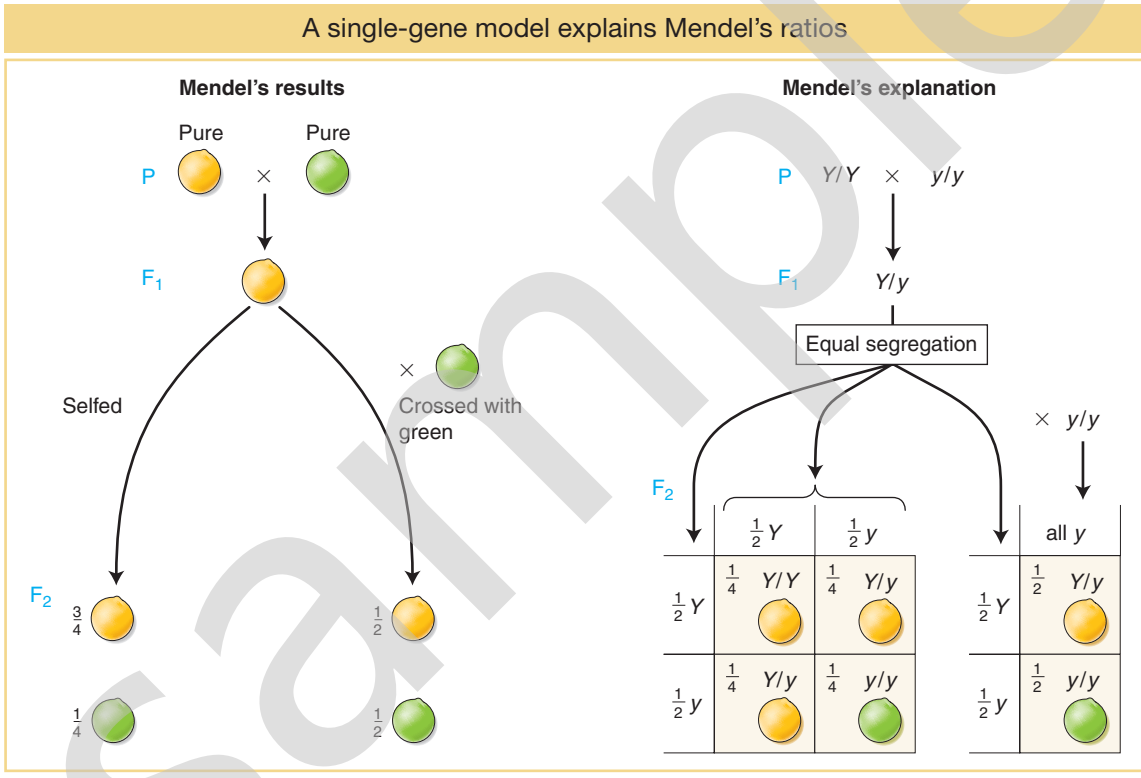


FIGURE 2-5 Mendel's results (*left*) are explained by a single-gene model (*right*) that postulates the equal segregation of the members of a gene pair into gametes.

genotypic ratio, which corresponds to a 1:1 *phenotypic* ratio of yellow-seeded to green-seeded plants. This is illustrated in the right-hand panel of Figure 2-5.

Note that, in defining the allele pairs that underlay his phenotypes, Mendel had identified a gene that radically affects pea color. This identification was not his prime interest, but we can see how finding single-gene inheritance patterns is a process of gene discovery, identifying individual genes that influence a biological property.

KEY CONCEPT All 1:1, 3:1, and 1:2:1 genetic ratios are diagnostic of single-gene inheritance and are based on equal segregation in a heterozygote.

Mendel's research in the mid-nineteenth century was not noticed by the international scientific community until similar observations were independently published by several other researchers in 1900. Soon research in many species of plants, animals, fungi, and algae showed that Mendel's law of equal segregation was applicable to all sexual eukaryotes and, in all cases, was based on the chromosomal segregations taking place in meiosis, a topic that we turn to in the next section.

2.2 The Chromosomal Basis of Single-Gene Inheritance Patterns

Mendel's view of equal segregation was that the members of a gene pair segregated equally *in gamete formation.* He did not know about the subcellular events that take place when cells divide in the course of gamete formation. Now we understand that gene pairs are located on chromosome pairs and that it is the members of a chromosome pair that actually segregate, carrying the genes with them. The members of a gene pair are segregated as an inevitable consequence.

Single-gene inheritance in diploids

When cells divide, so must the nucleus and its main contents, the chromosomes. To understand gene segregation, we must first understand and contrast the two types of nuclear divisions that take place in eukaryotic cells. When somatic (body) cells divide to increase their number, the accompanying nuclear division is called **mitosis,** a programmed stage of all eukaryotic cell-division cycles (Figure 2-6). Mitosis can take place in diploid or haploid cells. As a result, one progenitor cell becomes two genetically identical cells. Hence,

$$\text{either } 2n \longrightarrow 2n + 2n$$
$$\text{or } \; n \longrightarrow n + n$$

This "trick" of constancy is accomplished when each chromosome replicates to make two identical copies of itself, with underlying DNA replication. The two identical copies, which are often visually discernible, are called sister **chromatids.** Then, each copy is pulled to opposite ends of the cell. When the cell divides, each daughter cell has the same chromosomal set as its progenitor.

In addition, most eukaryotes have a sexual cycle, and, in these organisms, specialized diploid cells called **meiocytes** are set aside that divide to produce sex cells such as sperm and egg in plants and animals, or sexual spores in fungi or algae. Two sequential cell divisions take place, and the two nuclear divisions that accompany them are called **meiosis.** Because there are two divisions, four cells are produced from each progenitor cell. Meiosis takes place only in diploid cells, and the resulting gametes (sperm and eggs in animals and plants) are haploid. Hence, the net result of meiosis is

$$2n \longrightarrow n + n + n + n$$

This overall halving of chromosome number during meiosis is achieved through one replication and two divisions. As with mitosis, each chromosome replicates once, but in meiosis the replicated chromosomes (sister chromatids) remain attached. One of each of the replicated chromosome pairs is pulled to opposite ends of the cell, and division occurs. At the second division, the sister chromatids separate and are pulled to opposite ends of the cell.

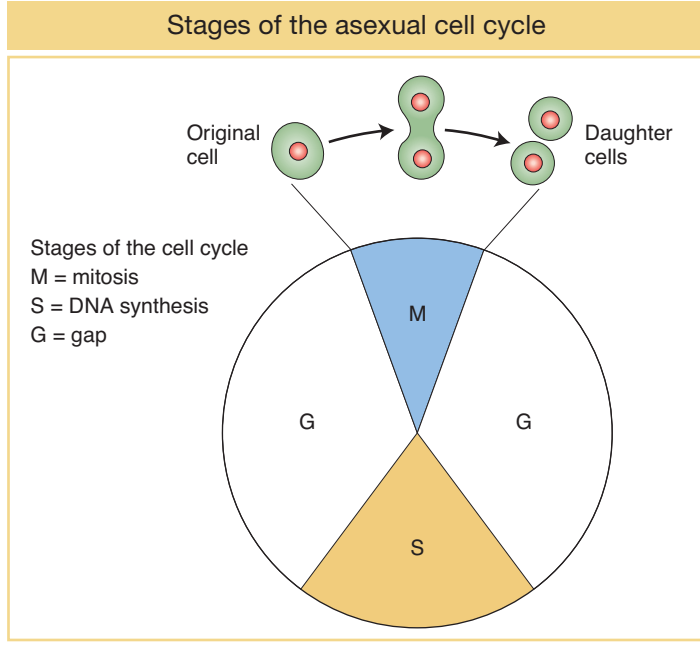


FIGURE 2-6

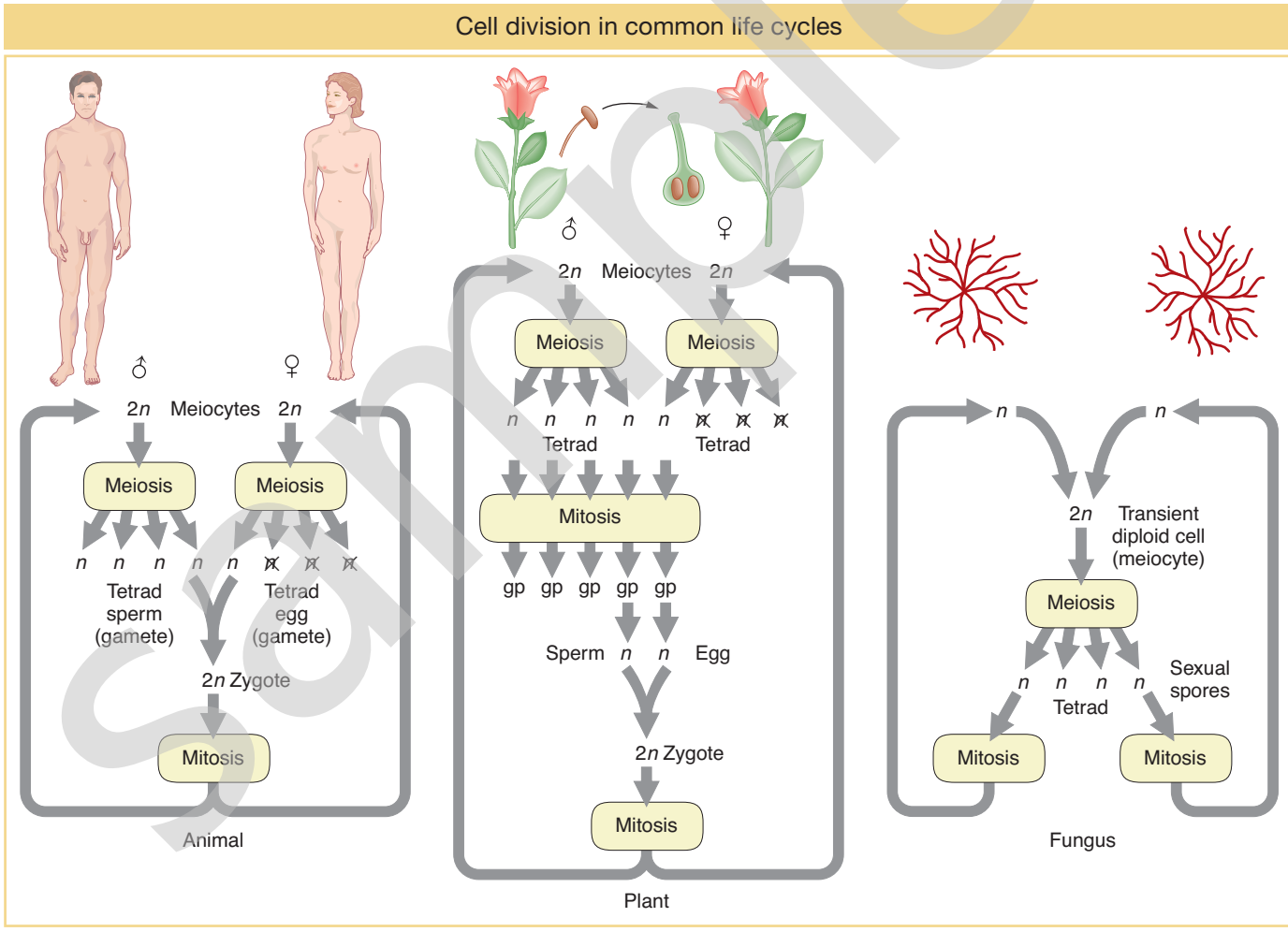


FIGURE 2-7 The life cycles of humans, plants, and fungi, showing the points at which mitosis and meiosis take place. Note that in the females of humans and many plants, three cells of the meiotic tetrad abort. The abbreviation *n* indicates a haploid cell, 2*n* a diploid cell; gp stands for gametophyte, the name of the small structure composed of haploid cells that will produce gametes. In many plants such as corn, a nucleus from the male gametophyte fuses with two nuclei from the female gametophyte, giving rise to a triploid (3*n*) cell, which then replicates to form the endosperm, a nutritive tissue that surrounds the embryo (which is derived from the 2*n* zygote).

The location of the meiocytes in animal, plant, and fungal life cycles is shown in Figure 2-7.

The basic *genetic* features of mitosis and meiosis are summarized in Figure 2-8. To make comparison easier, both processes are shown in a diploid cell. Notice, again, that mitosis takes place in one cell division, and the two resulting "daughter" cells have the same genomic content as that of the "mother" (progenitor) cell. The first key process to note is a premitotic chromosome replication. At the DNA level, this stage is the synthesis, or S, phase (see Figure 2-6), at which the DNA is replicated. The replication produces pairs of identical sister chromatids, which become visible at the beginning of mitosis. When a cell divides, each member of a pair of sister chromatids is pulled into a daughter cell, where it assumes the role of a fully fledged chromosome. Hence, each daughter cell has the same chromosomal content as the original cell.

Before meiosis, as in mitosis, chromosome replication takes place to form sister chromatids, which become visible at meiosis. The centromere appears not to divide at this stage, whereas it does in mitosis. Also in contrast with mitosis, the homologous pairs of sister chromatids now unite to form a bundle of four homologous chromatids. This joining of the homologous pairs is called *synapsis,* and it relies on the properties of a macromolecular assemblage called the synaptonemal complex (SC), which runs down the center of the pair. Replicate sister chromosomes are together called a **dyad** (from the Greek word for "two"). The unit comprising the pair of synapsed dyads is called a **bivalent.** The four chromatids that

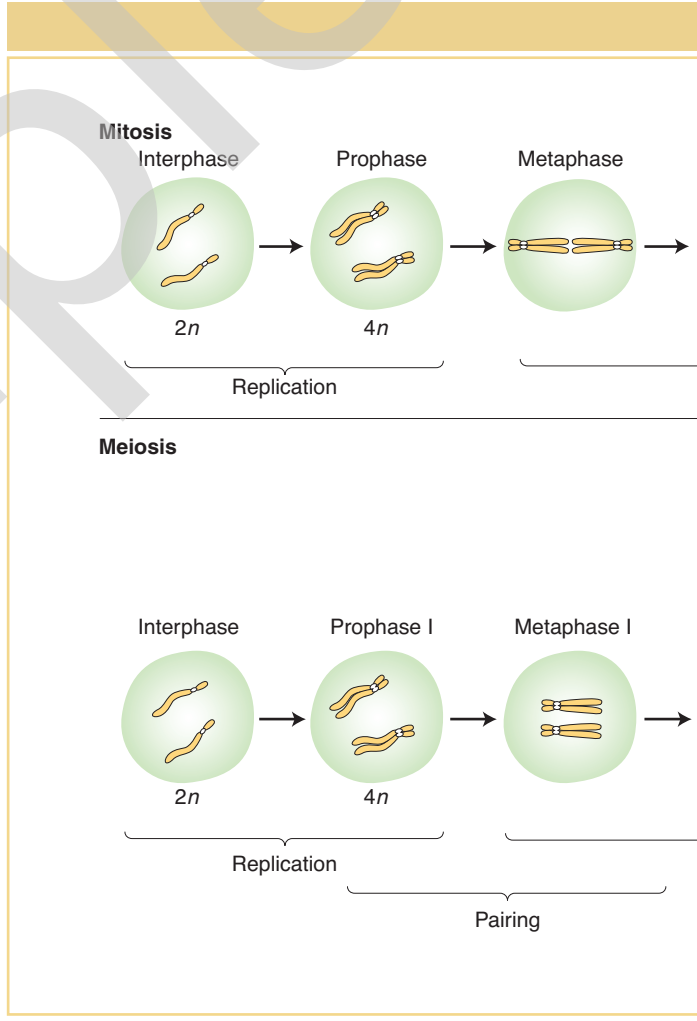
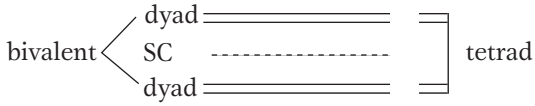


FIGURE 2-8 Simplified representation of mitosis and meiosis in diploid cells (2*n*, diploid; *n*, haploid). (Detailed versions are shown in Appendix 2-1, page 83.)

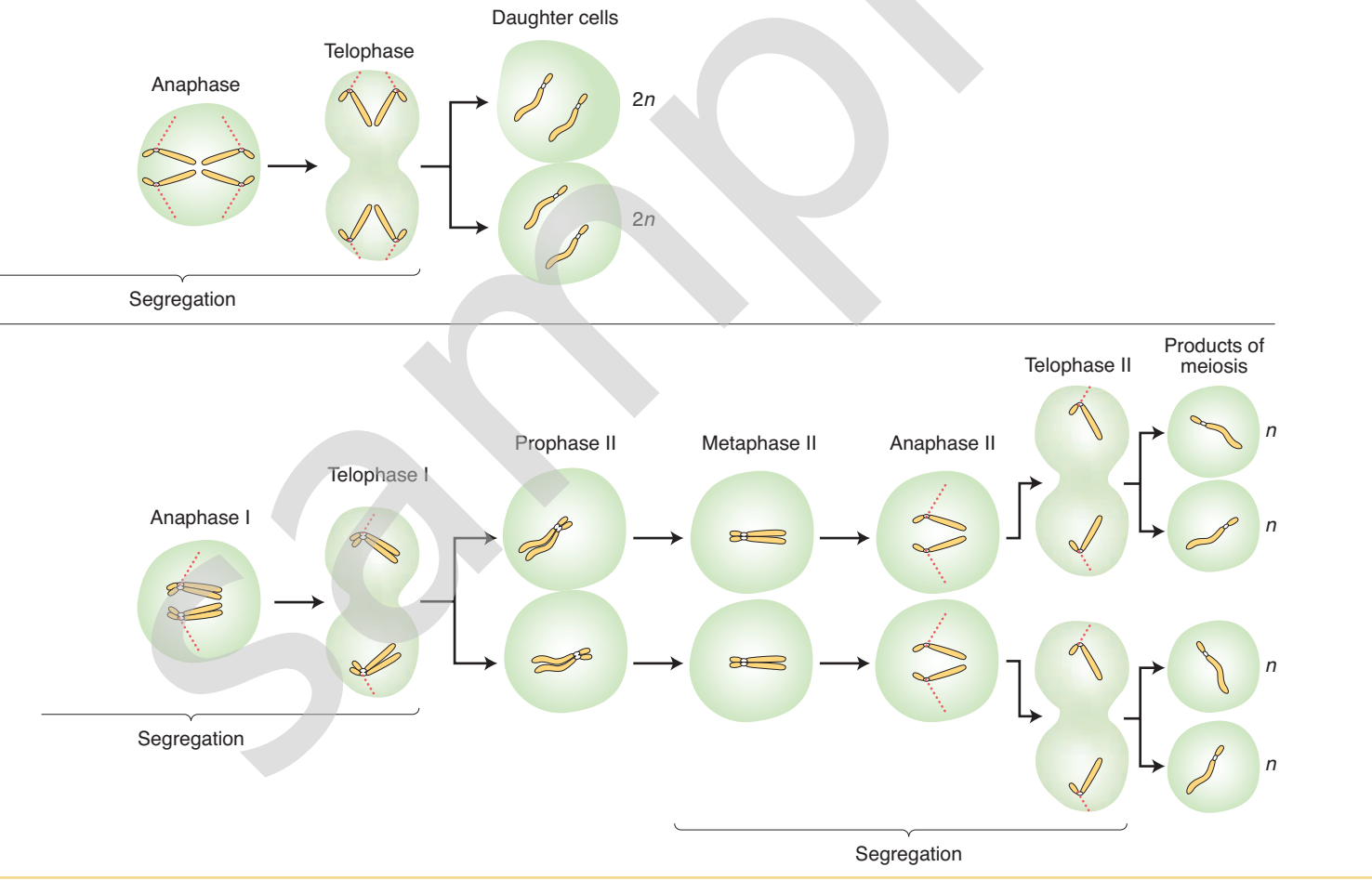
make up a bivalent are called a **tetrad** (Greek for "four"), to indicate that there are four homologous units in the bundle.



(A parenthetical note. The process of *crossing over* takes place at this tetrad stage. Crossing over changes the combinations of alleles of several different genes but does not directly affect single-gene inheritance patterns; therefore, we will postpone its detailed coverage until Chapter 4. For the present, it is worth noting that, apart from its allele-combining function, crossing over is also known to be a crucial event that must take place in order for proper chromosome segregation in the first meiotic division.)

The bivalents of all chromosomes move to the cell's equator, and, when the cell divides, one dyad moves into each new cell, pulled by spindle fibers attached near the centromeres. In the second cell division of meiosis, the centromeres divide and each member of a dyad (each member of a pair of chromatids) moves into a daughter cell. Hence, although the process starts with the same genomic content as that for mitosis, the two successive segregations result in four haploid cells. Each of the four haploid cells that constitute the four **products of meiosis** contains one member of a tetrad; hence, the group of four cells is sometimes called a tetrad, too. Meiosis can be summarized as follows:

Key stages of meiosis and mitosis



Start: → two homologs

Replication: → two dyads

Pairing: → tetrad

First division: → one dyad to each daughter cell

Second division: → one chromatid to each daughter cell

Research in cell biology has shown that the spindle fibers that pull apart chromosomes are polymers of the molecule tubulin. The pulling apart is caused mainly by a depolymerization and hence shortening of the fibers at the point where they are attached to the chromosomes.

The behavior of chromosomes during meiosis clearly explains Mendel's law of equal segregation. Consider a heterozygote of general type *A*/*a*. We can simply follow the preceding summary while considering what happens to the alleles of this gene:

Start: one homolog carries *A* and one carries *a*

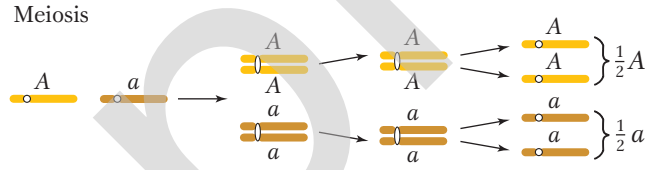
Replication: one dyad is *AA* and one is *aa*

Pairing: tetrad is *A*/*A*/*a*/*a*

First-division products: one cell *AA*, the other cell *aa* (crossing over can mix these types of products up, but the overall ratio is not changed)

Second-division products: four cells, two of type *A* and two of type *a*

Hence, the products of meiosis from a heterozygous meiocyte A/a are $\frac{1}{2}A$ and $\frac{1}{2}a$, precisely the equal ratio that is needed to explain Mendel's first law.



Note that we have focused on the broad genetic aspects of meiosis necessary to explain single-gene inheritance. More complete descriptions of the detailed stages of mitosis and meiosis are presented in Appendices 2-1 and 2-2 at the end of this chapter.

Single-gene inheritance in haploids

We have seen that the cellular basis of the law of equal segregation is the segregation of chromosomes in the first division of meiosis. In the discussion so far, the evidence for the equal segregation of alleles in meiocytes of both plants and animals is *indirect,* based on the observation that crosses show the appropriate ratios of progeny expected under equal segregation. Recognize that the gametes in these studies (such as Mendel's) must have come from *many different* meiocytes. However, in some organisms, their special life cycle allows the examination of the products of one single meiocyte. These organisms are called haploids, of which good examples are most fungi and algae. They spend most of their lives in the haploid state but can mate, in the process forming a transient diploid cell that becomes the meiocyte. In some species, the four products of a single meiosis are temporarily held together in a type of sac.

Baker's yeast, *Saccharomyces cerevisiae* (a fungus), provides a good example (see the yeast Model Organism box in Chapter 12). In fungi, there are simple forms of sexes called *mating types.* In *S. cerevisiae,* there are two mating types, and a successful cross can only occur between strains of different mating types. Let's look at a cross that includes a yeast mutant. Normal wild-type yeast colonies are white, but, occasionally, red mutants arise owing to a mutation in a gene in the biochemical pathway that synthesizes adenine. Let's use the red mutant to investigate equal segregation in a single meiocyte. We can call the mutant allele r for *red.* What symbol can we use for the normal, or wild-type, allele? In experimental genetics, the wild-type allele for any gene is generally designated by a plus sign, +. This sign is attached as a superscript to the symbol invented for the mutant allele. Hence, the wild-type allele in this example would be designated r^+, but a simple + is often used as shorthand. To see single-gene segregation, the red mutant is crossed with wild type. The cross would be

$$r^+ \times r$$

When two cells of opposite mating type fuse, a diploid cell is formed, and it is this cell that becomes the meiocyte. In the present example, the diploid meiocyte would be heterozygous, r^+/r. Replication and segregation of r^+ and r would give a tetrad of two meiotic products (spores) of genotype r^+ and two of r, all contained within a membranous sac called an **ascus.** Hence,

$$r^+/r \longrightarrow \left\{ \begin{matrix} r^+ \\ r^+ \\ r \\ r \end{matrix} \right. \text{ tetrad in ascus}$$

The details of the process are shown in Figure 2-9. If the four spores from one ascus are isolated (representing a tetrad of chromatids) and used to generate four yeast cultures, then equal segregation within one meiocyte is revealed directly as two white cultures and two red. If we analyzed the random spores

from many meiocytes, we would find about 50 percent red and 50 percent white.

Note the simplicity of haploid genetics: a cross requires the analysis of only one meiosis; in contrast, a diploid cross requires a consideration of meiosis in both the male and the female parent. This simplicity is an important reason for using haploids as model organisms. Another reason is that, in haploids, all alleles are expressed in the phenotype because there is no masking of recessives by dominant alleles on the other homolog.

2.3 The Molecular Basis of Mendelian Inheritance Patterns

Of course, Mendel had no idea of the molecular nature of the concepts he was working with. In this section, we can begin putting some of Mendel's concepts into a molecular context. Let's begin with alleles. We have used the concept of *alleles* without defining them at the molecular level. What are the *structural differences* between wild-type and mutant alleles at the DNA level of a gene? What are the *functional differences* at the protein level? Mutant alleles can be used to study single-gene inheritance without needing to understand their structural or functional nature. However, because a primary reason for embarking on single-gene inheritance is ultimately to investigate a gene's function, we must come to grips with the molecular nature of wild-type and mutant alleles at both the structural and the functional level.

Structural differences between alleles at the molecular level

Mendel proposed that genes come in different forms we now call alleles. What are alleles at the molecular level? When alleles such as *A* and *a* are examined at the DNA level by using modern technology, they are generally found to be identical in most of their sequences and differ only at one or several nucleotides of the hundreds or thousands of nucleotides that make up the gene. Therefore, we see that the alleles are truly different versions of the same gene. The following diagram represents the DNA of two alleles of one gene; the letter *x* represents a difference in the nucleotide sequence:

If the nucleotide sequence of an allele changes as the result of a rare chemical "accident," a new mutant allele is created. Such changes can occur anywhere along the nucleotide sequence of a gene. For example, a mutation could be a change in the identity of a single nucleotide or the deletion of one or more nucleotides or even the addition of one or more nucleotides.

There are many ways that a gene can be changed by mutation. For one thing, the mutational damage can occur at any one of many different sites. We can represent the situation as

Demonstration of equal segregation within one meiocyte in the yeast *S. cerevisiae*

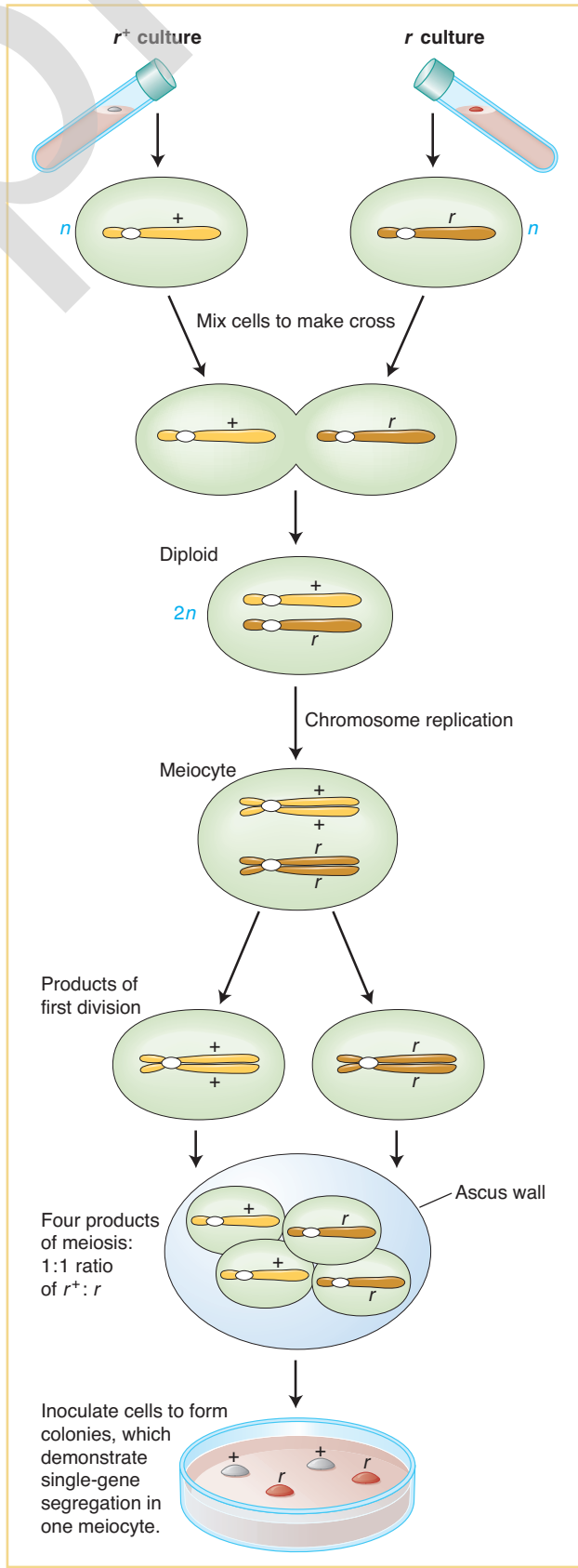
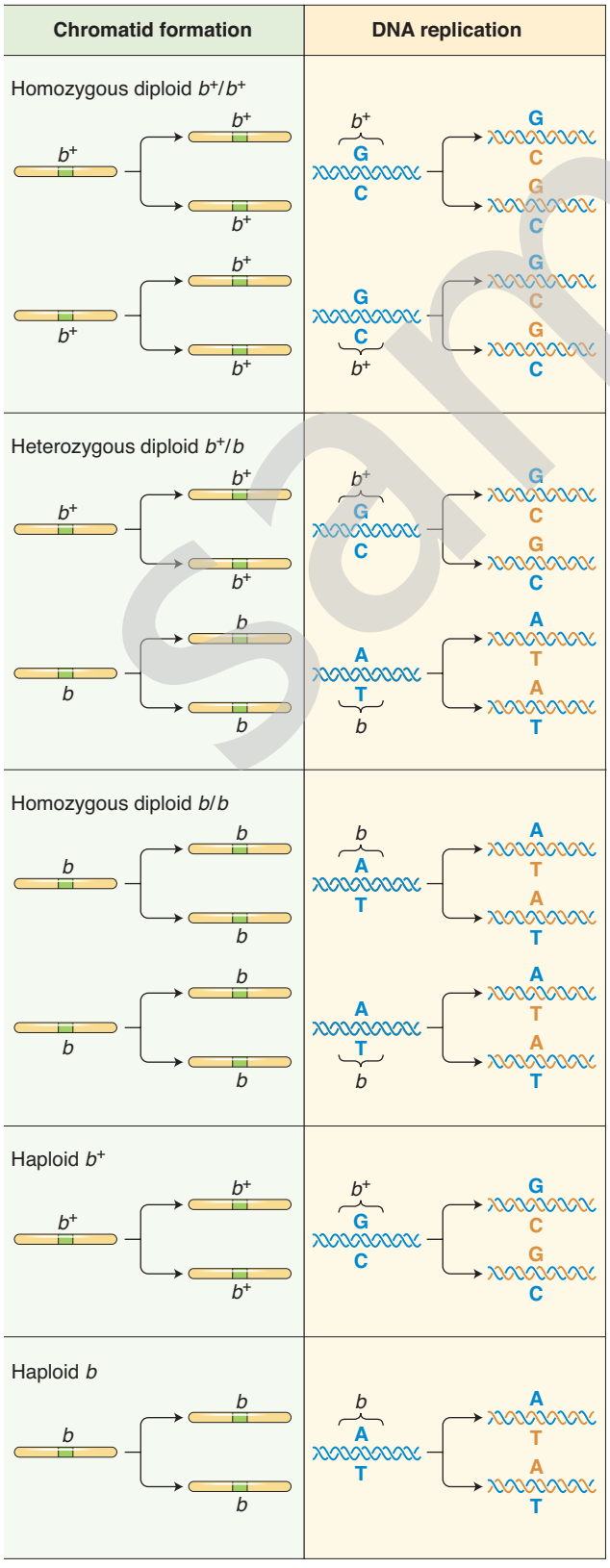
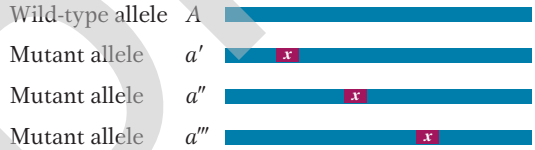


FIGURE 2-9 One ascus isolated from the cross + × *r* leads to two cultures of + and two of *r*.

DNA molecules replicate to form identical chromatids



follows, where dark blue indicates the normal wild-type DNA sequence and red with the letter *x* represents the altered sequence:



Molecular aspects of gene transmission

Replication of alleles during the S phase What happens to alleles at the molecular level during cell division? We know that the primary genomic component of each chromosome is a DNA molecule. This DNA molecule is replicated during the S phase, which precedes both mitosis and meiosis. As we will see in Chapter 7, replication is an accurate process and so all the genetic information is duplicated, whether wild type or mutant. For example, if a mutation is the result of a change in a single nucleotide pair—say, from GC (wild type) to AT (mutant)—then in a heterozygote, replication will be as follows:

homolog GC ⟶ replication ⟶ chromatid GC, chromatid GC

homolog AT ⟶ replication ⟶ chromatid AT, chromatid AT

DNA replication before mitosis in a haploid and a diploid are shown in Figure 2-10. This type of illustration serves to remind us that, in our considerations of the mechanisms of inheritance, it is essentially DNA molecules that are being moved around in the dividing cells.

Meiosis and mitosis at the molecular level The replication of DNA during the S phase produces two copies of each allele, *A* and *a*, that can now be segregated into separate cells. Nuclear division visualized at the DNA level is shown in Figure 2-11.

Demonstrating chromosome segregation at the molecular level We have interpreted single-gene phenotypic inheritance patterns in relation to the segregation of chromosomal DNA at meiosis. Is there any way to show DNA segregation directly (as opposed to phenotypic segregation)? The most straightforward approach would be to sequence the

FIGURE 2-10 Each chromosome divides longitudinally into two chromatids (*left*); at the molecular level (*right*), the single DNA molecule of each chromosome replicates, producing two DNA molecules, one for each chromatid. Also shown are various combinations of a gene with wild-type allele b^+ and mutant form *b*, caused by the change in a single base pair from GC to AT. Notice that, at the DNA level, the two chromatids produced when a chromosome replicates are always identical with each other and with the original chromosome.

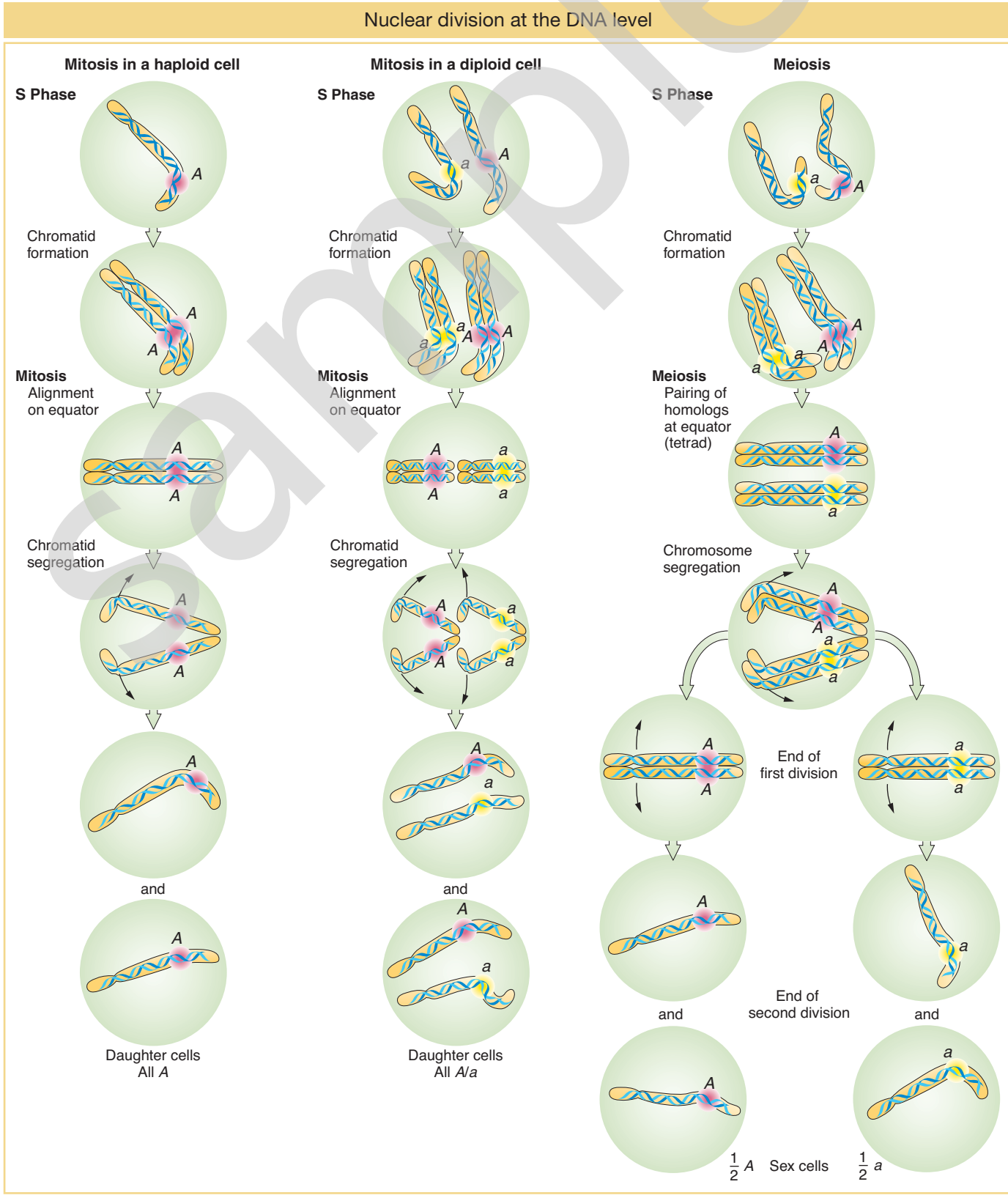


FIGURE 2-11 DNA and gene transmission in mitosis and meiosis in eukaryotes. The S phase and the main stages of mitosis and meiosis are shown. Mitotic divisions (*left and middle*) conserve the genotype of the original cell. At the right, the two successive meiotic divisions that take place during the sexual stage of the life cycle have the net effect of halving the number of chromosomes. The alleles *A* and *a* of one gene are used to show how genotypes are transmitted in cell division.

alleles (say, *A* and *a*) in the parents and the meiotic products: the result would be that one-half of the products would have the *A* DNA sequence and one-half would have the *a* DNA sequence. The same would be true for any DNA sequence that differed in the inherited chromosomes, including those not necessarily inside alleles correlated with known phenotypes such as red and white flowers. Thus, we see the rules of segregation enunciated by Mendel apply not only to genes but to any stretch of DNA along a chromosome.

KEY CONCEPT Mendelian inheritance is shown by any segment of DNA on a chromosome: by genes and their alleles and by molecular markers not necessarily associated with any biological function.

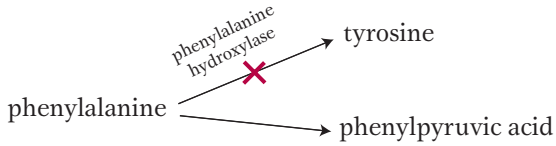
Alleles at the molecular level

At the molecular level, the primary phenotype of a gene is the protein it produces. What are the functional differences between proteins that explain the different effects of wild-type and mutant alleles on the properties of an organism?

Let's explore the topic by using the human disease phenylketonuria (PKU). We shall see in a later section on pedigree analysis that the PKU phenotype is inherited as a Mendelian recessive. The disease is caused by a defective allele of the gene that encodes the liver enzyme phenylalanine hydroxylase (PAH). This enzyme normally converts phenylalanine in food into the amino acid tyrosine:

$$\text{phenylalanine} \xrightarrow{\text{phenylalanine hydroxylase}} \text{tyrosine}$$

However, a mutation in the gene encoding this enzyme may alter the amino acid sequence in the vicinity of the enzyme's active site. In this case, the enzyme cannot bind phenylalanine (its substrate) or convert it into tyrosine. Therefore, phenylalanine builds up in the body and is converted instead into phenylpyruvic acid. This compound interferes with the development of the nervous system, leading to mental retardation.



Babies are now routinely tested for this processing deficiency at birth. If the deficiency is detected, phenylalanine can be withheld with the use of a special diet and the development of the disease arrested.

The PAH enzyme is made up of a single type of protein. What changes have occurred in the mutant form of the PKU gene's DNA, and how can such change at the DNA level affect protein function and produce the disease phenotype? Sequencing of the mutant alleles from many PKU patients has revealed a plethora of mutations at different sites along the gene, mainly in the protein-encoding regions, or the exons; the results are summarized in Figure 2-12. They represent a range of DNA changes, but most are small changes affecting only one nucleotide pair among the thousands that constitute the gene. What these alleles have in common is that they encode a defective protein that no longer has normal PAH activity. By changing one or more amino acids, the mutations all inactivate some essential part of the protein encoded by the gene. The effect of the mutation on the function of the gene depends on where within the gene the mutation occurs. An important functional region of the gene is that encoding an enzyme's active site; so this region is very sensitive to mutation. In addition, a minority of mutations are found to be in introns, and these mutations often prevent the normal processing of the primary RNA transcript.

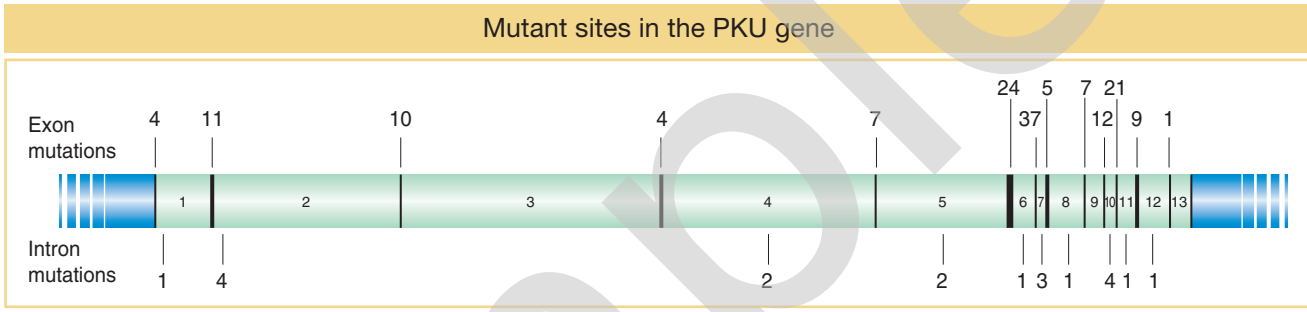


FIGURE 2-12 Many mutations of the human phenylalanine hydroxylase gene that cause enzyme malfunction are known. The number of mutations in the exons, or protein-encoding regions (black), are listed above the gene. The number of mutations in the intron regions (green, numbered 1 through 13) that alter splicing are listed below the gene. [*Data from C. R. Scriver,* Ann. Rev. Genet. *28, 1994, 141–165.*]

Some of the general consequences of mutation at the protein level are shown in Figure 2-13. Many of the mutant alleles are of a type generally called **null alleles:** the proteins encoded by them completely lack PAH function. Other mutant alleles reduce the level of enzyme function; they are sometimes called **leaky mutations,** because some wild-type function seems to "leak" into the mutant phenotype. DNA sequencing often detects changes that have no functional impact at all, so they are functionally wild type. Hence, we see that the terms *wild type* and *mutant* sometimes have to be used carefully.

KEY CONCEPT Most mutations that alter phenotype alter the amino acid sequence of the gene's protein product, resulting in reduced or absent function.

We have been pursuing the idea that finding a set of genes that impinge on the biological property under investigation is an important goal of genetics, because it defines the components of the system. However, finding the *precise* way in which mutant alleles lead to mutant phenotypes is often challenging, requiring not only the identification of the protein products of these genes, but also detailed cellular and physiological studies to measure the effects of the mutations. Furthermore,

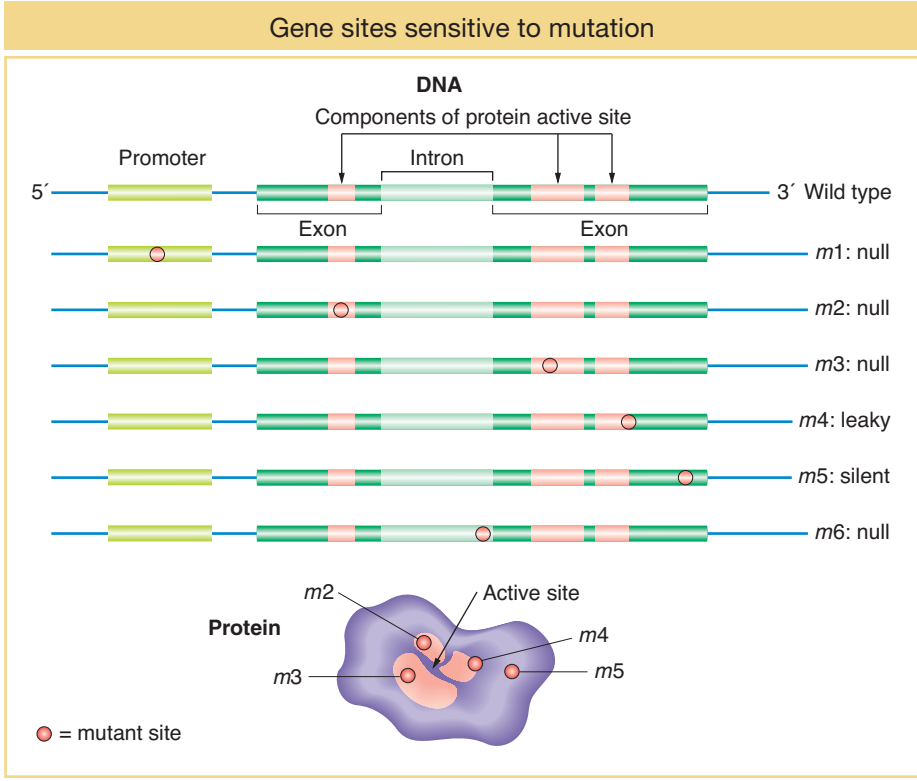


FIGURE 2-13 Mutations in the parts of a gene encoding enzyme active sites lead to enzymes that do not function (null mutations). Mutations elsewhere in the gene may have no effect on enzyme function (silent mutations). Promoters are sites important in transition initiation.

finding how the set of genes *interacts* is a second level of challenge and a topic that we will pursue later, starting in Chapter 6.

Dominance and recessiveness With an understanding of how genes function through their protein products, we can better understand dominance and recessiveness. Dominance was defined earlier in this chapter as the phenotype shown by a heterozygote. Hence, formally, it is the *phenotype* that is dominant or recessive, but, in practice, geneticists more often apply the term to alleles. This formal definition has no molecular content, but both dominance and recessiveness can have simple explanations at the molecular level. We introduce the topic here, to be revisited in Chapter 6.

How can alleles be dominant? How can they be recessive? Recessiveness is observed in null mutations in genes that are functionally **haplosufficient,** loosely meaning that one gene copy has enough function to produce a wild-type phenotype. Although a wild-type diploid cell normally has two fully functional copies of a gene, one copy of a haplosufficient gene provides enough gene product (generally a protein) to carry out the normal transactions of the cell. In a heterozygote (say, +/*m,* where *m* is a null), the single functional copy encoded by the + allele provides enough protein product for normal cellular function. In a simple example, assume a cell needs a minimum of 10 protein units to function normally. Each wild-type allele can produce 12 units. Hence, a homozygous wild type +/+ will produce 24 units. The heterozygote +/*m* will produce 12 units, in excess of the 10-unit minimum, and hence the mutant allele is recessive as it has no impact in the heterozygote.

Other genes are **haploinsufficient.** In such cases, a null mutant allele will be dominant because, in a heterozygote (+/*P*), the single wild-type allele cannot provide enough product for normal function. As another example, let's assume the cell needs a minimum of 20 units of this protein, and the wild-type allele produces only 12 units. A homozygous wild type +/+ makes 24 units, which is over the minimum. However, a heterozygote involving a null mutation (+/*P*) produces only 12; hence, the presence of the mutant allele in the heterozygote results in an inadequate supply of product and a mutant phenotype ensues.

In some cases, mutation results in a *new function* for the gene. Such mutations can be dominant because, in a heterozygote, the wild-type allele cannot mask this new function.

From the above brief considerations, we see that *phenotype,* the description or measurement that we track during Mendelian inheritance, is an emergent property based on the nature of alleles and the way in which the gene functions normally and abnormally. The same can be said for the descriptions dominant and recessive that we apply to a phenotype.

2.4 Some Genes Discovered by Observing Segregation Ratios

Recall that one general aim of genetic analysis today is to dissect a biological property by discovering the set of single genes that affect it. We learned that an important way to identify these genes is by the phenotypic segregation ratios generated by their mutations—most often 1:1 and 3:1 ratios, both of which are based on equal segregation as defined by Gregor Mendel.

Let's look at some examples that extend the Mendelian approach into a modern experimental setting. Typically, the researcher is confronted by an array of interesting mutant phenotypes that affect the property of interest (such as those depicted in Figure 2-1) and now needs to know whether they are inherited as single-mutant alleles. Mutant alleles can be either dominant or recessive, depending on their action; so the question of dominance also needs to be considered in the analysis.

The standard procedure is to cross the mutant with wild type. (If the mutant is sterile, then another approach is needed.) First, we will consider three simple cases that cover most of the possible outcomes:

1. A fertile flower mutant with no pigment in the petals (for example, white petaled in contrast with the normal red)

2. A fertile fruit-fly mutant with short wings

3. A fertile mold mutant that produces excess hyphal branches (hyperbranching)

A gene active in the development of flower color

To begin the process, the white-flowered plant is crossed with the normal wild-type red. All the F_1 plants are red flowered, and, of 500 F_2 plants sampled, 378 are red flowered and 122 are white flowered. If we acknowledge the existence of sampling error, these F_2 numbers are very close to a $\frac{3}{4} : \frac{1}{4}$ or 3:1, ratio. Because this ratio indicates single-gene inheritance, we can conclude that the mutant is caused by a recessive alteration in a single gene. According to the general rules of gene nomenclature, the mutant allele for white petals might be called *alb* for *albino* and the wild-type allele would be alb^+ or just +. (The conventions for allele nomenclature vary somewhat among organisms: some of the variations are shown in Appendix A on nomenclature.) We surmise that the wild-type allele plays an essential role in producing the colored petals of the plant, a property that is almost certainly necessary for attracting pollinators to the flower. The gene might be implicated in the biochemical synthesis of the pigment or in the part of the signaling system that tells the cells of the flower to start making pigment or in a number of other possibilities that require further investigation. At the purely genetic level, the crosses made would be represented symbolically as

P	$+/+ \times alb/alb$
F_1	all $+/alb$
F_2	$\frac{1}{4}$ $+/+$
	$\frac{1}{2}$ $+/alb$
	$\frac{1}{4}$ alb/alb

P	*alb*	*alb*
+	+/*alb*	+/*alb*
+	+/*alb*	+/*alb*

All F_1 are red

F_1	+	*alb*
+	+/+	+/*alb*
alb	+/*alb*	*alb*/*alb*

$\frac{3}{4}$ of F_2 are red, $\frac{1}{4}$ are white

or graphically as in the grids on the right (see also Figure 2-5). This type of grid showing gametes and gametic fusions is called a *Punnett square,* named after an early geneticist, Reginald C. Punnett. They are useful devices for explaining genetic ratios. We shall encounter more in later discussions.

A gene for wing development

In the fruit-fly example, the cross of the mutant short-winged fly with wild-type long-winged stock yielded 788 progeny, classified as follows:

196 short-winged males

194 short-winged females

197 long-winged males

201 long-winged females

In total, there are 390 short- and 398 long-winged progeny, very close to a 1:1 ratio. The ratio is the same within males and females, again within the bounds of sampling error. Hence, from these results, the "short wings" mutant was very likely produced by a dominant mutation. Note that, for a dominant mutation to be expressed, only a single "dose" of mutant allele is necessary; so, in most cases,

P	+	*SH*
+	+/+	*SH*/+
+	+/+	*SH*/+

F_1	+	+
+	+/+	+/+
+	+/+	+/+

F_1	+	*SH*
+	+/+	*SH*/+
SH	*SH*/+	*SH*/*SH*

when the mutant first shows up in the population, it will be in the heterozygous state. (This is not true for a recessive mutation such as that in the preceding plant example, which must be homozygous to be expressed and must have come from the selfing of an unidentified heterozygous plant in the preceding generation.)

When long-winged progeny were interbred, all of their progeny were long winged, as expected of a recessive wild-type allele. When the short-winged progeny were interbred, their progeny showed a ratio of three-fourths short to one-fourth long.

Dominant mutations are represented by uppercase letters or words: in the present example, the mutant allele might be named *SH,* standing for "short." Then the crosses would be represented symbolically as

$$\begin{aligned} &\text{P} && +/+ \times SH/+ \\ &\text{F}_1 && \tfrac{1}{2}\ +/+ \\ & && \tfrac{1}{2}\ SH/+ \end{aligned}$$

$$\begin{aligned} &\text{F}_1 && +/+ \times +/+ \\ & && \text{all } +/+ \end{aligned}$$

$$\begin{aligned} &\text{F}_1 && SH/+ \times SH/+ \\ & && \tfrac{1}{4}\ SH/SH \\ & && \tfrac{1}{2}\ SH/+ \\ & && \tfrac{1}{4}\ +/+ \end{aligned}$$

or graphically as shown in the grids on the left.

This analysis of the fly mutant identifies a gene that is part of a subset of genes that, in wild-type form, are crucial for the normal development of a wing. Such a result is the starting point of further studies that would focus on the precise developmental and cellular ways in which the growth of the wing is arrested, which, once identified, reveal the time of action of the wild-type allele in the course of development.

A gene for hyphal branching

A hyperbranching fungal mutant (such as the button-like colony in Figure 2-1) was crossed with a wild-type fungus with normal sparse branching. In a sample of 300 progeny, 152 were wild type and 148 were hyperbranching, very close to a 1:1 ratio. We infer from this single-gene inheritance ratio that the hyperbranching mutation is of a single gene. In haploids, assigning dominance is usually not possible, but, for convenience, we can call the hyperbranching allele *hb* and the wild type hb^+ or +. The cross must have been

$$\begin{aligned} &\text{P} && + \times hb \\ &\text{Diploid meiocyte} && +/hb \\ &\text{F}_1 && \tfrac{1}{2}\ + \\ & && \tfrac{1}{2}\ hb \end{aligned}$$

The mutation and inheritance analysis has uncovered a gene whose wild-type allele is essential for normal control of branching, a key function in fungal dispersal and nutrient acquisition. Now the mutant needs to be investigated to see the location in the normal developmental sequence at which the mutant produces a block. This information will reveal the time and place in the cells at which the normal allele acts.

Sometimes, the severity of a mutant phenotype renders the organism sterile, unable to go through the sexual cycle. How can the single-gene inheritance of

sterile mutants be demonstrated? In a diploid organism, a sterile recessive mutant can be propagated as a heterozygote and then the heterozygote can be selfed to produce the expected 25 percent homozygous recessive mutants for study. A sterile dominant mutant is a genetic dead end and cannot be propagated sexually, but, in plants and fungi, such a mutant can be easily propagated asexually.

What if a cross between a mutant and a wild type does not produce a 3:1 or a 1:1 ratio as discussed here, but some other ratio? Such a result can be due to the interactions of several genes or to an environmental effect. Some of these possibilities are discussed in Chapter 6.

Predicting progeny proportions or parental genotypes by applying the principles of single-gene inheritance

We can summarize the direction of analysis of gene discovery as follows:

Observe phenotypic ratios in progeny →
Deduce genotypes of parents (*A*/*A*, *A*/*a*, or *a*/*a*)

However, the same principle of inheritance (essentially Mendel's law of equal segregation) can also be used to predict phenotypic ratios in the progeny of parents of *known genotypes*. These parents would be from stocks maintained by the researcher. The types and proportions of the progeny of crosses such as *A*/*A* × *A*/*a*, *A*/*A* × *a*/*a*, *A*/*a* × *A*/*a*, and *A*/*a* × *a*/*a* can be easily predicted. In summary,

Cross parents of known genotypes → Predict phenotypic ratios in progeny

This type of analysis is used in general breeding to synthesize genotypes for research or for agriculture. It is also useful in predicting likelihoods of various outcomes in human matings in families with histories of single-gene diseases.

After single-gene inheritance has been established, an individual showing the dominant phenotype but of *unknown genotype* can be tested to see if the genotype is homozygous or heterozygous. Such a test can be performed by crossing the individual (of phenotype *A*/?) with a recessive tester strain *a*/*a*. If the individual is heterozygous, a 1:1 ratio will result ($\frac{1}{2}$ *A*/*a* and $\frac{1}{2}$ *a*/*a*); if the individual is homozygous, all progeny will show the dominant phenotype (all *A*/*a*). In general, the cross of an individual of unknown heterozygosity (for one gene or more) with a fully recessive parent is called a **testcross,** and the recessive individual is called a **tester.** We will encounter testcrosses many times throughout subsequent chapters; they are very useful in deducing the meiotic events taking place in more complex genotypes such as dihybrids and trihybrids. The use of a fully recessive tester means that meiosis in the tester parent can be ignored because all of its gametes are recessive and do not contribute to the phenotypes of the progeny. An alternative test for heterozygosity (useful if a recessive tester is not available and the organism can be selfed) is simply to self the unknown: if the organism being tested is heterozygous, a 3:1 ratio will be found in the progeny. Such tests are useful and common in routine genetic analysis.

KEY CONCEPT The principles of inheritance (such as the law of equal segregation) can be applied in two directions: (1) inferring genotypes from phenotypic ratios and (2) predicting phenotypic ratios from parents of known genotypes.

2.5 Sex-Linked Single-Gene Inheritance Patterns

The chromosomes that we have been analyzing so far are autosomes, the "regular" chromosomes that form most of the genomic set. However, many animals and plants have a special pair of chromosomes associated with sex. The sex

chromosomes also segregate equally, but the phenotypic ratios seen in progeny are often different from the autosomal ratios.

Sex chromosomes

Most animals and many plants show sexual dimorphism; in other words, individuals are either male or female. In most of these cases, sex is determined by a special pair of **sex chromosomes.** Let's look at humans as an example. Human body cells have 46 chromosomes: 22 homologous pairs of autosomes plus 2 sex chromosomes. Females have a pair of identical sex chromosomes called the **X chromosomes.** Males have a nonidentical pair, consisting of one X and one Y. The **Y chromosome** is considerably shorter than the X. Hence, if we let A represent autosomal chromosomes, we can write

$$\text{females} = 44\text{A} + \text{XX}$$

$$\text{males} = 44\text{A} + \text{XY}$$

At meiosis in females, the two X chromosomes pair and segregate like autosomes, and so each egg receives one X chromosome. Hence, with regard to sex chromosomes, the gametes are of only one type and the female is said to be the **homogametic sex.** At meiosis in males, the X and the Y chromosomes pair over a short region, which ensures that the X and Y separate so that there are two types of sperm, half with an X and the other half with a Y. Therefore, the male is called the **heterogametic sex.**

The inheritance patterns of genes on the sex chromosomes are different from those of autosomal genes. Sex-chromosome inheritance patterns were first investigated in the early 1900s in the laboratory of the great geneticist Thomas Hunt Morgan, using the fruit fly *Drosophila melanogaster* (see the Model Organism box on page 56). This insect has been one of the most important research organisms in genetics; its short, simple life cycle contributes to its usefulness in this regard. Fruit flies have three pairs of autosomes plus a pair of sex chromosomes, again referred to as X and Y. As in mammals, *Drosophila* females have the constitution XX and males are XY. However, the mechanism of sex determination in *Drosophila* differs from that in mammals. In *Drosophila,* the *number of X chromosomes* in relation to the autosomes determines sex: two X's result in a female, and one X results in a male. In mammals, the *presence of the Y chromosome* determines maleness and the absence of a Y determines femaleness. However, it is important to note that, despite this somewhat different basis for sex determination, the single-gene inheritance patterns of genes on the sex chromosomes are remarkably similar in *Drosophila* and mammals.

Vascular plants show a variety of sexual arrangements. **Dioecious species** are those showing animal-like sexual dimorphism, with female plants bearing flowers containing only ovaries and male plants bearing flowers containing only anthers (Figure 2-14). Some, but not all, dioecious plants have a nonidentical pair of chromosomes associated with (and almost certainly determining) the sex of the plant. Of the species with nonidentical sex chromosomes, a large proportion have an XY system. For example, the dioecious plant *Melandrium album* has 22 chromosomes per cell: 20 autosomes plus 2 sex chromosomes, with XX females and XY males. Other dioecious plants have no visibly different pair of chromosomes; they may still have sex chromosomes but not visibly distinguishable types.

Sex-linked patterns of inheritance

Cytogeneticists divide the X and Y chromosomes into homologous and differential regions. Again, let's use humans as an example (Figure 2-15). The differential regions, which contain most of the genes, have no counterparts on the other sex

chromosome. Hence, in males, the genes in the differential regions are said to be **hemizygous** ("half zygous"). The differential region of the X chromosome contains many hundreds of genes; most of these genes do not take part in sexual function, and they influence a great range of human properties. The Y chromosome contains only a few dozen genes. Some of these genes have counterparts on the X chromosome, but most do not. The latter type take part in male sexual function. One of these genes, *SRY,* determines maleness itself. Several other genes are specific for sperm production in males.

In general, genes in the differential regions are said to show inheritance patterns called **sex linkage.** Mutant alleles in the differential region of the X chromosome show a single-gene inheritance pattern called **X linkage.** Mutant alleles of the few genes in the differential region of the Y chromosome show **Y linkage.** A gene that is sex linked can show phenotypic ratios that are different in each sex. In this respect, sex-linked inheritance patterns contrast with the inheritance patterns of genes in the autosomes, which are the same in each sex. If the genomic location of a gene is unknown, a sex-linked inheritance pattern indicates that the gene lies on a sex chromosome.

The human X and Y chromosomes have two short homologous regions, one at each end (see Figure 2-15). In the sense that these regions are homologous, they are autosomal-like, and so they are called **pseudoautosomal regions 1 and 2.** One or both of these regions pairs in meiosis and undergoes crossing over (see Chapter 4 for details of crossing over). For this reason, the X and the Y chromosomes can act as a pair and segregate into equal numbers of sperm.

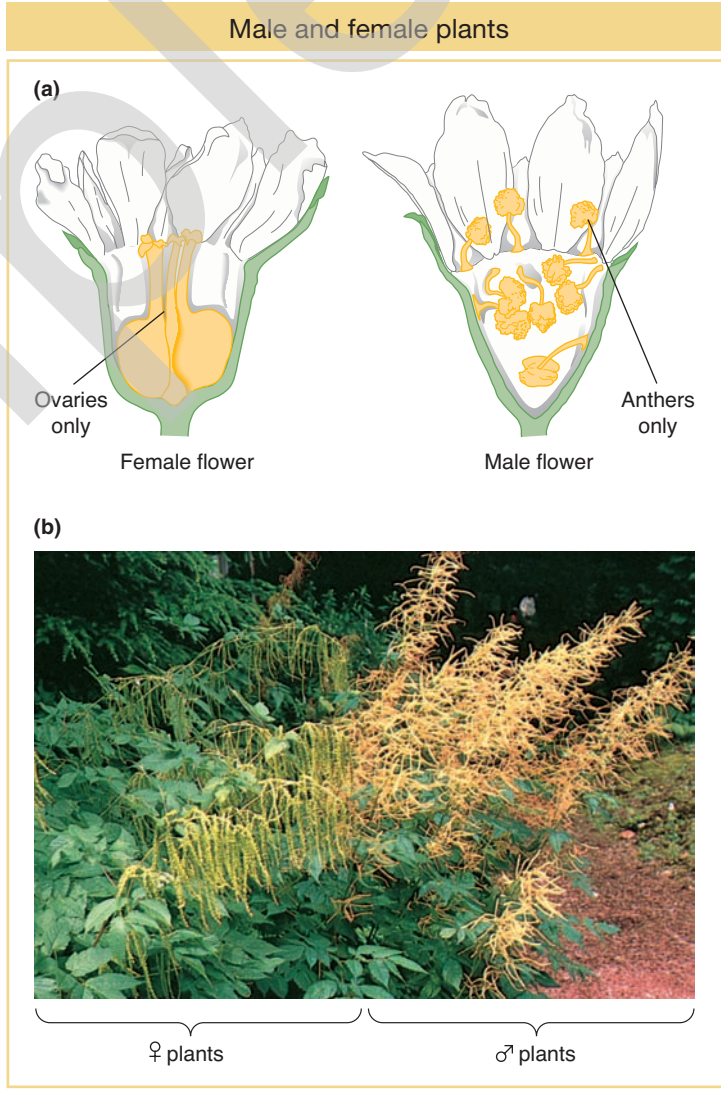


FIGURE 2-14 Examples of two dioecious plant species are (a) *Osmaronia dioica* and (b) *Aruncus dioicus.* [*(a) Leslie Bohm; (b) Anthony Griffiths.*]

X-linked inheritance

For our first example of X linkage, we turn to eye color in *Drosophila.* The wild-type eye color of *Drosophila* is dull red, but pure lines with white eyes are available

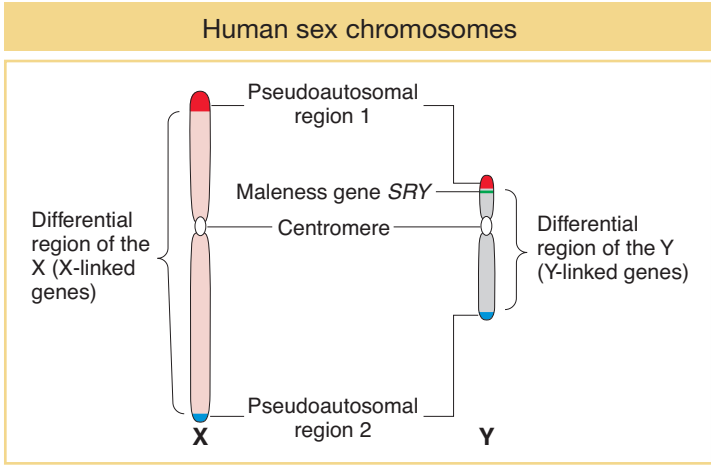


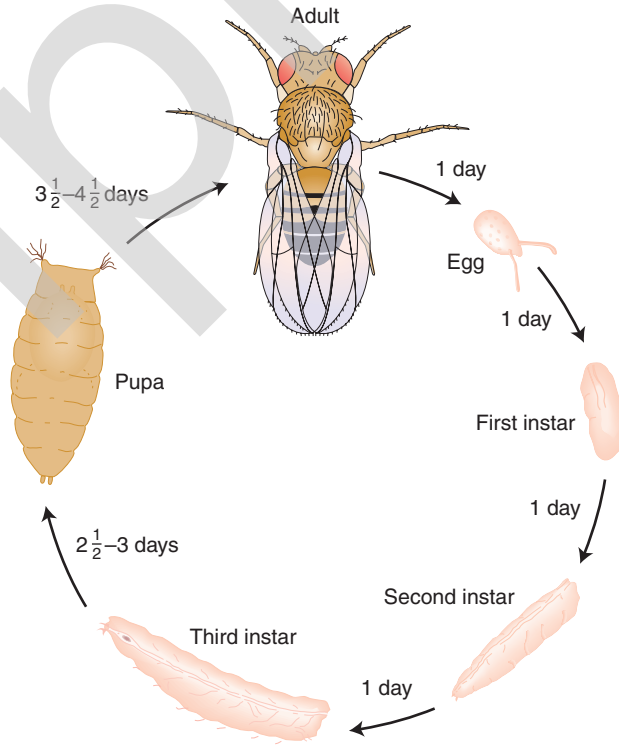
FIGURE 2-15 Human sex chromosomes contain a differential region and two pairing regions. The regions were located by observing where the chromosomes paired up in meiosis and where they did not.

MODEL ORGANISM *Drosophila*

Drosophila melanogaster was one of the first model organisms to be used in genetics. It is readily available from ripe fruit, has a short life cycle, and is simple to culture and cross. Sex is determined by X and Y sex chromosomes (XX = female, XY = male), and males and females are easily distinguished. Mutant phenotypes regularly arise in lab populations, and their frequency can be increased by treatment with mutagenic radiation or chemicals. It is a diploid organism, with four pairs of homologous chromosomes ($2n = 8$). In salivary glands and certain other tissues, multiple rounds of DNA replication without chromosomal division result in "giant chromosomes," each with a unique banding pattern that provides geneticists with landmarks for the study of chromosome mapping and rearrangement. There are many species and races of *Drosophila,* which have been important raw material for the study of evolution.

Time flies like an arrow; fruit flies like a banana.
(Groucho Marx)

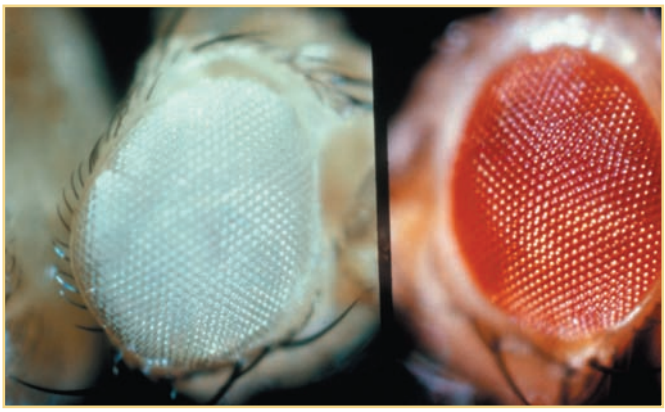
Drosophila melanogaster, the common fruit flt.
[© *blickwinkel/Alamy.*]



Life cycle of *Drosophila melanogaster.*

FIGURE 2-16 The red-eyed fly is wild type, and the white-eyed fly is a mutant. [*Science Source/Getty Images.*]

White-eyed and red-eyed *Drosophila*



(Figure 2-16). This phenotypic difference is determined by two alleles of a gene located on the differential region of the X chromosome. The mutant allele in the present case is *w* for white eyes (the lowercase letter indicates that the allele is recessive), and the corresponding wild-type allele is w^+. When white-eyed males are crossed with red-eyed females, all the F_1 progeny have red eyes, suggesting that the allele for white eyes is recessive. Crossing these red-eyed F_1 males and females produces a 3:1 F_2 ratio of red-eyed to white-eyed flies, but *all the white-eyed flies are males.* This inheritance pattern, which shows a clear difference between the sexes, is explained in Figure 2-17. The basis of the inheritance pattern is that all the F_1 flies receive a wild-type allele from their mothers, but the F_1 females also receive a white-eye allele from their fathers. Hence, all F_1 females are heterozygous wild type (w^+/w), and the F_1 males are hemizygous wild type (w^+). The F_1 females pass on the white-eye allele to half their sons, who express it, and to half their daughters, who do not express it, because they must inherit the wild-type allele from their fathers.

The reciprocal cross gives a different result; that is, the cross between white-eyed females and red-eyed males gives an F_1 in which all the females are red eyed but all the males are white eyed. In this case, every female inherited the dominant w^+ allele from the father's X chromosome, whereas every male inherited the recessive *w* allele from its mother.

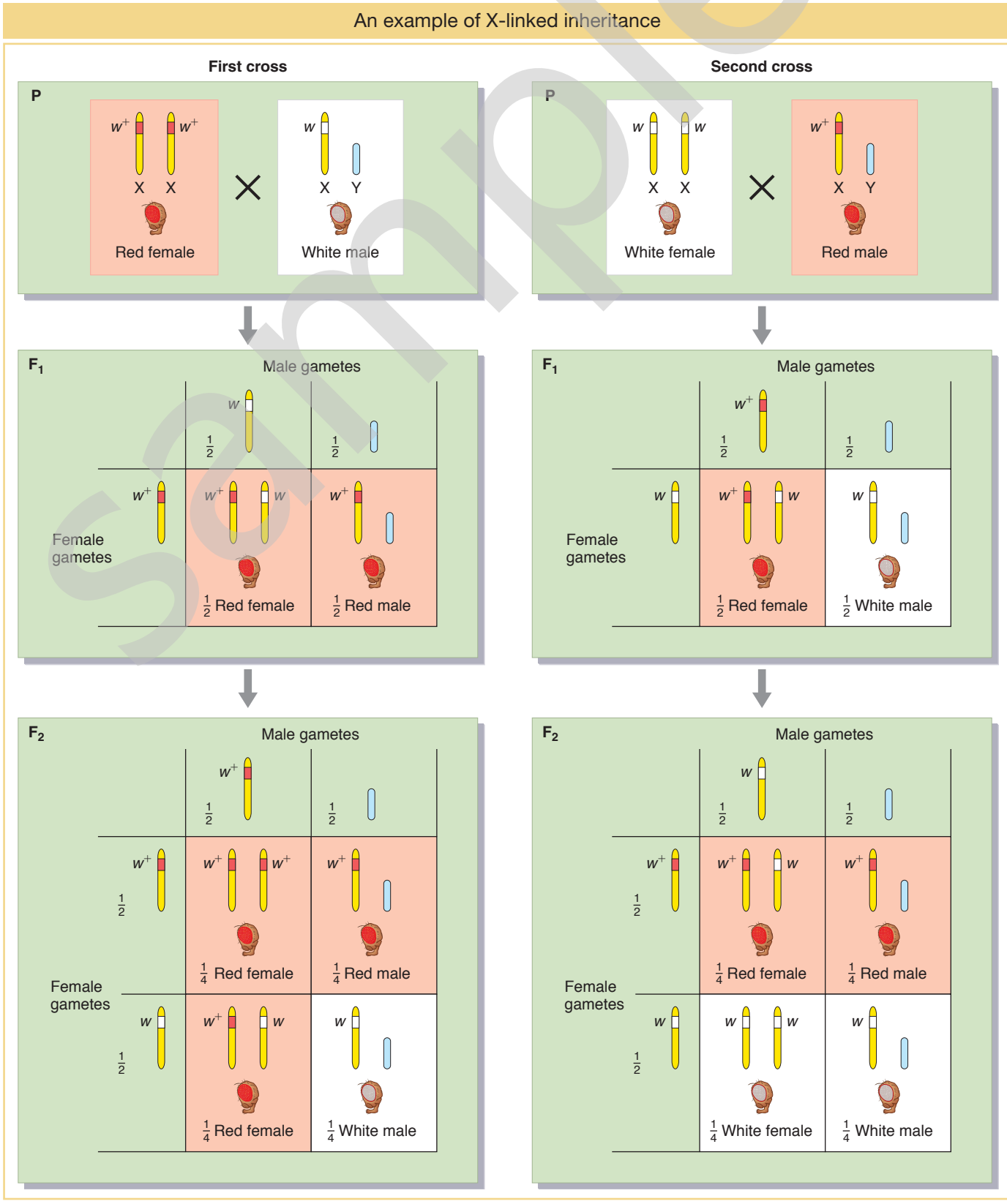


FIGURE 2-17 Reciprocal crosses between red-eyed (red) and white-eyed (white) *Drosophila* give different results. The alleles are X linked, and the inheritance of the X chromosome explains the phenotypic ratios observed, which are different from those of autosomal genes. (In *Drosophila* and many other experimental systems, a superscript plus sign is used to designate the normal, or wild-type, allele. Here, w^+ encodes red eyes and w encodes white eyes.)

The F_2 consists of one-half red-eyed and one-half white-eyed flies of both sexes. Hence, in sex linkage, we see examples not only of different ratios in different sexes, but also of differences between reciprocal crosses.

Note that *Drosophila* eye color has nothing to do with sex determination, and so we have an illustration of the principle that genes on the sex chromosomes are not necessarily related to sexual function. The same is true in humans: in the discussion of pedigree analysis later in this chapter, we shall see many X-linked genes, yet few could be construed as being connected to sexual function.

The abnormal allele associated with white eye color in *Drosophila* is recessive, but abnormal alleles of genes on the X chromosome that are dominant also arise, such as the *Drosophila* mutant hairy wing (*Hw*). In such cases, the wild-type allele (Hw^+) is recessive. The dominant abnormal alleles show the inheritance pattern corresponding to that of the wild-type allele for red eyes in the preceding example. The ratios obtained are the same.

KEY CONCEPT Sex-linked inheritance regularly shows different phenotypic ratios in the two sexes of progeny, as well as different ratios in reciprocal crosses.

Historically, in the early decades of the twentieth century, the demonstration by Morgan of X-linked inheritance of *white eyes* in *Drosophila* was a key piece of evidence that suggested that genes are indeed located on chromosomes, because an inheritance pattern was correlated with one specific chromosome pair. The idea became known as "the chromosome theory of inheritance." At that period in history, it had recently been shown that, in many organisms, sex is determined by an X and a Y chromosome and that, in males, these chromosomes segregate equally at meiosis to regenerate equal numbers of males and females in the next generation. Morgan recognized that the inheritance of alleles of the eye-color gene is exactly parallel to the inheritance of X chromosomes at meiosis; hence, the gene was likely to be on the X chromosome. The inheritance of *white eyes* was extended to *Drosophila* lines that had abnormal numbers of sex chromosomes. With the use of this novel situation, it was still possible to predict gene-inheritance patterns from the segregation of the abnormal chromosomes. That these predictions proved correct was a convincing test of the chromosome theory.

Other genetic analyses revealed that, in chickens and moths, sex-linked inheritance could be explained only if the female was the heterogametic sex. In these organisms, the female sex chromosomes were designated ZW and males were designated ZZ.

2.6 Human Pedigree Analysis

Human matings, like those of experimental organisms, provide many examples of single-gene inheritance. However, controlled experimental crosses cannot be made with humans, and so geneticists must resort to scrutinizing medical records in the hope that informative matings have been made (such as monohybrid crosses) that could be used to infer single-gene inheritance. Such a scrutiny of records of matings is called **pedigree analysis.** A member of a family who first comes to the attention of a geneticist is called the **propositus.** Usually, the phenotype of the propositus is exceptional in some way; for example, the propositus might have some type of medical disorder. The investigator then traces the history of the phenotype through the history of the family and draws a family tree, or pedigree, by using the standard symbols given in Figure 2-18.

To see single-gene inheritance, the patterns in the pedigree have to be interpreted according to Mendel's law of equal segregation, but humans usually have few children and so, because of this small progeny sample size, the expected

3:1 and 1:1 ratios are usually not seen unless many similar pedigrees are combined. The approach to pedigree analysis also depends on whether one of the contrasting phenotypes is a rare disorder or both phenotypes of a pair are common (in which case they are said to be "morphs" of a polymorphism). Most pedigrees are drawn for medical reasons and therefore concern medical disorders that are almost by definition rare. In this case, we have two phenotypes: the presence and the absence of the disorder. Four patterns of single-gene inheritance are revealed in pedigrees. Let's look, first, at recessive disorders caused by recessive alleles of single autosomal genes.

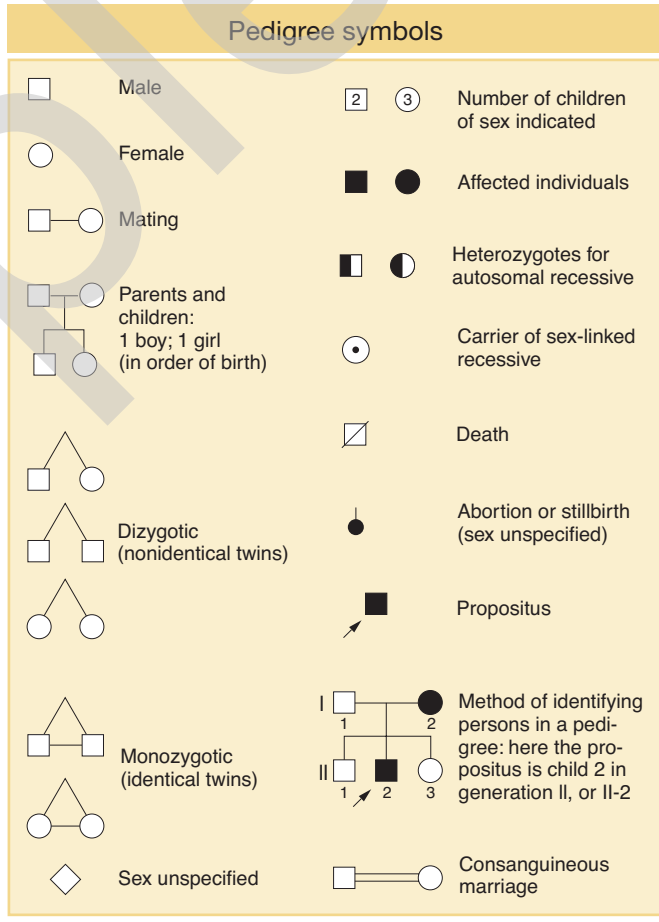
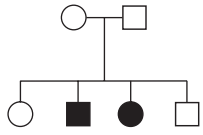


FIGURE 2-18 A variety of symbols are used in human pedigree analysis.

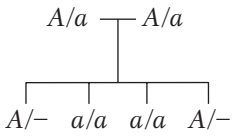
Autosomal recessive disorders

The affected phenotype of an autosomal recessive disorder is inherited as a recessive allele; hence, the corresponding unaffected phenotype must be inherited as the corresponding dominant allele. For example, the human disease phenylketonuria (PKU), discussed earlier, is inherited in a simple Mendelian manner as a recessive phenotype, with PKU determined by the allele *p* and the normal condition determined by *P*. Therefore, people with this disease are of genotype *p/p,* and people who do not have the disease are either *P/P* or *P/p.* Recall that the term *wild type* and its allele symbols are not used in human genetics because wild type is impossible to define.

What patterns in a pedigree would reveal autosomal recessive inheritance? The two key points are that (1) generally the disorder appears in the progeny of unaffected parents and (2) the affected progeny include both males and females. When we know that both male and female progeny are affected, we can infer that we are most likely dealing with simple Mendelian inheritance of a gene on an autosome, rather than a gene on a sex chromosome. The following typical pedigree illustrates the key point that affected children are born to unaffected parents:



From this pattern, we can deduce a simple monohybrid cross, with the recessive allele responsible for the exceptional phenotype (indicated in black). Both parents must be heterozygotes—say, *A/a;* both must have an *a* allele because each contributed an *a* allele to each affected child, and both must have an *A* allele because they are phenotypically normal. We can identify the genotypes of the children (shown left to right) as *A/−*, *a/a, a/a,* and *A/−*. Hence, the pedigree can be rewritten as follows:



This pedigree does not support the hypothesis of X-linked recessive inheritance, because, under that hypothesis, an affected daughter must have a heterozygous mother (possible) and a hemizygous father, which is clearly

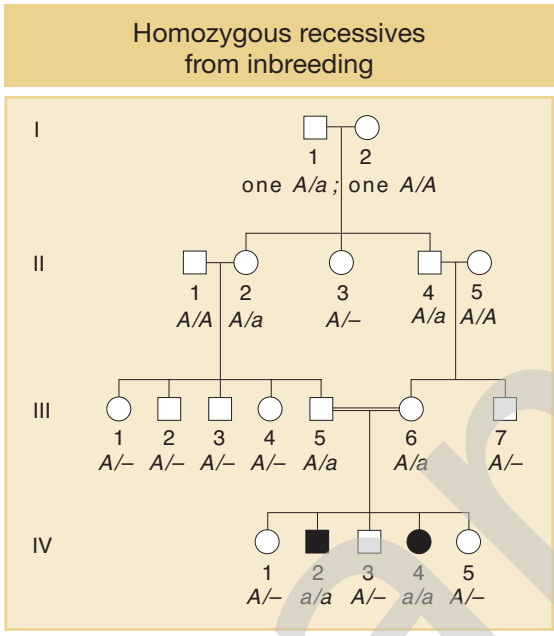


FIGURE 2-19 Pedigree of a rare recessive phenotype determined by a recessive allele *a*. Gene symbols are normally not included in pedigree charts, but genotypes are inserted here for reference. Persons II-1 and II-5 marry into the family; they are assumed to be normal because the heritable condition under scrutiny is rare. Note also that it is not possible to be certain of the genotype in some persons with normal phenotype; such persons are indicated by *A*/—. Persons III-5 and III-6, who generate the recessives in generation IV, are first cousins. They both obtain their recessive allele from a grandparent, either I-1 or I-2.

impossible because the father would have expressed the phenotype of the disorder.

Notice that, even though Mendelian rules are at work, Mendelian ratios are not necessarily observed in single families because of small sample size, as predicted earlier. In the preceding example, we observe a 1:1 phenotypic ratio in the progeny of a monohybrid cross. If the couple were to have, say, 20 children, the ratio would be something like 15 unaffected children and 5 with PKU (a 3:1 ratio), but, in a small sample of 4 children, any ratio is possible, and all ratios are commonly found.

The family pedigrees of autosomal recessive disorders tend to look rather bare, with few black symbols. A recessive condition shows up in groups of affected siblings, and the people in earlier and later generations tend not to be affected. To understand why this is so, it is important to have some understanding of the genetic structure of populations underlying such rare conditions. By definition, if the condition is rare, most people do not carry the abnormal allele. Furthermore, most of those people who do carry the abnormal allele are heterozygous for it rather than homozygous. The basic reason that heterozygotes are much more common than recessive homozygotes is that, to be a recessive homozygote, both parents must have the *a* allele, but, to be a heterozygote, only one parent must have it.

The birth of an affected person usually depends on the rare chance union of unrelated heterozygous parents. However, inbreeding (mating between relatives, sometimes referred to as *consanguinity* in humans) increases the chance that two heterozygotes will mate. An example of a marriage between cousins is shown in Figure 2-19. Individuals III-5 and III-6 are first cousins and produce two homozygotes for the rare allele. You can see from Figure 2-19 that an ancestor who is a heterozygote may produce many descendants who also are heterozygotes. Hence, two cousins can carry the *same* rare recessive allele inherited from a common ancestor. For two *unrelated* persons to be heterozygous, they would have to inherit the rare allele from *both* their families. Thus, matings between relatives generally run a higher risk of producing recessive disorders than do matings between nonrelatives. For this reason, first-cousin marriages contribute a large proportion of people with recessive diseases in the population.

Some other examples of human recessive disorders are shown in Figure 2-20. Cystic fibrosis is a disease inherited on chromosome 7 according to Mendelian rules as an autosomal recessive phenotype. Its most important symptom is the secretion of large amounts of mucus into the lungs, resulting in death from a combination of effects but usually precipitated by infection of the respiratory tract. The mucus can be dislodged by mechanical chest thumpers, and pulmonary infection can be prevented by antibiotics; thus, with treatment, cystic fibrosis patients can live to adulthood. The cystic fibrosis gene (and its mutant allele) was one of the first human disease genes to be isolated at the DNA level, in 1989. This line of research eventually revealed that the disorder is caused by a defective protein that normally transports chloride ions across the cell membrane. The resultant alteration of the salt balance changes the constitution of the lung mucus. This new understanding of gene function in affected and unaffected persons has given hope for more effective treatment.

Human albinism also is inherited in the standard autosomal recessive manner. The mutant allele is of a gene that normally synthesizes the brown or black pigment melanin, normally found in skin, hair, and the retina of the eye (Figure 2-21).

KEY CONCEPT In human pedigrees, an autosomal recessive disorder is generally revealed by the appearance of the disorder in the male and female progeny of unaffected parents.

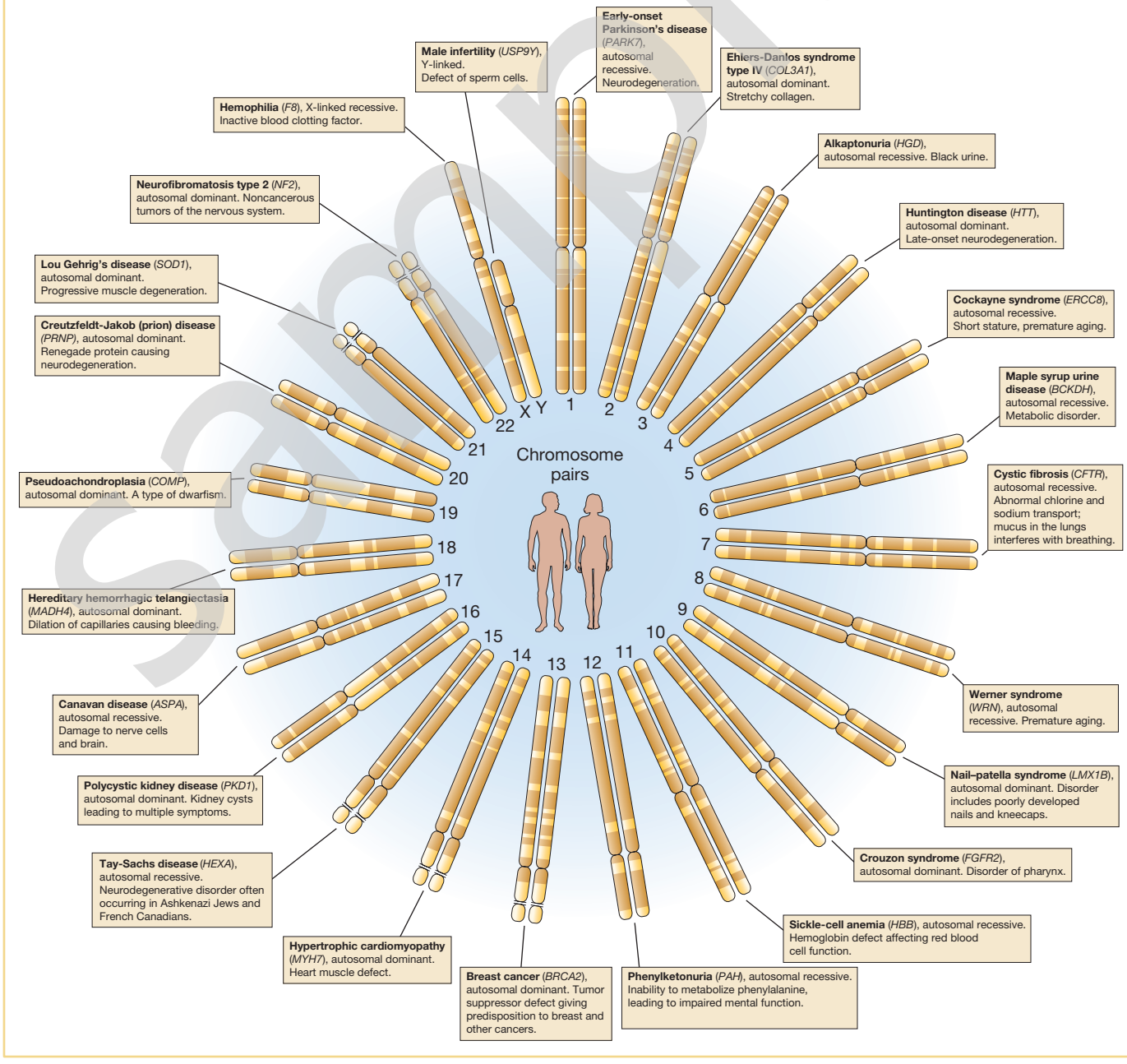


FIGURE 2-20 The positions of the genes mutated in some single-gene diseases, shown in the 23 pairs of chromosomes in a human being. Each chromosome has a characteristic banding pattern. X and Y are the sex chromosomes (XX in women and XY in men). Genes associated with each disease are shown in parentheses.

Autosomal dominant disorders

What pedigree patterns are expected from autosomal dominant disorders? Here, the normal allele is recessive, and the defective allele is dominant. It may seem paradoxical that a rare disorder can be dominant, but remember that dominance and recessiveness are simply properties of how alleles act in heterozygotes and are not defined in reference to how common they are in the population. A good

A mutant gene causes albinism

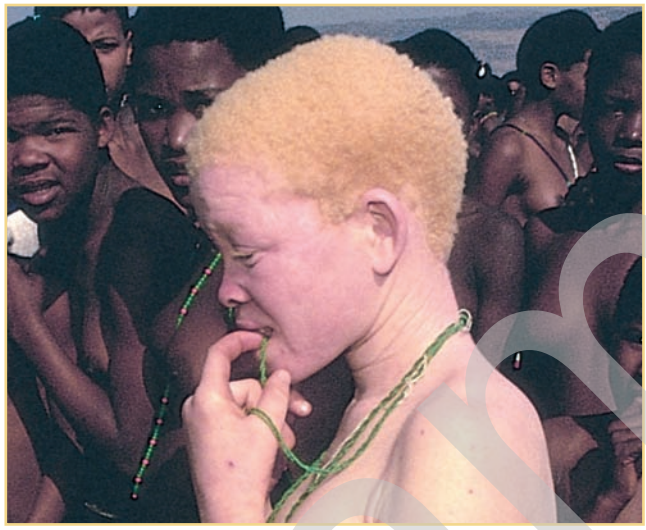


FIGURE 2-21 A nonfunctional version of a skin-pigment gene results in lack of pigment. In this case, both members of the gene pair are mutated. [*Yves GELLIE/Gamma-Rapho/Getty Images.*]

example of a rare dominant phenotype that shows single-gene inheritance is pseudoachondroplasia, a type of dwarfism (Figure 2-22). In regard to this gene, people with normal stature are genotypically *d/d,* and the dwarf phenotype could be in principle *D/d* or *D/D.* However, the two "doses" of the *D* allele in the *D/D* genotype are believed to produce such a severe effect that this genotype is lethal. If this belief is generally true, all dwarf individuals are heterozygotes.

In pedigree analysis, the main clues for identifying an autosomal dominant disorder with Mendelian inheritance are that the phenotype tends to appear in every generation of the pedigree and that affected fathers or mothers transmit the phenotype to both sons and daughters. Again, the equal representation of both sexes among the affected offspring rules out inheritance through the sex chromosomes. The phenotype appears in every generation because, generally, the abnormal allele carried by a person must have come from a parent in the preceding generation. (Abnormal alleles can also arise de novo by mutation. This possibility must be kept in mind for disorders that interfere with reproduction because, here, the condition is unlikely to have been inherited from an affected parent.) A typical pedigree for a dominant disorder is shown in Figure 2-23. Once again, notice that Mendelian ratios are not necessarily observed in families. As with recessive disorders, persons bearing one copy of the rare *A* allele (*A/a*) are much more common than those bearing two copies (*A/A*); so most affected people are heterozygotes, and virtually all matings that produce progeny with dominant disorders are $A/a \times a/a$. Therefore, if the

Pseudoachondroplasia phenotype

FIGURE 2-22 The human pseudoachondroplasia phenotype is illustrated here by a family of five sisters and two brothers. The phenotype is determined by a dominant allele, which we can call *D,* that interferes with the growth of long bones during development. This photograph was taken when the family arrived in Israel after the end of World War II. [*Bettmann/CORBIS.*]