21ST CENTURY SCIENCE

NEW THINKING ABOUT GENETICS

EDITED BY KARA ROGERS

Britannica Educational Publishing

NEW THINKING ABOUT
GENETICS

21st Century Science

New Thinking About
Genetics

Edited by Kara Rogers, Senior Editor, Biomedical Sciences

in association with
ROSEN
EDUCATIONAL SERVICES

Published in 2011 by Britannica Educational Publishing
(a trademark of Encyclopædia Britannica, Inc.)
in association with Rosen Educational Services, LLC
29 East 21st Street, New York, NY 10010.

Copyright © 2011 Encyclopædia Britannica, Inc. Britannica, Encyclopædia Britannica, and the Thistle logo are registered trademarks of Encyclopædia Britannica, Inc. All rights reserved.

Rosen Educational Services materials copyright © 2011 Rosen Educational Services, LLC. All rights reserved.

Distributed exclusively by Rosen Educational Services.
For a listing of additional Britannica Educational Publishing titles, call toll free (800) 237-9932.

First Edition

Britannica Educational Publishing
Michael I. Levy: Executive Editor
J.E. Luebering: Senior Manager
Marilyn L. Barton: Senior Coordinator, Production Control
Steven Bosco: Director, Editorial Technologies
Lisa S. Braucher: Senior Producer and Data Editor
Yvette Charboneau: Senior Copy Editor
Kathy Nakamura: Manager, Media Acquisition
Kara Rogers: Senior Editor, Biomedical Sciences

Rosen Educational Services
Alexandra Hanson-Harding: Editor
Nelson Sá: Art Director
Matthew Cauli: Designer, Cover Design
Cindy Reiman: Photography Manager
Introduction by David Nagle

Library of Congress Cataloging-in-Publication Data

New thinking about genetics / edited by Kara Rogers.—1st ed.
p.; cm.—(21st century science)
"In association with Britannica Educational Publishing, Rosen Educational Services."
Includes bibliographical references and index.
ISBN 978-1-61530-169-0 (eBook)
1. Genetics—Popular works. I. Rogers, Kara. II. Series: 21st century science.
[DNLM: 1. Genetic Phenomena. 2. Philosophy. QU 450 N5325 2010]
QH437.N49 2010
576.5—dc22

2009044215

On the cover: *In this composite image, a scientist works on a sample of human genetic material in front of a backdrop of models of DNA (deoxyribonucleic acid). DNA is an organic chemical that codes genetic information so that inherited traits can be transmitted to new generations of an organism.* Frederick Florin/AFP/Getty Images (pipette and hand); www.istockphoto.com / Nicolas Hansen (DNA).

Pp. 5, 21, 38, 57, 85, 123, 158, 190, 220, 262, 266, 268 © www.istockphoto.com/Mark Evans; p. 12 © www.istockphoto.com/Nicolas Hansen; p. 20 © www.istockphoto.com/Mads Abildgaard.

Contents

Introduction 12

Chapter 1: The Code of Life 21

Ancient Theories of Pangenesis and Blood in Heredity 22

Preformation and Natural Selection 24

The Work of Gregor Mendel 25

Theoretical Interpretation 27

Rediscovery 28

How the Gene Idea Became Reality 29

Early Molecular Genetics 30

DNA and the Genetic Code 31

Francis Crick 33

James Watson 34

Maurice Wilkins 36

Chapter 2: The Physical Basis of Heredity 38

Genes 38

Alleles 39

Genotype 40

Phenotype 40

Chromosomes 41

Sex Chromosomes 44

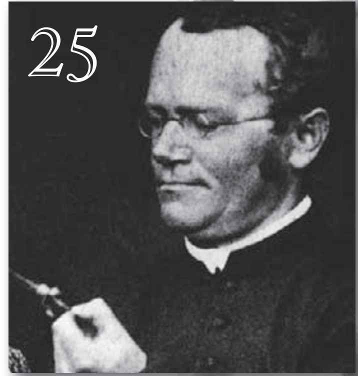
Chromosomes During Cell Division: Mitosis 45

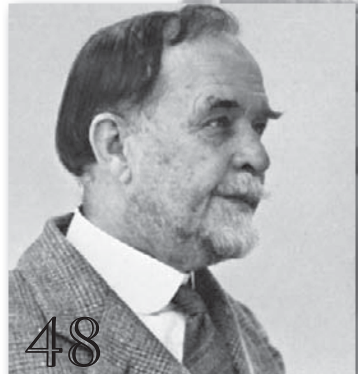
Chromosomes During Cell Division: Meiosis 46

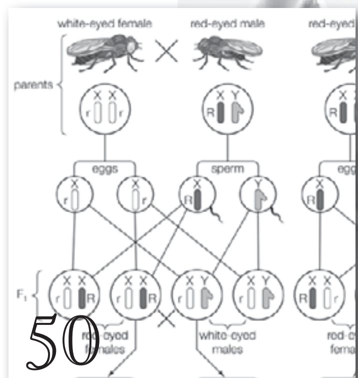
Linkage Groups 47

Thomas Hunt Morgan 48

Chromosomal Aberrations 52

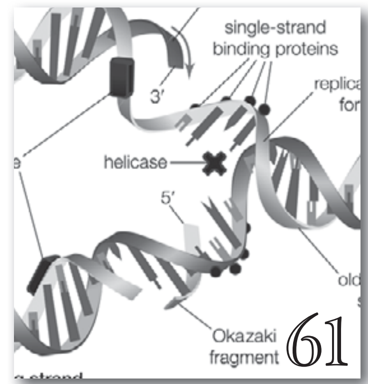


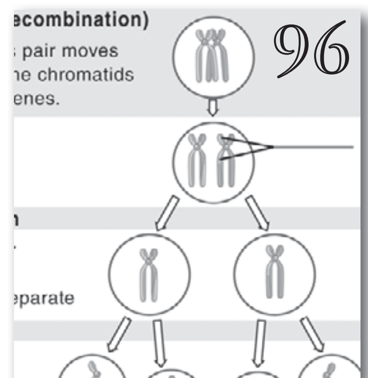




Chapter 3: DNA as the Agent of Heredity 57
Structure and Composition of DNA 59
DNA Replication 60
Expression of the Genetic Code 64
Transcription 67
Translation 69
Gene Mutation 71
Mechanisms of Mutation 72
Repair of Mutation 75
Regulation of Gene Expression 76
RNA Interference and Gene Silencing 79
Repetitive DNA 81
Barbara McClintock 82

Chapter 4: The Study of Genetics 85
Classical Genetics 85
Cytogenetics 86
Cytogenetic Techniques 87
Fluorescence In Situ Hybridization 87
Microbial Genetics 88
Molecular Genetics 89
Molecular Techniques 90
Polymerase Chain Reaction 90
DNA Fingerprinting 93
Immunogenetics 95
Genomics 95
Sequencing and Bioinformatic Analysis of Genomes 97
Functional Genomics 98
Gene Identification by Microarray Genomic Analysis 98
Comparative Genomics 99





Population Genetics 100
- Experimental Breeding 101
- Mathematical Techniques 102
- Hardy-Weinberg Law 102
- Genetic Drift 103
- Founder Principle 104

Behaviour Genetics 104
- Early History of Behaviour Genetics 104
- Methods of Behaviour Genetics Study 105

Study of Human Genetics 108

Epigenetics 110

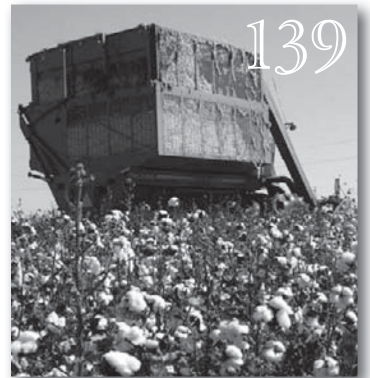
Stem Cells 112
- Embryonic Stem Cells 114
- Adult Stem Cells 115
- Induced Pluripotent Stem Cells 118

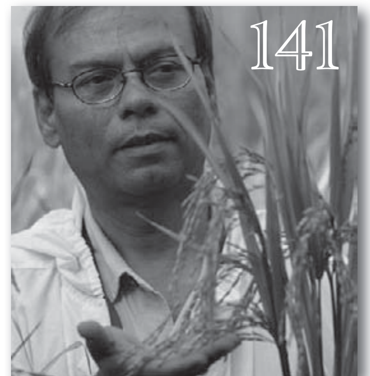
RNA Interference 120

Chapter 5: Genetic Engineering 123

Recombinant DNA Technology 125
- DNA Cloning 125
- Creating the Clone 127
- Isolating the Clone 131
- DNA Sequencing 132
- In Vitro Mutagenesis 133
- Gene Therapy 135
- Reverse Genetics 135
- Diagnostics 136
- Protein Manufacture 136

Genetically Modified Organisms 137
- GMOs in Agriculture 139
- GMOs in Medicine and Research 143
- Role of GMOs in Environmental Management 144

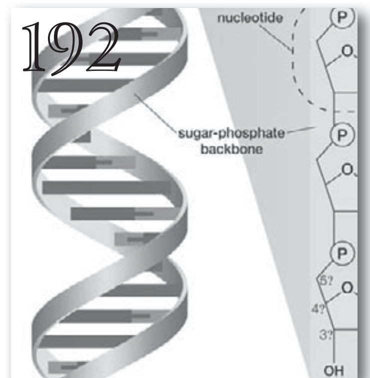




Sociopolitical Relevance of GMOs 145
Cloning 146
Early Cloning Experiments 147
Reproductive Cloning 149
Therapeutic Cloning 153
Ethical Controversies of Cloning 155

Chapter 6: Genetics Applied to Plants and Animals 158
Plant Breeding 158
Goals 160
Evaluation of Plants 161
Methods of Plant Breeding 163
Animal Breeding 174
Breeding and Variation 175
Breeding 178
Selection 181
Breeding Systems 187

Chapter 7: Human Genetics 190
The Human Chromosomes 190
The Human Genome 191
Role of the Human Genome in Research 194
Origins of the Human Genome 195
Social Impacts of Human Genome Research 197
The Genetics of Human Blood 198
Blood Types 199
Serum Proteins 200
Hemoglobin 201
The Genetics of Antibody Formation 202
The Genetics of Cellular Immunity 204

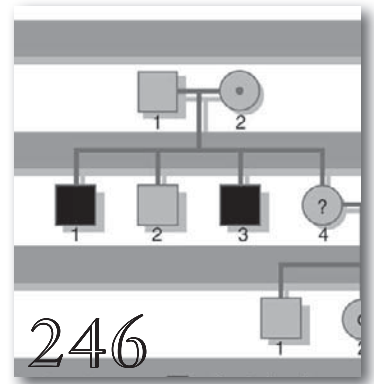


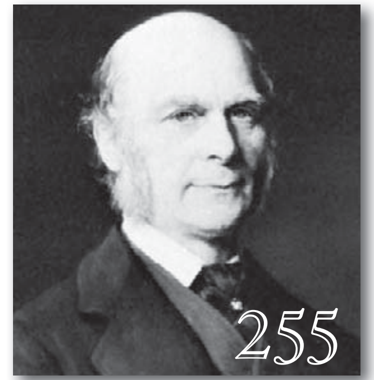
Influence of the Environment 205
Fraternal Twins 206
Identical Twins 207
Diagnosis of Twin Types 208
Inferences from Twin Studies 208
Genetics and the Concept of Race 212
Modern Scientific Explanations of Human Biological Variation 212
The Scientific Debate Over "Race" 217

Chapter 8: Genetic Diseases of Humans 220
Classes of Genetic Disease 221
Diseases Caused by Chromosomal Aberrations 221
Abnormalities of the Sex Chromosomes 223
Diseases of Autosomal Dominant Inheritance 226
Diseases of Autosomal Recessive Inheritance 229
Repeat Expansions 231
Mitochondrial DNA Mutations 232
Imprinted Gene Mutations 232
Diseases Caused by Multifactorial Inheritance 234
Victor McKusick 235
Genetics of Cancer 235
Telomeres 239
Genetic Damage from Environmental Agents 240
Management of Genetic Disease 242
Genetic Counseling 243
Calculating Risks of Known Carriers 243

Estimating Probability: Bayes's Theorem 245
Prenatal Diagnosis 248
Genetic Testing 252
Eugenics 255
The Future of Genetics in Medicine 259

Glossary 262
For Further Reading 266
Index 268





INTRODUCTION

Today, almost everyone has heard about DNA analysis from crime shows on television. It's how the police catch bad guys. But it can also be a way to find good guys. Since the 1990s, the remains of U.S. soldiers have been positively identified through enhanced procedures using mitochondrial DNA, which provides information about maternal lineage. In 2008 and 2009, for instance, the remains of U.S. Vietnam pilots who have been missing in action since the late 1960s and Operation Desert Storm pilots who have been missing since 1991 were positively identified using mitochondrial DNA analysis. The same process is also being used to identify the remains of more than 250 British and Australian World War I veterans discovered in a mass grave outside of Fromelles, France, and to successfully identify remains from conflicts dating back to the American Civil War. This book helps to unravel the mysteries of the science of DNA, chromosomes, and genes, as well as to bring to the forefront current methods and theories of genetic inquiry.

Humans have long been aware that various plants and animals seemed to have similarities in form and function to the "parents." Ancient peoples used this understanding in daily life; for example, mating male and female animals with superior characteristics to create new generations with improved physical traits. These early observations underlie a process that many centuries later came to be known as heredity.

Even into the 18th century ideas were somewhat general as to the actual process of heredity. Believers in preformation, the idea that a small, complete "homunculus" (Latin for "little human") were battling over whether the little fellow existed in the sperm or the egg. The theory of epigenesis seemed to make better sense to those who believed that the egg was just a sack containing no

structures, only a kind of "jelly" that developed through a series of steps.

Modern understanding of heredity began in the mid-1860s, with the work of a man whom many call the father of genetics, an Austrian monk named Gregor Mendel. Mendel provided the first mathematical foundation of the science of genetics through his work with garden peas. He followed a single characteristic or trait of the peas rather than a more general view, and he used exact counts rather than estimates. Mendel's classical experimentation methods are used today for gene discovery and assembly to affect biological properties of interest.

The components that Mendel believed held basic heredity data are known today to be genes and chromosomes. The role of these components in genetic inheritance was established in the early part the 20th century. However, the mystery of DNA (deoxyribonucleic acid), which was first discovered in 1869 and of which genes and chromosomes are composed, was not resolved until later, in 1953, when biochemists Francis Crick, James Watson, and Maurice Wilkins determined the molecule's now famous double-helix shape and inferred the process of genetic coding within cellular DNA. The trio was eventually awarded the Nobel Prize for Physiology or Medicine in 1962 for their discovery, which is widely considered one of the defining moments in genetics and generally in science.

While Mendel's classical genetics was a huge step forward in scientific reasoning, once Watson, Crick, and Wilkins mapped the structure of DNA, many more avenues could be explored using much more advanced techniques and equipment. Beyond the classical genetics of Mendel, there are several other important areas of study in the larger field of genetics. They include disciplines that involve the isolation and examination of

smaller components, such as in cytogenetics (the microscopic study of chromosomes and genes), microbial genetics (using simpler bacteria or viruses), and molecular genetics. In addition, there are fields such as genomics and population genetics that investigate very broad concepts and rely on various scientific methodologies.

And yet, genetics does not exist strictly in the world of science. It touches all of human thought and experience and has sociopolitical, ethical, and moral dimensions as well. But modern techniques such as genetic engineering take genetic manipulation to a whole new level. Genetic engineering, which involves artificially modifying DNA, is a primary means by which genetics influences human existence in modern times. For instance, the production of genetically modified organisms (GMOs) has provided many benefits to people. Essentially, GMOs are organisms engineered by genome alteration to reflect what are considered to be "beneficial" aspects. An example of this is a certain kind of rice that has been changed so that it contains more iron to aid iron-deficient diets.

On occasion, the promise of genetic engineering has been offset by possible negative consequences, creating controversy. While the benefits of GMOs may be seen in increased food production and in the production of more nutrient-fortified crops, concerns have arisen about the safety of these products. The potential to stimulate an allergic response in some persons, the development of insecticide-resistant "superbugs," the increased use of chemicals in fields of herbicide-resistant crops, and other worries have caused tensions in global trade and other areas that demand a cautious approach when dealing with GMOs.

Cloning, in which an organisms possesses the same DNA as the unit or individual from which it is derived, is another controversial area of genetics. While cloning is a natural process in some biological systems, its development

and use as a tool for genetics and reproductive research occurred gradually, over the course of decades in the 20th century. For many years, success in cloning animals was limited to experiments involving "lower" classes of organisms, such as amphibians, and it was generally thought that mammalian cloning was beyond the reach of scientists. However, in the 1990s, the successful cloning of Dolly, a female sheep, proved otherwise.

Shortly after the breakthrough in mammalian cloning, another important advance in genetics took place—the sequencing of the human genome. In recent years, scientists have studied the human genome in order to better understand what defines humans genetically as a species and what defines human populations culturally and historically. Experts known from studying twins that many factors in personality are purely genetic, whereas others are determined by environmental and cultural factors. They also known that humans and Neanderthals, a group of archaic humans, have very similar genomes, with about 99.5 percent similarity in sequence.

The study of genetics also helps experts understand human migration patterns. They know that humans migrated out of Africa approximately 60,000 years ago, moving to Europe, Asia, and eventually the Americas. Experts have learned that different skin colours and other physical features, such as different body builds, helped humans to adapt to different environments over the millennia. And yet many experts have also come to believe that the concept of race seems to be outliving its usefulness as the extensive intermingling of peoples across the globe means that people have the races of several continents in their blood—that all people regardless of their physical variations are capable of learning any kind of cultural behavior.

One unfortunate characteristic that humans share is a vulnerability to genetic-based diseases. Genetic defects cause a variety of diseases and disorders in children and adults. In the 20th and 21st centuries, a number of advancements were made in the understanding of various genetic abnormalities. Some genetic syndromes are caused by sex-linked chromosomes, whereas other problems are caused by autosomal (non-sex-linked) chromosomes, including Huntington disease and PKU (phenylketonuria). Some diseases have multifactorial causes, such as cancers that can be exacerbated by genetic factors. Scientists estimate that genetic factors cause at least 30 percent of cases of the eye tumour retinoblastoma. Some breast and colorectal cancers are influenced by genetic factors as well.

But gene damage doesn't just come from parents. Chromosomes can also be damaged by environmental factors. The decomposition of ozone has led to skin cancer, X rays and chemicals can cause chromosome abnormalities or mutations, and certain viruses, such as HIV (human immunodeficiency virus), Epstein-Barr, and hepatitis B and C, can cause genetic-level damage that encourage the growth of cancers.

One of the fastest-growing fields in medicine is genetic counseling, due to the growth of knowledge in how to look at and test for various maladies. Scientists can help potential parents evaluate their chances of passing along a genetic disease to their child even before they conceive.

Dealing with genes and the human species can be tricky, however. One controversial area is eugenics. Although the idea of breeding better, healthier people started in idealism, eugenics slowly began to become discredited, especially when the Nazis used the idea to support extermination of entire races. From eugenics to cloning,

there have been ethical considerations in the field of genetics that in many cases have trumped pure science. In those areas that are generally agreed to be of social benefit, however, advances in genetics have accelerated exponentially and look to hold great promise for advancement in medical procedures and pharmacology.

In October 2009, both the Nobel Prize for Chemistry and the Nobel Prize for Physiology or Medicine were awarded for gene-based research. Americans Venkatraman Ramakrishnan and Thomas Steitz and Israeli Ada Yonath shared the Nobel Prize in Chemistry for mapping out the position of the thousands of atoms that make up ribosomes. Ribosomes use information from DNA to make proteins needed for life. The scientists' research will help experts develop new antibiotics. The Nobel Prize for Physiology or Medicine was given to Americans Elizabeth H. Blackburn, Carol W. Greider, and Jack W. Szostak for their work on telomeres, structures on the ends of chromosomes that protect DNA from degrading when it replicates. This research may help scientists find new ways to fight cancer. It's no wonder that many scientists believe that as the 20th century was the century of physics, the 21st will be the century of biology.

CHAPTER 1

THE CODE OF LIFE

Since the dawn of civilization, humankind has recognized the influence of heredity and has applied its principles to the improvement of cultivated crops and domestic animals. A Babylonian tablet more than 6,000 years old, for example, shows pedigrees of horses and indicates possible inherited characteristics. Other old carvings show cross-pollination of date palm trees. Most of the mechanisms of heredity, however, remained a mystery until the 19th century, when genetics as a systematic science began.

Genetics arose out of the identification of genes, the fundamental units responsible for heredity. Genetics may be defined as the study of genes at all levels, including the ways in which they act in the cell and the ways in which they are transmitted from parents to offspring. Modern genetics focuses on the chemical substance that genes are made of, called deoxyribonucleic acid, or DNA, and the ways in which it affects the chemical reactions that constitute the living processes within the cell. Gene action depends on interaction with the environment. Green plants, for example, have genes containing the information necessary to synthesize the photosynthetic pigment chlorophyll that gives them their green colour. Chlorophyll is synthesized in an environment containing light because the gene for chlorophyll is expressed only when it interacts with light. If a plant is placed in a dark environment, chlorophyll synthesis stops because the gene is no longer expressed.

Genetics as a scientific discipline stemmed from the work of Gregor Mendel in the middle of the 19th century. Mendel suspected that traits were inherited as discrete

units, and, although he knew nothing of the physical or chemical nature of genes at the time, his units became the basis for the development of the present understanding of heredity. All present research in genetics can be traced back to Mendel's discovery of the laws governing the inheritance of traits. The word *gene*, coined in 1909 by Danish botanist Wilhelm Johannsen, has given genetics its name.

Genetics forms one of the central pillars of biology and overlaps with many other areas such as agriculture, medicine, and biotechnology.

ANCIENT THEORIES OF PANGENESIS AND BLOOD IN HEREDITY

Although scientific evidence for patterns of genetic inheritance did not appear until Mendel's work, history shows that humankind must have been interested in heredity since ancient times. Curiosity must first have been based on human family resemblances, such as similarity in body structure, voice, gait, and gestures. Such notions were instrumental in the establishment of family and royal dynasties. Early nomadic tribes were interested in the qualities of the animals that they herded and domesticated and, undoubtedly, bred selectively. The first human settlements that practiced farming appear to have selected crop plants with favourable qualities. Ancient tomb paintings show racehorse breeding pedigrees containing clear depictions of the inheritance of several distinct physical traits in the horses. Despite this interest, the first recorded speculations on heredity did not exist until the time of the ancient Greeks; some aspects of their ideas are still considered relevant today.

Hippocrates (*c.* 460–*c.* 375 BCE), known as the father of medicine, believed in the inheritance of acquired characteristics, and, to account for this, he devised the hypothesis known as pangenesis. He postulated that all organs of the body of a parent gave off invisible "seeds," which were like miniaturized building components and were transmitted during sexual intercourse, reassembling themselves in the mother's womb to form a baby.

Aristotle (384–322 BCE) emphasized the importance of blood in heredity. He thought that the blood supplied generative material for building all parts of the adult body, and he reasoned that blood was the basis for passing on this generative power to the next generation. In fact, he believed that the male's semen was purified blood and that a woman's menstrual blood was her equivalent of semen. These male and female contributions united in the womb to produce a baby. The blood contained some type of hereditary essences, but he believed that the baby would develop under the influence of these essences, rather than being built from the essences themselves.

Aristotle's ideas about the role of blood in procreation were probably the origin of the still prevalent notion that somehow the blood is involved in heredity. Today people still speak of certain traits as being "in the blood" and of "blood lines" and "blood ties." The Greek model of inheritance, in which a teeming multitude of substances was invoked, differed from that of the Mendelian model. Mendel's idea was that distinct differences between individuals are determined by differences in single yet powerful hereditary factors. These single hereditary factors were identified as genes. Copies of genes are transmitted through sperm and egg and guide the development of the offspring. Genes are also responsible for reproducing the distinct features of both parents that are visible in their children.

PREFORMATION AND NATURAL SELECTION

In the two millennia between the lives of Aristotle and Mendel, few new ideas were recorded on the nature of heredity. In the 17th and 18th centuries the idea of preformation was introduced. Scientists using the newly developed microscopes imagined that they could see miniature replicas of human beings inside sperm heads. French biologist Jean-Baptiste Lamarck invoked the idea of "the inheritance of acquired characters," not as an explanation for heredity but as a model for evolution.

Lamarck lived at a time when the fixity of species was taken for granted, yet he maintained that this fixity was found only in a constant environment. He enunciated the law of use and disuse, which states that when certain organs become specially developed as a result of some environmental need, then that state of development is hereditary and can be passed on to progeny. He believed that in this way, over many generations, giraffes could arise from deerlike animals that had to keep stretching their necks to reach high leaves on trees.

British naturalist Alfred Russel Wallace originally postulated the theory of evolution by natural selection. However, Charles Darwin's observations during his circumnavigation of the globe aboard the HMS *Beagle* (1831–36) provided evidence for natural selection and his suggestion that humans and animals shared a common ancestry. Many scientists at the time believed in a hereditary mechanism that was a version of the ancient Greek idea of pangenesis, and Darwin's ideas did not appear to fit with the theory of heredity that sprang from the experiments of Mendel.

THE WORK OF GREGOR MENDEL

Before Gregor Mendel, theories for a hereditary mechanism were based largely on logic and speculation, not on experimentation. In his monastery garden, Mendel carried out a large number of cross-pollination experiments between variants of the garden pea, which he obtained as pure breeding lines. He crossed peas with yellow seeds to those with green seeds and observed that the progeny seeds (the first generation, F_1) were all yellow. When the F_1 individuals were self-pollinated or crossed among themselves, their progeny (F_2) showed a ratio of 3:1 (3/4 yellow and 1/4 green). He deduced that, since the F_2 generation contained some green individuals, the determinants of greenness must have been present in the F_1 generation, although they were not expressed because yellow is dominant over green.

Gregor Mendel (1822–1884) pioneered the type of scientific experimentation that is used in modern genetics. Here he can be seen working in his laboratory. Authenticated News/Hulton Archive/Getty Images

From the precise mathematical 3:1 ratio (of which he found several other examples), he deduced not only the existence of discrete hereditary units (genes) but also that the units were present in pairs in the pea

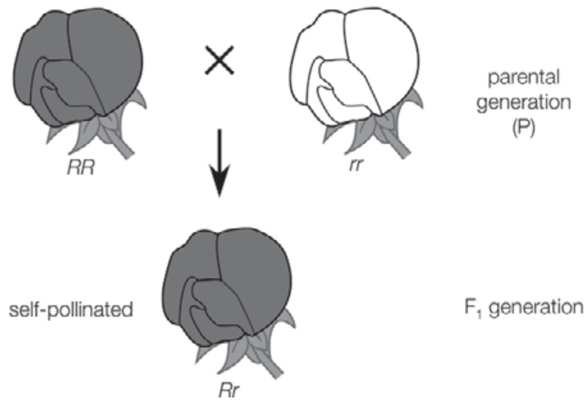
plant and that the pairs separated during gamete formation (gametes are mature sexual reproductive cells). Hence, the two original lines of pea plants were proposed to be *YY* (yellow) and *yy* (green). The gametes from these were *Y* and *y*, thereby producing an F_1 generation of *Yy* that were yellow in colour because of the dominance of *Y*. In the F_1 generation, half the gametes were *Y* and the other half were *y*, making the F_2 generation produced from random mating 1/4 *Yy*, 1/2 *YY*, and 1/4 *yy*, thus explaining the 3:1 ratio. The forms of the pea colour genes, *Y* and *y*, are called alleles. An allele is one member of a pair of genes located in a particular place on a particular chromosome. The dominant allele determines hereditary variation in an organism.

Mendel also analyzed pure lines that differed in pairs of characters, such as seed colour (yellow versus green) and seed shape (round versus wrinkled). The cross of yellow round seeds with green wrinkled seeds resulted in an F_1 generation that were all yellow and round, revealing the dominance of the yellow and round traits. However, the F_2 generation produced by self-pollination of F_1 plants showed a ratio of 9:3:3:1 (9/16 yellow round, 3/16 yellow wrinkled, 3/16 green round, and 1/16 green wrinkled; note that a 9:3:3:1 ratio is simply two 3:1 ratios combined). From this result and others like it, he deduced the independent assortment of separate gene pairs at gamete formation.

Mendel's success can be attributed in part to his classic experimental approach. He chose his experimental organism well and performed many controlled experiments to collect data. From his results, he developed brilliant explanatory hypotheses and went on to test these hypotheses experimentally. Mendel's methodology established a prototype for genetics that is still used today for gene discovery and understanding the genetic properties of inheritance.

Theoretical Interpretation

Mendel related his results to the cell theory of fertilization, according to which a new organism is generated from the fusion of two cells. Mendel proposed that in order for a cross-pollinated hybrid to produce pure breeding forms of



Cross of a purple-flowered (shown here in gray) and a white-flowered strain of peas. R *stands for the gene for purple flowers and* r *for the gene for white flowers.* Encyclopædia Britannica, Inc.

both the dominant and the recessive type, the hybrid must form germ cells bearing the potential to yield either type. Thus, there had to be some temporary accommodation of the two differing characters in the hybrid as well as a separation process in the formation of the germ cells (in this case, the pollen and egg cells). This has since been described as the law of segregation, or the doctrine of the purity of the germ cells. Since one pollen cell fuses with one egg cell, all possible combinations of the differing pollen and egg cells would yield just the results suggested by Mendel's theory.

Mendel first presented his results in two separate lectures in 1865 to the Natural Science Society in Brünn. His paper *Experiments on Plant Hybrids* was published in the society's journal, *Verhandlungen des naturforschenden Vereines* in Brünn, the following year. It attracted little attention, although many libraries received it and reprints were sent out. The tendency of those who read it was to conclude that Mendel had simply demonstrated more accurately what was already widely assumed—namely, that hybrid progeny revert to their originating forms. They overlooked the potential for variability and the evolutionary implications that his demonstration of the recombination of traits made possible. Most notably, Swiss botanist Karl Wilhelm von Nägeli actually corresponded with Mendel, despite remaining skeptical as to the significance of his results and doubting that the germ cells in hybrids could be pure.

Rediscovery

In 1900, Dutch botanist and geneticist Hugo de Vries, German botanist and geneticist Carl Erich Correns, and Austrian botanist Erich Tschermak von Seysenegg independently reported results of hybridization experiments

similar to Mendel's, though each later claimed not to have known of Mendel's work while doing their own experiments. However, both de Vries and Correns had read Mendel earlier—Correns even made detailed notes on the subject—but had forgotten. De Vries had a diversity of results in 1899, but it was not until he reread Mendel in 1900 that he was able to select and organize his data into a rational system. Tschermak had not read Mendel before obtaining his results, and his first account of his data offers an interpretation in terms of hereditary potency. He described the 3:1 ratio as an "unequal valency" (*Wertigkeit*). In subsequent papers he incorporated the Mendelian theory of segregation and the purity of the germ cells into his text.

HOW THE GENE IDEA BECAME REALITY

In Great Britain, biologist William Bateson became the leading proponent of Mendel's theory. Around him gathered an enthusiastic band of followers. However, Darwinian evolution was assumed to be based chiefly on the selection of small, blending variations, whereas Mendel worked with clearly nonblending variations. Bateson soon found that championing Mendel aroused opposition from Darwinians. He and his supporters were called Mendelians, and their work was considered irrelevant to evolution. It took some three decades before the Mendelian theory was sufficiently developed to find its rightful place in evolutionary theory.

The distinction between a characteristic and its determinant was not consistently made by Mendel or by his successors, the early Mendelians. In 1909, Danish botanist and geneticist Wilhelm Johannsen clarified this point and named the determinants genes. Four years later,

American zoologist and geneticist Thomas Hunt Morgan located the genes on the chromosomes (threadlike strands made of tightly compacted DNA and proteins), and the popular picture of them as beads on a string emerged. This discovery had implications for Mendel's claim of an independent transmission of traits, for genes close together on the same chromosome are not transmitted independently.

Moreover, as genetic studies pushed the analysis down to smaller and smaller dimensions, the Mendelian gene appeared to fragment. Molecular genetics has thus challenged any attempts to achieve a unified conception of the gene as the elementary unit of heredity. Today the gene is defined in several ways, depending upon the nature of the investigation. Genetic material can be synthesized, manipulated, and hybridized with genetic material from other species, but to fully understand its functions in the whole organism, an understanding of Mendelian inheritance is necessary. As the architect of genetic experimental and statistical analysis, Mendel remains the acknowledged father of genetics.

EARLY MOLECULAR GENETICS

In 1908, British physician Archibald Garrod proposed the important idea that the human disease alkaptonuria, and certain other hereditary diseases, were caused by inborn errors of metabolism, providing for the first time evidence that linked genes with molecular action at the cell level. Molecular genetics did not begin in earnest until 1941 when American geneticist George Beadle and American biochemist Edward Tatum showed that the genes they were studying in the fungus *Neurospora crassa* acted by coding for catalytic proteins called enzymes. Subsequent studies in other organisms extended this idea to show that

genes generally code for proteins. Soon afterward, American bacteriologist Oswald Avery, Canadian American geneticist Colin M. MacLeod, and American biologist Maclyn McCarty showed that bacterial genes are made of DNA, a finding that was later extended to all organisms.

DNA AND THE GENETIC CODE

A major landmark was attained in 1953 when American geneticist and biophysicist James D. Watson and British biophysicists Francis Crick and Maurice Wilkins devised a double helix model for DNA structure. In this model, DNA can be likened to a spiraling staircase or a twisting ladder. The DNA double helix consists of two intertwined sugar-phosphate chains, with the flat base pairs forming the steps between them. This model showed that DNA was capable of self-replication by separating its complementary strands and using them as templates for the synthesis of new DNA molecules. Each of the intertwined strands of DNA was proposed to be a chain of chemical groups called nucleotides, of which there were known to be four types. Because proteins are strings of amino acids, it was proposed that a specific nucleotide sequence of DNA could contain a code for an amino acid sequence and hence protein structure. In 1955, American molecular biologist Seymour Benzer, extending earlier studies in fruit flies (*Drosophila)*, showed that the mutant sites within a gene could be mapped in relation to each other. His linear map indicated that the gene itself is a linear structure.

In 1958 the strand-separation method for DNA replication (called the semiconservative method) was demonstrated experimentally for the first time by American molecular biologist Matthew Meselson and American geneticist Franklin W. Stahl. In 1961, Crick and South African biologist Sydney Brenner showed that the genetic

Francis Crick (left) *and James Watson in Cambridge in 1953, prior to their discovery of the molecular structure of DNA.* Courtesy of the James D. Watson Collection, Cold Spring Harbor Laboratory Archives

code must be read in triplets of nucleotides, called codons. American geneticist Charles Yanofsky showed that the positions of mutant sites within a gene matched perfectly the positions of altered amino acids in the amino acid sequence of the corresponding protein. In 1966 the complete genetic code of all 64 possible triplet coding units (codons), and the specific amino acids they code for, was deduced by American biochemists Marshall Nirenberg and Har Gobind Khorana. Subsequent studies in many organisms showed that the double helical structure of DNA, the mode of its replication, and the genetic code are the same in virtually all organisms, including plants, animals, fungi, bacteria, and viruses. In 1961, French biologist François Jacob and French biochemist Jacques Monod established the prototypical model for gene regulation by

showing that bacterial genes can be turned on (initiating transcription into ribonucleic acid [RNA] and protein synthesis) and off through the binding action of regulatory proteins to a region just upstream of the coding region of the gene.

FRANCIS CRICK

(b. June 8, 1916, Northampton, Northamptonshire, Eng.—d. July 28, 2004, San Diego, Calif., U.S.)

British biophysicist Francis Crick contributed to the determination of the molecular structure of DNA, the chemical substance ultimately responsible for the hereditary control of life functions. Crick, along with James Watson and Maurice Wilkins, received the 1962 Nobel Prize for Physiology or Medicine for this accomplishment, which became a cornerstone of genetics and was widely regarded as one of the most important discoveries of 20th-century biology.

During World War II, Crick interrupted his education to work as a physicist in the development of magnetic mines for use in naval warfare, but afterward he turned to biology at the Strangeways Research Laboratory, University of Cambridge (1947). Interested in pioneering efforts to determine the three-dimensional structures of large molecules found in living organisms, he transferred to the university's Medical Research Council Unit at the Cavendish Laboratories in 1949.

In 1951, when the American biologist James Watson arrived at the laboratory, it was known that the mysterious nucleic acids, especially DNA, played a central role in the hereditary determination of the structure and function of each cell. Watson convinced Crick that knowledge of DNA's three-dimensional structure would make its hereditary role apparent. Using the X-ray diffraction studies of DNA done by Wilkins and X-ray diffraction pictures produced by Rosalind Franklin, Watson and Crick were able to construct the double helix molecular model, which is consistent with the known physical and chemical properties of DNA. Watson and Crick theorized that if the two helical strands were separated, each would serve as a template (pattern) for the formation, from small molecules in the cell, of a new

sister strand identical to its former partner. This copying process explained replication of the gene and, eventually, the chromosome, known to occur in dividing cells. Their model also indicated that the sequence of bases along the DNA molecule spells some kind of code "read" by a cellular mechanism that translates it into the specific proteins responsible for a cell's particular structure and function.

By 1961 Crick had evidence to show that each group of three bases (a codon) on a single DNA strand designates the position of a specific amino acid on the backbone of a protein molecule. He also helped to determine which codons code for each of the 20 amino acids normally found in protein and thus helped clarify the way in which the cell eventually uses the DNA "message" to build proteins. From 1977 until his death, Crick held the position of distinguished professor at the Salk Institute for Biological Studies in San Diego, California, where he conducted research on the neurological basis of consciousness. His book *Of Molecules and Men* (1966) discusses the implications of the revolution in molecular biology. *What Mad Pursuit: A Personal View of Scientific Discovery* was published in 1988. In 1991 Crick received the Order of Merit.

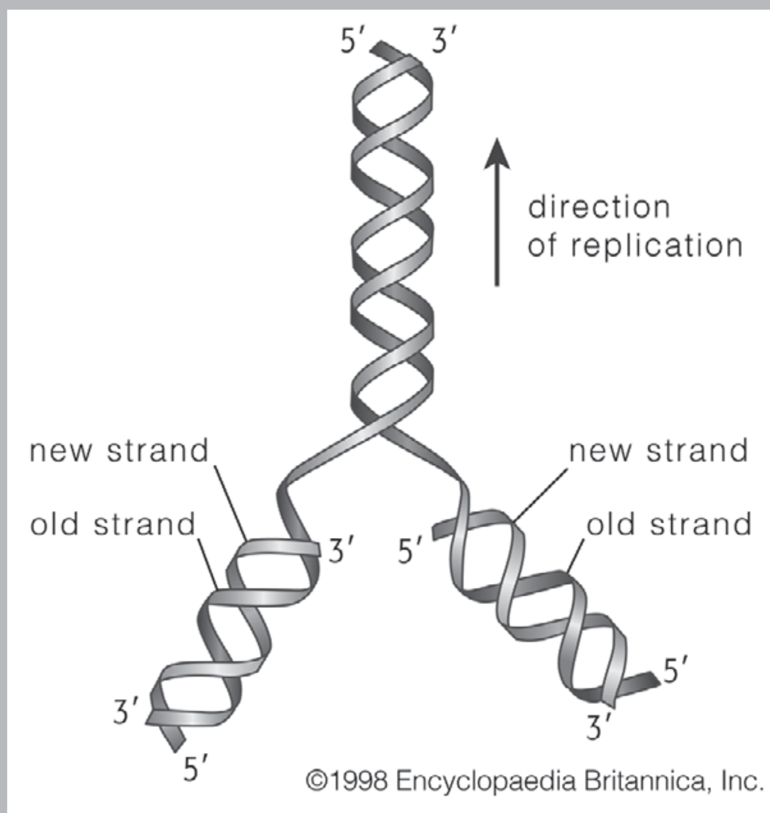
JAMES WATSON

(b. April 6, 1928, Chicago, Ill., U.S.)

American geneticist and biophysicist James Watson played a crucial role in the discovery of the molecular structure of DNA. For this accomplishment he was awarded the 1962 Nobel Prize for Physiology or Medicine with Francis Crick and Maurice Wilkins.

Watson enrolled at the University of Chicago when only 15 and graduated in 1947. From his virus research at Indiana University (Ph.D., 1950), and from the experiments of Canadian-born American bacteriologist Oswald Avery, which proved that DNA affects hereditary traits, Watson became convinced that the gene could be understood only after something was known about nucleic acid molecules. He learned that scientists working in the Cavendish Laboratories at the University of Cambridge were using photographic patterns made by X rays that had been shot through protein crystals to study the structure of protein molecules.

After working at the University of Copenhagen, where he first determined to investigate DNA, he did research at the Cavendish



The initial proposal of the structure of DNA by James Watson and Francis Crick was accompanied by a suggestion on the means of replication.

Laboratories (1951–53). There Watson learned X-ray diffraction techniques and worked with Crick on the problem of DNA structure. In 1952 he determined the structure of the protein coat surrounding the tobacco mosaic virus but made no dramatic progress with DNA. Suddenly, in the spring of 1953, Watson saw that the essential DNA components—four organic bases—must be linked in definite pairs. This discovery was the key factor that enabled Watson and Crick to formulate the double helix molecular model for DNA. Watson and Crick's model also showed how the DNA molecule could duplicate itself. Thus it became known how genes, and eventually chromosomes, duplicate themselves. Watson and Crick published their epochal discovery in two papers in the British journal *Nature* in

April–May 1953. Their research answered one of the fundamental questions in genetics.

Watson subsequently taught at Harvard University (1955–76), where he served as professor of biology (1961–76). He conducted research on the role of nucleic acids in the synthesis of proteins. In 1965 he published *Molecular Biology of the Gene*, one of the most extensively used modern biology texts. He later wrote *The Double Helix* (1968), an informal and personal account of the DNA discovery and the roles of the people involved in it, which aroused some controversy. In 1968 Watson assumed the leadership of the Laboratory of Quantitative Biology at Cold Spring Harbor, Long Island, N.Y., and made it a world centre for research in molecular biology. He concentrated its efforts on cancer research. In 1981 his *The DNA Story* (written with John Tooze) was published. From 1988 to 1992 at the National Institutes of Health, Watson helped direct the Human Genome Project, a project to map and decipher all the genes in the human chromosomes, but he eventually resigned because of alleged conflicts of interests involving his investments in private biotechnology companies.

In early 2007 Watson's own genome was sequenced and made publicly available on the Internet. He was the second person in history to have a personal genome sequenced in its entirety. In October of the same year, he sparked controversy by making a public statement alluding to the idea that the intelligence of Africans might not be the same as that of other peoples and that intellectual differences among geographically separated peoples might arise over time as a result of genetic divergence. Watson's remarks were immediately denounced as racist. Though he denied this charge, he resigned from his position at Cold Spring Harbor and formally announced his retirement less than two weeks later.

MAURICE WILKINS

(b. Dec. 15, 1916, Pongaroa, N.Z.—d. Oct. 6, 2004, London, Eng.)

New Zealand-born British biophysicist Maurice Wilkins investigated DNA using a technique called X-ray diffraction. His studies proved crucial to the determination of DNA's molecular structure by James Watson and Francis Crick. For this work the three scientists were jointly awarded the 1962 Nobel Prize for Physiology or Medicine.

Wilkins, the son of a physician (who was originally from Dublin), was educated at King Edward's School in Birmingham, England, and St. John's College in Cambridge. His doctoral thesis, completed for the University of Birmingham in 1940, contained his original formulation of the electron-trap theory of phosphorescence and thermoluminescence. He participated for two years during World War II in the Manhattan Project at the University of California in Berkeley, working on mass spectrograph separation of uranium isotopes for use in the atomic bomb.

Upon his return to Great Britain, Wilkins lectured at the University of St. Andrews in Scotland. In 1946 he joined the Medical Research Council's Biophysics Unit at King's College in London. In 1955 he became its deputy director, and from 1970 to 1980 he served as the unit's director. There he began the series of investigations that led ultimately to his X-ray diffraction studies of DNA. Wilkins headed a group that included Rosalind Franklin, a crystallographer who produced DNA pictures that also aided the work of Crick and Watson. Wilkins later applied X-ray diffraction techniques to the study of ribonucleic acid (RNA).

At King's College proper, Wilkins was professor of molecular biology (1963–70) and of biophysics (1970–81) and emeritus professor thereafter. While there he published literature on light microscopy techniques for cytochemical research. His autobiography, *The Third Man of the Double Helix*, was published in 2003.

CHAPTER 2
THE PHYSICAL BASIS OF HEREDITY

When Gregor Mendel formulated his laws of heredity, he postulated a particulate nature for the units of inheritance. What exactly these particles were he did not know. Today scientists understand not only the physical location of hereditary units (i.e., the genes) but their molecular composition as well. The unraveling of the physical basis of heredity makes up one of the most fascinating chapters in the history of biology. To understand how heredity works, it is important to understand the various genetic components of which we and all living things are made.

GENES

Genes are units of hereditary information. Each gene occupies a fixed position (locus) on a chromosome. Genes achieve their effects by directing the synthesis of proteins.

In eukaryotes (organisms with a clearly defined nucleus, such as animals and plants), genes are contained within the cell nucleus. The mitochondria (in animals) and the chloroplasts (in plants) also contain small subsets of genes distinct from the genes found in the nucleus. In prokaryotes (organisms lacking a distinct nucleus, such as bacteria), genes are contained in a single chromosome that is free-floating in the cell cytoplasm. Many bacteria also contain plasmids—extrachromosomal genetic elements with a small number of genes.

Genes are composed of DNA, except in some viruses, which have genes consisting of a closely related compound called ribonucleic acid (RNA). A DNA molecule is

composed of two chains of nucleotides that wind about each other to resemble a twisted ladder. The sides of the ladder are made up of sugars and phosphates, and the rungs are formed by bonded pairs of nitrogenous bases. These bases are adenine (A), guanine (G), cytosine (C), and thymine (T). An A on one chain bonds to a T on the other (thus forming an A–T ladder rung); similarly, a C on one chain bonds to a G on the other. If the bonds between the bases are broken, the two chains unwind, and free nucleotides within the cell attach themselves to the exposed bases of the now-separated chains. The free nucleotides line up along each chain according to the base-pairing rule—A bonds to T, C bonds to G. This process results in the creation of two identical DNA molecules from one original and is the method by which hereditary information is passed from one generation of cells to the next.

Alleles

Alleles, also called allelomorphs, are any one of two or more genes that may occur alternatively at a given site (locus) on a chromosome. Alleles may occur in pairs, or there may be multiple alleles affecting the expression (phenotype) of a particular trait. If the paired alleles are the same, the organism is said to be homozygous for that trait; if they are different, the organism is heterozygous. A dominant allele will override the traits of a recessive allele in a heterozygous pairing. In some traits, however, alleles may be codominant—i.e., neither acts as dominant or recessive. An example is the human ABO blood system; persons with type AB blood have one allele for A and one for B. (Persons with neither are type O.)

Most traits are determined by more than two alleles. Multiple forms of the allele may exist, though only two will attach to the designated gene site during meiosis (a

stage of cell division in the process of sexual reproduction). Also, some traits are controlled by two or more gene sites. Both possibilities multiply the number of alleles involved. All genetic traits are the result of the interactions of alleles. Mutation, crossing over, and environmental conditions selectively change the frequency of phenotypes (and thus their alleles) within a population.

Genotype

The genotype is the genetic constitution of an organism. The genotype determines the hereditary potentials and limitations of an individual from embryonic formation through adulthood. Among organisms that reproduce sexually, an individual's genotype comprises the entire complex of genes inherited from both parents. It can be demonstrated mathematically that sexual reproduction virtually guarantees that each individual will have a unique genotype (except for those individuals, such as identical twins, who are derived from the same fertilized egg).

Phenotype

The phenotype represents the observable characteristics of an organism, such as shape, size, colour, and behaviour, that result from the interaction of its genotype with the environment. The common type of a group of physically similar organisms is sometimes also known as the phenotype.

The phenotype may change constantly throughout the life of an individual because of environmental changes and the physiological and morphological changes associated with aging. Different environments can influence the development of inherited traits (for example, as size is affected by available food supply) and alter expression by

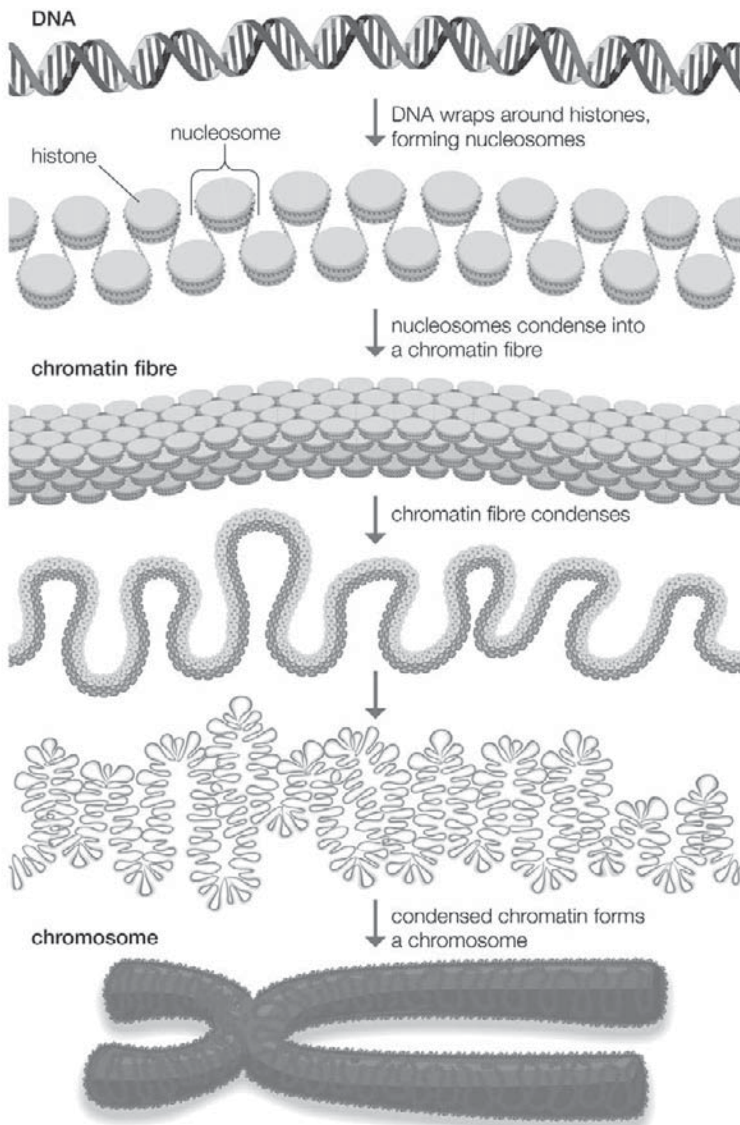
similar genotypes (for example, twins maturing in dissimilar families). Furthermore, all inherited possibilities in the genotype are not expressed in the phenotype because some are the result of latent, recessive, or inhibited genes.

CHROMOSOMES

The microscopic, threadlike part of the cell that carries hereditary information in the form of genes is known as a chromosome. The structure and location of chromosomes is one of the chief differences between the two basic types of cells—prokaryotic cells and eukaryotic cells. Among organisms with prokaryotic cells (i.e., bacteria and blue-green algae), chromosomes consist entirely of DNA. The single chromosome of a prokaryotic cell is not enclosed within a nuclear membrane. Among all other organisms (i.e., the eukaryotes), the chromosomes are contained in a membrane-bound cell nucleus. The chromosomes of a eukaryotic cell consist primarily of DNA attached to a protein core. They also contain RNA. Among both prokaryotes and eukaryotes, the arrangement of components in the DNA molecules determines the genetic information.

Every species has a characteristic number of chromosomes (chromosome number). In addition, the number of chromosomes within the nucleus is usually constant in all individuals of a given species—for example, 46 in the human; 40 in the house mouse; 8 in the fruit fly (*Drosophila melanogaster*); 20 in corn (maize); 24 in the tomato; and 48 in the potato.

In species that reproduce asexually, the chromosome number is the same in all the cells of the organism. Among sexually reproducing organisms, the number of chromosomes in the body (somatic) cells is diploid ($2n$; a pair of each chromosome), twice the haploid ($1n$) number found



DNA wraps around proteins called histones to form units known as nucleosomes. These units condense into a chromatin fibre, which condenses further to form a chromosome. Encyclopædia Britannica, Inc.

in the sex cells, or gametes. The haploid number is produced during meiosis. During fertilization, two gametes combine to produce a zygote, a single cell with a diploid set of chromosomes.

Somatic cells reproduce by dividing, a process called mitosis. Between cell divisions the genetic material (chromatin) is diffused throughout the nucleus in a tangled network of filaments called chromonemata. These long filaments are formed from the uncoiled chromosomes and probably provide a large surface area, thereby facilitating DNA synthesis. During this phase, DNA duplicates itself in preparation for cell division.

At the onset of cell division, the chromonemata coil up and are surrounded by a protein sheath, forming a tiny rod, or chromosome. Each chromosome actually consists of a set of duplicate chromatids that are held together by the centromere. The centromere is the point of attachment to the spindle fibres (part of a structure that pulls the chromatids to opposite ends of the cell). During the middle stage in cell division, the centromere duplicates, and the chromatid pair separates; each chromatid becomes a separate chromosome at this point. The cell divides, and both of the daughter cells have a complete (diploid) set of chromosomes. The chromosomes uncoil in the new cells, again forming the diffuse network of filaments.

Among many organisms that have separate sexes, there are two basic types of chromosomes: sex chromosomes and autosomes. Autosomes control the inheritance of all the characteristics except the sex-linked ones, which are controlled by the sex chromosomes. Humans have 22 pairs of autosomes and one pair of sex chromosomes. All act in the same way during cell division.

Chromosome breakage is the physical breakage of subunits of a chromosome. It is usually followed by

reunion (frequently at a foreign site, resulting in a chromosome unlike the original). Breakage and reunion of homologous chromosomes during meiosis is the basis for the classical model of crossing over, which results in unexpected types of offspring of a mating.

Sex Chromosomes

The sex chromosomes determine whether an individual is male or female. The sex chromosomes of humans and other mammals are designated as X and Y. Individuals having two X chromosomes (XX) are female; individuals having one X chromosome and one Y chromosome (XY) are male.

The X chromosome resembles a large autosomal chromosome with a long and a short arm. The Y chromosome has one long arm and a very short second arm. This path to maleness or femaleness originates at the moment of meiosis, when a cell divides to produce gametes, or sex cells having half the normal number of chromosomes. During meiosis the male XY sex-chromosome pair separates and passes on an X or a Y to separate gametes; the result is that one-half of the gametes (sperm) that are formed contains the X chromosome and the other half contains the Y chromosome. The female has two X chromosomes, and all female egg cells normally carry a single X. The eggs fertilized by X-bearing sperm become females (XX), whereas those fertilized by Y-bearing sperm become males (XY).

Unlike the paired autosomes, in which each member normally carries alleles of the same genes, the paired sex chromosomes do not carry an identical complement of genetic information. The X chromosome, being larger, carries many more genes than does the Y. Traits controlled

by genes found only on the X chromosome are said to be sex-linked. Recessive sex-linked traits, such as hemophilia and red-green colour blindness, occur far more frequently in men than in women. This is because the male who inherits the recessive allele on his X chromosome has no allele on his Y chromosome to counteract its effects. The female, on the other hand, must inherit the recessive allele on both of her X chromosomes in order to fully display the trait.

A woman who inherits the recessive allele for a sex-linked disorder on one of her X chromosomes may, however, show a limited expression of the trait. The reason for this is that, in each somatic cell of a normal female, one of the X chromosomes is randomly deactivated. This deactivated X chromosome can be seen as a small, dark-staining structure—the Barr body—in the cell nucleus.

The effects of genes carried only on the Y chromosome are, of course, expressed only in males. Most of these genes are the so-called maleness determiners, which are necessary for development of the testes in the fetus.

Chromosomes During Cell Division: Mitosis

Mitosis is a process of cell duplication, or reproduction, during which one cell gives rise to two genetically identical daughter cells. Strictly applied, the term *mitosis* is used to describe the duplication and distribution of chromosomes.

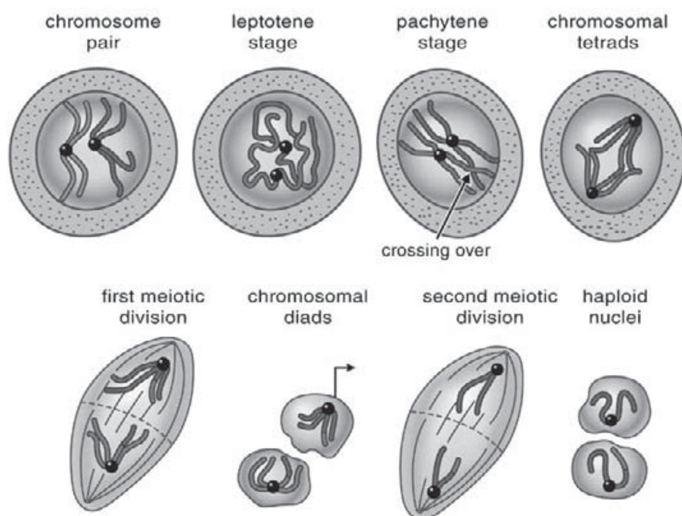
When the chromosomes condense during cell division, they have already undergone replication. At this point, each chromosome consists of two identical replicas, called chromatids, joined at a point called the centromere. During mitosis the sister chromatids separate, one going to each daughter cell. Chromosomes thus meet the first

criterion for being the repository of genes: they are replicated, and a full copy is passed to each daughter cell during mitosis.

Chromosomes During Cell Division: Meiosis

Meiosis is the division of a germ cell involving two fissions of the nucleus and giving rise to four gametes, or sex cells, each possessing half the number of chromosomes of the original cell. During meiosis, each chromosome (consisting of two chromatids) becomes paired with a physically similar chromosome. These homologous chromosomes then separate, with one member of each pair going to a different cell.

Meiosis involves multiple stages. At the leptotene stage the chromosomes appear as long, thin threads. At pachytene they pair, the corresponding portions of the two chromosomes lying side by side. The chromosomes



Behaviour of chromosomes at meiosis.

then duplicate and contract into paired chromatids. At this stage the pair of chromosomes is known as a tetrad, as it consists of four chromatids. Also at this stage an extremely important event occurs: portions of the maternal and paternal chromosomes are exchanged. This exchange process, called crossing over, results in chromatids that include both paternal and maternal genes and consequently introduces new genetic combinations. Since each chromosome carries many genes, and the chromosome pairs exchange segments through the process of crossing over, the genes rather than the chromosomes are the units of Mendelian segregation.

The first meiotic division separates the chromosomal tetrads, with the paternal chromosome (whose chromatids now contain some maternal genes) going to one cell and the maternal chromosome (containing some paternal genes) going to another cell. During the second meiotic division the chromatids separate. The original diploid cell has thus given rise to four haploid gametes. Not only has a reduction in chromosome number occurred, but the resulting single member of each homologous chromosome pair may be a new combination (through crossing over) of genes present in the original diploid cell.

Linkage Groups

All of the genes on a single chromosome belong to a linkage group, and as such, they are inherited as a group—that is, during cell division they act and move as a unit rather than independently. The existence of linkage groups is the reason some traits do not comply with Mendel's law of independent assortment (recombination of genes and the traits they control); i.e., the principle applies only if genes are located on different chromosomes. Variation in the gene composition of a chromosome can occur when a

Thomas Hunt Morgan. Courtesy of the California Institute of Technology, Pasadena

THOMAS HUNT MORGAN

(b. Sept. 25, 1866, Lexington, Ky., U.S.—d. Dec. 4, 1945, Pasadena, Calif.)

American zoologist and geneticist Thomas Hunt Morgan was known for his experimental research with the fruit fly (*Drosophila*) by which he established the chromosome theory of heredity. He showed that genes are linked in a series on chromosomes and are responsible for identifiable, hereditary traits. Morgan's work played a key role in establishing the field of genetics. He received the Nobel Prize for Physiology or Medicine in 1933.

Experiments in Embryology

During the period 1893–1910, Morgan applied experimental techniques to fundamental problems of embryology. In order to identify causally related events during development, he analyzed such problems as the formation of embryos from separated blastomeres (early embryonic cells) and fertilization in nucleated and nonnucleated egg fragments. As examples of the effects of physical factors, he analyzed the way in which the spatial orientation of eggs affects their future development and the action of salt concentration on the development of fertilized and unfertilized eggs. In 1904 he accepted an invitation to assume the professorship of experimental zoology at Columbia University, where, during the next 24 years, he conducted most of his important research in heredity.

Like most embryologists and many biologists at the turn of the century, Morgan found the Darwinian theory of evolution lacking in plausibility. It was difficult to conceive of the development of complex adaptations simply by an accumulation of slight chance variations.

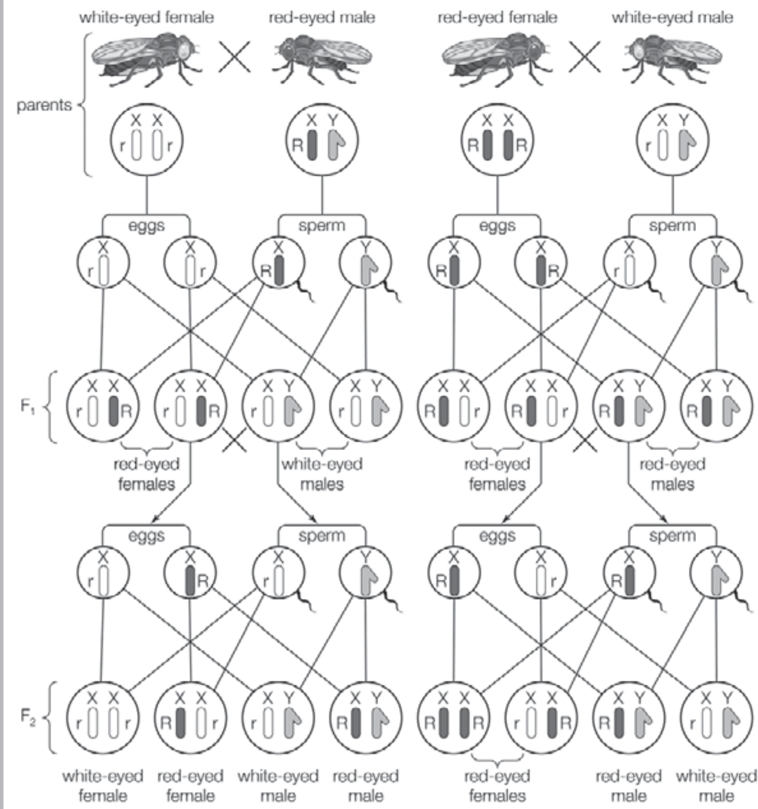
Moreover, Darwin had provided no mechanism of heredity to account for the origin or transmission of variations, except his early and hypothetical theory of pangenesis. Although Morgan believed that evolution itself was a fact, the mechanism of natural selection proposed by Darwin seemed incomplete because it could not be put to an experimental test.

Morgan had quite different objections to the Mendelian and chromosome theories. Both theories attempted to explain biological phenomena by postulating units or material entities in the cell that somehow control developmental events. To Morgan this was too reminiscent of the preformation theory—the idea that the fully formed adult is present in the egg or sperm—that had dominated embryology in the 18th and early 19th centuries. Although Morgan admitted that the chromosomes might have something to do with heredity, he argued in 1909 and 1910 that no single chromosome could carry specific hereditary traits. He also claimed that Mendelian theory was purely hypothetical: although it could account for and even predict breeding results, it could not describe the true processes of heredity. That each pair of chromosomes separates, with the individual chromosomes then going into different sperm or egg cells in exactly the same manner as Mendelian factors, did not seem to be sufficient proof to Morgan for claiming that the two processes had anything to do with each other.

Work on *Drosophila*

Morgan apparently began breeding *Drosophila* in 1908. In 1909 he observed a small but discrete variation known as white-eye in a single male fly in one of his culture bottles. Aroused by curiosity, he bred the fly with normal (red-eyed) females. All of the offspring (F_1) were red-eyed. Brother-sister matings among the F_1 generation produced a second generation (F_2) with some white-eyed flies, all of which were males. To explain this curious phenomenon, Morgan developed the hypothesis of sex-limited—today called sex-linked—characters, which he postulated were part of the X chromosome of females.

Other genetic variations arose in Morgan's stock, many of which were also found to be sex-linked. Because all the sex-linked characters were usually inherited together, Morgan became convinced that the X chromosome carried a number of discrete hereditary units, or factors.



Sex-linked inheritance of white eyes in Drosophila *flies.* Encyclopædia Britannica, Inc.

He adopted the term *gene*, which was introduced by the Danish botanist Wilhelm Johannsen in 1909, and concluded that genes were possibly arranged in a linear fashion on chromosomes. Much to his credit, Morgan rejected his skepticism about both the Mendelian and chromosome theories when he saw from two independent lines of evidence—breeding experiments and cytology—that one could be treated in terms of the other.

In collaboration with American geneticists Alfred Henry Sturtevant, Calvin Blackman Bridges, and Hermann Joseph Muller, who were graduates at Columbia, Morgan quickly developed the *Drosophila* work into a large-scale theory of heredity. Particularly important in this work was the demonstration that each Mendelian gene could be assigned a specific position along a linear chromosome

"map." Further cytological work showed that these map positions could be identified with precise chromosome regions, thus providing definitive proof that Mendel's factors had a physical basis in chromosome structure. A summary and presentation of the early phases of this work was published by Morgan, Sturtevant, Bridges, and Muller in 1915 as the influential book *The Mechanism of Mendelian Heredity*. To varying degrees Morgan also accepted the Darwinian theory by 1916.

In 1928 Morgan was invited to organize the division of biology of the California Institute of Technology. He was also instrumental in establishing the Marine Laboratory on Corona del Mar as an integral part of Caltech's biology training program. In subsequent years, Morgan and his coworkers, including a number of postdoctoral and graduate students, continued to elaborate on the many features of the chromosome theory of heredity.

In 1924 Morgan received the Darwin Medal, and in 1939 he was awarded the Copley Medal by the Royal Society of London. Among Morgan's most important books are *Evolution and Adaptation* (1903), *A Critique of the Theory of Evolution* (1916), *Heredity and Sex* (1913), and *The Theory of the Gene*.

chromosome breaks, and the sections join with the partner chromosome if it has broken in the same places. This exchange of genes between chromosomes—crossing over—usually occurs during meiosis, when the total number of chromosomes is halved.

Sex linkage is the tendency of a characteristic to be linked to one sex. For example, the X chromosome in *Drosophila* flies and humans carries a complete set of genes; the Y chromosome has only a few genes. Eggs of females carry an X chromosome; sperm of males may carry an X or a Y. An egg fertilized by a sperm with an X chromosome results in a female; one fertilized by a sperm with a Y chromosome results in a male. In offspring with the XY chromosome pair, any trait carried by the X chromosome will appear unless there is a corresponding gene (allele) on the Y chromosome.

Chromosomal Aberrations

The chromosome set of a species remains relatively stable over long periods of time. However, within populations there can be found abnormalities involving the structure or number of chromosomes. These alterations arise spontaneously from errors in the normal processes of the cell. Their consequences are usually deleterious, giving rise to individuals who are unhealthy or sterile, though in rare cases alterations provide new adaptive opportunities that allow evolutionary change to occur. In fact, the discovery of visible chromosomal differences between species has given rise to the belief that radical restructuring of chromosome architecture has been an important force in evolution.

Deletions

The simplest, but perhaps most damaging, structural change is a deletion—the complete loss of a part of one chromosome. In a haploid cell this is lethal because part of the essential genome is lost. However, even in diploid cells deletions are generally lethal or have other serious consequences. In a diploid a heterozygous deletion results in a cell that has one normal chromosome set and another set that contains a truncated chromosome. Such cells show genomic imbalance, which increases in severity with the size of the deletion. Another potential source of damage is that any recessive, deleterious, or lethal alleles that are in the normal counterpart of the deleted region will be expressed in the phenotype.

In humans, cri-du-chat ("cry of the cat") syndrome is caused by a heterozygous deletion at the tip of the short arm of chromosome 5. This condition is the result of a deletion arising in parental germinal tissues or even in sex cells. The manifestations of this deletion, in addition to

the "cat cry" that gives the syndrome its name, include severe intellectual disability and an abnormally small head.

Duplications

A heterozygous duplication (an extra copy of some chromosome region) also results in a genomic imbalance with deleterious consequences. Small duplications within a gene can arise spontaneously. Larger duplications can be caused by crossovers following asymmetrical chromosome pairing or by meiotic irregularities resulting from other types of altered chromosome structures. If a duplication becomes homozygous, it can provide the organism with an opportunity to acquire new genetic functions through mutations within the duplicate copy.

Inversions

An inversion occurs when a chromosome breaks in two places and the region between the break rotates 180° before rejoining with the two end fragments. If the inverted segment contains the centromere (i.e., the point where the two chromatids are joined), the inversion is said to be pericentric; if not, it is called paracentric. Inversions do not result in a gain or loss of genetic material, and they have deleterious effects only if one of the chromosomal breaks occurs within an essential gene or if the function of a gene is altered by its relocation to a new chromosomal neighbourhood (called the position effect).

Individuals who are heterozygous for inversions produce aberrant meiotic products along with normal products. The only way uninverted and inverted segments can pair is by forming an inversion loop. If no crossovers occur in the loop, half of the gametes will be normal and the other half will contain an inverted chromosome. If a crossover does occur within the loop of a paracentric inversion, a chromosome bridge and an acentric

chromosome (i.e., a chromosome without a centromere) will be formed, and this will give rise to abnormal meiotic products carrying deletions, which are inviable. In a pericentric inversion, a crossover within the loop does not result in a bridge or an acentric chromosome, but inviable products are produced carrying a duplication and a deletion.

Translocations

If a chromosome break occurs in each of two nonhomologous chromosomes and the two breaks rejoin in a new arrangement, the new segment is called a translocation. A cell bearing a heterozygous translocation has a full set of genes and will be viable unless one of the breaks causes damage within a gene or if there is a position effect on gene function. However, once again the pairing properties of the chromosomes at meiosis result in aberrant meiotic products. Specifically, half of the products are deleted for one of the chromosome regions that changed positions and half of the products are duplicated for the other. These duplications and deletions usually result in inviability, so translocation heterozygotes are generally semisterile ("half-sterile").

Polyploidy

An individual with additional chromosome sets is called a polyploid. Individuals with three sets of chromosomes (triploids, $3n$) or four sets of chromosomes (tetraploids, $4n$) are polyploid derivatives of the basic diploid ($2n$) constitution. Polyploids with odd numbers of sets (e.g., triploids) are sterile because homologous chromosomes pair only two by two, and the extra chromosome moves randomly to a cell pole, resulting in highly unbalanced, nonfunctional meiotic products. It is for this reason that triploid watermelons are seedless. However, polyploids

with even numbers of chromosome sets can be fertile if orderly two-by-two chromosome pairing occurs.

Though two organisms from closely related species frequently hybridize, the chromosomes of the fusing partners are different enough that the two sets do not pair at meiosis, resulting in sterile offspring. However, if by chance the number of chromosome sets in the hybrid accidentally duplicates, a pairing partner for each chromosome will be produced, and the hybrid will be fertile. These chromosomally doubled hybrids are called allotetraploids. Bread wheat, which is hexaploid ($6n$) due to several natural spontaneous hybridizations, is an example of an allotetraploid. Some polyploid plants are able to produce seeds through an asexual type of reproduction called apomixis; in such cases, all progeny are identical to the parent. Polyploidy does arise spontaneously in humans, but all polyploids either abort in utero or die shortly after birth.

Aneuploidy

Some cells have an abnormal number of chromosomes that is not a whole multiple of the haploid number. This condition is called aneuploidy. Most aneuploids arise by nondisjunction, a failure of homologous chromosomes to separate at meiosis. When a gamete of this type is fertilized by a normal gamete, the zygotes formed will have an unequal distribution of chromosomes. Such genomic imbalance results in severe abnormalities or death. Only aneuploids involving small chromosomes tend to survive and even then only with an aberrant phenotype.

The most common form of aneuploidy in humans results in Down syndrome, a suite of specific disorders in individuals possessing an extra chromosome 21 (trisomy 21). The symptoms of Down syndrome include intellectual disability, severe disorders of internal organs such as the

heart and kidneys, up-slanted eyes, an enlarged tongue, and abnormal dermal ridge patterns on the fingers, palms, and soles. Other forms of aneuploidy in humans result from abnormal numbers of sex chromosomes. Turner syndrome is a condition in which females have only one X chromosome. Symptoms may include short stature, webbed neck, kidney or heart malformations, underdeveloped sex characteristics, or sterility. Klinefelter syndrome is a condition in which males have one extra female sex chromosome, resulting in an XXY pattern. (Other, less frequent, chromosomal patterns include XXXY, XXXXY, XXYY, and XXXYY.) Symptoms of Klinefelter syndrome may include sterility, a tall physique, lack of secondary sex characteristics, breast development, and learning disabilities.

CHAPTER 3
DNA as the Agent of Heredity

In 1869 Swiss chemist Johann Friedrich Miescher extracted a substance containing nitrogen and phosphorus from cell nuclei. The substance was originally called nuclein, but it is now known as DNA. DNA is the chemical component of the chromosomes that is chiefly responsible for their staining properties in microscopic preparations.

Since the chromosomes of eukaryotes, whose cells possess a clearly defined nucleus, contain a variety of proteins in addition to DNA, the question naturally arose whether the nucleic acids or the proteins, or both together, were the carriers of the genetic information. Until the early 1950s most biologists were inclined to believe that the proteins were the chief agents of heredity. Nucleic acids contain only four different unitary building blocks, but proteins are made up of 20 different amino acids. Proteins therefore appeared to have a greater diversity of structure, and the diversity of the genes seemed at first likely to rest on the diversity of the proteins.

Evidence that DNA acts as the carrier of the genetic information was first firmly demonstrated by exquisitely simple microbiological studies. In 1928 English bacteriologist Frederick Griffith was studying two strains of the bacterium *Streptococcus pneumoniae*; one strain was lethal to mice (virulent) and the other was harmless (avirulent). Griffith found that mice inoculated with either the heat-killed virulent bacteria or the living avirulent bacteria remained free of infection, but mice inoculated with a mixture of both became infected and died. It seemed as if some chemical “transforming principle” had transferred

from the dead virulent cells into the avirulent cells and changed them. In 1944 American bacteriologist Oswald T. Avery and his coworkers found that the transforming factor was DNA. Avery and his research team obtained mixtures from heat-killed virulent bacteria and inactivated either the proteins, polysaccharides (sugar subunits), lipids, DNA, or RNA and added each type of preparation individually to avirulent cells. The only molecular class whose inactivation prevented transformation to virulence was DNA. Therefore, it seemed that DNA, because it could transform, must be the hereditary material.

A similar conclusion was reached from the study of bacteriophages, viruses that attack and kill bacterial cells. From a host cell infected by one bacteriophage, hundreds of bacteriophage progeny are produced. In 1952 American biologists Alfred D. Hershey and Martha Chase prepared two populations of bacteriophage particles. In one population, the outer protein coat of the bacteriophage was labeled with a radioactive isotope; in the other, the DNA was labeled. After allowing both populations to attack bacteria, Hershey and Chase found that only when DNA was labeled did the progeny bacteriophage contain radioactivity. Therefore, they concluded that DNA is injected into the bacterial cell, where it directs the synthesis of numerous complete bacteriophages at the expense of the host. In other words, in bacteriophages DNA is the hereditary material responsible for the fundamental characteristics of the virus.

Today the genetic makeup of most organisms can be transformed using externally applied DNA, in a manner similar to that used by Avery for bacteria. Transforming DNA is able to pass through cellular and nuclear membranes and then integrate into the chromosomal DNA of the recipient cell. Furthermore, using modern DNA

technology, it is possible to isolate the section of chromosomal DNA that constitutes an individual gene, manipulate its structure, and reintroduce it into a cell to cause changes that show beyond doubt that the DNA is responsible for a large part of the overall characteristics of an organism. For reasons such as these, it is now accepted that, in all living organisms, with the exception of some viruses, genes are composed of DNA.

STRUCTURE AND COMPOSITION OF DNA

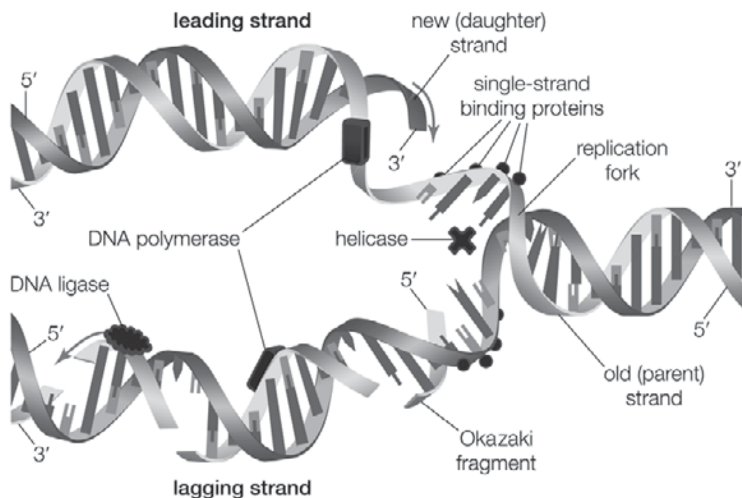
The remarkable properties of the nucleic acids, which qualify these substances to serve as the carriers of genetic information, have captured the attention of many scientists. The groundwork was laid by pioneer biochemists who found that nucleic acids are long chainlike molecules, the backbones of which consist of repeated sequences of phosphate and sugar linkages—ribose sugar in RNA and deoxyribose sugar in DNA. Attached to the sugar links in the backbone are two kinds of nitrogenous bases: purines and pyrimidines. The purines are adenine (A) and guanine (G) in both DNA and RNA; the pyrimidines are cytosine (C) and thymine (T) in DNA and cytosine (C) and uracil (U) in RNA. A single purine or pyrimidine is attached to each sugar, and the entire phosphate-sugar-base subunit is called a nucleotide. The nucleic acids extracted from different species of animals and plants have different proportions of the four nucleotides. Some are relatively richer in adenine and thymine, while others have more guanine and cytosine. However, it was found by biochemist Erwin Chargaff that the amount of A is always equal to T, and the amount of G is always equal to C.

The findings of Chargaff suggested to Watson and Crick that adenine was somehow paired with thymine and that guanine was paired with cytosine. Using this information, as well as the information from the X-ray diffraction studies of Wilkins and Rosalind Franklin, Watson and Crick came up with the now-famous model showing DNA as a double helix composed of two intertwined chains of nucleotides, in which the adenines of one chain are linked to the thymines of the other, and the guanines in one chain are linked to the cytosines of the other. The structure resembles a ladder that has been twisted into a spiral shape: the sides of the ladder are composed of sugar and phosphate groups, and the rungs are made up of the paired nitrogenous bases. By making a wire model of the structure, it became clear that the only way the model could conform to the requirements of the molecular dimensions of DNA was if A always paired with T and G with C; in fact, the A-T and G-C pairs showed a satisfying lock-and-key fit. Although most of the bonds in DNA are strong covalent bonds, the A-T and G-C bonds are weak hydrogen bonds. However, multiple hydrogen bonds along the centre of the molecule confer enough stability to hold the two strands together.

The two strands of the double helix model are antiparallel; that is, the nucleotides are arranged in opposite orientation. One DNA strand runs from 5' → 3' (five prime to three prime), whereas the other runs from 3' → 5'.

DNA REPLICATION

The structure of DNA determined in 1953 suggested at least three different ways that DNA might self-replicate. The experiments of Matthew Meselson and Franklin Stahl on the bacterium *Escherichia coli* in 1958 suggested that



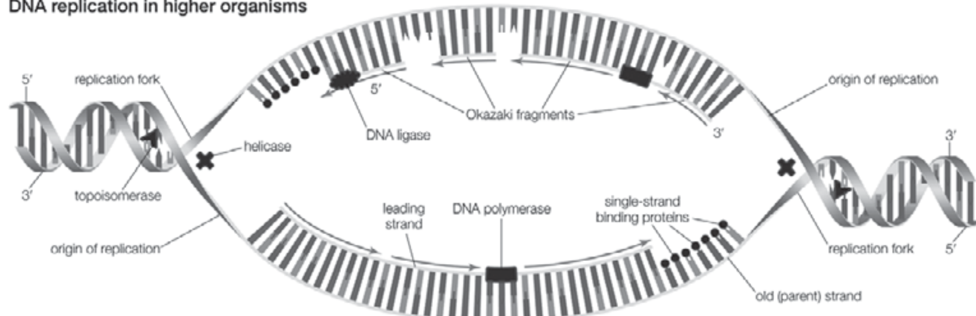
In semiconservative DNA replication an existing DNA molecule is separated into two template strands. New nucleotides align with and bind to the nucleotides of the existing strands, thus forming two DNA molecules that are identical to the original DNA molecule. Encyclopædia Britannica, Inc.

DNA replicates semiconservatively. Meselson and Stahl grew bacterial cells in the presence of ^{15}N, a heavy isotope of nitrogen, so that the DNA of the cells contained ^{15}N. These cells were then transferred to a medium containing the normal isotope of nitrogen, ^{14}N, and allowed to go through cell division. The researchers were able to demonstrate that, in the DNA molecules of the daughter cells, one strand contained only ^{15}N, and the other strand contained ^{14}N. This is precisely what is expected by the semiconservative mode of replication, in which the original DNA molecules should separate into two template strands containing ^{15}N, and the newly aligned nucleotides should all contain ^{14}N.

The hooking together of free nucleotides in the newly synthesized strand takes place one nucleotide at a time in

the 5' → 3' direction. An incoming free nucleotide pairs with the complementary nucleotide on the template strand, and then the 5' end of the free nucleotide is covalently joined to the 3' end of a nucleotide already in place. The process is then repeated. The result is a nucleotide chain, referred to chemically as a nucleotide polymer or a polynucleotide. Of course the polymer is not a random polymer; its nucleotide sequence has been directed by the nucleotide sequence of the template strand. It is this templating process that enables hereditary information to be replicated accurately and passed down through the generations. In a very real way, human DNA has been replicated in a direct line of descent from the first vertebrates that evolved hundreds of millions of years ago.

DNA replication starts at a site on the DNA called the origin of replication. In higher organisms, replication begins at multiple origins of replication and moves along the DNA in both directions outward from each origin, creating two replication "forks." The events at both replication forks are identical. In order for DNA to replicate, however, the two strands of the double helix first must be unwound from each other. A class of enzymes called DNA topoisomerases removes helical twists by cutting a



DNA replication in higher organisms begins at multiple origins of replication and progresses in two directions. Encyclopædia Britannica, Inc.

DNA strand and then resealing the cut. Enzymes called helicases then separate the two strands of the double helix, exposing two template surfaces for the alignment of free nucleotides.

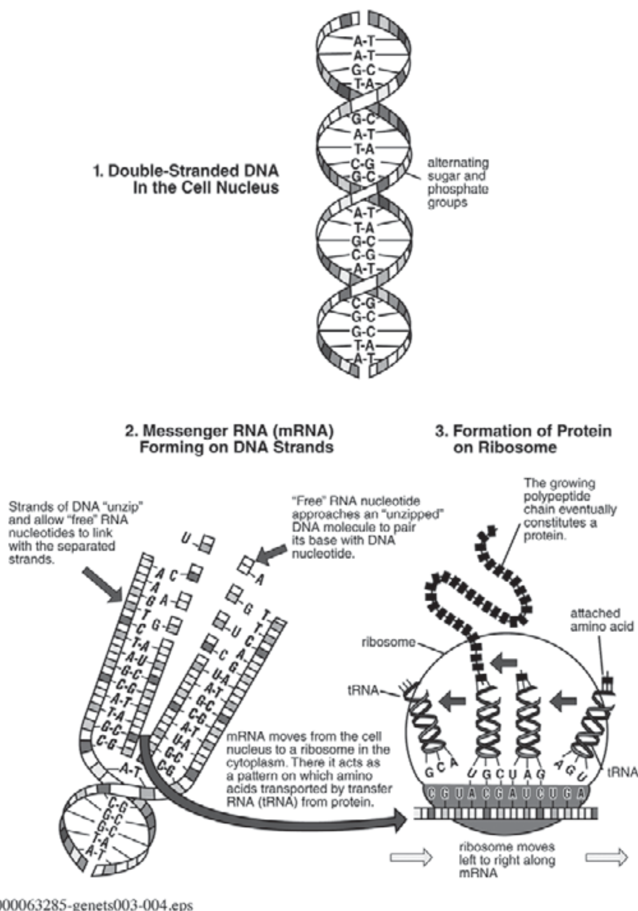
Beginning at the origin of replication, a complex enzyme called DNA polymerase moves along the DNA molecule, pairing nucleotides on each template strand with free complementary nucleotides. Because of the antiparallel nature of the DNA strands, new strand synthesis is different on each template. On the 3' → 5' template strand, polymerization proceeds in the 5' → 3' direction, and this growing strand is called the leading strand. However, polymerization must be carried out differently on the 5' → 3' template strand because nucleotides cannot be assembled in the 3' → 5' direction. Here short sequences of RNA are polymerized on the template. These sequences act as primers to which the DNA polymerase can add nucleotides in the 5' → 3' direction but in the opposite direction in which synthesis is proceeding on the lagging strand. The DNA polymerase hence makes short segments of DNA called Okazaki fragments in the "wrong" direction. For this reason the strand synthesized on the 5' → 3' template strand is called the lagging strand. Later, the RNA primers are removed and the Okazaki fragments are joined. This RNA priming system cannot be used to synthesize the very end of the 3' → 5' strand; once the last RNA primer is removed, synthesis cannot continue over the remaining gap. To overcome this obstacle, the enzyme telomerase adds multiple copies of a nucleotide sequence to the end of the DNA strand to allow completion of replication. Despite the peculiar events on the lagging strand, the entire DNA strand is eventually polymerized, and the two daughter DNA molecules thus produced are identical.

EXPRESSION OF THE GENETIC CODE

DNA represents a type of information that is vital to the shape and form of an organism. It contains instructions in a coded sequence of nucleotides, and this sequence interacts with the environment to produce form—the living organism with all of its complex structures and functions.

The form of an organism is largely determined by protein. A large proportion of the visible parts of an organism is protein; for example, hair, muscle, and skin are made up largely of protein. Other chemical compounds that make up the human body, such as carbohydrates, fats, and more-complex chemicals, are either synthesized by catalytic proteins (enzymes) or are deposited at specific times and in specific tissues under the influence of proteins. For example, the black-brown skin pigment melanin is synthesized by enzymes and deposited in special skin cells called melanocytes. Genes exert their effect mainly by determining the structure and function of the many thousands of different proteins, which in turn determine the characteristics of an organism. Generally, it is true to say that each protein is coded for by one gene, bearing in mind that the production of some proteins requires the cooperation of several genes.

Proteins are polymeric molecules; that is, they are made up of chains of monomeric elements, as is DNA. In proteins, the monomers are amino acids. Organisms generally contain 20 different types of amino acids, and the distinguishing factors that make one protein different from another are its length and specific amino acid sequence, which are determined by the number and sequence of nucleotide pairs in DNA. In other words, there is a colinearity (i.e., parallel structure) between the polymer that is DNA and the polymer that is protein.



DNA in the cell nucleus carries a genetic code, which consists of sequences of adenine (A), thymine (T), guanine (G), and cytosine (C) (Figure 1). RNA, which contains uracil (U) instead of thymine, carries the code to protein-making sites in the cell. To make RNA, DNA pairs its bases with those of the "free" nucleotides (Figure 2). Messenger RNA (mRNA) then travels to the ribosomes in the cell cytoplasm, where protein synthesis occurs (Figure 3). The base triplets of transfer RNA (tRNA) pair with those of mRNA and at the same time deposit their amino acids on the growing protein chain. Finally, the synthesized protein is released to perform its task in the cell or elsewhere in the body. Encyclopædia Britannica, Inc.

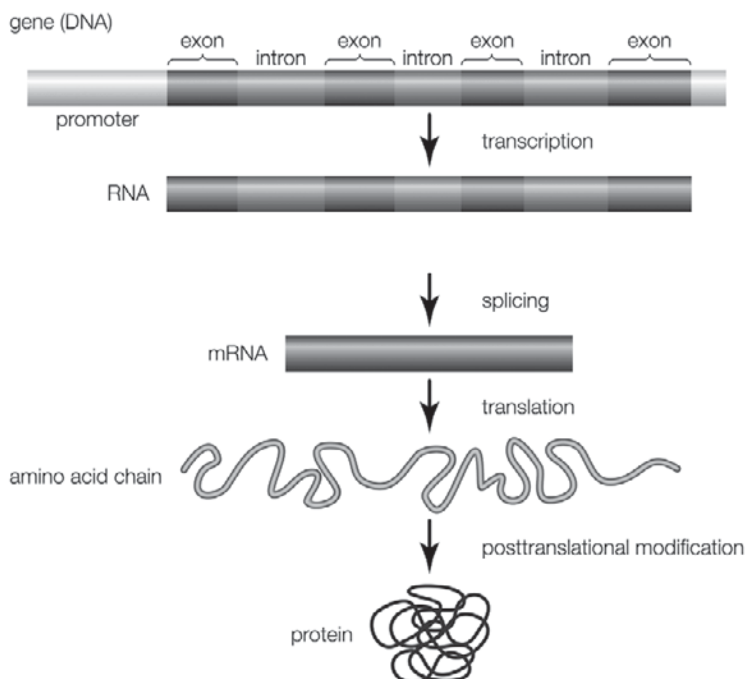
Hence, genetic information flows from DNA into protein. However, this is not a single-step process. First, the nucleotide sequence of DNA is copied into the nucleotide sequence of single-stranded RNA in a process called transcription. Transcription of any one gene takes place at the chromosomal location of that gene. Whereas the unit of replication is a whole chromosome, the transcriptional unit is a relatively short segment of the chromosome, the gene. The active transcription of a gene depends on the need for the activity of that particular gene in a specific tissue or at a given time.

The nucleotide sequence in RNA faithfully mirrors that of the DNA from which it was transcribed. The uracil in RNA has exactly the same hydrogen-bonding properties as thymine, so there are no changes at the information level. For most RNA molecules, the nucleotide sequence is converted into an amino acid sequence, a process called translation. In prokaryotes, translation begins during the transcription process, before the full RNA transcript is made. In eukaryotes, transcription finishes, and the RNA molecule passes from the nucleus into the cytoplasm, where translation takes place.

The genome of a type of virus called a retrovirus (of which human immunodeficiency virus, or HIV, is an example) is composed of RNA instead of DNA. In a retrovirus, RNA is reverse transcribed into DNA, which can then integrate into the chromosomal DNA of the host cell that the retrovirus infects. The synthesis of DNA is catalyzed by the enzyme reverse transcriptase. The existence of reverse transcriptase shows that genetic information is capable of flowing from RNA to DNA in exceptional cases. Since it is believed that life arose in an RNA world, it is likely that the evolution of reverse transcriptase was an important step in the transition to the present DNA world.

Transcription

Transcription is the synthesis of RNA from DNA. During transcription, only one strand of DNA is usually copied. This is called the template strand, and the RNA molecules produced are single-stranded messenger RNAs (mRNAs). The DNA strand that would correspond to the mRNA is called the coding or sense strand. In eukaryotes (organisms that possess a nucleus) the initial product of transcription is called a pre-mRNA. Pre-mRNA is extensively edited through splicing before the mature mRNA is



Genes are made up of promoter regions and alternating regions of introns (noncoding sequences) and exons (coding sequences). The production of a functional protein involves the transcription of the gene from DNA into RNA, the removal of introns and splicing together of exons, the translation of the spliced RNA sequences into a chain of amino acids, and the posttranslational modification of the protein molecule. Encyclopædia Britannica, Inc.

produced and ready for translation by the ribosome, the cellular organelle that serves as the site of protein synthesis. Transcription of any one gene takes place at the chromosomal location of that gene, which is a relatively short segment of the chromosome. The active transcription of a gene depends on the need for the activity of that particular gene in a specific cell or tissue or at a given time.

Small segments of DNA are transcribed into RNA by the enzyme RNA polymerase, which achieves this copying in a strictly controlled process. The first step is to recognize a specific sequence on DNA called a promoter that signifies the start of the gene. The two strands of DNA become separated at this point, and RNA polymerase begins copying from a specific point on one strand of the DNA using a special type of sugar-containing nucleoside called ribonucleoside 5'-triphosphate to begin the growing chain. Additional ribonucleoside triphosphates are used as the substrate, and, by cleavage of their high-energy phosphate bond, ribonucleoside monophosphates are incorporated into the growing RNA chain. Each successive ribonucleotide is directed by the complementary base pairing rules of DNA. For example, a C (cytosine) in DNA directs the incorporation of a G (guanine) into RNA. Likewise, a G in DNA is copied into a C in RNA, a T (thymine) into an A (adenine), and an A into a U (uracil; RNA contains U in place of the T of DNA). Synthesis continues until a termination signal is reached, at which point the RNA polymerase drops off the DNA, and the RNA molecule is released.

Ahead of many genes in prokaryotes (organisms that lack a nucleus), there are signals called "operators" where specialized proteins called repressors bind to the DNA just upstream of the start point of transcription and prevent access to the DNA by RNA polymerase. These

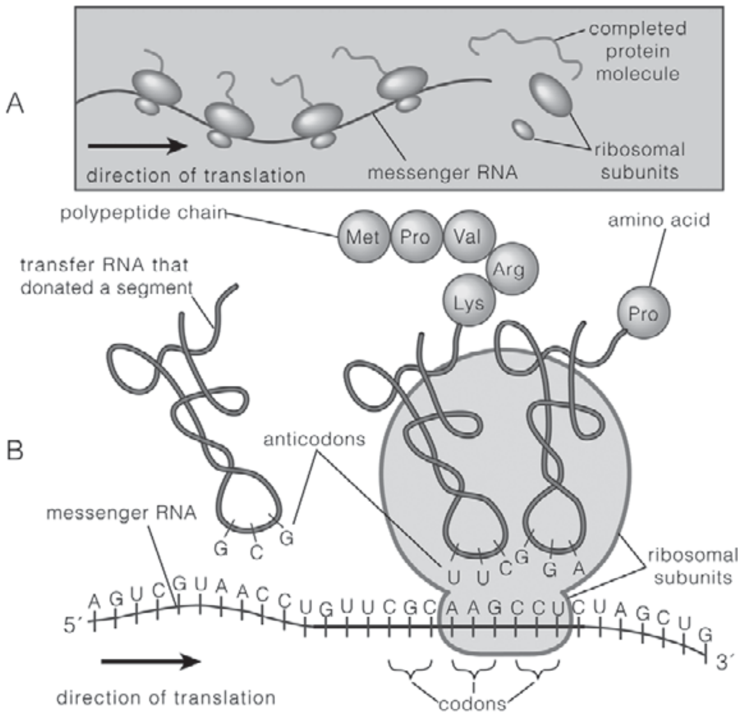
repressor proteins thus prevent transcription of the gene by physically blocking the action of the RNA polymerase. Typically, repressors are released from their blocking action when they receive signals from other molecules in the cell indicating that the gene needs to be expressed. Ahead of some prokaryotic genes are signals to which activator proteins bind to stimulate transcription.

Transcription in eukaryotes is more complicated than in prokaryotes. First, the RNA polymerase of higher organisms is a more complicated enzyme than the relatively simple five-subunit enzyme of prokaryotes. In addition, there are many more accessory factors that help to control the efficiency of the individual promoters. These accessory proteins are called transcription factors and typically respond to signals from within the cell that indicate whether transcription is required. In many human genes, several transcription factors may be needed before transcription can proceed efficiently. A transcription factor can cause either repression or activation of gene expression in eukaryotes.

Translation

Translation is the synthesis of protein from RNA. Each protein is made up of an amino acid chain that is generated from mRNA. These chains are folded into helices, zigzags, and other shapes to form proteins and are sometimes associated with other amino acid chains.

The specific amounts of amino acids in a protein and their sequence determine the protein's unique properties; for example, muscle protein and hair protein contain the same 20 amino acids, but the sequences of these amino acids in the two proteins are quite different. If the nucleotide sequence of mRNA is thought of as a written message,



Synthesis of protein.

it can be said that this message is read by the translation apparatus in "words" of three nucleotides, starting at one end of the mRNA and proceeding along the length of the molecule. Each of these three-letter words, or codons, stands for a specific amino acid, so if the message in mRNA is 900 nucleotides long, which corresponds to 300 codons, it will be translated into a chain of 300 amino acids.

Translation takes place on ribosomes—complex particles in the cell that contain RNA and protein. In prokaryotes the ribosomes are loaded onto the mRNA while transcription is still ongoing. The mRNA sequence

is read three bases at a time from its 5' end toward its 3' end, and one amino acid is added to the growing chain from its respective transfer RNA (tRNA), until the complete protein chain is assembled. Translation stops when the ribosome encounters a termination codon, normally UAG, UAA, or UGA. Special release factors associate with the ribosome in response to these codons, and the newly synthesized protein, tRNAs, and mRNA all dissociate. The ribosome then becomes available to interact with another mRNA molecule.

Any one mRNA is translated by several ribosomes along its length, each one at a different stage of translation. In eukaryotes ribosomes that produce proteins to be used in the same cell are not associated with membranes. However, proteins that must be exported to another location in the organism are synthesized on ribosomes located on the outside of flattened membranous chambers called the endoplasmic reticulum (ER). A completed amino acid chain is extruded into the inner cavity of the ER. Subsequently, the ER transports the proteins via small vesicles to another cytoplasmic organelle called the Golgi apparatus, which in turn buds off more vesicles that eventually fuse with the cell membrane. The protein is then released from the cell.

GENE MUTATION

Given the complexity of DNA and the vast number of cell divisions that take place within the lifetime of a multicellular organism, copying errors are likely to occur. If unrepaired, such errors will change the sequence of the DNA bases and alter the genetic code. Mutation is the random process whereby genes change from one allelic form to another. Scientists who study mutation

use the most common genotype found in natural populations, called the wild type, as the standard against which to compare a mutant allele. Mutation can occur in two directions; mutation from wild type to mutant is called a forward mutation, and mutation from mutant to wild type is called a back mutation, or reversion.

Mechanisms of Mutation

Mutations arise from changes to the DNA of a gene. These changes can be quite small, affecting only one nucleotide pair, or they can be relatively large, affecting hundreds or thousands of nucleotides. Mutations in which one base is changed are called point mutations—for example, substitution of the nucleotide pair AT by GC, CG, or TA. Base substitutions can have different consequences at the protein level. Some base substitutions are "silent," meaning that they result in a new codon that codes for the same amino acid as the wild type codon at that position or a codon that codes for a different amino acid that happens to have the same properties as those in the wild type. Substitutions that result in a functionally different amino acid are called "missense" mutations; these can lead to alteration or loss of protein function. A more severe type of base substitution, called a "nonsense" mutation, results in a stop codon in a position where there was not one before, which causes the premature termination of protein synthesis and, more than likely, a complete loss of function in the finished protein.

Another type of point mutation that can lead to drastic loss of function is a frameshift mutation, the addition or deletion of one or more DNA bases. In a protein-coding gene, the sequence of codons starting with AUG and ending with a termination codon is called the reading

frame. If a nucleotide pair is added to or subtracted from this sequence, the reading frame from that point will be shifted by one nucleotide pair, and all of the codons downstream will be altered. The result will be a protein whose first section (before the mutational site) is that of the wild type amino acid sequence, followed by a tail of functionally meaningless amino acids. Large deletions of many codons will not only remove amino acids from a protein but may also result in a frameshift mutation if the number of nucleotides deleted is not a multiple of three. Likewise, an insertion of a block of nucleotides will add amino acids to a protein and perhaps also have a frameshift effect.

A number of human diseases are caused by the expansion of a trinucleotide pair repeat. For example, fragile-X syndrome, the most common type of inherited intellectual disability in humans, is caused by the repetition of up to 1,000 copies of a CGG repeat in a gene on the X chromosome.

The impact of a mutation depends upon the type of cell involved. In a haploid cell, any mutant allele will most likely be expressed in the phenotype of that cell. In a diploid cell, a dominant mutation will be expressed over the wild type allele, but a recessive mutation will remain masked by the wild type. If recessive mutations occur in both members of one gene pair in the same cell, the mutant phenotype will be expressed. Mutations in germinal cells (i.e., reproductive cells) may be passed on to successive generations. However, mutations in somatic (body) cells will exert their effect only on that individual and will not be passed on to progeny.

The impact of an expressed somatic mutation depends upon which gene has been mutated. In most cases, the somatic cell with the mutation will die, an event that is generally of little consequence in a multicellular organism.

However, mutations in a special class of genes called proto-oncogenes can cause uncontrolled division of that cell, resulting in a group of cells that constitutes a cancerous tumour.

Mutations can affect gene function in several different ways. First, the structure and function of the protein coded by that gene can be affected. For example, enzymes are particularly susceptible to mutations that affect the amino acid sequence at their active site (i.e., the region that allows the enzyme to bind with its specific substrate). This may lead to enzyme inactivity; a protein is made, but it has no enzymatic function. Second, some nonsense or frameshift mutations can lead to the complete absence of a protein. Third, changes to the promoter region of the gene can result in gene malfunction by interfering with transcription. In this situation, protein production is either inhibited or it occurs at an inappropriate time because of alterations somewhere in the regulatory region. Fourth, mutations within introns (noncoding sequences of DNA) that affect the specific nucleotide sequences that direct intron splicing (removal of introns from RNA) may result in an mRNA that still contains an intron. When translated, this extra RNA will almost certainly be meaningless at the protein level, and its extra length will lead to a functionless protein. Any mutation that results in a lack of function for a particular gene is called a "null" mutation. Less-severe mutations are called "leaky" mutations because some normal function still "leaks through" into the phenotype.

Most mutations occur spontaneously and have no known cause. The synthesis of DNA is a cooperative venture of many different interacting cellular components, and occasionally mistakes occur that result in mutations. Like many chemical structures, the bases of DNA are able to exist in several conformations called

isomers. The keto form of a DNA base is the normal form that gives the molecule its standard base-pairing properties. However, the keto form occasionally changes spontaneously to the enol form, which has different base-pairing properties. For example, the keto form of cytosine pairs with guanine (its normal pairing partner), but the enol form of cytosine pairs with adenine. During DNA replication, this adenine base will act as the template for thymine in the newly synthesized strand. Therefore, a CG base pair will have mutated to a TA base pair. If this change results in a functionally different amino acid, then a missense mutation may result. Another spontaneous event that can lead to mutation is depurination, the complete loss of a purine base (adenine or guanine) at some location in the DNA. The resulting gap can be filled by any base during subsequent replications.

Ionizing radiation, some chemicals, and certain viruses are capable of acting as mutagens—agents that can increase the rate at which mutations occur. Some mutagens have been implicated as a cause of cancer. For example, ultraviolet (UV) radiation from the sun is known to cause skin cancer, and cigarette smoke is a primary cause of lung cancer.

Repair of Mutation

A variety of mechanisms exists for repairing copying errors caused by DNA damage. One of the best-studied systems is the repair mechanism for damage caused by ultraviolet radiation. Ultraviolet radiation joins adjacent thymines, creating thymine dimers, which, if not repaired, may cause mutations. Special repair enzymes either cut the bond between the thymines or excise the bonded dimer and replace it with two single thymines. If both of these repair methods fail, a third method allows the DNA replication

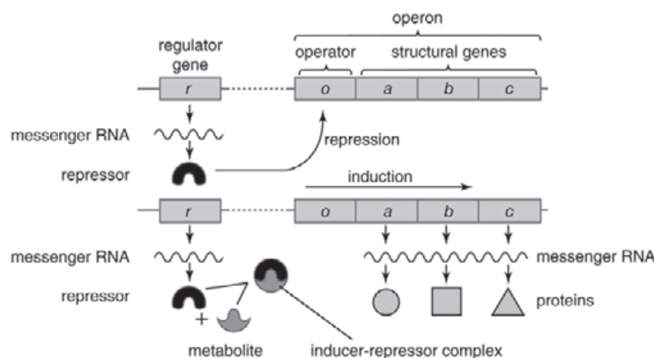
process to bypass the dimer; however, it is this bypass system that causes most mutations because bases are then inserted at random opposite the thymine dimer. Xeroderma pigmentosum, a severe hereditary disease of humans, is caused by a mutation in a gene coding for one of the thymine dimer repair enzymes. Individuals with this disease are highly susceptible to skin cancer.

Reverse mutation from the aberrant state of a gene back to its normal, or wild type, state can result in a number of possible molecular changes at the protein level. True reversion is the reversal of the original nucleotide change. However, phenotypic reversion can result from changes that restore a different amino acid with properties identical to the original. Second-site changes within a protein can also restore normal function. For example, an amino acid change at a site different from that altered by the original mutation can sometimes interact with the amino acid at the first mutant site to restore a normal protein shape. Also, second-site mutations at other genes can act as suppressors, restoring wild type function. For example, mutations in the anticodon region of a tRNA gene can result in a tRNA that sometimes inserts an amino acid at an erroneous stop codon; if the original mutation is caused by a stop codon, which arrests translation at that point, then a tRNA anticodon change can insert an amino acid and allow translation to continue normally to the end of the mRNA. Alternatively, some mutations at separate genes open up a new biochemical pathway that circumvents the block of function caused by the original mutation.

REGULATION OF GENE EXPRESSION

Not all genes in a cell are active in protein production at any given time. Gene action can be switched on or off in

Genetics



Model of the operon and its relation to the regulator gene. Encyclopaedia Britannica, Inc.

response to the cell's stage of development and external environment. In multicellular organisms, different kinds of cells express different parts of the genome. In other words, a skin cell and a muscle cell contain exactly the same genes, but the differences in structure and function of these cells result from the selective expression and repression of certain genes.

In prokaryotes and eukaryotes, most gene-control systems are positive, meaning that a gene will not be transcribed unless it is activated by a regulatory protein. However, some bacterial genes show negative control. In this case the gene is transcribed continuously unless it is switched off by a regulatory protein. An example of negative control in prokaryotes involves three adjacent genes used in the metabolism of the sugar lactose by *E. coli*. The part of the chromosome containing the genes concerned is divided into two regions, one that includes the structural genes (i.e., those genes that together code for protein structure) and another that is a regulatory region. This overall unit is called an operon. If lactose is not present in a cell, transcription of the genes that code for the

lactose-processing enzymes—β-galactosidase, permease, and transacetylase—is turned off. This is achieved by a protein called the lac repressor, which is produced by the repressor gene and binds to a region of the operon called the operator. Such binding prevents RNA polymerase, which initially binds at the adjacent promoter, from moving into the coding region. If lactose enters the cell, it binds to the lac repressor and induces a change of shape in the repressor so that it can no longer bind to the DNA at the operon. Consequently, the RNA polymerase is able to travel from the promoter down the three adjacent protein-coding regions, making one continuous transcript. This three-gene transcript is subsequently translated into three separate proteins.

Although the operon model has proved a useful model of gene regulation in bacteria, different regulatory mechanisms are employed in eukaryotes. First, there are no operons in eukaryotes, and each gene is regulated independently. Furthermore, the series of events associated with gene expression in higher organisms is much more complex than in prokaryotes and involves multiple levels of regulation.

In order for a gene to produce a functional protein, a complex series of steps must occur. Some type of signal must initiate the transcription of the appropriate region along the DNA, and, finally, an active protein must be made and sent to the appropriate location to perform its specific task. Regulation can be exerted at many different places along this pathway. The fundamental level of control is the rate of transcription. Transcription itself is also a complex process with many different components, and each one is a potential point of control. Regulatory proteins called activators or enhancers are needed for the transcription of genes at a specific time or in a certain cell.

Thus, control is positive (not negative as in the lac operon) in that these proteins are necessary for the promotion of transcription. Activators bind to specific regions of the DNA in the upstream regulatory region, some very distant from the binding of the initiation complex.

Following the transcription of DNA into RNA, a process of editing and splicing takes place in which the noncoding nucleotide sequences called introns are excised from the primary transcript, resulting in functional mRNA. For most genes this is a routine step in the production of mRNA, but in some genes there are alternative ways to splice the primary transcript, resulting in different mRNAs, which in turn result in different proteins.

Some genes are controlled at the translational and post-translational levels. One type of translational control is the storage of uncapped mRNA to meet future demands for protein synthesis. In other cases, control is exerted through the stability or instability of mRNA. The rate of translation of some mRNAs can also be regulated. Post-translationally, certain proteins (e.g., insulin) are synthesized in an inactive form and must be chemically modified to become active. Other proteins are targeted to specific locations inside the cell (e.g., mitochondria) by means of highly specific amino acid sequences at their ends, called leader sequences; when the protein reaches its correct site, the leader segment is cut off, and the protein begins to function. Post-translational control is also exerted through mRNA and protein degradation.

RNA INTERFERENCE AND GENE SILENCING

RNA interference, or RNAi, is a unique regulatory system occurring within eukaryotic cells that controls the activity

of genes. RNAi functions specifically to silence, or deactivate, genes.

The ability of interfering RNA to silence genes was discovered in the 1990s by American scientists Andrew Z. Fire and Craig C. Mello, who shared the 2006 Nobel Prize for Physiology or Medicine for their work. Fire and Mello successfully inhibited the expression of specific genes by introducing short double-stranded RNA (dsRNA) segments into the cells of nematodes (*Caenorhabditis elegans*). The dsRNA segments underwent enzymatic processing that enabled them to attach to molecules of mRNA possessing complementary nucleotide sequences. The attachment of the two RNAs inhibited the translation of the mRNA molecules into proteins.

Although Fire and Mello's work involved the experimental introduction of interfering RNA into cells, gene silencing by RNAi is a natural genetic mechanism in eukaryotes that takes place following transcription. Special microRNA (miRNA) segments, each of which is approximately 20 nucleotides in length, are encoded by the genomes of eukaryotic organisms. Each miRNA is produced from a precursor transcript (pre-miRNA). After the pre-miRNA migrates from the nucleus into the cytoplasm, it is cleaved into a mature miRNA by an enzyme known as DICER. The mature miRNA molecule then binds to an RNA-induced silencing complex (RISC), which contains multiple proteins, including a ribonuclease enzyme. The miRNA nucleotide sequence directs the protein complex to bind to a complementary sequence of mRNA. Once bound to the mRNA, the miRNA-RISC complex then enzymatically cleaves targeted sites on the mRNA molecule, thereby inhibiting the translation of the gene into a protein, which effectively silences the gene.

RNAi plays an important role not only in regulating genes but also in mediating cellular defense against

infection by RNA viruses, including influenza viruses and rhabdoviruses, a group that contains the causative agent of rabies. In fact, a number of plants and animals have evolved antiviral RNAi genes that encode short segments of RNA molecules with sequences that are complementary to viral sequences. This complementarity enables interfering RNA produced by the cell to bind to and inactivate specific RNA viruses.

RNAi is also an innate mechanism by which cells can suppress the activity of transposons, or “jumping genes.” Certain types of transposable elements are able to produce mobile copies of themselves, which subsequently are inserted into various regions of the genome, giving rise to repetitive sequences of DNA. These insertions generally are of little concern in terms of human health. However, some insertions lead to increased or decreased gene activity and can give rise to disease. For example, certain types of cancer and Duchenne muscular dystrophy, a hereditary muscle-wasting disorder, are associated with insertions of transposons.

REPETITIVE DNA

One major difference between the genomes of prokaryotes and eukaryotes is that most eukaryotes contain repetitive DNA, with the repeats either clustered or spread out between the unique genes. There are several categories of repetitive DNA: (1) single copy DNA, which contains the structural genes (protein-coding sequences), (2) families of DNA, in which one gene somehow copies itself, and the repeats are located in small clusters (tandem repeats) or spread throughout the genome (dispersed repeats), and (3) satellite DNA, which contains short nucleotide sequences repeated as many as thousands of times. Such repeats are often found clustered in tandem near the centromeres

BARBARA MCCLINTOCK

(b. June 16, 1902, Hartford, Conn., U.S.—d. Sept. 2, 1992, Huntington, N.Y.)

American scientist Barbara McClintock is known for her discovery in the 1940s and 1950s of mobile genetic elements, or transposons ("jumping genes").

McClintock, whose father was a physician, took great pleasure in science as a child and evidenced early the independence of mind and action that she would exhibit throughout her life. After attending high school, she enrolled as a biology major at Cornell University in 1919. She received a B.S. in 1923, a master's degree two years later, and, having specialized in cytology, genetics, and zoology, a Ph.D. in 1927. During graduate school she began the work that would occupy her

Barbara McClintock in Stockholm, where she received the Nobel Prize for Physiology or Medicine in 1983. Keystone/Hulton Archive/Getty Images

entire life: the chromosomal analysis of corn (maize). She used a microscope and a staining technique that allowed her to examine, identify, and describe individual corn chromosomes.

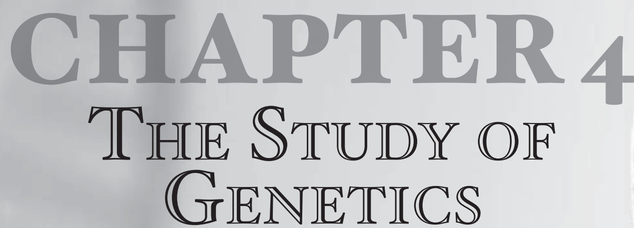
In 1931 McClintock and a colleague, Harriet Creighton, published "A Correlation of Cytological and Genetical Crossing-over in *Zea mays*," a paper that established that chromosomes formed the basis of genetics. Based on her experiments and publications during the 1930s, McClintock was elected vice president of the Genetics Society of America in 1939 and president of the Genetics Society in 1944. She received a Guggenheim Fellowship in 1933 to study in Germany, but she left early owing to the rise of Nazism. When she returned to Cornell, her alma mater, she found that the university would not hire a female professor. The Rockefeller Foundation funded her research at Cornell (1934–36) until she was hired by the University of Missouri (1936–41).

In 1941 McClintock moved to Long Island, New York, to work at the Cold Spring Harbor Laboratory, where she spent the rest of her professional life. In the 1940s, by observing and experimenting with variations in the coloration of kernels of corn, she discovered that genetic information is not stationary. By tracing pigmentation changes in corn and using a microscope to examine that plant's large chromosomes, she isolated two genes that she called "controlling elements." These genes controlled the genes that were actually responsible for pigmentation. McClintock found that the controlling elements could move along the chromosome to a different site, and that these changes affected the behaviour of neighbouring genes. She suggested that these transposable elements were responsible for new mutations in pigmentation or other characteristics.

McClintock's work was ahead of its time and was for many years considered too radical—or was simply ignored—by her fellow scientists. Deeply disappointed with her colleagues, she stopped publishing the results of her work and ceased giving lectures, though she continued doing research. Not until the late 1960s and 1970s, after biologists had determined that the genetic material was DNA, did members of the scientific community begin to verify her early findings. When recognition finally came, McClintock was inundated with awards and honours, most notably the 1983 Nobel Prize for Physiology or Medicine. She was the first woman to be the sole winner of this award.

(i.e., the attachment points for the nuclear spindle fibres that move chromosomes during cell division).

Microsatellite DNA is composed of tandem repeats of two nucleotide pairs that are dispersed throughout the genome. Minisatellite DNA, sometimes called variable number tandem repeats (VNTRs), is composed of blocks of longer repeats also dispersed throughout the genome. There is no known function for satellite DNA, nor is it known how the repeats are created. Finally, transposons, though they are generally inactive, also contribute to repetitive DNA.



CHAPTER 4
The Study of Genetics

The study of genetics forms one of the most rapidly advancing areas of biological research. Since Gregor Mendel's studies of the transmission of traits in peas, scientists have uncovered the hereditary properties of genes and chromosomes, the chemical structure of DNA, and the mechanisms underlying the translation of the genetic code into proteins. Geneticists today continue to make vital discoveries about the genomes of the organisms that constitute life on Earth. Much of this basic research has led to improved insight into the genetic abnormalities that give rise to disease, as well as to the development of new ways in which the genetic code can be manipulated experimentally to learn more about the functions of individual genes.

The areas of study encompassed by genetics cover a broad range of topics—from classical genetics, the foundation of the field, to comparative genomics, the basis of understanding evolutionary relationships between organisms, to epigenetics, which has provided startling discoveries about the influence of environmental factors on genes and heredity. The study of genetics is facilitated by a variety of tools as well, some of which, such as DNA fingerprinting and the polymerase chain reaction, have applications beyond the realm of research.

CLASSICAL GENETICS

Classical genetics, which remains the foundation for all other areas in genetics, is concerned primarily with the method by which genetic traits—classified as dominant

(always expressed), recessive (subordinate to a dominant trait), intermediate (partially expressed), or polygenic (due to multiple genes)—are transmitted in plants and animals. These traits may be sex-linked (resulting from the action of a gene on the sex chromosomes) or autosomal (resulting from the action of a gene on a chromosome other than a sex chromosome). Classical genetics began with Mendel's study of inheritance in garden peas and continues with studies of inheritance in many different plants and animals. Today a prime reason for performing classical genetics is for gene discovery—the finding and assembling of a set of genes that affects a biological property of interest.

CYTOGENETICS

Cytogenetics, the microscopic study of chromosomes, blends the skills of cytologists, who study the structure and activities of cells, with those of geneticists, who study genes. Cytologists discovered chromosomes and the way in which they duplicate and separate during cell division at about the same time that geneticists began to understand the behaviour of genes at the cellular level. The close correlation between the two disciplines led to their combination.

Plant cytogenetics early became an important subdivision of cytogenetics because, as a general rule, plant chromosomes are larger than those of animals. Animal cytogenetics became important after the development of the so-called squash technique, in which entire cells are pressed flat on a piece of glass and observed through a microscope; the human chromosomes were numbered using this technique.

Today there are multiple ways to attach molecular labels to specific genes and chromosomes, as well as to

specific RNAs and proteins, that make these molecules easily discernible from other components of cells, thereby greatly facilitating cytogenetics research.

Cytogenetic Techniques

Cytogenetics focuses on the microscopic examination of genetic components of the cell, including chromosomes, genes, and gene products. Some cytogenetic techniques involve placing cells in paraffin wax, slicing thin sections, and preparing them for microscopic study. The squash technique involves simply squashing entire cells and studying their contents. Dyes that selectively stain various parts of the cell are used; the genes, for example, may be located by selectively staining the DNA of which they are composed. Radioactive and fluorescent tags are valuable in determining the location of various genes and gene products in the cell. Tissue-culture techniques are often used to grow cells before squashing; white blood cells can be grown from samples of human blood and studied with the squash technique. One major application of cytogenetics in humans is in diagnosing abnormal chromosomal complements such as Down syndrome (caused by an extra copy of chromosome 21) and Klinefelter syndrome (occurring in males with an extra X chromosome). Some diagnosis is prenatal, performed on cell samples from amniotic fluid or the placenta.

Fluorescence In Situ Hybridization

Fluorescence in situ hybridization (FISH) is a technique that employs fluorescent probes for the detection of specific DNA sequences in chromosomes. FISH has a much higher rate of sensitivity and specificity than other

genetic diagnostic tests such as karyotyping and thus can be used to detect a variety of structural abnormalities in chromosomes, including small genetic deletions involving just one to five genes. It is also useful in detecting moderate-sized deletions such as those causing Prader-Willi syndrome, a rare genetic disorder characterized by a rounded face, low forehead, and intellectual disability. FISH also provides results more quickly than karyotyping because no cell culture is required.

FISH is commonly used for preimplantation genetic diagnosis (PGD) during in vitro fertilization. PGD involves obtaining a single cell from an embryo in the blastocyst stage of development. This single cell can then be analyzed using FISH. One problem with using FISH for PGD is that a single cell is scant material for diagnosis; therefore, a large array of tests cannot be performed. Similarly, if the test fails for any technical reason, it cannot be repeated.

MICROBIAL GENETICS

Microorganisms were generally ignored by the early geneticists because they are small in size and were thought to lack variable traits and the sexual reproduction necessary for a mixing of genes from different organisms. After it was discovered that microorganisms have many different physical and physiological characteristics that are amenable to study, they became objects of great interest to geneticists because of their small size and the fact that they reproduce much more rapidly than larger organisms. Bacteria became important model organisms in genetic analysis, and many discoveries of general interest in genetics arose from their study. Bacterial genetics is the centre of cloning technology used in basic research.

Viral genetics is another key part of microbial genetics. The genetics of viruses that attack bacteria were the first to be elucidated. Since then, studies and findings of viral genetics have been applied to viruses pathogenic on plants and animals, including humans. Viruses are also used as vectors (agents that carry and introduce modified genetic material into an organism) in DNA technology.

Physiological techniques, directed at exploring functional properties or organisms, are also used in microbial genetics investigations. In microorganisms, most genetic variations involve some important cell function. For example, some strains of *E. coli* are able to synthesize the vitamin thiamin from simple compounds; others, which lack an enzyme necessary for this synthesis, cannot survive unless thiamin is already present. The two strains can be distinguished by placing them on a thiamin-free mixture: those that grow have the gene for the enzyme, those that fail to grow do not. The technique is also applied to human cells, since many inherited human abnormalities are caused by a faulty gene that fails to produce a vital enzyme; albinism, which results from an inability to produce the pigment melanin in the skin, hair, or iris of the eyes, is an example of an enzyme deficiency in humans.

MOLECULAR GENETICS

Molecular genetics is the study of the molecular structure of DNA, its cellular activities (including its replication), and its influence in determining the overall makeup of an organism. Molecular genetics relies heavily on genetic engineering (recombinant DNA technology), which can be used to modify organisms by adding foreign DNA, thereby forming transgenic organisms. Since the early 1980s, these techniques have been used extensively in basic biological

research and are also fundamental to the biotechnology industry, which is devoted to the manufacture of agricultural and medical products. Transgenesis forms the basis of gene therapy, the attempt to cure genetic disease by addition of normally functioning genes from exogenous sources.

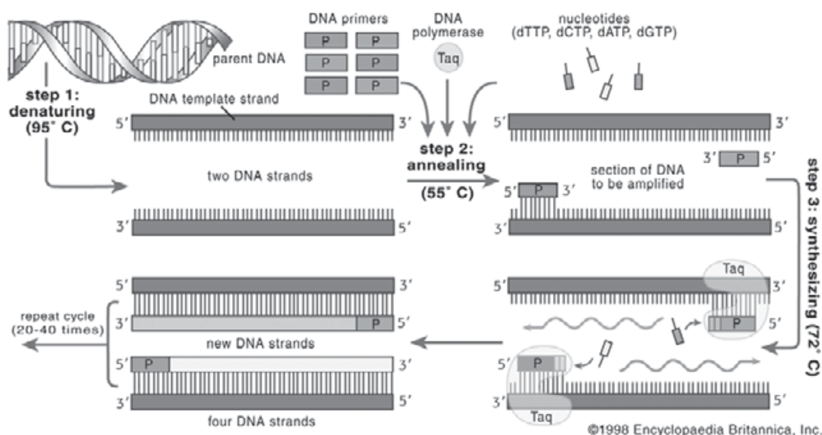
Molecular Techniques

Molecular genetics techniques are deeply involved with the direct study of DNA, which typically uses recombinant DNA technology. The DNA of any gene of interest from a donor organism (such as a human) can be cut out of a chromosome and inserted into a vector to make recombinant DNA, which can then be amplified and manipulated, studied, or used to modify the genomes of other organisms by transgenesis. A fundamental step in recombinant DNA technology is amplification. This is carried out by inserting the recombinant DNA molecule into a bacterial cell, which replicates and produces many copies of the bacterial genome and the recombinant DNA molecule (constituting a DNA clone).

A collection of large numbers of clones of recombinant donor DNA molecules is called a genomic library. Such libraries are the starting point for sequencing entire genomes such as the human genome. Today genomes can be scanned for small molecular variants called single nucleotide polymorphisms, or SNPs ("snips"), which act as chromosomal tags to associated specific regions of DNA that have a property of interest and may be involved in a human disease or disorder.

Polymerase Chain Reaction

The polymerase chain reaction, or PCR, is a technique used to make numerous copies of a specific segment of



The three-step process of the polymerase chain reaction.

DNA quickly and accurately. PCR enables investigators to obtain the large quantities of DNA that are required for various experiments and procedures in molecular biology, forensic analysis, evolutionary biology, and medical diagnostics.

PCR was developed in 1983 by Kary B. Mullis, an American biochemist who won the Nobel Prize for Chemistry in 1993 for his invention. Before the development of PCR, the methods used to amplify, or generate copies of, recombinant DNA fragments were time-consuming and labour-intensive. In contrast, a machine designed to carry out PCR reactions can complete many rounds of replication, producing billions of copies of a DNA fragment, in just one or two hours.

The PCR technique is based on the natural processes a cell uses to replicate a new DNA strand and requires only a few biological ingredients. The integral component is the template DNA—the DNA that contains the region to be copied, such as a gene. As little as one DNA molecule can serve as a template. The only information needed for this fragment to be replicated is the sequence of two short regions of nucleotides (the subunits of DNA) at

either end of the region of interest. These two short template sequences must be known so that two primers—short stretches of nucleotides that correspond to the template sequences—can be synthesized. The primers bind, or anneal, to the template at their complementary sites and serve as the starting point for copying. DNA synthesis at one primer is directed toward the other, resulting in replication of the desired intervening sequence. Also needed are free nucleotides used to build the new DNA strands and a DNA polymerase, an enzyme that does the building by sequentially adding on free nucleotides according to the instructions of the template.

PCR is a three-step process that is carried out in repeated cycles. The initial step is the denaturation, or separation, of the two strands of the DNA molecule. This is accomplished by heating the starting material to temperatures of about 95 °C (203 °F). Each strand is a template on which a new strand is built. In the second step the temperature is reduced to about 55 °C (131 °F) so that the primers can anneal to the template. In the third step the temperature is raised to about 72 °C (162 °F), and the DNA polymerase begins adding nucleotides onto the ends of the annealed primers. At the end of the cycle, which typically lasts less than five minutes, the temperature is raised and the process begins again. The number of copies doubles after each cycle. Usually 25 to 30 cycles produce a sufficient amount of DNA.

In the original PCR procedure, one problem was that the DNA polymerase had to be replenished after every cycle because it is not stable at the high temperatures needed for denaturation. This problem was solved in 1987 with the discovery of a heat-stable DNA polymerase called *Taq,* an enzyme isolated from the thermophilic bacterium *Thermus aquaticus,* which inhabits hot springs. *Taq* polymerase also led to the invention of the PCR machine.

Because DNA from a wide range of sources can be amplified, the technique has been applied to many fields. PCR is used to diagnose genetic disease and to detect low levels of viral infection. In forensic medicine it is used to analyze minute traces of blood and other tissues in order to identify the donor by his or her genetic "fingerprint." The technique has also been used to amplify DNA fragments found in preserved tissues, such as those of a 40,000-year-old frozen woolly mammoth or of a 7,500-year-old human found in a peat bog.

DNA Fingerprinting

DNA fingerprinting, also known as DNA typing, is a method of isolating and making images of sequences of DNA. The technique was developed in 1984 by the British geneticist Alec Jeffreys, after he noticed the existence of certain sequences of DNA (called minisatellites) that do not contribute to the function of a gene but are repeated within the gene and in other genes of a DNA sample. Jeffreys also determined that each organism has a unique pattern of these minisatellites, the only exception being multiple individuals from a single zygote (e.g., identical twins).

The procedure for creating a DNA fingerprint consists of first obtaining a sample of cells containing DNA (e.g., from skin, blood, or hair), extracting the DNA, and purifying it. The DNA is then cut at specific points along the strand with substances called restriction enzymes. This produces fragments of varying lengths that are sorted by placing them on a gel and then subjecting the gel to an electric current (electrophoresis): the shorter the fragment the more quickly it will move toward the positive pole (anode). The sorted, double-stranded DNA fragments are then subjected to a blotting technique in which

they are split into single strands and transferred to a nylon sheet. The fragments undergo autoradiography in which they are exposed to DNA probes—pieces of synthetic DNA that have been made radioactive and that bind to the minisatellites. A piece of X-ray film is then exposed to the fragments, and a dark mark is produced at any point where a radioactive probe has become attached. The resultant pattern of these marks can then be analyzed.

An early use of DNA fingerprinting was in legal disputes, notably to help solve crimes and to determine paternity. The technique was challenged, however, over concerns about sample contamination, faulty preparation procedures, and erroneous interpretation of the results. Efforts were made to improve reliability, and today the technique has been refined through the use of more specific and more sensitive probes and better blotting membranes. It also has been recognized that DNA fingerprinting, similar to other DNA analysis techniques, is limited by the quality of the sample obtained. DNA samples that are degraded or collected postmortem typically produce less reliable results than do samples that are obtained from a living individual.

If only a small amount of DNA is available for fingerprinting, PCR may be used to create thousands of copies of a DNA segment. Once an adequate amount of DNA has been produced, the exact sequence of nucleotide pairs in a segment of DNA can be determined using one of several biomolecular sequencing methods. Automated equipment has greatly increased the speed of DNA sequencing and has made available many practical applications, including pinpointing segments of genes that cause genetic diseases, mapping the human genome, engineering drought-resistant plants, and producing biological drugs from genetically altered bacteria.

IMMUNOGENETICS

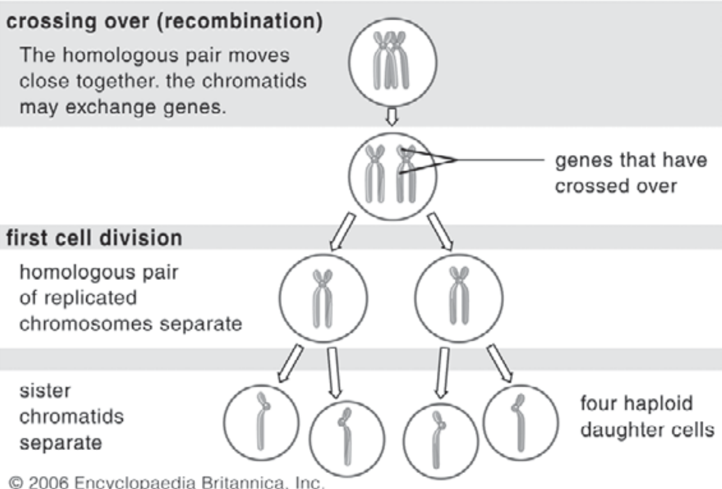
Many substances are antigenic; i.e., when introduced into a vertebrate body, they stimulate the production of specific proteins called antibodies. Various antigens exist in red blood cells, including those that make up the major blood groups of humans (A, B, AB, O). These and other antigens are genetically determined; their study constitutes immunogenetics. Blood antigens of humans include inherited variations, and the particular combination of antigens in an individual is almost as unique as fingerprints and has been used in such areas as paternity testing (although this approach has been largely supplanted by DNA-based techniques).

Immunological techniques are used in blood group determinations in blood transfusions, in organ transplants, and in determining Rhesus incompatibility in childbirth. Specific antigens of the human leukocyte antigen (HLA) genes are correlated with human diseases and disease predispositions. Antibodies also have a genetic basis, and their seemingly endless ability to match any antigen presented is based on special types of DNA shuffling processes between antibody genes. Immunology is also useful in identifying specific recombinant DNA clones that synthesize a specific protein of interest.

GENOMICS

Genomics is the study of the structure, function, and inheritance of the genome (entire set of genetic material) of an organism. A major part of genomics is determining the sequence of molecules that make up the genomic DNA content of an organism.

Every organism contains a basic set of chromosomes, unique in number and size for every species, that includes



During meiosis, an event known as chromosomal crossing over sometimes occurs as a part of recombination. In this process, a region of one chromosome is exchanged for a region of another chromosome, thereby producing unique chromosomal combinations that further divide into haploid daughter cells.

the complete set of genes plus any DNA between them. While the term *genome* was not brought into use until 1920, the existence of genomes has been known since the late 19th century, when chromosomes were first observed as stained bodies visible under the microscope. The initial discovery of chromosomes was then followed in the 20th century by the mapping of genes on chromosomes based on the frequency of exchange of parts of chromosomes by a process called chromosomal crossing over, an event that occurs as a part of the normal process of recombination and the production of sex cells (gametes) during meiosis. The genes that could be mapped by chromosomal crossing over were mainly those for which mutant phenotypes (visible manifestations of an organism's genetic composition) had been observed, only a small proportion of the total genes in the genome. The discipline of genomics

arose when the technology became available to deduce the complete nucleotide sequence of genomes, sequences generally in the range of billions of nucleotide pairs.

Sequencing and Bioinformatic Analysis of Genomes

Genomic sequences are usually determined using automatic sequencing machines. In a typical experiment to determine a genomic sequence, genomic DNA first is extracted from a sample of cells of an organism and then is broken into many random fragments. These fragments are cloned in a DNA vector (carrier) that is capable of carrying large DNA inserts. Because the total amount of DNA that is required for sequencing and additional experimental analysis is several times the total amount of DNA in an organism's genome, each of the cloned fragments is amplified individually by replication inside a living bacterial cell, which reproduces rapidly and in great quantity to generate many bacterial clones. The cloned DNA is then extracted from the bacterial clones and is fed into the sequencing machine. The resulting sequence data are stored in a computer. When a large enough number of sequences from many different clones is obtained, the computer ties them together using sequence overlaps. The result is the genomic sequence, which is then deposited in a publicly accessible database.

A complete genomic sequence in itself is of limited use; the data must be processed to find the genes and, if possible, their associated regulatory sequences. The need for these detailed analyses has given rise to the field of bioinformatics, in which computer programs scan DNA sequences looking for genes, using algorithms based on the known features of genes, such as unique triplet sequences of nucleotides known as start and stop codons

that span a gene-sized segment of DNA or sequences of DNA that are known to be important in regulating adjacent genes. Once candidate genes are identified, they must be annotated to ascribe potential functions. Such annotation is generally based on known functions of similar gene sequences in other organisms, a type of analysis made possible by evolutionary conservation of gene sequence and function across organisms as a result of their common ancestry. However, after annotation there is still a subset of genes for which functions cannot be deduced; these functions gradually become revealed with further research.

Functional Genomics

Analysis of genes at the functional level is one of the main uses of genomics, an area known generally as functional genomics. Determining the function of individual genes can be done in several ways. Classical, or forward, genetic methodology starts with a randomly obtained mutant of interesting phenotype and uses this to find the normal gene sequence and its function. Reverse genetics starts with the normal gene sequence (as obtained by genomics), induces a targeted mutation into the gene, then, by observing how the mutation changes phenotype, deduces the normal function of the gene. The two approaches, forward and reverse, are complementary. Often a gene identified by forward genetics has been mapped to one specific chromosomal region, and the full genomic sequence reveals a gene in this position with an already annotated function.

Gene Identification by Microarray Genomic Analysis

Genomics has greatly simplified the process of finding the complete subset of genes that is relevant to some specific

temporal or developmental event of an organism. For example, microarray technology allows a sample of the DNA of a clone of each gene in a whole genome to be laid out in order on the surface of a special chip, which is basically a small thin piece of glass that is treated in such a way that DNA molecules firmly stick to the surface.

For any specific developmental stage of interest (e.g., the growth of root hairs in a plant or the production of a limb bud in an animal), the total RNA is extracted from cells of the organism, labeled with a fluorescent dye, and used to bathe the surfaces of the microarrays. As a result of specific base pairing, the RNAs present bind to the genes from which they were originally transcribed and produce fluorescent spots on the chip's surface. Hence, the total set of genes that were transcribed during the biological function of interest can be determined. Note that forward genetics can aim at a similar goal of assembling the subset of genes that pertain to some specific biological process. The forward genetic approach is to first induce a large set of mutations with phenotypes that appear to change the process in question, followed by attempts to define the genes that normally guide the process. However, the technique can only identify genes for which mutations produce an easily recognizable mutant phenotype, and so genes with subtle effects are often missed.

Comparative Genomics

A further application of genomics is in the study of evolutionary relationships. Using classical genetics, evolutionary relationships can be studied by comparing the chromosome size, number, and banding patterns between populations, species, and genera. However, if full genomic sequences are available, comparative genomics brings to bear a resolving power that is much greater than that of

classical genetics methods and allows much more subtle differences to be detected. This is because comparative genomics allows the DNAs of organisms to be compared directly and on a small scale.

Overall, comparative genomics has shown high levels of similarity between closely related animals, such as humans and chimpanzees, and, more surprisingly, similarity between seemingly distantly related animals, such as humans and insects. Comparative genomics applied to distinct populations of humans has shown that the human species is a genetic continuum, and the differences between populations are restricted to a very small subset of genes that affect superficial appearance such as skin colour. Furthermore, because DNA sequence can be measured mathematically, genomic analysis can be quantified in a very precise way to measure specific degrees of relatedness. Genomics has detected small-scale changes, such as the existence of surprisingly high levels of gene duplication and mobile elements within genomes.

POPULATION GENETICS

The study of genes in populations of animals, plants, and microbes provides information on past migrations, evolutionary relationships and extents of mixing among different varieties and species, and methods of adaptation to the environment. Statistical methods are used to analyze gene distributions and chromosomal variations in populations.

Population genetics is based on the mathematics of the frequencies of alleles and of genetic types in populations. For example, the Hardy-Weinberg formula, $p^2 + 2pq + q^2 = 1$, predicts the frequency of individuals with the respective homozygous dominant (*AA*), heterozygous (*Aa*), and

homozygous recessive (*aa*) genotypes in a randomly mating population. Selection, mutation, and random changes can be incorporated into such mathematical models to explain and predict the course of evolutionary change at the population level. These methods can be used on alleles of known phenotypic effect, such as the recessive allele for albinism, or on DNA segments of any type of known or unknown function.

Human population geneticists have traced the origins and migration and invasion routes of modern humans, *Homo sapiens*. DNA comparisons between the present peoples on the planet have pointed to an African origin of *H. sapiens*. Tracing specific forms of genes has allowed geneticists to deduce probable migration routes out of Africa to the areas colonized today. Similar studies show to what degree present populations have been mixed by recent patterns of travel.

EXPERIMENTAL BREEDING

Genetically diverse lines of organisms can be crossed in such a way to produce different combinations of alleles in one line. For example, parental lines are crossed, producing an F_1 generation, which is then allowed to undergo random mating to produce offspring that have purebreeding genotypes (i.e., *AA*, *bb*, *cc*, or *DD*). This type of experimental breeding is the origin of new plant and animal lines, which are an important part of making laboratory stocks for basic research. When applied to commerce, transgenic commercial lines produced experimentally are called genetically modified organisms (GMOs). Many of the plants and animals used by humans today (e.g., cows, pigs, chickens, sheep, wheat, corn [maize], potatoes, and rice) have been bred in this way.

Mathematical Techniques

Because much of genetics is based on quantitative data, mathematical techniques are used extensively in genetics. The laws of probability are applicable to crossbreeding and are used to predict frequencies of specific genetic constitutions in offspring. Geneticists also use statistical methods to determine the significance of deviations from expected results in experimental analyses. In addition, population genetics is based largely on mathematical logic—for example, the Hardy-Weinberg equilibrium and its derivatives.

Bioinformatics uses computer-centred statistical techniques to handle and analyze the vast amounts of information accumulating from genome sequencing projects. The computer program scans the DNA looking for genes, determining their probable function based on other similar genes, and comparing different DNA molecules for evolutionary analysis. Bioinformatics has made possible the discipline of systems biology, treating and analyzing the genes and gene products of cells as a complete and integrated system.

Hardy-Weinberg Law

The Hardy-Weinberg law is an algebraic equation that describes the genetic equilibrium within a population. It was discovered independently in 1908 by German physician Wilhelm Weinberg and British mathematician Godfrey Harold Hardy.

The science of population genetics is based on this principle, which may be stated as follows: in a large, random-mating population, the proportion of dominant and recessive genes present tends to remain constant from

generation to generation unless outside forces act to change it. In such a way even the rarest forms of genes, which one would assume would disappear, are preserved. The outside forces that can disrupt this natural equilibrium are selection, mutation, and migration. The discovery of this law was especially significant in affirming natural selection as the primary mechanism of evolution. If the proportions of gene forms in a population do not change, the rate of evolution will be zero. Individual variations occur because of the various genetic combinations that result from random mating of individuals, but nonrandom, or selective, mating must occur for natural selection to take place. Certain gene-controlled traits are selected for or selected against by the partners involved. Over a long period of time, this selective pressure will change the frequency of appearance of certain gene forms, and the traits they control will become commoner or rarer in the population.

Medical geneticists can use the Hardy-Weinberg law to calculate the probability of human matings that may result in offspring affected by genetic disorders. The law is also useful in determining whether the number of harmful mutations in a population is increasing as a result of radiation from industrial processes, medical techniques, and fallout.

Genetic Drift

Genetic drift, also called genetic sampling error (or Sewall Wright effect), is a change in the gene pool of a small population that takes place strictly by chance. Genetic drift can result in genetic traits being lost from a population or becoming widespread in a population without respect to the survival or reproductive value of the alleles involved. A

random statistical effect, genetic drift can occur only in small, isolated populations in which the gene pool is small enough that chance events can change its makeup substantially. In larger populations, any specific allele is carried by so many individuals that it is almost certain to be transmitted by some of them unless it is biologically unfavourable.

Founder Principle

The founder principle is used to explain whether a daughter population or migrant population may differ in genetic composition from its parent population because the founders of the daughter population were not a representative sample of the parent population. For example, if only blue-eyed inhabitants of a town whose residents included brown-eyed people decided to found a new town, their descendants would all be blue-eyed.

BEHAVIOUR GENETICS

Behaviour genetics, or psychogenetics, is the study of the influence of an organism's genetic composition on its behaviour and the interaction of heredity and environment insofar as they affect behaviour. The question of the determinants of behavioral abilities and disabilities has commonly been referred to as the "nature-nurture" controversy.

Early History of Behaviour Genetics

The relationship between behaviour and genetics, or heredity, dates to the work of the English scientist Sir Francis Galton (1822–1911), who coined the phrase "nature and nurture." Galton studied the families of outstanding

men of his day and concluded, like his cousin Charles Darwin, that mental powers run in families. Galton became the first to use twins in genetic research and pioneered many of the statistical methods of analysis that are in use today. In 1918 British statistician and geneticist Ronald Aylmer Fisher published a paper that showed how Gregor Mendel's laws of inheritance applied to complex traits influenced by multiple genes and environmental factors.

The first human behavioral genetic research on intelligence and mental illness began in the 1920s, when environmentalism (the theory that behaviour is a result of nongenetic factors such as various childhood experiences) became popular and before Nazi Germany's abuse of genetics made the notion of hereditary influence abhorrent. Although genetic research on human behaviour continued throughout the following decades, it was not until the 1970s that a balanced view came to prevail in psychiatry that recognized the importance of nature as well as nurture. In psychology, this reconciliation did not take hold until the 1980s. Much behavioral genetic research today focuses on identifying specific genes that affect behavioral dimensions, such as personality and intelligence, and disorders, such as autism, hyperactivity, depression, and schizophrenia.

Methods of Behaviour Genetics Study

Quantitative genetic methods are used to estimate the net effect of genetic and environmental factors on individual differences in any complex trait, including behavioral traits. In addition, molecular genetic methods are used to identify specific genes responsible for genetic influence. Research is carried out in both animals and humans; however, studies using animal models tend to provide

more accurate data than studies in humans because both genes and environment can be manipulated and controlled in the laboratory.

By mating related animals such as siblings for many generations, nearly pure strains are obtained in which all offspring are genetically highly similar. It is possible to screen for genetic influence on behaviour by comparing the behaviour of different inbred strains raised in the same laboratory environment. Another method, known as selective breeding, evaluates genetic involvement by attempting to breed for high and low extremes of a trait for several generations. Both methods have been applied to a wide variety of animal behaviours, especially learning and behavioral responses to drugs, and this research provides evidence for widespread influence of genes on behaviour.

Because genes and environments cannot be manipulated in the human species, two quasi-experimental methods are used to screen for genetic influence on individual differences in complex traits such as behaviour. The twin method relies on the accident of nature that results in identical (monozygotic) twins or fraternal (dizygotic) twins. Monozygotic twins are like clones, genetically identical to each other because they came from the same fertilized egg. Dizygotic twins, on the other hand, developed from two eggs that happened to be fertilized at the same time. Like other siblings, dizygotic twins are only half as similar genetically as monozygotic twins. To the extent that behavioral variability is caused by environmental factors, dizygotic twins should be as similar for the behavioral trait as are monozygotic twins because both types of twins are reared by the same parents in the same place at the same time. If the trait is influenced by genes, then dizygotic twins ought to be less similar than

Several sets of twins. Jochen Luebke/AFP/Getty Images

monozygotic twins. For schizophrenia, for example, the concordance (risk of one twin's being schizophrenic if the other is) is about 45 percent for monozygotic twins and about 15 percent for dizygotic twins. For intelligence as assessed by IQ tests, the correlation, an index of resemblance (0.00 indicates no resemblance and 1.00 indicates perfect resemblance), is 0.85 for monozygotic twins and 0.60 for dizygotic twins for studies throughout the world of more than 10,000 pairs of twins. The twin method has been robustly defended as a rough screen for genetic influence on behaviour.

The adoption method is a quasi-experimental design that relies on a social accident in which children are

adopted away from their biological (birth) parents early in life, thus cleaving the effects of nature and nurture. Because the twin and adoption methods are so different, greater confidence is warranted when results from these two methods converge on the same conclusion—as they usually do. An influential adoption study of schizophrenia in 1966 by American behavioral geneticist Leonard Heston showed that children adopted away from their schizophrenic biological mothers at birth were just as likely to become schizophrenic (about 10 percent) as were children reared by their schizophrenic biological mothers. A 20-year study begun in the 1970s in the United States of intelligence of adopted children and their biological and adoptive parents showed increasing similarity from infancy to childhood to adolescence between the adopted children and their biological parents but no resemblance between the adopted children and their adoptive parents.

In contrast to traditional molecular genetic research that focused on rare disorders caused by a single genetic mutation, molecular genetic research on complex behavioral traits and common behavioral disorders is much more difficult because multiple genes are involved and each gene has a relatively small effect. However, some genes identified in animal models have contributed to an improved understanding of complex human behavioral disorders such as reading disability, hyperactivity, autism, and dementia.

STUDY OF HUMAN GENETICS

Some geneticists specialize in the hereditary processes of human genetics. Most of the emphasis is on understanding and treating genetic disease and genetically influenced

ill health, areas collectively known as medical genetics. One broad area of activity is laboratory research dealing with the mechanisms of human gene function and malfunction and investigating pharmaceutical and other types of treatments. Since there is a high degree of evolutionary conservation between organisms, research on model organisms—such as bacteria, fungi, and fruit flies (*Drosophila*)—which are easier to study, often provides important insights into human gene function.

Many single-gene diseases, caused by mutant alleles of a single gene, have been discovered. Two well-characterized single-gene diseases include phenylketonuria (PKU) and Tay-Sachs disease. Other diseases, such as heart disease, schizophrenia, and depression, are thought to have more complex heredity components that involve a number of different genes. These diseases are the focus of a great deal of research that is being carried out today. In addition, abnormalities in chromosomes have been identified by studies employing techniques such as chromosomal banding. Individual chromosomes are identified by the banding patterns revealed by different staining techniques. Segments of chromosomes or chromosomes that are aberrant in number and morphology may be precisely identified.

Another broad area of activity is clinical genetics, which centres on advising parents of the likelihood of their children being affected by genetic disease caused by mutant genes and abnormal chromosome structure and number. Such genetic counseling is based on examining individual and family medical records and on diagnostic procedures that can detect unexpressed, abnormal forms of genes. Counseling is carried out by physicians with a particular interest in this area or by specially trained nonphysicians.

EPIGENETICS

Epigenetics is the study of the chemical modification of specific genes or gene-associated proteins of an organism. Epigenetic modifications can define how the information in genes is expressed and used by cells. The term *epigenetics* came into general use in the early 1940s, when British embryologist Conrad Waddington used it to describe the interactions between genes and gene products, which direct development and give rise to an organism's phenotype (observable characteristics). Since then, information revealed by epigenetics studies has revolutionized the fields of genetics and developmental biology. Specifically, researchers have uncovered a range of possible chemical modifications to DNA and to proteins called histones that associate tightly with DNA in the nucleus. These modifications can determine when or even if a given gene is expressed in a cell or organism.

It is clear that at least some epigenetic modifications are heritable, passed from parents to offspring, although they are not inherited by the same mechanism as is typical genetic information. Typical genetic information is encoded in the sequences of nucleotides that make up the DNA; this information is therefore passed from generation to generation as faithfully as the DNA replication process is accurate. Epigenetic information, however, is inherited only if the chemical modifications that constitute it are regenerated on newly synthesized DNA or proteins. Some forms of epigenetic modification are faithfully transmitted; however, others may be "erased" or "reset," depending on a variety of factors.

The principal type of epigenetic modification that is understood is methylation (addition of a methyl group). Methylation can be transient and can change rapidly during the life span of a cell or organism, or it can be essentially permanent once set early in the development

of the embryo. Other largely permanent chemical modifications also play a role; these include histone acetylation (addition of an acetyl group), ubiquitination (the addition of a ubiquitin protein), and phosphorylation (the addition of a phosphoryl group). The specific location of a given chemical modification can also be important. For example, certain histone modifications distinguish actively expressed regions of the genome from regions that are not highly expressed. These modifications may correlate with chromosome banding patterns generated by staining procedures common in karyotype analyses. Similarly, specific histone modifications may distinguish actively expressed genes from genes that are poised for expression or genes that are repressed in different kinds of cells.

Epigenetic changes not only influence the expression of genes in plants and animals but also enable the differentiation of pluripotent stem cells (cells having the potential to become any of many different kinds of cells). In other words, epigenetic changes allow cells that all share the same DNA and are ultimately derived from one fertilized egg to become specialized—for example, as liver cells, brain cells, or skin cells.

As the mechanisms of epigenetics become better understood, researchers recognize that the epigenome—chemical modification at the level of the genome—also influences a wide range of biomedical conditions. This new perception has opened the door to a deeper understanding of normal and abnormal biological processes and has offered the possibility of novel interventions that might prevent or ameliorate certain diseases.

Epigenetic contributions to disease fall into two classes. One class involves genes that are themselves regulated epigenetically, such as the imprinted (parent-specific) genes associated with Angelman syndrome or Prader-Willi

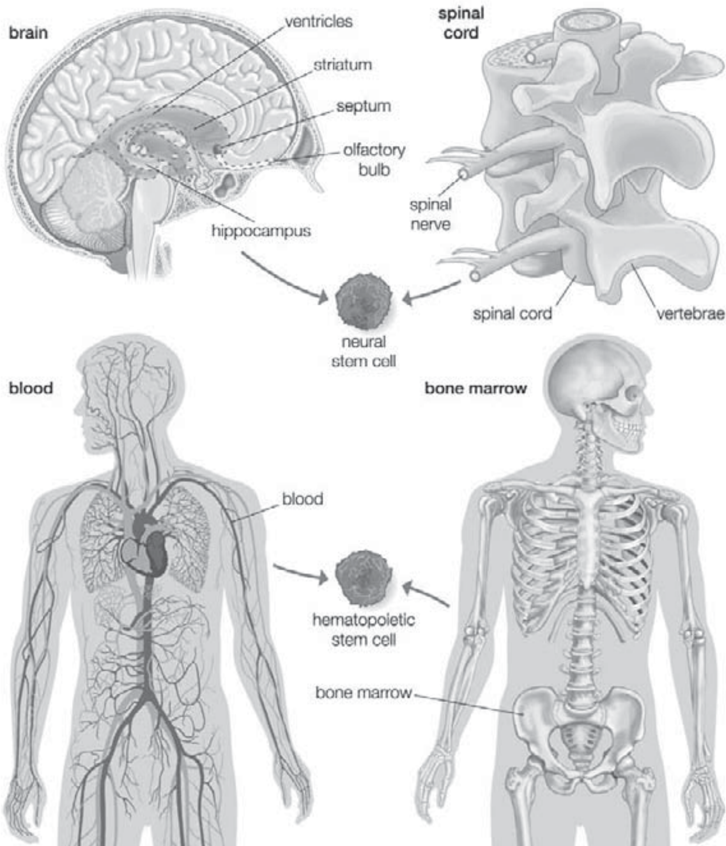
syndrome. Clinical outcomes in cases of these syndromes depend on the degree to which an inherited normal or mutated gene is or is not expressed. The other class involves genes whose products participate in the epigenetic machinery and thereby regulate the expression of other genes. For example, the gene *MECP2* (methyl CpG binding protein 2) encodes a protein that binds to specific methylated regions of DNA and contributes to the silencing of those sequences. Mutations that impair the *MECP2* gene can lead to Rett syndrome.

Many tumours and cancers are believed to involve epigenetic changes attributable to environmental factors. These changes include a general decrease in methylation, which is thought to contribute to the increased expression of growth-promoting genes, punctuated by gene-specific increases in methylation that are thought to silence tumour-suppressor genes. Epigenetic signaling attributed to environmental factors has also been associated with some characteristics of aging by researchers that studied the apparently unequal aging rates in genetically identical twins.

One of the most promising areas of epigenetic investigation involves stem cells. Researchers have understood for some time that epigenetic mechanisms play a key role in defining the "potentiality" of stem cells. As those mechanisms become clearer, it may become possible to intervene and effectively alter the developmental state and even the tissue type of given cells. The implications of this work for future clinical regenerative intervention for conditions ranging from trauma to neurodegenerative disease are profound.

STEM CELLS

A stem cell is an undifferentiated cell that can divide to produce some offspring cells that continue as stem cells



Neural and hematopoietic stem cells have tremendous potential in the development of therapies for certain diseases, such as diabetes and Parkinson disease. Neural stem cells occur in the spinal cord and in specific regions of the brain, and hematopoietic stem cells occur in the blood and bone marrow. Encyclopædia Britannica, Inc.

and some cells that are destined to differentiate (become specialized). Stem cells are an ongoing source of the differentiated cells that make up the tissues and organs of animals and plants. There is great interest in stem cells because they have potential in the development of therapies for replacing defective or damaged cells resulting from a

variety of disorders and injuries, such as Parkinson disease, heart disease, and diabetes. There are two major types of stem cells: embryonic stem cells and adult stem cells, which are also called tissue stem cells.

Embryonic Stem Cells

Embryonic stem cells (often referred to as ES cells) are stem cells that are derived from the inner cell mass of a mammalian embryo at a very early stage of development, when it is composed of a hollow sphere of dividing cells (a blastocyst). Embryonic stem cells from human embryos and from embryos of certain other mammalian species can be grown in tissue culture.

Extensive experience with mouse embryonic stem cells made it possible for scientists to grow human embryonic stem cells from early human embryos, and the first human stem cell line was created in 1998. Human embryonic stem cells are in many respects similar to mouse embryonic stem cells. The human embryonic stem cells form a wide variety of differentiated tissues in vitro, and they form teratomas when grafted into immunocompetent mice. It is not known whether the cells can colonize all the tissues of a human embryo, but it is presumed from their other properties that they are indeed pluripotent cells, and they therefore are regarded as a possible source of differentiated cells for cell therapy—the replacement of a patient's defective cell type with healthy cells. Large quantities of cells, such as dopamine-secreting neurons for the treatment of Parkinson disease and insulin-secreting pancreatic beta cells for the treatment of diabetes, could be produced from embryonic stem cells for cell transplantation. Cells for this purpose have previously been obtainable only from sources in very limited supply, such

as the pancreatic beta cells obtained from the cadavers of human organ donors.

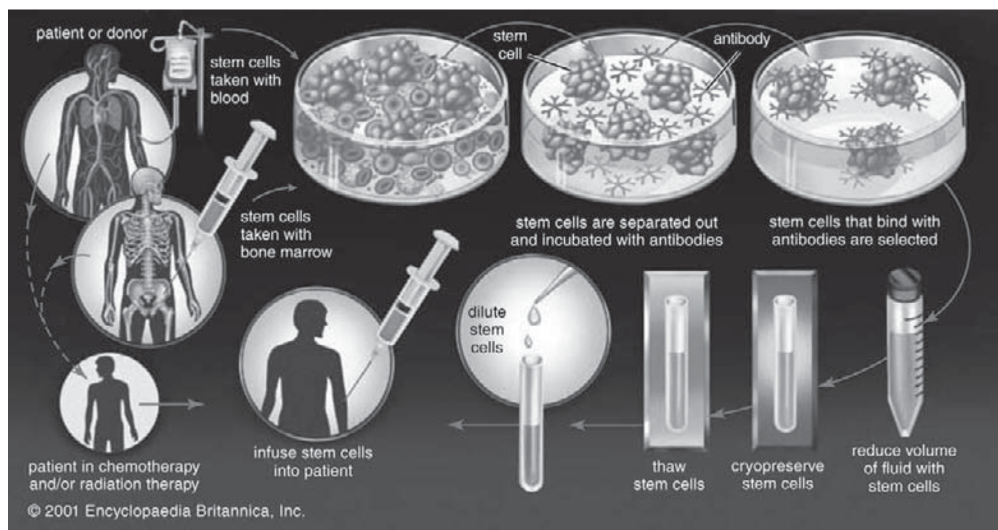
The use of human embryonic stem cells evokes ethical concerns, because the blastocyst-stage embryos are destroyed in the process of obtaining the stem cells. The embryos from which stem cells have been obtained are produced through in vitro fertilization, and people who consider preimplantation human embryos to be human beings generally believe that such work is morally wrong. Others accept it because they regard the blastocysts to be simply balls of cells, and human cells used in laboratories have not previously been accorded any special moral or legal status. Moreover, it is known that none of the cells of the inner cell mass are exclusively destined to become part of the embryo itself—all of the cells contribute some or all of their cell offspring to the placenta, which also has not been accorded any special legal status. The divergence of views on this issue is illustrated by the fact that the use of human embryonic stem cells is allowed in some countries and prohibited in others.

Adult Stem Cells

Some tissues in the adult body, such as the epidermis of the skin, the lining of the small intestine, and bone marrow, undergo continuous cellular turnover. They contain stem cells, which persist indefinitely, and a much larger number of "transit amplifying cells," which arise from the stem cells and divide a finite number of times until they become differentiated. The stem cells exist in niches formed by other cells, which secrete substances that keep the stem cells alive and active. Some types of tissue, such as liver tissue, show minimal cell division or undergo cell division only when injured. In such tissues there is

probably no special stem-cell population, and any cell can participate in tissue regeneration when required.

Bone marrow contains cells called hematopoietic stem cells, which generate all the cell types of the blood and the immune system. Hematopoietic stem cells are also found in small numbers in peripheral blood and in larger numbers in umbilical cord blood. In bone marrow, hematopoietic stem cells are anchored to osteoblasts of the trabecular bone and to blood vessels. They generate progeny that can become lymphocytes, granulocytes, red blood cells, and certain other cell types, depending on the balance of growth factors in their immediate environment.



High doses of chemotherapy or radiation destroy not only cancer cells but also bone marrow, which is rich in blood-forming stem cells. In order to replace damaged marrow, stem cells are harvested from either the blood or the bone marrow of the cancer patient before therapy; cells also may be taken from a genetically compatible donor. In order to remove unwanted cells, such as tumour cells, from the sample, it is incubated with antibodies that bind only to stem cells. The fluid that contains the selected cells is reduced in volume and frozen until needed. The fluid is then thawed, diluted, and reinfused into the patient's body. Once in the bloodstream, the stem cells travel to the bone marrow, where they implant themselves and begin producing healthy cells.

Work with experimental animals has shown that transplants of hematopoietic stem cells can occasionally colonize other tissues, with the transplanted cells becoming neurons, muscle cells, or epithelia. The degree to which transplanted hematopoietic stem cells are able to colonize other tissues is exceedingly small. Despite this, the use of hematopoietic stem cell transplants is being explored for conditions such as heart disease or autoimmune disorders. It is an especially attractive option for those opposed to the use of embryonic stem cells.

Bone marrow transplants (also known as bone marrow grafts) represent a type of stem cell therapy that is in common use. They are used to allow cancer patients to survive otherwise lethal doses of radiation therapy or chemotherapy that destroy the stem cells in bone marrow. For this procedure, the patient's own marrow is harvested before the cancer treatment and is then reinfused into the body after treatment. The hematopoietic stem cells of the transplant colonize the damaged marrow and eventually repopulate the blood and the immune system with functional cells. Bone marrow transplants are also often carried out between individuals (allograft). In this case the grafted marrow has some beneficial antitumour effect. Risks associated with bone marrow allografts include rejection of the graft by the patient's immune system and reaction of immune cells of the graft against the patient's tissues (graft-versus-host disease).

Bone marrow is a source for mesenchymal stem cells (sometimes called marrow stromal cells, or MSCs), which are precursors to non-hematopoietic stem cells that have the potential to differentiate into several different types of cells, including cells that form bone, muscle, and connective tissue. In cell cultures, bone-marrow-derived mesenchymal stem cells demonstrate pluripotency when exposed to substances that influence cell differentiation.

Harnessing these pluripotent properties has become highly valuable in the generation of transplantable tissues and organs. In 2008 scientists used mesenchymal stem cells to bioengineer a section of trachea that was transplanted into a woman whose upper airway had been severely damaged by tuberculosis. The stem cells were derived from the woman's bone marrow, cultured in a laboratory, and used for tissue engineering. In the engineering process, a donor trachea was stripped of its interior and exterior cell linings, leaving behind a trachea "scaffold" of connective tissue. The stem cells derived from the recipient were then used to recolonize the interior of the scaffold, and normal epithelial cells, also isolated from the recipient, were used to recolonize the exterior of the trachea. The use of the recipient's own cells to populate the trachea scaffold prevented immune rejection and eliminated the need for immunosuppression therapy. The transplant, which was successful, was the first of its kind.

Research has shown that there are also stem cells in the brain. In mammals very few new neurons are formed after birth, but some neurons in the olfactory bulbs and in the hippocampus are continually being formed. These neurons arise from neural stem cells, which can be cultured in vitro in the form of neurospheres—small cell clusters that contain stem cells and some of their progeny. This type of stem cell is being studied for use in cell therapy to treat Parkinson disease and other forms of neurodegeneration or traumatic damage to the central nervous system.

Induced Pluripotent Stem Cells

Due to the ethical and moral issues surrounding the use of embryonic stem cells, scientists have searched for ways to reprogram adult somatic (body) cells. Studies of cell fusion, in which differentiated adult somatic cells grown in

culture with embryonic stem cells fuse with the stem cells and acquire embryonic stem-cell-like properties, led to the idea that specific genes could reprogram differentiated adult cells. An advantage of cell fusion is that it relies on existing embryonic stem cells instead of eggs. However, fused cells stimulate an immune response when transplanted into humans, which leads to transplant rejection. As a result, research has become increasingly focused on the genes and proteins capable of reprogramming adult cells to a pluripotent state.

In order to make adult cells pluripotent without fusing them to embryonic stem cells, regulatory genes that induce pluripotency must be introduced into the nuclei of adult cells. To do this, adult cells are grown in cell culture, and specific combinations of regulatory genes are inserted into retroviruses (viruses that convert RNA into DNA), which are then introduced to the culture medium. The retroviruses transport the RNA of the regulatory genes into the nuclei of the adult cells, where the genes are then incorporated into the DNA of the cells. About 1 out of every 10,000 cells acquires embryonic stem cell properties. Although the mechanism is still uncertain, it is clear that some of the genes confer embryonic stem cell properties by means of the regulation of numerous other genes. Adult cells that become reprogrammed in this way are known as induced pluripotent stem (iPS) cells.

Similar to embryonic stem cells, induced pluripotent stem cells can be stimulated to differentiate into select types of cells that could in principle be used for disease-specific treatments. In addition, the generation of induced pluripotent stem cells from the adult cells of patients affected by genetic diseases can be used to model the diseases in the laboratory. For example, in 2008 researchers isolated skin cells from a child with an inherited neurological disease called spinal muscular atrophy and then

reprogrammed these cells into induced pluripotent stem cells. The reprogrammed cells retained the disease genotype of the adult cells and were stimulated to differentiate into motor neurons that displayed functional insufficiencies associated with spinal muscular atrophy. By recapitulating the disease in the laboratory, scientists were able to study closely the cellular changes that occurred as the disease progressed. Such models promise not only to improve scientists' understanding of genetic diseases but also to facilitate the development of new therapeutic strategies tailored to each type of genetic disease.

Patient-specific induced pluripotent stem cells are unlikely to be rejected by the immune system, which makes them highly valuable in terms of their therapeutic application. However, before induced pluripotent stem cells can be used to treat human diseases, researchers must find a way to introduce the active reprogramming genes without using retroviruses, which can cause diseases such as leukemia in humans. A possible alternative to the use of retroviruses to transport regulatory genes into the nuclei of adult cells is the use of plasmids, which are less tumourigenic than viruses.

RNA INTERFERENCE

In the 1990s American scientists Andrew Z. Fire and Craig C. Mello discovered that genes can be silenced by segments of double-stranded RNA (dsRNA) that are introduced into cells in tissue culture. Gene silencing by dsRNA makes use of the naturally occurring cell machinery that is involved in the processing of microRNA (miRNA) in eukaryotic cells. For example, each dsRNA is cleaved into small pieces by the DICER enzyme. These pieces are called short interfering RNAs (siRNAs) and are about 20

to 25 nucleotides in length. Similar to miRNA, siRNA binds to the RNA-induced silencing complex (RISC) and cleaves targeted sequences of messenger RNA (mRNA).

There are different types of synthetic dsRNAs that can be employed to disrupt gene function. Commonly used molecules include siRNA, which bypasses DICER cleavage, and small hairpin RNA (shRNA), which actually is one RNA strand containing two unique siRNA segments that is folded into a double strand, with the adjacent nucleotides joined through heating (annealing) rather than complementary base pairing. This creates a structure that resembles a hairpin because it has a tight loop at one end. Inside a cell, the shRNA is cleaved into its two component siRNAs by DICER.

RNAi is an exceptionally powerful genetics research tool. Synthetic dsRNAs are designed to prevent the expression of specific genes, thereby enabling scientists to manipulate the activity of genes in order to better understand their functions. In addition, abnormally overactive genes contribute to certain human diseases, and silencing this activity using RNAi has become an important area of medical research. Today, RNAi is being explored as a form of treatment for a variety of diseases, including macular degeneration, hepatitis, AIDS, Huntington disease, and cancer.

In macular degeneration, RNA sequences that block the production of a protein called vascular endothelial growth factor (VEGF) in cells of the retina can inhibit the excess growth of retinal blood vessels, which leak and lead to vision loss. RNAi treatments for macular degeneration involve the injection of “naked RNA” into the eye. The term *naked RNA* is used to distinguish this approach from those that employ viral vectors to introduce dsRNA into diseased cells. Interfering RNAs incorporated into

vectors are being studied for their effectiveness in slowing tumour growth. For example, mRNA transcripts of genes known to be overactive in certain forms of cancer serve as useful targets for RNAi-based treatments, which can silence overactive genes and slow disease progression.

Factors such as ensuring that interfering RNAs reach the cells and that the viral vectors themselves do not give rise to dangerous side effects has complicated the development of RNAi therapies. Furthermore, sequence similarities between genes can result in the binding of dsRNAs to otherwise properly functioning genes. This can result in the silencing of healthy genes vital to normal cell function. Nevertheless, the technique remains promising for applications in medicine.

CHAPTER 5

Genetic Engineering

The field of genetic engineering encompasses the artificial manipulation, modification, and recombination of DNA or other nucleic acid molecules in order to modify an organism or population of organisms. The term *genetic engineering* initially meant any of a wide range of techniques for the modification or manipulation of organisms through the processes of heredity and reproduction. As such, the term embraced both artificial selection and all the interventions of biomedical techniques, among them artificial insemination, in vitro fertilization (e.g., "test-tube" babies), sperm banks, cloning, and gene manipulation. Today, however, the term denotes the narrower field of recombinant DNA technology, or gene cloning, in which DNA molecules from two or more sources are combined either within cells or in vitro and are then inserted into host organisms in which they are able to propagate. Gene cloning is used to produce new genetic combinations that are of value to science, medicine, agriculture, or industry.

DNA is the carrier of genetic information; it achieves its effects by directing the synthesis of proteins. Most recombinant DNA technology involves the insertion of foreign genes into the plasmids of common laboratory strains of bacteria. Plasmids are small rings of DNA; they are not part of the bacterium's chromosome (the main repository of the organism's genetic information). Nonetheless, they are capable of directing protein synthesis, and, like chromosomal DNA, they are reproduced and passed on to the bacterium's progeny. Thus, by incorporating foreign DNA (for example, a mammalian gene) into a bacterium, researchers can obtain an almost limitless number of copies of the inserted gene. Furthermore, if the inserted gene is operative (i.e., if it directs protein

synthesis), the modified bacterium will produce the protein specified by the foreign DNA.

A key step in the development of genetic engineering was the discovery of restriction enzymes in 1968 by the Swiss microbiologist Werner Arber. However, type II restriction enzymes, which are essential to genetic engineering for their ability to cleave a specific site within the DNA (as opposed to type I restriction enzymes, which cleave DNA at random sites), were not identified until 1969, when the American molecular biologist Hamilton O. Smith purified this enzyme. Drawing on Smith's work, the American molecular biologist Daniel Nathans helped advance the technique of DNA recombination in 1970–71 and demonstrated that type II enzymes could be useful in genetic studies. Genetic engineering itself was pioneered in 1973 by the American biochemists Stanley N. Cohen and Herbert W. Boyer, who were among the first to cut DNA into fragments, rejoin different fragments, and insert the new genes into *E. coli* bacteria, which then reproduced.

Genetic engineering has advanced the understanding of many theoretical and practical aspects of gene function and organization. Through recombinant DNA techniques, bacteria have been created that are capable of synthesizing human insulin, human growth hormone, alpha interferon, a hepatitis B vaccine, and other medically useful substances. Plants may be genetically adjusted to enable them to fix nitrogen, and genetic diseases can possibly be corrected by replacing "bad" genes with "normal" ones. Nevertheless, special concern has been focused on such achievements for fear that they might result in the introduction of unfavourable and possibly dangerous traits into microorganisms that were previously free of them—e.g., resistance to antibiotics, production of toxins, or a tendency to cause disease.

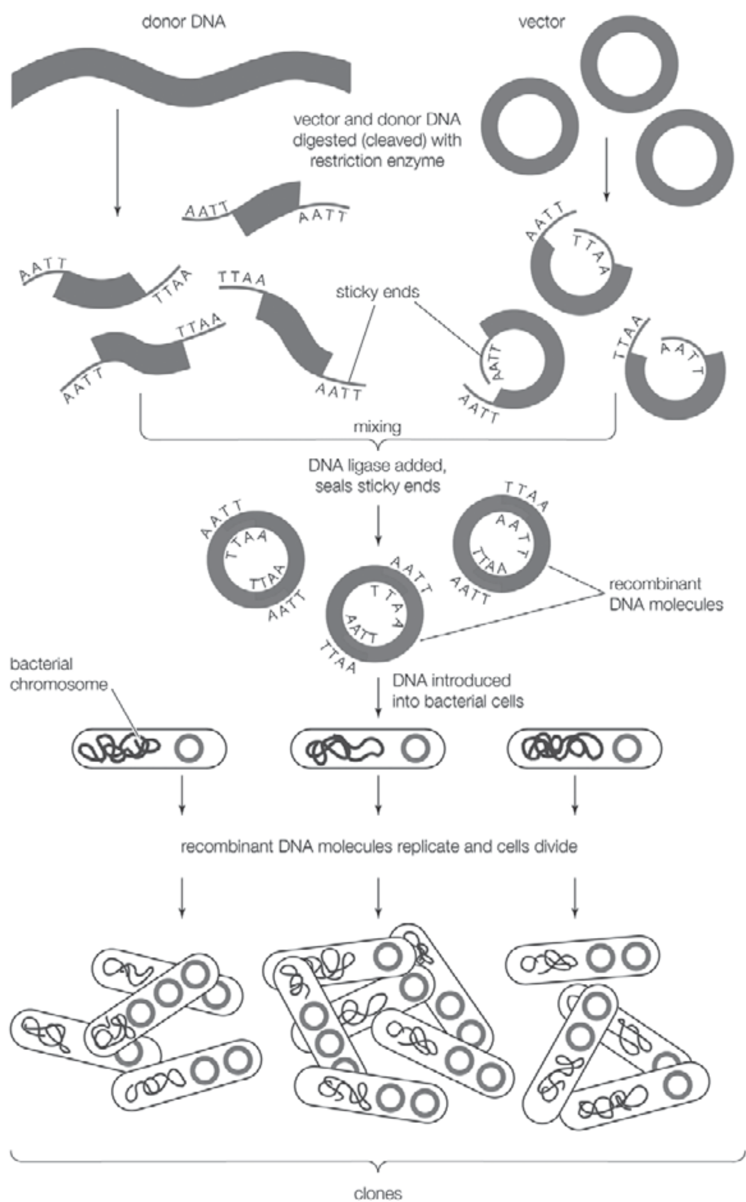
The "new" microorganisms created by recombinant DNA research were deemed patentable in 1980, and in 1986 the U.S. Department of Agriculture approved the sale of the first living genetically altered organism—a virus, used as a pseudorabies vaccine, from which a single gene had been cut. Since then hundreds of patents have been awarded for genetically altered bacteria and plants.

RECOMBINANT DNA TECHNOLOGY

Recombinant DNA technology involves the joining together of DNA molecules from two different species that are inserted into a host organism to produce new genetic combinations that are of value to science, medicine, agriculture, and industry. Since the focus of all genetics is the gene, the fundamental goal of laboratory geneticists is to isolate, characterize, and manipulate genes. Although it is relatively easy to isolate a sample of DNA from a collection of cells, finding a specific gene within this DNA sample can be compared to finding a needle in a haystack. Consider the fact that each human cell contains approximately 2 metres (6 feet) of DNA. Therefore, a small tissue sample will contain many kilometres of DNA. However, recombinant DNA technology has made it possible to isolate one gene or any other segment of DNA, enabling researchers to determine its nucleotide sequence, study its transcripts, mutate it in highly specific ways, and reinsert the modified sequence into a living organism.

DNA Cloning

In biology a clone is a group of individual cells or organisms descended from one progenitor. This means that the members of a clone are genetically identical because cell



Steps involved in the engineering of a recombinant DNA molecule. Encyclopædia Britannica, Inc.

replication produces identical daughter cells each time. The use of the word *clone* has been extended to recombinant DNA technology, which has provided scientists with the ability to produce many copies of a single fragment of DNA, such as a gene, creating identical copies that constitute a DNA clone. In practice the procedure is carried out by inserting a DNA fragment into a small DNA molecule and then allowing this molecule to replicate inside a simple living cell such as a bacterium. The small replicating molecule is called a DNA vector (carrier). The most commonly used vectors are plasmids (circular DNA molecules that originated from bacteria), viruses, and yeast cells. Plasmids are not a part of the main cellular genome, but they can carry genes that provide the host cell with useful properties, such as drug resistance, mating ability, and toxin production. They are small enough to be conveniently manipulated experimentally, and, furthermore, they will carry extra DNA that is spliced into them.

Creating the Clone

The steps in cloning are as follows. DNA is extracted from the organism under study and is cut into small fragments of a size suitable for cloning. Most often this is achieved by cleaving the DNA with a restriction enzyme. Restriction enzymes are extracted from several different species and strains of bacteria, in which they act as defense mechanisms against viruses. They can be thought of as "molecular scissors," cutting the DNA at specific target sequences. The most useful restriction enzymes make staggered cuts; that is, they leave a single-stranded overhang at the site of cleavage. These overhangs are very useful in cloning because the unpaired nucleotides will pair with other overhangs made using the same restriction enzyme. So, if

the donor DNA and the vector DNA are both cut with the same enzyme, there is a strong possibility that the donor fragments and the cut vector will splice together because of the complementary overhangs. The resulting molecule is called recombinant DNA. It is recombinant in the sense that it is composed of DNA from two different sources. Thus, it is a type of DNA that would be impossible naturally and is an artifact created by DNA technology.

The next step in the cloning process is to cut the vector with the same restriction enzyme used to cut the donor DNA. Vectors have target sites for many different restriction enzymes, but the most convenient ones are those that occur only once in the vector molecule. This is because the restriction enzyme then merely opens up the vector ring, creating a space for the insertion of the donor DNA segment. Cut vector DNA and donor DNA are mixed in a test tube, and the complementary ends of both types of DNA unite randomly. Of course, several types of unions are possible: donor fragment to donor fragment, vector fragment to vector fragment, and, most important, vector fragment to donor fragment, which can be selected for. Recombinant DNA associations form spontaneously in the above manner, but these associations are not stable because, although the ends are paired, the sugar-phosphate backbone of the DNA has not been sealed. This is accomplished by the application of an enzyme called DNA ligase, which seals the two segments, forming a continuous and stable double helix.

The mixture should now contain a population of vectors each containing a different donor insert. This solution is mixed with live bacterial cells that have been specially treated to make their cells more permeable to DNA. Recombinant molecules enter living cells in a

process called transformation. Usually, only a single recombinant molecule will enter any individual bacterial cell. Once inside, the recombinant DNA molecule replicates like any other plasmid DNA molecule, and many copies are subsequently produced. Furthermore, when the bacterial cell divides, all of the daughter cells receive the recombinant plasmid, which again replicates in each daughter cell.

The original mixture of transformed bacterial cells is spread out on the surface of a growth medium in a flat dish (Petri dish) so that the cells are separated from one another. These individual cells are invisible to the naked eye, but as each cell undergoes successive rounds of cell division, visible colonies form. Each colony is a cell clone, but it is also a DNA clone because the recombinant vector has now been amplified by replication during every round of cell division. Thus, the Petri dish, which may contain many hundreds of distinct colonies, represents a large number of clones of different DNA fragments. This collection of clones is called a DNA library. By considering the size of the donor genome and the average size of the inserts in the recombinant DNA molecule, a researcher can calculate the number of clones needed to encompass the entire donor genome, or, in other words, the number of clones needed to constitute a genomic library.

Another type of library is a cDNA library. Creation of a cDNA library begins with messenger RNA (mRNA) instead of DNA. MRNA carries encoded information from DNA to ribosomes for translation into protein. To create a cDNA library, these mRNA molecules are treated with the enzyme reverse transcriptase, which is used to make a DNA copy of an mRNA. The resulting DNA molecules are called complementary DNA (cDNA). A cDNA library represents a sampling of the

transcribed genes, whereas a genomic library includes untranscribed regions.

Both genomic and cDNA libraries are made without regard to obtaining functional cloned donor fragments. Genomic clones do not necessarily contain full-length copies of genes. Furthermore, genomic DNA from eukaryotes, whose cells have a clearly defined nucleus, contain introns, which are regions of DNA that are not translated into protein and cannot be processed by bacterial cells. This means that even full-sized genes are not translated in their entirety. In addition, eukaryotic regulatory signals are different from those used by prokaryotes, such as bacteria, whose cells lack a clearly defined nucleus. However, it is possible to produce expression libraries by slicing cDNA inserts immediately adjacent to a bacterial promoter region on the vector; in these expression libraries, eukaryotic proteins are made in bacterial cells, which allows several important technological applications.

Several bacterial viruses (bacteriophages) have also been used as vectors. The most commonly used is the lambda phage. The central part of the lambda genome is not essential for the virus to replicate in *E. coli*, so this can be excised using an appropriate restriction enzyme, and inserts from donor DNA can be spliced into the gap. In fact, when the phage repackages DNA into its protein capsule, it includes only DNA fragments the same length of the normal phage genome.

Vectors are chosen depending on the total amount of DNA that must be included in a library. Cosmids are engineered vectors that are hybrids of plasmid and phage lambda; however, they can carry larger inserts than either pUC plasmids (plasmids engineered to produce a very high number of DNA copies but that can accommodate only small inserts) or lambda phage alone. Bacterial

artificial chromosomes (BACs) are vectors based on F-factor (fertility factor) plasmids of *E. coli* and can carry much larger amounts of DNA. Yeast artificial chromosomes (YACs) are vectors based on autonomously replicating plasmids of *Saccharomyces cerevisiae* (baker's yeast). In yeast (a eukaryotic organism) a YAC behaves like a yeast chromosome and segregates properly into daughter cells. These vectors can carry the largest inserts of all and are used extensively in cloning large genomes such as the human genome.

Isolating the Clone

In general, cloning is undertaken in order to obtain the clone of one particular gene or DNA sequence of interest. The next step after cloning, therefore, is to find and isolate that clone among other members of the library. If the library encompasses the whole genome of an organism, then somewhere within that library will be the desired clone. There are several ways of finding it, depending on the specific gene concerned. Most commonly, a cloned DNA segment that shows homology to the sought gene is used as a probe. For example, if a mouse gene has already been cloned, then that clone can be used to find the equivalent human clone from a human genomic library.

Bacterial colonies constituting a library are grown in a collection of Petri dishes. Then a porous membrane is laid over the surface of each plate, and cells adhere to the membrane. The cells are ruptured, and DNA is separated into single strands—all on the membrane. The probe is also separated into single strands and labeled, often with radioactive phosphorus. A solution of the radioactive probe is then used to bathe the membrane. The single-stranded probe DNA will adhere only to the DNA of the

clone that contains the equivalent gene. The membrane is dried and placed against a sheet of radiation-sensitive film, and somewhere on the films a black spot will appear, announcing the presence and location of the desired clone. The clone can then be retrieved from the original Petri dishes.

DNA Sequencing

Once a segment of DNA has been cloned, its nucleotide sequence can be determined. The nucleotide sequence is the most fundamental level of knowledge of a gene or genome. It is the blueprint that contains the instructions for building an organism, and no understanding of genetic function or evolution could be complete without obtaining this information.

Knowledge of the sequence of a DNA segment has many uses, and some examples follow. First, it can be used to find genes, segments of DNA that code for a specific protein or phenotype. If a region of DNA has been sequenced, it can be screened for characteristic features of genes. For example, open reading frames (ORFs)—long sequences that begin with a start codon (three adjacent nucleotides; the sequence of a codon dictates amino acid production) and are uninterrupted by stop codons (except for one at their termination)—suggest a protein-coding region. Also, human genes are generally adjacent to so-called CpG islands—clusters of cytosine and guanine, two of the nucleotides that make up DNA. If a gene with a known phenotype (such as a disease gene in humans) is known to be in the chromosomal region sequenced, then unassigned genes in the region will become candidates for that function. Second, homologous DNA sequences of different organisms can be compared in order to

plot evolutionary relationships both within and between species. Third, a gene sequence can be screened for functional regions. In order to determine the function of a gene, various domains can be identified that are common to proteins of similar function. For example, certain amino acid sequences within a gene are always found in proteins that span a cell membrane; such amino acid stretches are called transmembrane domains. If a transmembrane domain is found in a gene of unknown function, it suggests that the encoded protein is located in the cellular membrane. Other domains characterize DNA-binding proteins. Several public databases of DNA sequences are available for analysis by any interested individual.

The two basic sequencing approaches are the Maxam-Gilbert method, discovered by and named for American molecular biologists Allan M. Maxam and Walter Gilbert, and the Sanger method, discovered by English biochemist Frederick Sanger. In the most commonly used method, the Sanger method, DNA chains are synthesized on a template strand, but chain growth is stopped when one of four possible dideoxy nucleotides, which lack a 3' hydroxyl group, is incorporated, thereby preventing the addition of another nucleotide. A population of nested, truncated DNA molecules results that represents each of the sites of that particular nucleotide in the template DNA. These molecules are separated in a procedure called electrophoresis, and the inferred nucleotide sequence is deduced using a computer.

In Vitro Mutagenesis

Another use of cloned DNA is in vitro mutagenesis in which a mutation is produced in a segment of cloned DNA. The DNA is then inserted into a cell or organism,

and the effects of the mutation are studied. Mutations are useful to geneticists in enabling them to investigate the components of any biological process. However, traditional mutational analysis relied on the occurrence of random spontaneous mutations—a hit-or-miss method in which it was impossible to predict the precise type or position of the mutations obtained. In vitro mutagenesis, however, allows specific mutations to be tailored for type and for position within the gene. A cloned gene is treated in the test tube (in vitro) to obtain the specific mutation desired, and then this fragment is reintroduced into the living cell, where it replaces the resident gene.

One method of in vitro mutagenesis is oligonucleotide-directed mutagenesis. A specific point in a sequenced gene is pinpointed for mutation. An oligonucleotide, a short stretch of synthetic DNA of the desired sequence, is made chemically. For example, the oligonucleotide might have adenine in one specific location instead of guanine. This oligonucleotide is hybridized to the complementary strand of the cloned gene; it will hybridize despite the one base pair mismatch. Various enzymes are added to allow the oligonucleotide to prime the synthesis of a complete strand within the vector. When the vector is introduced into a bacterial cell and replicates, the mutated strand will act as a template for a complementary strand that will also be mutant, and thus a fully mutant molecule is obtained. This fully mutant cloned molecule is then reintroduced into the donor organism, and the mutant DNA replaces the resident gene.

Another version of in vitro mutagenesis is gene disruption, or gene knockout. Here, the resident functional gene is replaced by a completely nonfunctional copy. The advantage of this technique over random mutagenesis is that specific genes can be knocked out at will, leaving all other genes untouched by the mutagenic procedure.

Gene Therapy

Gene therapy is the introduction of a normal gene into an individual's genome in order to repair a mutation that causes a genetic disease. When a normal gene is inserted into a mutant nucleus, it most likely will integrate into a chromosomal site different from the defective allele; although this may repair the mutation, a new mutation may result if the normal gene integrates into another functional gene. If the normal gene replaces the mutant allele, there is a chance that the transformed cells will proliferate and produce enough normal gene product for the entire body to be restored to the undiseased phenotype.

So far, human gene therapy has been attempted only on somatic (body) cells for diseases such as cancer and severe combined immunodeficiency syndrome (SCIDS). Somatic cells cured by gene therapy may reverse the symptoms of disease in the treated individual, but the modification is not passed on to the next generation. Germinal gene therapy aims to place corrected cells inside the germ line (e.g., cells of the ovary or testis). If this is achieved, these cells will undergo meiosis and provide a normal gametic contribution to the next generation. Germinal gene therapy has been achieved experimentally in animals but not in humans.

Reverse Genetics

Recombinant DNA technology has made possible a type of genetics called reverse genetics. Traditionally, genetic research starts with a mutant phenotype, and, by Mendelian crossing analysis, a researcher is able to attribute the phenotype to a specific gene. Reverse genetics travels in precisely the opposite direction. Researchers begin with a gene of unknown function and use molecular

analysis to determine its phenotype. One important tool in reverse genetics is gene knockout. By mutating the cloned gene of unknown function and using it to replace the resident copy or copies, the resultant mutant phenotype will show which biological function this gene normally controls.

Diagnostics

Recombinant DNA technology has led to powerful diagnostic procedures useful in both medicine and forensics. In medicine these diagnostic procedures are used in counseling prospective parents as to the likelihood of having a child with a particular disease, and they are also used in the prenatal prediction of genetic disease in the fetus. Researchers look for specific DNA fragments that are located in close proximity to the gene that causes the disease of concern. These fragments, called restriction fragment length polymorphisms (RFLPs), often serve as effective "genetic markers." In forensics, DNA fragments called variable number tandem repeats (VNTRs), which are highly variable between individuals, are employed to produce what is called a "DNA fingerprint." DNA fingerprinting can be used to determine if blood or other body fluids left at the scene of a crime belongs to a suspect.

Protein Manufacture

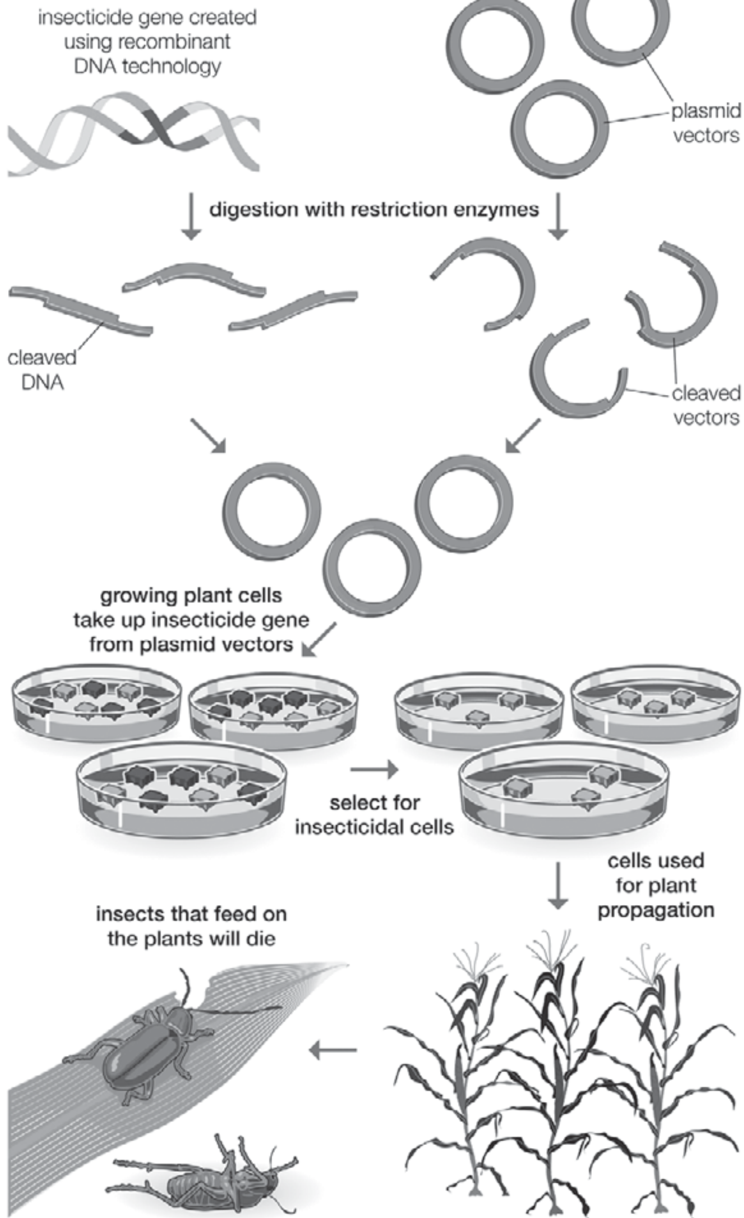
Recombinant DNA procedures have been used to convert bacteria into "factories" for the synthesis of foreign proteins. This technique is useful not only for preparing large amounts of protein for basic research but also for producing valuable proteins for medical use. For example, the genes for human proteins such as growth hormone, insulin, and blood-clotting factor can be commercially

manufactured. Another approach to producing proteins via recombinant DNA technology is to introduce the desired gene into the genome of an animal, engineered in such a way that the protein is secreted in the animal's milk, facilitating harvesting.

GENETICALLY MODIFIED ORGANISMS

A genetically modified organism (GMO) contains a genome that has been engineered in the laboratory in order to favour the expression of desired physiological traits or the production of desired biological products. In conventional livestock production, crop farming, and even pet breeding, it has long been the practice to breed select individuals of a species in order to produce offspring that have desirable traits. In genetic modification, however, recombinant genetic technologies are employed to produce organisms whose genomes have been precisely altered at the molecular level, usually by the inclusion of genes from unrelated species of organisms that code for traits that would not be obtained easily through conventional selective breeding.

GMOs are produced using scientific methods that include recombinant DNA technology and reproductive cloning. Reproductive cloning technology generates offspring that are genetically identical to the parent by the transfer of an entire donor nucleus into the enucleated cytoplasm of a host egg. The first animal produced using this cloning technique was a sheep named Dolly, born in 1996. Since then a number of other animals, including pigs, horses, and dogs, have been generated using reproductive cloning technology. Recombinant DNA technology, on the other hand, involves the insertion of one or more individual genes from an organism of



Genetically modified organisms are produced using scientific methods that include recombinant DNA technology. Encyclopædia Britannica, Inc.

one species into the DNA of another. Whole-genome replacement, involving the transplantation of one bacterial genome into the "cell body," or cytoplasm, of another microorganism, has been reported, although this technology is still limited to basic scientific applications.

GMOs produced through genetic technologies have become a part of everyday life, entering into society through agriculture, medicine, research, and environmental management. However, while GMOs have benefited human society in many ways, some disadvantages exist; therefore, the production of GMOs remains a highly controversial topic in many parts of the world.

GMOs in Agriculture

Genetically modified (GM) foods were first approved for human consumption in the United States in 1995, and by 1999 almost 50 percent of the corn, cotton, and soybeans planted in the United States were GM. The introduction

A cotton picker harvesting genetically modified cotton. Scott Olson/ Getty Images

of these crops dramatically increased per area crop yields and, in some cases, reduced the use of chemical insecticides. For example, the application of wide-spectrum insecticides declined in many areas growing plants, such as potatoes, cotton, and corn, that were endowed with a gene from the bacterium *Bacillus thuringiensis*, which produces a natural insecticide called Bt toxin.

Field studies conducted in India in which Bt cotton was compared with non-Bt cotton demonstrated a 30–80 percent increase in yield from the GM crop. This increase was attributed to marked improvement in the GM plants' ability to overcome bollworm infestation, which was otherwise common. Studies of Bt cotton production in Arizona, U.S., demonstrated only small gains in yield—about 5 percent—with an estimated cost reduction of $25–65 (USD) per acre due to decreased pesticide applications. In China, a seven-year study of farms planting Bt cotton demonstrated initial success of the GM crop, with farmers who had planted Bt cotton reducing their pesticide use by 70 percent and increasing their earnings by 36 percent. However, after four years, the benefits of Bt cotton eroded as populations of insect pests other than bollworm increased, and farmers once again were forced to spray broad-spectrum pesticides. While the problem was not Bt-resistant bollworms, as had been feared initially, it nonetheless became clear that much more research was needed for communities to realize sustainable and environmentally responsible benefits from planting GM crops.

Other GM plants were engineered for resistance to a specific chemical herbicide, rather than resistance to a natural predator or pest. Herbicide-resistant crops (HRC) have been available since the mid-1980s; these crops enable effective chemical control of weeds, since only the HRC plants can survive in fields treated with the corresponding herbicide. However, because these crops

encourage increased application of chemicals to the soil, rather than decreased application, they remain controversial with regard to their environmental impact.

By 2002 more than 60 percent of processed foods consumed in the United States contained at least some GM ingredients. Despite the concerns of some consumer groups, especially in Europe, numerous scientific panels, including the U.S. Food and Drug Administration, have concluded that consumption of GM foods is safe, even in cases involving GM foods with genetic material from very distantly related organisms. Indeed, foods containing GM ingredients do not require special labeling in the United States, although some groups have continued to lobby to change this ruling. By 2006, although the majority of GM crops were still grown in the Americas, GM plants tailored for production and consumption in other parts of the world were in field tests. For example, sweet potatoes intended for Africa were modified for resistance to sweet potato feathery mottle virus (SPFMV) by inserting into the sweet potato genome a gene

Plant biotechnologist Dr. Swapan Datta examines golden rice, a genetically modified food high in beta carotene, which has produced controversy, despite its health benefits to much of the world's population. David Greedy/Getty Images

encoding a viral coat protein from the strain of virus that causes SPFMV. The premise for this modification was based on earlier studies in other plants such as tobacco in which introduction of viral coat proteins rendered plants resistant to the virus.

The so-called "golden" rice intended for Asia was genetically modified to produce almost 20 times the beta-carotene of previous varieties. Golden rice was created by modifying the rice genome to include a gene from the daffodil *Narcissus pseudonarcissus* that produces an enzyme known as phyotene synthase and a gene from the bacterium *Erwinia uredovora* that produces an enzyme called phyotene desaturase. The introduction of these genes enabled beta-carotene, which is converted to vitamin A in the human liver, to accumulate in the rice endosperm—the edible part of the rice plant—thereby increasing the amount of beta-carotene available for vitamin A synthesis in the body.

Another form of modified rice was generated to help combat iron deficiency, which impacts close to 30 percent of the world population. This GM crop was engineered by introducing into the rice genome a ferritin gene from the common bean, *Phaseolus vulgaris*, that produces a protein capable of binding iron, as well as a gene from the fungus *Aspergillus fumigatus* that produces an enzyme capable of digesting compounds that increase iron bioavailability via digestion of phytate (an inhibitor of iron absorption). The iron-fortified GM rice was engineered to overexpress an existing rice gene that produces a cysteine-rich metallothioneinlike (metal-binding) protein that enhances iron absorption.

A variety of other crops modified to endure the weather extremes common in other parts of the globe are also in production.

GMOs in Medicine and Research

GMOs have emerged as one of the mainstays of biomedical research since the 1980s. For example, GM animal models of human genetic diseases enabled researchers to test novel therapies and to explore the roles of candidate risk factors and modifiers of disease outcome. GM microbes, plants, and animals also revolutionized the production of complex pharmaceuticals by enabling the generation of safer and cheaper vaccines and therapeutics. Pharmaceutical products range from recombinant hepatitis B vaccine produced by GM baker's yeast to injectable insulin (for diabetics) produced in GM *E. coli* bacteria and to factor VIII (for hemophiliacs) and tissue plasminogen activator (tPA, for heart attack or stroke patients), both of which are produced in GM mammalian cells grown in laboratory culture. Furthermore, GM plants that produce "edible vaccines" are under development. Such plants, which are engineered to express antigens derived from microbes or parasites that infect the digestive tract, might someday offer a safe, cheap, and painless way to provide vaccines worldwide, without concern for the availability of refrigeration or sterile needles. Novel DNA vaccines may be useful in the struggle to prevent diseases that have proved resistant to traditional vaccination approaches, including HIV/AIDS, tuberculosis, and cancer.

Genetic modification of insects has become an important area of research, especially in the struggle to prevent parasitic diseases. For example, GM mosquitoes have been developed that express a small protein called SM1, which blocks entry of the malaria parasite, *Plasmodium*, into the mosquito's gut. This results in the disruption of the parasite's life cycle and renders the mosquito malaria-resistant. Introduction of these GM mosquitoes into the

wild may someday help eradicate transmission of the malaria parasite without widespread use of harmful chemicals such as DDT or disruption of the normal food chain.

Finally, genetic modification of humans, or so-called gene therapy, is becoming a treatment option for diseases ranging from rare metabolic disorders to cancer. Coupling stem cell technology with recombinant DNA methods may someday allow stem cells derived from a patient to be modified in the laboratory to introduce a desired gene. For example, a normal beta-globin gene may be introduced into the DNA of bone marrow-derived hematopoietic stem cells from a patient with sickle cell anemia, and introduction of these GM cells into the patient could cure the disease without the need for a matched donor.

Role of GMOs in Environmental Management

Another application of GMOs is in the management of environmental issues. For example, some bacteria can produce biodegradable plastics, and the transfer of this ability to microbes that can be easily grown in the laboratory may enable the wide-scale "greening" of the plastics industry. Zeneca, a British company, developed a microbially produced biodegradable plastic called Biopol. This plastic is made using a GM bacterium, *Ralstonia eutropha*, to convert glucose and a variety of organic acids into a flexible polymer. GMOs endowed with the bacterially encoded ability to metabolize oil and heavy metals may provide efficient bioremediation strategies.

Genetic modification technologies may help save endangered species such as the giant panda, whose genome was sequenced in an international effort led by the Beijing Genomics Institute at Shenzhen. Genetic studies of the

Tian Tian, a giant panda at the National Zoo in Washington, D.C., chews on bamboo. Karen Bleier/AFP/Getty Images

panda genome may provide insight into why pandas have such low rates of reproductive success in captivity. A likely set of genes to consider for future genetic modification, should the goals of panda conservation warrant it, is the major histocompatibility complex (MHC). The MHC genes play an important role in regulating immune function and also influence behaviours and physiological patterns associated with reproduction.

Sociopolitical Relevance of GMOs

While GMOs offer many potential benefits to society, the potential risks associated with them have fueled controversy, especially in the food industry. Many skeptics warn about the dangers that GM crops may pose to human

health. For example, genetic manipulation may potentially alter the allergenic properties of crops. However, the more-established risk involves the potential spread of engineered crop genes to native flora and the possible evolution of insecticide-resistant "superbugs." In 1998 the European Union (EU) addressed such concerns by implementing strict GMO labeling laws and a moratorium on the growth and import of GM crops. In addition, the stance of the EU on GM crops has led to trade disputes with the United States, which, by comparison, has accepted GM foods very openly. Other countries, such as Canada, China, Argentina, and Australia, also have open policies on GM foods, but some African states have rejected international food aid containing GM crops.

The use of GMOs in medicine and research has produced a debate that is more philosophical in nature. For example, while genetic researchers believe they are working to cure disease and ameliorate suffering, many people worry that current gene therapy approaches may one day be applied to produce "designer" children or to lengthen the natural human life span. Similar to many other technologies, gene therapy and the production and application of GMOs can be used to address and resolve complicated scientific, medical, and environmental issues, but they must be used wisely.

CLONING

Cloning is the process of generating a genetically identical copy of a cell or an organism. Cloning happens all the time in nature—for example, when a cell replicates itself asexually without any genetic alteration or recombination. Prokaryotic organisms (organisms lacking a cell nucleus), such as bacteria and yeasts, create genetically identical duplicates of themselves using binary fission or budding.

In eukaryotic organisms such as humans, all the cells that undergo mitosis, such as skin cells and cells lining the gastrointestinal tract, are clones; the only exceptions are gametes (eggs and sperm), which undergo meiosis and genetic recombination.

In biomedical research cloning is broadly defined to mean the duplication of any kind of biological material for scientific study, such as a piece of DNA or an individual cell. For example, segments of DNA are replicated exponentially by a process known as polymerase chain reaction, or PCR, a technique that is used widely in basic biological research. The type of cloning that is the focus of much ethical controversy involves the generation of cloned embryos, particularly those of humans, which are genetically identical to the organisms from which they are derived, and the subsequent use of these embryos for research, therapeutic, or reproductive purposes.

Early Cloning Experiments

Reproductive cloning was originally carried out by artificial "twinning," or embryo splitting, which was first performed on a salamander embryo in 1902 by German embryologist Hans Spemann. In 1928, Spemann, who was later awarded the Nobel Prize for Physiology or Medicine (1935) for his research on embryonic development, theorized about another cloning procedure known as nuclear transfer. This procedure was performed in 1952 by American scientists Robert W. Briggs and Thomas J. King, who used DNA from frog embryonic cells to generate cloned tadpoles. A decade later, British biologist John Gurdon successfully carried out nuclear transfer using DNA from adult frog cells.

Advancements in the field of molecular biology led to the development of techniques that allowed scientists

to manipulate cells and to detect chemical markers that signal changes within cells. With the advent of recombinant DNA technology in the 1970s, it became possible for scientists to create transgenic clones—clones with genomes containing pieces of DNA from other organisms. Beginning in the 1980s mammals such as sheep were cloned from early and partially differentiated embryonic cells. In 1996 British developmental biologist Ian Wilmut generated a cloned sheep, named Dolly, by means of nuclear transfer involving an enucleated embryo and a differentiated cell nucleus. This technique, which was later refined and became known as somatic cell nuclear transfer (SCNT), represented an extraordinary advance in the science of cloning because it resulted in the creation of a genetically identical clone of an already grown sheep. It also indicated that it was possible for the DNA in differentiated somatic cells to revert to an undifferentiated embryonic stage, thereby reestablishing pluripotency—the potential of an embryonic cell to grow into any one of the numerous different types of mature body cells that make up a complete organism. The realization that the DNA of somatic cells could be reprogrammed to a pluripotent state significantly impacted research into therapeutic cloning and the development of stem cell therapies.

Soon after the generation of Dolly, a number of other animals were cloned by SCNT, including pigs, goats, rats, mice, dogs, horses, and mules. Despite these successes, the birth of a viable SCNT primate clone has not been achieved. In 2001 a team of scientists cloned a rhesus monkey through a process called embryonic cell nuclear transfer, which is similar to SCNT except that it uses DNA from an undifferentiated embryo. In 2007 macaque monkey embryos were cloned by SCNT; however, these clones lived only to the blastocyst stage of embryonic

development. Likewise, SCNT has been carried out with very limited success in humans.

Reproductive Cloning

Reproductive cloning involves the implantation of a cloned embryo into a real or an artificial uterus. The embryo develops into a fetus that is then carried to term. Reproductive cloning experiments were performed for more than 40 years through the process of embryo splitting, in which a single early-stage two-cell embryo is manually divided into two individual cells and then grows as two identical embryos. Reproductive cloning techniques underwent significant change in the 1990s, following the birth of Dolly, who was generated through the process of SCNT. This process entails the removal of the entire nucleus from a somatic cell of an organism, followed by insertion of the nucleus into an egg cell that has had its own nucleus removed (enucleation). Once the somatic nucleus is inside the egg, the egg is stimulated with a mild electrical current and begins dividing. Thus, a cloned embryo, essentially an embryo of an identical twin of the original organism, is created. The SCNT process has undergone significant refinement since the 1990s, and procedures have been developed to prevent damage to eggs during nuclear extraction and somatic cell nuclear insertion. For example, the use of polarized light to visualize an egg cell's nucleus facilitates the extraction of the nucleus from the egg, resulting in a healthy, viable egg and thereby increasing the success rate of SCNT.

Reproductive cloning using SCNT is considered very harmful since fetuses of the cloned embryos rarely survive gestation and usually are born with birth defects. Wilmut's team of scientists needed 227 tries to create Dolly. Likewise, attempts to produce a macaque monkey clone

in 2007 involved 100 cloned embryos, implanted into 50 female macaque monkeys, none of which gave rise to a viable pregnancy. In January 2008, scientists at Stemagen, a stem cell research and development company in California, announced that they had cloned five human embryos by means of SCNT and that the embryos had matured to the stage at which they could have been implanted in a womb. However, the scientists destroyed the embryos after five days, presumably because of ethical reasons and the need to perform molecular analyses on the embryos.

Somatic Cell Nuclear Transfer

SCNT is an extraordinarily delicate technique, and the cellular and genetic processes that underlie its success are only just beginning to be understood. The transfer of a somatic cell nucleus to the cytoplasm of an enucleated egg represents a dramatic change for the nucleus because once inside the egg, the somatic nucleus is reprogrammed by egg cytoplasmic factors to become a zygote (fertilized egg) nucleus.

In most SCNT experiments, the egg is allowed to develop to the blastocyst stage, at which point a culture of embryonic stem cells (ESCs) can be created from the inner cell mass of the blastocyst. Although mouse and monkey ESCs have been made using SCNT, human ESCs, which have valuable applications in both medicine and research, have not been generated with this technique.

Today, the most practical application of SCNT is in the reproductive cloning of farm animals that have exceptional qualities, such as the ability to produce large quantities of milk. Reproductive cloning is accomplished by implanting an SCNT-derived blastocyst into the uterus of a surrogate mother, in which the embryo develops into a fetus carried to term. The technique also could be used

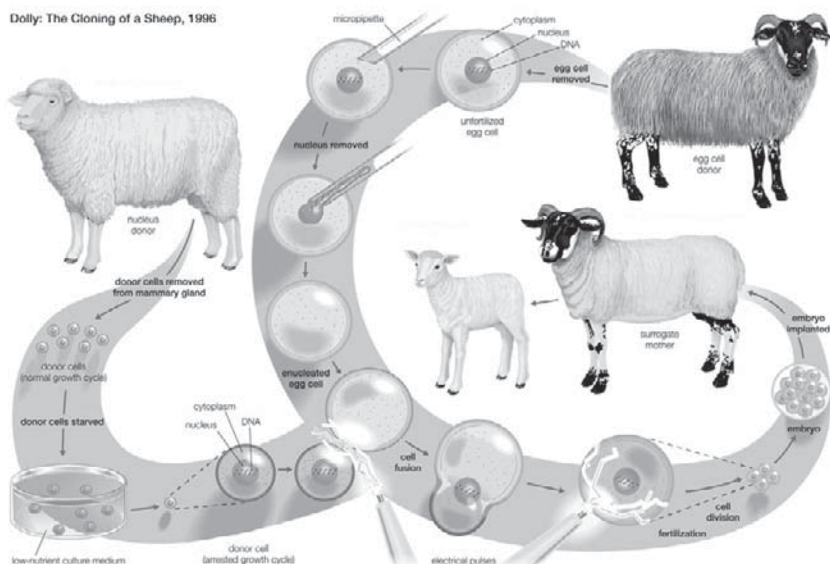
to resurrect extinct species; for example, cells collected from a frozen woolly mammoth could be used as nuclear donors for enucleated elephant eggs. Proof of principle for such "resurrection" was provided by an experiment in which mice were cloned using somatic cell nuclei derived from a mouse that was frozen for more than 15 years.

Dolly

Dolly, a female Finn Dorset sheep that lived from 1996 to 2003, was the first successfully cloned mammal and was produced by British developmental biologist Ian Wilmut and colleagues of the Roslin Institute, near Edinburgh. The announcement in February 1997 of the world's first clone of an adult animal was a milestone in science, dispelling decades of presumption that adult mammals could not be cloned and igniting a debate concerning the many possible uses and misuses of mammalian cloning technology.

In the 1990s the concept of mammalian clones, even humans, was not completely new. Naturally occurring genetic clones, or individuals genetically identical to one another, had long been recognized in the form of monozygotic (identical) twins. Unlike Dolly, however, such clones are derived, as their scientific name indicates, from a single zygote, or fertilized egg. Moreover, clones had been generated previously in the laboratory, but only from embryonic cells or from the adult cells of plants and "lower" animals such as frogs.

Decades of attempts to clone mammals from existing adults had met with repeated failure, which led to the presumption that something special and irreversible must happen to the DNA of mammalian cells during the animal's development. Indeed, until 1997 it had been generally accepted dogma that adult mammalian cells are no longer genetically totipotent, or capable of giving rise to all of the



Dolly the sheep was successfully cloned in 1996 by fusing the nucleus from a mammary-gland cell of a Finn Dorset ewe into an enucleated egg cell taken from a Scottish Blackface ewe. Carried to term in the womb of another Scottish Blackface ewe, Dolly was a genetic copy of the Finn Dorset ewe. Encyclopædia Britannica, Inc.

different cell and tissue types (e.g., liver, brain, and bone) required for making a complete and viable animal. It was presumed that somatic-cell differentiation, the process by which a single fertilized egg is converted into all of the different cell types found in an adult, involved some irreversible, likely epigenetic step. That Dolly remained alive and well long after her birth—that she had a functional heart, liver, brain, and other organs, all derived genetically from the nuclear DNA of an adult mammary-gland cell—proved otherwise. At the very minimum, the specific tissue from which Dolly's nuclear DNA was derived must have been totipotent. By extension, it was reasonable to suggest that the nuclear DNA of other adult tissues also remains totipotent. With the successful creation of Dolly, this speculation became a testable hypothesis.

Dolly did not spring from the laboratory bench fully formed but developed to term normally in the uterus of a Scottish Blackface ewe. Although Dolly's nuclear genome was derived from a mammary-gland cell taken from an adult Finn Dorset ewe, that nucleus had to be fused by electrical pulses with an unfertilized egg cell, the nucleus of which had been removed. The "host" egg cytoplasm was taken from a Scottish Blackface ewe, and later another Scottish Blackface ewe served as the surrogate mother. Furthermore, in order for the mammary gland cell nucleus and genomic DNA to be accepted and functional within the context of the host egg, the donor cell first had to be induced to abandon the normal cycle of growth and division and enter a quiescent stage. To do this, researchers deliberately withheld nutrients from the cells. The importance of this step had been determined experimentally, and although a number of hypotheses had been raised to explain its necessity, which, if any, of them was correct remained unclear. Nevertheless, starting with a collection of donor cell nuclei and host egg cytoplasms, a number of fused couplets successfully formed embryos; these were transferred to surrogate ewes. Of 13 recipient ewes, one became pregnant, and 148 days later, which is essentially normal gestation for a sheep, Dolly was born.

On Feb. 14, 2003, Dolly was euthanized by veterinarians after being found to suffer from progressive lung disease. Her body was preserved and displayed at the National Museum of Scotland in Edinburgh.

Therapeutic Cloning

Therapeutic cloning is intended to use cloned embryos for the purpose of extracting stem cells from them, without ever implanting the embryos in a womb. Therapeutic cloning enables the cultivation of stem cells that are

genetically identical to a patient. The stem cells could be stimulated to differentiate into any of the more than 200 cell types in the human body. The differentiated cells then could be transplanted into the patient to replace diseased or damaged cells without the risk of rejection by the immune system. These cells could be used to treat a variety of conditions, including Alzheimer disease, Parkinson disease, diabetes mellitus, stroke, and spinal cord injury. In addition, stem cells could be used for in vitro (laboratory) studies of normal and abnormal embryo development or for testing drugs to see if they are toxic or cause birth defects.

Although stem cells have been derived from the cloned embryos of animals such as mice, the generation of stem cells from cloned primate embryos has proved exceptionally difficult. For example, in 2007 stem cells successfully derived from cloned macaque embryos were able to differentiate into mature heart cells and brain neurons. However, the experiment started with 304 egg cells and resulted in the development of only two lines of stem cells, one of which had an abnormal Y chromosome. Likewise, the production of stem cells from human embryos has been fraught with the challenge of maintaining embryo viability.

In 2001 scientists at Advanced Cell Technology, a research company in Massachusetts, successfully transferred DNA from human cumulus cells, which are cells that cling to and nourish human eggs, into eight enucleated eggs. Of these eight eggs, three developed into early-stage embryos (containing four to six cells); however, the embryos survived only long enough to divide once or twice. In 2004 South Korean researcher Hwang Woo Suk claimed to have cloned human embryos using SCNT and to have extracted stem cells from the embryos. However, this later proved

to be a fraud; Hwang had fabricated evidence and had actually carried out the process of parthenogenesis, in which an unfertilized egg begins to divide with only half a genome. The following year, a team of researchers from the University of Newcastle upon Tyne was able to grow a cloned human embryo to the 100-cell blastocyst stage using DNA from embryonic stem cells. Although they did not generate a line of stem cells from the blastocyst, this was a major step forward in therapeutic cloning research. However, because embryonic stem cells can only be isolated and cloned from a recently fertilized embryo and not a grown patient, the procedure was considered to have only narrow clinical application.

Progress in research on therapeutic cloning in humans has been slow relative to the advances made in reproductive cloning in animals. This is primarily due to the technical challenges and ethical controversy arising from the procuring of human eggs solely for research purposes and the lack of understanding of how somatic cell DNA is reprogrammed into an undifferentiated embryonic state during SCNT. In addition, the development of induced pluripotent stem cells, which are derived from somatic cells that have been reprogrammed to an embryonic state through the introduction of specific genetic factors into the cell nuclei, has challenged the use of cloning methods and of human eggs.

Ethical Controversies of Cloning

Human reproductive cloning remains universally condemned, primarily for the psychological, social, and physiological risks associated with cloning. A cloned embryo intended for implantation into a womb requires thorough molecular testing to fully determine whether an

embryo is healthy and whether the cloning process is complete. In addition, as demonstrated by 100 failed attempts to generate a cloned macaque in 2007, a viable pregnancy is not guaranteed. Because the risks associated with reproductive cloning in humans introduce a very high likelihood of loss of life, the process is considered unethical. There are other philosophical issues that also have been raised concerning the nature of reproduction and human identity that reproductive cloning might violate. Concerns about eugenics, the once popular notion that the human species could be improved through the selection of individuals possessing desired traits, also have surfaced, since cloning could be used to breed "better" humans, thus violating principles of human dignity, freedom, and equality.

There also exists controversy over the ethics of therapeutic and research cloning. Some individuals and groups have an objection to therapeutic cloning because it is considered the manufacture and destruction of a human life, even though that life has not developed past the embryonic stage. Those who are opposed to therapeutic cloning believe that the technique supports and encourages acceptance of the idea that human life can be created and expended for any purpose. However, those who support therapeutic cloning believe that there is a moral imperative to heal the sick and to seek greater scientific knowledge. Many of these supporters believe that therapeutic and research cloning should be not only allowed but also publicly funded, similar to other types of disease and therapeutics research. Most supporters also argue that the embryo demands special moral consideration, requiring regulation and oversight by funding agencies. In addition, it is important to many philosophers and policy makers that women and couples not be exploited for the purpose of obtaining their embryos or eggs.

There are laws and international conventions that attempt to uphold certain ethical principles and regulations concerning cloning. In 2005 the United Nations passed a nonbinding Declaration on Human Cloning that calls upon member states "to adopt all measures necessary to prohibit all forms of human cloning inasmuch as they are incompatible with human dignity and the protection of human life." This does provide leeway for member countries to pursue therapeutic cloning. The United Kingdom, through its Human Fertilisation and Embryology Authority, issues licenses for creating human embryonic stem cells through nuclear transfer. These licenses ensure that human embryos are cloned for legitimate therapeutic and research purposes aimed at obtaining scientific knowledge about disease and human development. The licenses require the destruction of embryos by the 14th day of development, since this is when embryos begin to develop the primitive streak, the first indicator of an organism's nervous system.

The United States federal government has not passed any laws regarding human cloning due to disagreement within the legislative branch about whether to ban all cloning or to ban only reproductive cloning. The Dickey-Wicker amendment, attached to U.S. appropriations bills since 1995, has prevented the use of federal dollars to fund the harm or destruction of human embryos for research. It is presumed that nuclear transfer and any other form of cloning is subject to this restriction.

CHAPTER 6
Genetics Applied to Plants and Animals

Agriculture and animal husbandry apply genetic techniques to improve plants and animals, many of which have important commercial applications. Plant geneticists use special techniques to produce new species, such as hybrid grains (e.g., produced by crossing wheat and rye), and plants resistant to destruction by insect and fungal pests. Plant breeders use the techniques of budding and grafting to maintain desirable gene combinations originally obtained from crossbreeding. Transgenic plant cells can be made into plants by growing the cells in the presence of special hormones. In addition, the use of the chemical compound colchicine, which causes chromosomes to double in number, has resulted in many new varieties of fruits, vegetables, and flowers. Many transgenic lines of crop plants are commercially advantageous and have been introduced into the market.

Similar to plant breeders, animal breeders rely on breeding analysis and transgenic modification using recombinant DNA techniques. Animal breeders also use artificial insemination to propagate the genes of prize bulls. Prize cows can transmit their genes to hundreds of offspring by hormone treatment, which stimulates the release of many eggs that are collected, fertilized, and transplanted into foster mothers. Furthermore, reproductive cloning of mammals has enabled multiple identical copies of animals with certain desirable traits to be produced.

PLANT BREEDING

Plant breeding is the application of genetic principles to produce plants that are more useful to humans. This is

accomplished by selecting plants found to be economically or aesthetically desirable, first by controlling the mating of selected individuals, and then by selecting certain individuals among the progeny. Such processes, repeated over many generations, can change the hereditary makeup and value of a plant population far beyond the natural limits of previously existing populations.

Plant breeding is an ancient activity, dating to the very beginnings of agriculture. Probably soon after the earliest domestications of cereal grains, humans began to recognize degrees of excellence among the plants in their fields and saved seed from the best for planting new crops. Such tentative selective methods were the forerunners of early plant-breeding procedures. The results of these early trials were conspicuous. Most present-day varieties are so modified from their wild progenitors that they are unable to survive in nature. Indeed, in some cases, the cultivated forms are so strikingly different from existing wild relatives that it is difficult even to identify their ancestors. These remarkable transformations were accomplished by early plant breeders in a very short time from an evolutionary point of view, and the rate of change was probably greater than for any other evolutionary event.

Scientific plant breeding is a relatively recent development. For many centuries, the role of pollination and fertilization in the process of reproduction was not widely appreciated, and it was not until the early part of the 20th century that the laws of genetic inheritance were recognized and a beginning was made toward applying them to the improvement of plants. One of the major facts that has emerged during the short history of scientific breeding is that an enormous wealth of genetic variability exists in the plants of the world and that only a start has been made in tapping its potential.

Goals

The plant breeder usually has in mind an ideal plant that combines a maximum number of desirable characteristics. These characteristics may include resistance to diseases and insects; tolerance to heat and frost; appropriate size, shape, and time to maturity; and many other general and specific traits that contribute to improved adaptation to the environment, ease in growing and handling, greater yield, and better quality. The breeder of fancy show plants must also consider aesthetic appeal. Thus the breeder can rarely focus attention on any one characteristic but must take into account the manifold traits that make the plant more useful in fulfilling the purpose for which it is grown.

Increase of Yield

One of the aims of virtually every breeding project is to increase yield. This can often be brought about by selecting obvious morphological variants. One example is the selection of dwarf, early maturing varieties of rice. These dwarf varieties are sturdy and give a greater yield of grain. Furthermore, their early maturity frees the land quickly, often allowing an additional planting of rice or other crop the same year.

Another way of increasing yield is to develop varieties resistant to diseases and insects. In many cases the development of resistant varieties has been the only practical method of pest control. Perhaps the most important feature of resistant varieties is the stabilizing effect they have on production and hence on steady food supplies. Varieties tolerant to drought, heat, or cold provide the same benefit.

Modifications of Range and Constitution

Another common goal of plant breeding is to extend the area of production of a crop species. A good example is

the modification of grain sorghum since its introduction to the United States in the 19th century. Of tropical origin, grain sorghum was originally confined to the southern Plains area and the Southwest of the United States, but earlier maturing varieties were developed, and today it is an important crop even in regions with characteristically low temperatures and short growing seasons, such as Wisconsin, Minnesota, and North Dakota.

Development of crop varieties suitable for mechanized agriculture is a major goal of modern plant breeding. Uniformity of plant characters is very important in mechanized agriculture because field operations are much easier when the individuals of a variety are similar in time of germination, growth rate, size of fruit, and so on. Uniformity in maturity is, of course, essential when crops such as tomatoes and peas are harvested mechanically.

The nutritional quality of plants can be greatly improved by breeding. For example, it is possible to breed varieties of corn (maize) much higher in lysine than previously existing varieties. Breeding high-lysine maize varieties for those areas of the world where maize is the major source of this nutritionally essential amino acid has become a major goal in plant breeding.

In breeding ornamentals, attention is paid to factors such as longer blooming periods, improved keeping qualities of flowers, general thriftiness, and other features that contribute to usefulness and aesthetic appeal. Novelty itself is often a virtue in ornamentals, and the spectacular, even the bizarre, is often sought.

Evaluation of Plants

The appraisal of the value of plants so that the breeder can decide which individuals should be discarded and which

allowed to produce the next generation is a much more difficult task with some traits than with others.

Qualitative Characters

The easiest characters, or traits, to deal with are those involving discontinuous, or qualitative, differences that are governed by one or a few major genes. Many such inherited differences exist, and they frequently have profound effects on plant value and utilization. Examples are starchy versus sugary kernels (characteristic of field and sweet corn, respectively) and determinant versus indeterminant habit of growth in green beans (determinant varieties are adapted to mechanical harvesting). Such differences can be seen easily and evaluated quickly, and the expression of the traits remains the same regardless of the environment in which the plant grows. Traits of this type are termed highly heritable.

Quantitative Characters

In other cases, however, plant traits grade gradually from one extreme to another in a continuous series, and classification into discrete classes is not possible. Such variability is termed quantitative. Many traits of economic importance are of this type; e.g., height, cold and drought tolerance, time to maturity, and, in particular, yield. These traits are governed by many genes, each having a small effect. Although the distinction between the two types of traits is not absolute, it is nevertheless convenient to designate qualitative characters as those involving discrete differences and quantitative characters as those involving a graded series.

Quantitative characters are much more difficult for the breeder to control, for three main reasons: (1) the sheer numbers of the genes involved make hereditary

change slow and difficult to assess; (2) the variations of the traits involved are generally detectable only through measurement and exacting statistical analyses; and (3) most of the variations are due to the environment rather than to genetic endowment; for example, the heritability of certain traits is less than 5 percent, meaning that 5 percent of the observed variation is caused by genes and 95 percent is caused by environmental influences.

It follows that carefully designed experiments are required to distinguish plants that are superior because they carry desirable genes from those that are superior because they happen to grow in a favourable site.

Methods of Plant Breeding

Mating Systems

Plant mating systems devolve about the type of pollination, or transferal of pollen from flower to flower. A flower is self-pollinated (a "selfer") if pollen is transferred to it from any flower of the same plant and cross-pollinated (an "outcrosser" or "outbreeder") if the pollen comes from a flower on a different plant. About half of the more important cultivated plants are naturally cross-pollinated, and their reproductive systems include various devices that encourage cross-pollination; e.g., protandry (pollen shed before the ovules are mature, as in the carrot and walnut), dioecy (stamens and pistils borne on different plants, as in the date palm, asparagus, and hops), and genetically determined self-incompatibility (inability of pollen to grow on the stigma of the same plant, as in white clover, cabbage, and many other species).

Other plant species, including a high proportion of the most important cultivated plants such as wheat, barley,

rice, peas, beans, and tomatoes, are predominantly self-pollinating. There are relatively few reproductive mechanisms that promote self-pollination; the most positive of which is failure of the flowers to open (cleistogamy), as in certain violets. In barley, wheat, and lettuce the pollen is shed before or just as the flowers open; and in the tomato pollination follows opening of the flower, but the stamens form a cone around the stigma. In such species there is always a risk of unwanted cross-pollination.

In controlled breeding procedures it is imperative that pollen from the desired male parent, and no other pollen, reaches the stigma of the female parent. When stamens and pistils occur in the same flower, the anthers must be removed from flowers selected as females before pollen is shed. This is usually done with forceps or scissors. Protection must also be provided from “foreign” pollen. The most common method is to cover the flower with a plastic or paper bag. When the stigma of the female parent becomes receptive, pollen from the desired male parent is transferred to it, often by breaking an anther over the stigma, and the protective bag is replaced. The production of certain hybrids is, therefore, tedious and expensive because it often requires a series of delicate, exacting, and properly timed hand operations. When male and female parts occur in separate flowers, as in corn (maize), controlled breeding is easier.

A cross-pollinated plant, which has two parents, each of which is likely to differ in many genes, produces a diverse population of plants hybrid (heterozygous) for many traits. A self-pollinated plant, which has only one parent, produces a more uniform population of plants pure breeding (homozygous) for many traits. Thus, in contrast to outbreeders, self-breeders are likely to be highly homozygous and hence true breeding for a specified trait.

Breeding Self-Pollinated Species

The breeding methods that have proved successful with self-pollinated species are (1) mass selection; (2) pure-line selection; (3) hybridization, with the segregating generations handled by the pedigree method, the bulk method, or the backcross method; and (4) development of hybrid varieties.

Mass Selection

In mass selection, seeds are collected from (usually a few dozen to a few hundred) desirable appearing individuals in a population, and the next generation is sown from the stock of mixed seed. This procedure, sometimes referred to as phenotypic selection, is based on how each individual looks. Mass selection has been used widely to improve old "land" varieties, varieties that have been passed down from one generation of farmers to the next over long periods.

An alternative approach that has no doubt been practiced for thousands of years is simply to eliminate undesirable types by destroying them in the field. The results are similar whether superior plants are saved or inferior plants are eliminated: seeds of the better plants become the planting stock for the next season.

A modern refinement of mass selection is to harvest the best plants separately and to grow and compare their progenies. The poorer progenies are destroyed and the seeds of the remainder are harvested. It should be noted that selection is based not solely on the appearance of the parent plants but also on the appearance and performance of their progeny. Progeny selection is usually more effective than phenotypic selection when dealing with quantitative characters of low heritability. It should be

noted, however, that progeny testing requires an extra generation; hence gain per cycle of selection must be double that of simple phenotypic selection to achieve the same rate of gain per unit time.

Mass selection, with or without progeny testing, is perhaps the simplest and least expensive of plant-breeding procedures. It finds wide use in the breeding of certain forage species, which are not important enough economically to justify more detailed attention.

Pure-Line Selection

Pure-line selection generally involves three more or less distinct steps: (1) numerous superior appearing plants are selected from a genetically variable population; (2) progenies of the individual plant selections are grown and evaluated by simple observation, frequently over a period of several years; and (3) when selection can no longer be made on the basis of observation alone, extensive trials are undertaken, involving careful measurements to determine whether the remaining selections are superior in yielding ability and other aspects of performance. Any progeny superior to an existing variety is then released as a new "pure-line" variety. Much of the success of this method during the early 1900s depended on the existence of genetically variable land varieties that were waiting to be exploited. They provided a rich source of superior pure-line varieties, some of which are still represented among commercial varieties. In recent years the pure-line method has decreased in importance in the breeding of major cultivated species; however, the method is still widely used with the less important species that have not yet been heavily selected.

A variation of the pure-line selection method that dates back centuries is the selection of single-chance

variants, mutations or "sports" in the original variety. A very large number of varieties that differ from the original strain in characteristics such as colour, lack of thorns or barbs, dwarfness, and disease resistance have originated in this fashion.

Hybridization

During the 20th century planned hybridization between carefully selected parent plants dominated the breeding of self-pollinated species. The object of hybridization is to combine desirable genes found in two or more different varieties and to produce pure-breeding progeny superior in many respects to the parental types.

Efficiently managing the enormous numbers of genotypes that occur in the generations following hybridization is a challenging facet of plant breeding. As an example of the power of hybridization in creating variability, a cross between hypothetical wheat varieties differing by only 21 genes is capable of producing more than 10,000,000,000 different genotypes in the second generation. At spacings normally used by farmers, more than 50,000,000 acres would be required to grow a population large enough to permit every genotype to occur in its expected frequency. While the great majority of these second generation genotypes are hybrid (heterozygous) for one or more traits, it is statistically possible that 2,097,152 different pure-breeding (homozygous) genotypes can occur, each potentially a new pure-line variety. These numbers illustrate the importance of efficient techniques in managing hybrid populations, for which purpose the pedigree procedure is most widely used.

Pedigree breeding starts with the crossing of two genotypes, each of which have one or more desirable characters lacked by the other. If the two original parents

do not provide all of the desired characters, a third parent can be included by crossing it to one of the hybrid progeny of the first generation (F_1). In the pedigree method superior types are selected in successive generations, and a record is maintained of parent-progeny relationships.

The F_2 generation (progeny of the crossing of two F_1 individuals) affords the first opportunity for selection in pedigree programs. In this generation the emphasis is on the elimination of individuals carrying undesirable major genes. In the succeeding generations the hybrid condition gives way to pure breeding as a result of natural self-pollination, and families derived from different F_2 plants begin to display their unique character. Usually one or two superior plants are selected within each superior family in these generations. By the F_5 generation the pure-breeding condition (homozygosity) is extensive, and emphasis shifts almost entirely to selection between families. The pedigree record is useful in making these eliminations. At this stage each selected family is usually harvested in mass to obtain the larger amounts of seed needed to evaluate families for quantitative characters. This evaluation is usually carried out in plots grown under conditions that simulate commercial planting practice as closely as possible. When the number of families has been reduced to manageable proportions by visual selection, usually by the F_7 or F_8 generation, precise evaluation for performance and quality begins. The final evaluation of promising strains involves (1) observation, usually in a number of years and locations, to detect weaknesses that may not have appeared previously; (2) precise yield testing; and (3) quality testing. Many plant breeders test for five years at five representative locations before releasing a new variety for commercial production.

The bulk-population method of breeding differs from the pedigree method primarily in the handling of

generations following hybridization. The F_2 generation is sown at normal commercial planting rates in a large plot. At maturity the crop is harvested en masse, and the seeds are used to establish the next generation in a similar plot. No record of ancestry is kept. During the period of bulk propagation natural selection tends to eliminate plants having poor survival value. Two types of artificial selection also are often applied: (1) destruction of plants that carry undesirable major genes and (2) mass techniques such as harvesting when only part of the seeds are mature to select for early maturing plants or the use of screens to select for increased seed size. Single plant selections are then made and evaluated in the same way as in the pedigree method of breeding. The chief advantage of the bulk population method is that it allows the breeder to handle very large numbers of individuals inexpensively.

Often an outstanding variety can be improved by transferring to it some specific desirable character that it lacks. This can be accomplished by first crossing a plant of the superior variety to a plant of the donor variety, which carries the trait in question, and then mating the progeny back to a plant having the genotype of the superior parent. This process is called backcrossing. After five or six backcrosses the progeny will be hybrid for the character being transferred but like the superior parent for all other genes. Selfing (self-pollinating) the last backcross generation, coupled with selection, will give some progeny pure breeding for the genes being transferred. The advantages of the backcross method are its rapidity, the small number of plants required, and the predictability of the outcome. A serious disadvantage is that the procedure diminishes the occurrence of chance combinations of genes, which sometimes leads to striking improvements in performance.

Hybrid Varieties

The development of hybrid varieties differs from hybridization in that no attempt is made to produce a pure-breeding population; only the F_1 hybrid plants are sought. The F_1 hybrid of crosses between different genotypes is often much more vigorous than its parents. This hybrid vigour, or heterosis, can be manifested in many ways, including increased rate of growth, greater uniformity, earlier flowering, and increased yield, the last being of greatest importance in agriculture.

By far the greatest development of hybrid varieties has been in corn (maize), primarily because its male flowers (tassels) and female flowers (incipient ears) are separate and easy to handle, thus proving economical for the production of hybrid seed. The production of hand-produced F_1 hybrid seed of other plants, including ornamental flowers, has been economical only because greenhouse growers and home gardeners have been willing to pay high prices for hybrid seed.

A built-in cellular system of pollination control has made hybrid varieties possible in a wide range of plants, including many that are self-pollinating, such as sorghums. This system, called cytoplasmic male sterility, or cytosterility, prevents normal maturation or function of the male sex organs (stamens) and results in defective pollen or none at all. It obviates the need for removing the stamens either by hand or by machine. Cytosterility depends on the interaction between male sterile genes ($R + r$) and factors found in the cytoplasm of the female sex cell. The genes are derived from each parent in the normal Mendelian fashion, but the cytoplasm (and its factors) is provided by the egg only; therefore, the inheritance of cytosterility is determined by the female parent. All plants

with fertile cytoplasm produce viable pollen, as do plants with sterile cytoplasm but at least one *R* gene; plants with sterile cytoplasm and two *r* genes are male sterile (produce defective pollen).

The production of F_1 hybrid seed between two strains is accomplished by interplanting a sterile version of one strain (say A) in an isolated field with a fertile version of another strain (B). Since strain A produces no viable pollen, it will be pollinated by strain B, and all seeds produced on strain A plants must therefore be F_1 hybrids between the strains. The F_1 hybrid seeds are then planted to produce the commercial crop. Much of the breeder's work in this process is in developing the pure-breeding sterile and fertile strains to begin the hybrid seed production.

Breeding Cross-Pollinated Species

The most important methods of breeding cross-pollinated species are (1) mass selection; (2) development of hybrid varieties; and (3) development of synthetic varieties. Since cross-pollinated species are naturally hybrid (heterozygous) for many traits and lose vigour as they become purebred (homozygous), a goal of each of these breeding methods is to preserve or restore heterozygosity.

Mass Selection

Mass selection in cross-pollinated species takes the same form as in self-pollinated species; i.e., a large number of superior appearing plants are selected and harvested in bulk and the seed used to produce the next generation. Mass selection has proved to be very effective in improving qualitative characters, and, applied over many generations, it is also capable of improving quantitative characters, including yield, despite the low heritability of such characters. Mass selection has long been a major

method of breeding cross-pollinated species, especially in the economically less important species.

Hybrid Varieties

The outstanding example of the exploitation of hybrid vigour through the use of F_1 hybrid varieties has been with corn (maize). The production of a hybrid corn variety involves three steps: (1) the selection of superior plants; (2) selfing for several generations to produce a series of inbred lines, which although different from each other are each pure-breeding and highly uniform; and (3) crossing selected inbred lines. During the inbreeding process the vigour of the lines decreases drastically, usually to less than half that of field-pollinated varieties. Vigour is restored, however, when any two unrelated inbred lines are crossed, and in some cases the F_1 hybrids between inbred lines are much superior to open-pollinated varieties. An important consequence of the homozygosity of the inbred lines is that the hybrid between any two inbreds will always be the same. Once the inbreds that give the best hybrids have been identified, any desired amount of hybrid seed can be produced.

Pollination in corn (maize) is by wind, which blows pollen from the tassels to the styles (silks) that protrude from the tops of the ears. Thus controlled cross-pollination on a field scale can be accomplished economically by interplanting two or three rows of the seed parent inbred with one row of the pollinator inbred and detasselling the former before it sheds pollen. In practice most hybrid corn is produced from "double crosses," in which four inbred lines are first crossed in pairs (A × B and C × D) and then the two F_1 hybrids are crossed again (A × B) × (C × D). The double-cross procedure has the advantage that the commercial F_1 seed is produced on the highly productive

single cross A × B rather than on a poor-yielding inbred, thus reducing seed costs. Cytoplasmic male sterility may be used to eliminate detasselling of the seed parent, thus providing further economies in producing hybrid seed.

Much of the hybrid vigour exhibited by F_1 hybrid varieties is lost in the next generation. Consequently, seed from hybrid varieties is not used for planting stock but the farmer purchases new seed each year from seed companies.

Perhaps no other development in the biological sciences has had greater impact on increasing the quantity of food supplies available to the world's population than has the development of hybrid corn (maize). Hybrid varieties in other crops, made possible through the use of male sterility, have also been dramatically successful.

Synthetic Varieties

A synthetic variety is developed by intercrossing a number of genotypes of known superior combining ability—i.e., genotypes that are known to give superior hybrid performance when crossed in all combinations. (By contrast, a variety developed by mass selection is made up of genotypes bulked together without having undergone preliminary testing to determine their performance in hybrid combination.) Synthetic varieties are known for their hybrid vigour and for their ability to produce usable seed for succeeding seasons. Because of these advantages, synthetic varieties have become increasingly favoured in the growing of many species, such as the forage crops, in which expense prohibits the development or use of hybrid varieties.

Distribution and Maintenance of New Varieties

The benefits of superior new varieties obviously cannot be realized until sufficient seed has been produced to permit

commercial production. Although the primary function of the plant breeder is to develop new varieties, the breeder usually also carries out an initial small-scale seed increase. Seed thus produced is called breeders seed. The next stage is the multiplication of breeders seed to produce foundation seed. Production of foundation seed is usually carried out by seed associations or institutes, whose work is regulated by government agencies. The third step is the production of certified seed, the progeny of foundation seed, produced on a large scale by specialized seed growers for general sale to farmers and gardeners. Certified seed must be produced and handled in such a way as to meet the standards set by the certifying agency (usually a seed association). Seed associations are also usually responsible for maintaining the purity of new varieties once they have been released for commercial production.

The distribution of new varieties developed by commercial plant-breeding companies is often through seed associations, but many reputable companies market their products without following the official certification process. In some countries, particularly in Europe, new varieties can be patented for periods up to 15 years or more, during which time the breeder has an exclusive right to reproduce and sell the variety.

ANIMAL BREEDING

Animal breeding entails the controlled propagation of domestic animals in order to improve desirable qualities. Humanity has been modifying domesticated animals to better suit human needs for centuries. Selective breeding involves using knowledge from several branches of science, including genetics, statistics, reproductive physiology, and computer science.

Red Poll cow and calf. © J.C. Allen and Son

Breeding and Variation

English agriculturist Robert Bakewell was a very successful breeder of commercial livestock in the 18th century. His work was based on the traditional method of visual appraisal of the animals that he selected. Although he did not write about his methods, it is recorded that he traveled extensively by horseback and collected sheep and cattle that he considered useful. It is thought that he made wide outcrosses of diverse breeds, and then practiced inbreeding with the intent of fixing desirable characteristics in the crossbred animals. He was also the first to systematically let his animals for stud. For these reasons he is generally recognized as the first scientific breeder.

In animal breeding, a population is a group of interbreeding individuals—i.e., a breed or strain within a breed that is different in some aspects from other breeds or

strains. Typically, certain animals within a breed are designated as purebred. The essential difference between purebred and nonpurebred animals is that the genealogy of purebred animals has been carefully recorded, usually in a herd book, or studbook, kept by some sanctioning association. Purebred associations provide other services that are useful to their members to enhance their businesses.

Selective breeding utilizes the natural variations in traits that exist among members of any population. Breeding progress requires understanding the two sources of variation: genetics and environment. For some traits there is an interaction of genetics and the environment. Differences in the animals' environment, such as amount of feed, care, and even the weather, may have an impact on their growth, reproduction, and productivity. Such variations in performance because of the environment are not transmitted to the next generation. For

Thoroughbred stallion with dark bay coat. © Scott Smudsky

most traits measured in domestic animals, the environment has a larger impact on variation than do genetic differences. For example, only about 30 percent of the variation in milk production in dairy cattle can be attributed to genetic effects; the remainder of the variation is due to environmental effects. Thus, environmental factors must be considered and controlled in selecting breeding stock.

Genetic variation is necessary in order to make progress in breeding successive generations. Each gene, which is the basic unit of heredity, occupies a specific location, or locus, on a chromosome. Two or more genes may be associated with a specific locus and therefore with a specific trait. (Traits that can be observed directly, such as size, colour, shape, and so forth, make up an organism's phenotype.) These genes are known as alleles. If paired alleles are the same, the organism is called homozygous for that trait; if they are different, the organism is heterozygous. Typically, one of the alleles will be expressed to the exclusion of the other allele, in which case the two alleles are referred to as dominant and recessive, respectively. However, sometimes neither dominates, in which case the two alleles are called codominant.

Although no complete knowledge of the genetic makeup of any breed of livestock exists yet, genetic variations can be used for improving stock. Researchers partition total genetic variation into additive, dominance, and epistatic types of gene action. Additive variation is easiest to use in breeding because it is common and the effect of each allele at a locus just adds to the effect of other alleles at that same locus. Genetic gains made using additive genetic effects are permanent and cumulate from one generation to the next.

Although dominance variation is not more complex in theory, it is more difficult to control in practice because of

how one allele masks the effect of another. For example, let a indicate a locus, with a_1 and a_2 representing two possible alleles at that location. Then a_1a_1, a_1a_2 (which is identical to a_2a_1), and a_2a_2 are the three possible genotypes. If a_1 dominates a_2, the genotypes a_1a_2 and a_1a_1 cannot be outwardly distinguished. Thus, the inability to observe differences between a_1a_2 and a_1a_1 presents a major difficulty in using dominance variance in selective breeding.

Additive and dominance variations are caused by genes at one locus. Epistatic variation is caused by the joint effects of genes at two or more loci. There has been little deliberate use of this type of genetic variation in breeding because of the complex nature of identifying and controlling the relevant genes.

Breeding

Breeding Objectives

Breeding objectives can be discussed in terms of changing the genetic makeup of a population of animals, where population is defined as a recognized breed. Choice of

Black Angus bull. Henry Elder/Encyclopædia Britannica, Inc.

breeding goals and design of an effective breeding program is usually not an easy task. Complicating the implementation of a breeding program is the number of generations needed to reach the initial goals. Ultimately, breeding goals are dictated by market demand; however, it is not easy to predict what consumers will want several years in advance. Sometimes the marketplace demands a different product than was defined as desirable in the original breeding objective. When this happens, breeders have to adjust their program, which results in less-efficient selection than if the new breeding goal had been used from the beginning. For example, consumers want leaner beef that is tender. Thus, ranchers have changed their cattle-breeding programs to meet this new demand. These trends have gradually changed over the last few decades; for example, Angus cattle are particularly noted for the quality of beef produced. The use of ultrasound is now widespread in determining the fat and lean content of live animals, which will hasten the changing of carcass quality to meet consumer demands.

Additional complications arise from simultaneously trying to improve multiple traits and the difficulty of determining what part of the variation for each trait is under genetic control. In addition, some traits are genetically correlated, and this correlation may be positive or negative; that is, the traits may be complementary or antagonistic. Breeding methods depend on heritability and genetic correlations for desirable traits.

Heritability and Genetic Correlations in Breeding

Heritability is the proportion of the additive genetic variation to the total variation. Heritability is important because without genetic variation there can be no genetic change in the population. Alternatively, if heritability is high, genetic

change can be quite rapid, and simple means of selection are all that is needed. Using an increasing scale from 0 to 1, a heritability of 0.75 means that 75 percent of the total variance in a trait is controlled by additive gene action. With heritabilities this high, just the record of a single individual's traits can easily be used to create an effective breeding program.

Some general statements can be made about heritability, keeping in mind that exceptions exist. Traits related to fertility have low heritabilities. Examples include the average number of times that a cow must be bred before she conceives and the average number of pigs in a litter. Traits related to production have intermediate heritabilities. Examples include the amount of milk a cow produces, the rates of weight gain in steers and pigs, and the number of eggs laid by chickens. So-called quality traits tend to have higher heritabilities. Examples include the amount of fat a pig has over its back and the amount of protein in a cow's milk. The magnitude of heritability is one of the primary considerations in designing breeding programs.

Female pigs can have as many as 20 piglets in a litter. China holds the record for having the largest population of domestic pigs. The United States is second. © Corbis

Genetic correlation occurs when a single gene affects two traits. There may be many such genes that affect two or more traits. Genetic correlations can be positive or negative, which is indicated by assigning a number in the range from +1 to - 1, with 0 indicating no genetic correlation. A correlation of +1 means that the traits always occur together, while a correlation of - 1 means that having either trait always excludes having the other trait. Thus, the greater the displacement of the value from 0, the greater the correlation (positive or negative) between traits. The practical breeding consequence is that selection for one trait will pull along any positively correlated traits, even though there is no deliberate selection for them. For example, selecting for increased milk production also increases protein production. Another example is the selection for increased weight gain in broiler chickens, which also increases the fat content of the birds.

When traits have a negative genetic correlation, it is difficult to select simultaneously for both traits. For example, as milk production is increased in dairy cows through genetic selection, it is slightly more difficult for the high-producing cows to conceive. This negative correlation is partly due to the partitioning of the cows' nutrients between production and reproduction, with production being prioritized in early lactation. In the case of dairy cattle, milk production is on the order of 20,000 pounds per year and is increasing. This is a large metabolic demand, so nutrient demand is large to meet this need. Thus, selecting for improved fertility may result in a reduction in milk production or its rate of gain.

Selection

Methods of Selection

Types of selection are individual or mass selection, within and between family selection, sibling selection, and

progeny testing, with many variations. Within family selection uses the best individual from each family for breeding. Between family selection uses the whole family for selection. Mass selection uses records of only the candidates for selection. Mass selection is most effective when heritability is high and the trait is expressed early in life, in which case all that is required is observation and selection based on phenotypes. When mass selection is not appropriate, other methods of selection, which make use of relatives or progeny, can be used singularly or in combination. Modern technologies allow use of all these types of selection at the same time, which results in greater accuracy.

Elements Needed to Make Genetic Progress

Genetic gain per year (ΔG) depends on balancing several factors, as expressed in the equation

$$\Delta G = (A\sigma_g i)/I,$$

where A is the accuracy of selection, σ_g is the standard deviation of the additive genetic variation in the population, i is the selection intensity (proportion selected for further breeding), and I is the generation interval (age of breeding). The σg factor cannot be easily changed within a breed, though it can be changed by crossbreeding. The other quantities in the equation can be changed.

More complete pedigree records on candidates for breeding can increase the accuracy of selection, but waiting for candidates to reach full maturity in order to have better genetic data will increase the generation interval. Whether an increase in generation interval is justified by a more accurate selection process depends on individual circumstances. Selection intensity can also be increased, by narrowing the proportion of the population used in

breeding, but it should be done without increasing the generation interval. Because generation interval is the divisor in the genetic gain equation, anything that increases the generation interval has an unfavourable impact on genetic progress, all else being equal.

Evaluation of Animals

Methods of ranking animals for breeding purposes have changed as statistical and genetic knowledge has increased. Along with increases in breeding knowledge, advancements in computing have enabled breeders to quickly and easily process routine breeding evaluations, as well as to develop research needed to rank large populations of animals. Evaluating and ranking candidates for selection depends on equating their performance record to a statistical model. A performance record (y) can be expressed as

$$y = g + e + \varepsilon,$$

where g stands for genetic effects, e indicates known (categorized) environmental effects, and ε indicates random environmental effects.

The first task in estimating g is to statistically eliminate environmental effects, a process that involves setting up a system of equations to simultaneously solve for all of the genetic effects for the sires and cows. Information from relatives is included in g and increases the accuracy of evaluation of the candidates for selection. All relatives that are available can be incorporated in this type of evaluation. This model is called the animal model.

The animal model is used extensively in evaluating beef and dairy cattle, chickens, and pigs. To apply this model for evaluating large populations requires use of high-speed computers and extensive use of advanced mathematical techniques from numerical analysis. In

evaluating the dairy cattle in the United States, a system of equations with more than 25 million variables is needed.

Accuracy of Selection

In some cases the accuracy of selection for a trait can be measured using a calibrated tool or a scale. Thus, measurements of such traits can be replicated with high reliability. Alternatively, some traits are difficult to measure on an objective scale, in which case a well-designed subjective scoring method can be effective. An excellent example is hip dysplasia, a degenerative disease of the hip joints that is common in many large dog breeds. Apparently, hip dysplasia is not associated with a single allele, making its incidence very difficult to control. However, an index has been developed by radiologists that allows young dogs to be assigned a score indicating their likelihood of developing the disease as they age. In 1997 American animal geneticist E.A. Leighton reported that, in fewer than five

German shepherd (Alsatian). © Sally Anne Thompson/Animal Photography

generations of selection in a breeding experiment using these scores, the incidence of canine hip dysplasia in German shepherd dogs measured at 12 to 16 months of age had decreased from the breed average of 55 percent to 24 percent among the experimental population; in Labrador retrievers the incidence dropped from 30 to 10 percent.

Because close relatives share many genes, an examination of the relatives of a candidate for breeding can improve accuracy of selection. The more complete the genealogical record, or pedigree, the more effective the selection process. A pedigree is most useful when the heritability of a trait is relatively low, especially for traits that are expressed later in life or in only one sex.

Reproductive techniques can be used to increase the rate of genetic progress. In particular, for species that are mostly bred by artificial insemination, the best dams can be chosen and induced to superovulate, or release multiple eggs from their ovaries. These eggs are fertilized in the

Holstein-Friesian cow. Henry Elder/Encyclopædia Britannica, Inc.

uterus and then flushed out in a nonsurgical procedure that does not impair future conception of the donor female. Using this procedure, valuable females can produce more than one calf per year. Each embryo is implanted in a less-valuable host female to be carried through gestation. The sex of the embryos can be determined in utero at about 50 days of gestation. The normal gestation for Holstein-Friesian cattle is about 280 days, so this early determination of sex saves many days and allows the breeding program to be adjusted. In particular, the donor cow could be collected again, or another superior cow could be bred to produce males. Thus, these reproductive technologies reduce the generation interval and increase selection intensity by getting more than one male calf from superior females. Both superovulation and sex determination are commonly used procedures. Superovulation is also used when breeders want to increase the number of female calves from a valuable cow.

Progeny Testing

Progeny testing is used extensively in the beef and dairy cattle industry to aid in evaluating and selecting stock to be bred. Progeny testing is most useful when a high level of accuracy is needed for selecting a sire to be used extensively in artificial insemination. Progeny testing programs consist of choosing the best sires and dams in the population based on an animal model evaluation.

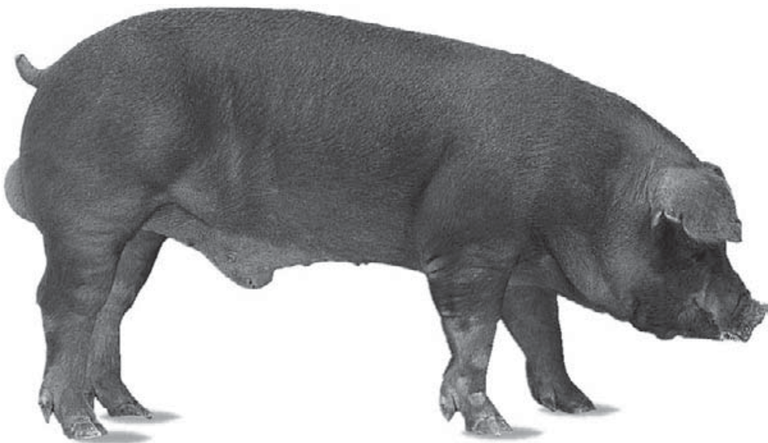
A description of progeny testing in dairy breeding provides a good example. The best 1 to 2 percent of the cows from the population are chosen as bull mothers, and the best progeny-tested bulls are chosen to produce another generation of sires. The parents are mated to complement any individual deficiencies. The accuracy of evaluation of bull mothers is typically about 40 percent, and of sires that produce young bulls the accuracy is more

than 80 percent. This is not as high as the industry wants for bulls to be used in artificial insemination. To reach greater accuracy, the next generation of sires is mated to enough cows in the population for each sire to produce 60 to 80 progeny. After the daughters of the young sires have a production record, the young sires are evaluated, and about the best 10 percent are used extensively to produce commercial cows. Some of the progeny-tested sires will have thousands of daughters before a superior sire is found to replace them. About 70 percent of dairy cattle are bred by artificial insemination, so these sires control the genetic destiny of dairy cattle. Consistently applying this selection procedure has been very successful.

Breeding Systems

Crossbreeding

Crossbreeding involves the mating of animals from two breeds. Normally, breeds are chosen that have complementary traits that will enhance the offsprings' economic value. An example is the crossbreeding of Yorkshire and



Duroc boar. Grant Heilman Photography

Duroc breeds of pigs. Yorkshires have acceptable rates of gain in muscle mass and produce large litters, and Durocs are very muscular and have other acceptable traits, so these breeds are complementary. Another example is Angus and Charolais beef cattle. Angus produce high-quality beef and Charolais are especially large, so crossbreeding produces an animal with acceptable quality and size.

The other consideration in crossbreeding is heterosis, or hybrid vigour, which is displayed when the offspring performance exceeds the average performance of the parent breeds. This is a common phenomenon in which increased size, growth rate, and fertility are displayed by crossbred offspring, especially when the breeds are more genetically dissimilar. Such increases generally do not increase in successive generations of crossbred stock, so purebred lines must be retained for crossbreeding and for continual improvement in the parent breeds. In general, there is more heterosis for traits with low heritability. In particular, heterosis is thought to be associated with the collective action of many genes having small effects individually but large effects cumulatively. Because of hybrid vigour, a high proportion of commercial pork and beef come from crossbred animals.

Inbreeding

Mating animals that are related causes inbreeding. Inbreeding is often described as "narrowing the genetic base" because the mating of related animals results in off-spring that have more genes in common. Inbreeding is used to concentrate desirable traits. Mild inbreeding has been used in some breeds of dogs and has been extensively used in laboratory mice and rats. For example, mice have been bred to be highly sensitive to compounds that might be detrimental or useful to humans. These mice are highly

inbred so that researchers can obtain the same response with replicated treatments.

Inbreeding is generally detrimental in domestic animals. Increased inbreeding is accompanied by reduced fertility, slower growth rates, greater susceptibility to disease, and higher mortality rates. As a result, producers try to avoid mating related animals. This is not always possible, though, when long-continued selection for the same traits is practiced within a small population, because parents of future generations are the best candidates from the last generation, and some inbreeding tends to accumulate. The rate of inbreeding can be reduced, but, if inbreeding depression becomes evident, some method of introducing more diverse genes will be needed. The most common method is some form of crossbreeding.

CHAPTER 7

Human Genetics

Human genetics is the study of the inheritance of characteristics by children from parents. Inheritance in humans does not differ in any fundamental way from that in other organisms.

The study of human heredity occupies a central position in genetics. Much of this interest stems from a basic desire to know more about the human species. An understanding of human heredity is also of critical importance in the prediction, diagnosis, and treatment of diseases that have a genetic component. The quest to determine the genetic basis of human health has given rise to the field of medical genetics. In general, medicine has given focus and purpose to human genetics, so that the terms *medical genetics* and *human genetics* are often considered synonymous.

THE HUMAN CHROMOSOMES

A new era in cytogenetics, the field of investigation concerned with studies of the chromosomes, began in 1956 with the discovery by Indonesian-born American geneticist Joe Hin Tjio and Swedish geneticist Albert Levan that human somatic cells contain 23 pairs of chromosomes. Since that time the field has advanced with amazing rapidity and has demonstrated that human chromosome aberrations rank as major causes of fetal death and of tragic human diseases, many of which are accompanied by varying forms of intellectual disability.

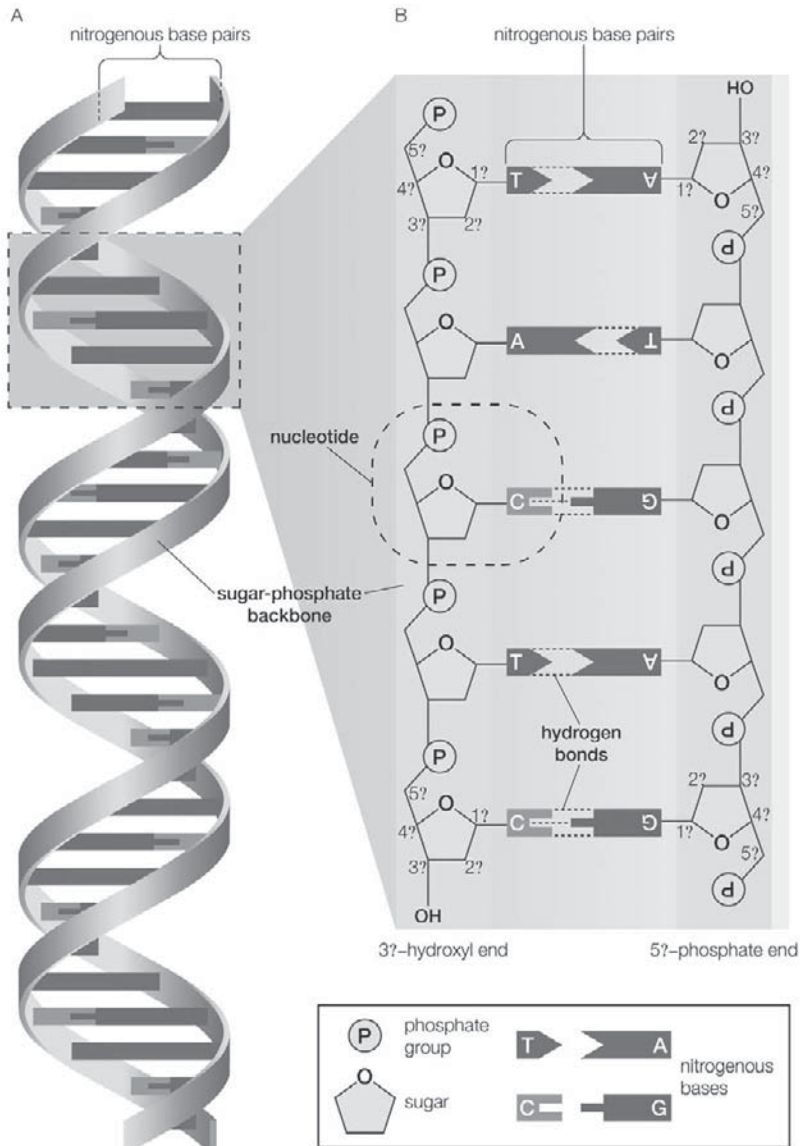
Since the chromosomes can be delineated only during mitosis, it is necessary to examine material in which there are many dividing cells. This can usually be accomplished by culturing cells from the blood or skin. After growth,

the cells are fixed on slides and then stained with a variety of DNA-specific dyes that permit the delineation and identification of the chromosomes. The Denver system of chromosome classification, established in 1959, identified the chromosomes by their length and the position of the centromeres. Since then the method has been improved by the use of special staining techniques that impart unique light and dark bands to each chromosome. These bands permit the identification of chromosomal regions that are duplicated, missing, or transposed to other chromosomes.

In a typical karyotype micrograph, which is an image of the chromosomes, the 46 human chromosomes (the diploid number) are arranged in homologous pairs, each consisting of one maternally derived and one paternally derived member. The chromosomes are all numbered except for the X and the Y chromosomes, which are the sex chromosomes. In humans, as in all mammals, the normal female has two X chromosomes and the normal male has one X chromosome and one Y chromosome. The female is thus the homogametic sex, as all her gametes normally have one X chromosome. The male is heterogametic, as he produces two types of gametes—one type containing an X chromosome and the other containing the Y chromosome. There is good evidence that the Y chromosome in humans, unlike that in *Drosophila*, is necessary (but not sufficient) for maleness.

THE HUMAN GENOME

The human genome contains approximately three billion base pairs of DNA, which make up the entire set of chromosomes of the human organism. The human genome includes the coding regions of DNA, which encode all the genes (about 25,000) of the human organism, as well as



The human genome is made up of approximately three billion base pairs of deoxyribonucleic acid (DNA). The bases of DNA are adenine (A), thymine (T), guanine (G), and cytosine (C). Encyclopædia Britannica, Inc.

the noncoding regions of DNA, which do not encode any genes. By 2003 the DNA sequence of the entire human genome was known.

The human genome, like the genomes of all other animals, is a collection of long polymers of DNA. These polymers are maintained in duplicate copy in the form of chromosomes in every human cell and encode in their sequence of constituent bases (guanine [G], adenine [A], thymine [T], and cytosine [C]) the details of the molecular and physical characteristics that form the corresponding organism. The sequence of these polymers, their organization and structure, and the chemical modifications they contain not only provide the machinery needed to express the information held within the genome but also provide the genome with the capability to replicate, repair, package, and otherwise maintain itself. In addition, the genome is essential for the survival of the human organism; without it no cell or tissue could live beyond a short period of time. For example, red blood cells (erythrocytes), which live for only about 120 days, and skin cells, which on average live for only about 17 days, must be renewed to maintain the viability of the human body, and it is within the genome that the fundamental information for the renewal of these cells, and many other types of cells, is found.

The human genome is not uniform. Excepting identical (monozygous) twins, no two humans on Earth share exactly the same genomic sequence. Further, the human genome is not static. Subtle and sometimes not so subtle changes arise with startling frequency. Some of these changes are neutral or even advantageous; these are passed from parent to child and eventually become commonplace in the population. Other changes may be detrimental, resulting in reduced survival or decreased fertility of those individuals who harbour them; these changes tend to be

rare in the population. The genome of modern humans, therefore, is a record of the trials and successes of the generations that have come before. Reflected in the variation of the modern genome is the range of diversity that underlies what are typical traits of the human species. There is also evidence in the human genome of the continuing burden of detrimental variations that sometimes lead to disease.

Knowledge of the human genome provides an understanding of the origin of the human species, the relationships between subpopulations of humans, and the health tendencies or disease risks of individual humans. Indeed, in the past 20 years knowledge of the sequence and structure of the human genome has revolutionized many fields of study, including medicine, anthropology, and forensics. With technological advances that enable inexpensive and expanded access to genomic information, the amount of and the potential applications for the information that is extracted from the human genome is extraordinary.

Role of the Human Genome in Research

Since the 1980s there has been an explosion in genetic and genomic research. The combination of the discovery of the polymerase chain reaction, improvements in DNA sequencing technologies, advances in bioinformatics (mathematical biological analysis), and increased availability of faster, cheaper computing power has given scientists the ability to discern and interpret vast amounts of genetic information from tiny samples of biological material. Further, methodologies such as fluorescence in situ hybridization (FISH) and comparative genomic hybridization (CGH) have enabled the detection of the

organization and copy number of specific sequences in a given genome.

Understanding the origin of the human genome is of particular interest to many researchers since the genome is indicative of the evolution of humans. The public availability of full or almost full genomic sequence databases for humans and a multitude of other species has allowed researchers to compare and contrast genomic information between individuals, populations, and species. From the similarities and differences observed, it is possible to track the origins of the human genome and to see evidence of how the human species has expanded and migrated to occupy the planet.

Origins of the Human Genome

Comparisons of specific DNA sequences between humans and their closest living relative, the chimpanzee, reveal 99 percent identity, although the homology drops to 96 percent if insertions and deletions in the organization of those sequences are taken into account. This degree of sequence variation between humans and chimpanzees is only about 10-fold greater than that seen between two unrelated humans. From comparisons of the human genome with the genomes of other species, it is clear that the genome of modern humans shares common ancestry with the genomes of all other animals on the planet and that the modern human genome arose between 150,000 and 300,000 years ago.

Ongoing collaboration between archaeologists, anthropologists, and molecular geneticists at the Max Planck Institute in Germany and the Lawrence Berkeley National Laboratory and the Joint Genome Institute in the United States has enabled sequence comparisons

between modern humans (*Homo sapiens*) and Neanderthals (*H. neanderthalensis*). The data obtained so far demonstrate that modern humans and Neanderthals share about 99.5 percent genome sequence identity. As more Neanderthal genomic sequence data becomes available, it will be particularly interesting to see what genetic differences distinguish the modern human species from Neanderthals. Research suggests that populations of *H. sapiens* split from *H. neanderthalensis* ancestral populations perhaps as recently as 370,000 years ago and likely shared a common ancestor some 500,000–700,000 years ago. Despite this, genomic studies have indicated that there was almost no interbreeding between *H. sapiens* and *H. neanderthalensis*. Thus, when Neanderthals, the last of the *Homo* relatives of modern humans, became extinct about 30,000 years ago, only modern humans were left to populate the Earth.

Comparing the DNA sequences of groups of modern humans from different continents also allows scientists to define the relationships and even the ages of these different populations. By combining these genetic data with archaeological and linguistic information, anthropologists have been able to discern the origins of *H. sapiens* in Africa and to track the timing and location of the waves of human migration out of Africa that led to the eventual spread of humans to other continents of the globe. For example, genetic evidence indicates that the first humans migrated out of Africa approximately 60,000 years ago, settling in southern Europe, the Middle East, southern Asia, and Australia. From there, subsequent and sequential migrations brought humans to northern Eurasia and across what was then a land bridge to North America and finally to South America.

The global distribution of human skin colour is a well-defined example of genetic variation in which differential selective pressures favoured different characteristics in

skin colour that conferred a survival advantage. Selective pressures for skin colour correlate with regional climate factors, such as latitude and sunlight. For example, the first populations of humans to settle in northern regions of the world were under selective pressure that favoured light skin colour to facilitate the absorption of sunlight, thereby preventing premature death from debilitating bone diseases.

As humans migrated across the continents, sequence variations arose that became differentially fixed in different populations. Some variations likely reflect what are called founder effects, changes in gene frequency that occur in small populations. Founder effects are generally characterized by genes that are expressed with increasing frequency from one generation to the next and can be traced back to the original founders of the population. Other variations reflect differential selective pressures at work. For example, populations living in equatorial climates were under strong selective pressure that favoured dark skin colour to protect against extreme sun exposure, thereby decreasing the deleterious health effects caused by sunburn and skin cancer. In contrast, populations migrating to more polar latitudes, where levels of sun exposure are relatively low, experienced strong selective pressure that favoured light skin colour, thereby facilitating the absorption of sunlight by the skin for the synthesis of vitamin D. In northern Europe and Scandinavia, therefore, individuals with genetic variations leading to lighter skin colour were less likely to become vitamin D deficient and suffer from the bone disease known as rickets.

Social Impacts of Human Genome Research

Databases have been compiled that list and summarize specific DNA variations that are common in certain

human populations but not in others. Because the underlying DNA sequences are passed from parent to child in a stable manner, these genetic variations provide a tool for distinguishing the members of one population from those of the other. Public genetic ancestry projects, in which small samples of DNA can be submitted and analyzed, have allowed individuals to trace the continental or even subcontinental origins of their most ancient ancestors.

The role of genetics in defining traits and health risks for individuals has been recognized for generations. Long before DNA or genomes were understood, it was clear that many traits tended to run in families and that family history was one of the strongest predictors of health or disease. Knowledge of the human genome has advanced that realization, enabling studies that have identified the genes and even specific sequence variations that contribute to a multitude of traits and disease risks. With this information in hand, health care professionals are able to practice predictive medicine, which translates in the best of scenarios to preventative medicine. Indeed, presymptomatic genetic diagnoses have enabled countless people to live longer and healthier lives. For example, mutations responsible for familial cancers of the breast and colon have been identified, enabling presymptomatic testing of individuals in at-risk families. Individuals who carry the mutant gene or genes are counseled to seek heightened surveillance. In this way, if and when cancer appears, these individuals can be diagnosed early, when the cancers are most effectively treated.

THE GENETICS OF HUMAN BLOOD

More is known about the genetics of the blood than about any other human tissue. One reason for this is that blood samples can be easily secured and subjected to

biochemical analysis without harm or major discomfort to the person being tested. Furthermore, many chemical properties of human blood display relatively simple patterns of inheritance.

Blood Types

Certain chemical substances within the red blood cells may serve as antigens, proteins that stimulate antibody production when introduced into other animals. For example, when cells that contain specific antigens are introduced into the body of an experimental animal such as a rabbit, the animal responds by producing antibodies in its own blood.

The antigens that are present on the surfaces of red blood cells form the basis of blood classification. Two of the more widely known blood group classifications in humans are the ABO and MNSs systems, which are characterized by the A, B, and O antigens and the M, N, S, and s antigens, respectively. In addition to the ABO and MNSs systems, geneticists have identified about 14 blood-type gene systems associated with other chromosomal locations. The best known of these is the Rh system. The Rh antigens are of particular importance in human medicine. Curiously, however, their existence was discovered in monkeys. When blood from the rhesus monkey (hence the designation Rh) is injected into rabbits, the rabbits produce so-called Rh antibodies that will agglutinate (adhere together) not only the red blood cells of the monkey but the cells of a large proportion of humans as well. Some people (Rh-negative individuals), however, lack the Rh antigen; the proportion of such persons varies from one human population to another. Akin to data concerning the ABO system, the evidence for Rh genes indicates that only a single chromosome locus (called *r*) is involved

and is located on chromosome 1. At least 35 Rh alleles are known for the *r* location; basically the Rh-negative condition is recessive.

A medical problem may arise when a woman who is Rh-negative carries a fetus that is Rh-positive. The first such child may have no difficulty, but later similar pregnancies may produce severely anemic newborn infants. Exposure to the red blood cells of the first Rh-positive fetus appears to immunize the Rh-negative mother, that is, she develops antibodies that may produce permanent (sometimes fatal) brain damage in any subsequent Rh-positive fetus. Damage arises from the scarcity of oxygen reaching the fetal brain because of the severe destruction of red blood cells. Measures are available for avoiding the severe effects of Rh incompatibility by transfusions to the fetus within the uterus; however, genetic counselling before conception is helpful so that the mother can receive Rh immunoglobulin immediately after her first and any subsequent pregnancies involving an Rh-positive fetus. This immunoglobulin effectively destroys the fetal red blood cells before the mother's immune system is stimulated. The mother thus avoids becoming actively immunized against the Rh antigen and will not produce antibodies that could attack the red blood cells of a future Rh-positive fetus.

Serum Proteins

Human serum, the fluid portion of the blood that remains after clotting, contains various proteins that have been shown to be under genetic control. Study of genetic influences has flourished since the development of precise methods for separating and identifying serum proteins. These move at different rates under the impetus of an electrical field (electrophoresis), as do proteins from many

other sources (e.g., muscle or nerve). Since the composition of a protein is specified by the structure of its corresponding gene, biochemical studies based on electrophoresis permit direct study of tissue substances that are only a metabolic step or two away from the genes themselves.

Electrophoretic studies have revealed that at least one-third of the human serum proteins occur in variant forms. Many of the serum proteins are polymorphic, occurring as two or more variants with a frequency of not less than 1 percent each in a population. Patterns of polymorphic serum protein variants have been used to determine whether twins are identical (as in assessing compatibility for organ transplants) or whether two individuals are related (as in resolving paternity suits). Whether or not the different forms have a selective advantage is not generally known.

Much attention in the genetics of substances in the blood has been centred on serum proteins called haptoglobins, transferrins (which transport iron), and gamma globulins (a number of which are known to immunize against infectious diseases). Haptoglobins appear to relate to two common alleles at a single chromosome locus; the mode of inheritance of the other two seems more complicated, about 18 kinds of transferrins having been described. Like blood-cell antigen genes, serum-protein genes are distributed worldwide in the human population in a way that permits their use in tracing the origin and migration of different groups of people.

Hemoglobin

Hundreds of variants of hemoglobin have been identified by electrophoresis, but relatively few are frequent enough to be called polymorphisms. Of the polymorphisms, the

alleles for sickle-cell and thalassemia hemoglobins produce serious disease in homozygotes, whereas others (hemoglobins C, D, and E) do not. The sickle-cell polymorphism confers a selective advantage on the heterozygote living in an environment where malaria is present; the thalassemia polymorphism provides a similar advantage.

THE GENETICS OF ANTIBODY FORMATION

One of the central problems in understanding the genetics of the immune system has been in explaining the genetic regulation of antibody production. Immunobiologists have demonstrated that the system can produce well over 1,000,000 specific antibodies, each corresponding to a particular antigen. It would be difficult to envisage that each antibody is encoded by a separate gene—such an arrangement would require a disproportionate share of the entire human genome. Recombinant DNA analysis has illuminated the mechanisms by which a limited number of immunoglobulin genes can encode this vast number of antibodies.

Each antibody molecule consists of several different polypeptide chains—the light chains (L) and the longer heavy chains (H). The latter determine to which of five different classes (IgM, IgG, IgA, IgD, or IgE) an immunoglobulin belongs. Both the L and H chains are unique among proteins in that they contain constant and variable parts. The constant parts have relatively identical amino acid sequences in any given antibody. The variable parts, on the other hand, have different amino acid sequences in each antibody molecule. It is the variable parts, then, that determine the specificity of the antibody.

Recombinant DNA studies of immunoglobulin genes in mice have revealed that the light-chain genes are encoded in four separate parts in germline DNA: a leader segment (L), a variable segment (V), a joining segment (J), and a constant segment (C). These segments are widely separated in the DNA of an embryonic cell, but in a mature B cell (a type of white blood cell of fundamental importance in the immune system) they are found in relative proximity (albeit separated by introns). The mouse has more than 200 light-chain variable region genes, only one of which will be incorporated into the proximal sequence that codes for the antibody production in a given B cell. Antibody diversity is greatly enhanced by this system, as the V and J segments rearrange and assort randomly in each B-cell precursor. The mechanisms by which this DNA rearrangement takes place are not clear, but transposons are undoubtedly involved. Similar combinatorial processes take place in the genes that code for the heavy chains; furthermore, both the light-chain and heavy-chain genes can undergo somatic mutations to create new antibody-coding sequences. The net effect of these combinatorial and mutational processes enables the coding of millions of specific antibody molecules from a limited number of genes. It should be stressed, however, that each B cell can produce only one antibody. It is the B cell population as a whole that produces the tremendous variety of antibodies in humans and other mammals.

Plasma cell tumours (myelomas) have made it possible to study individual antibodies since these tumours, which are descendants of a single plasma cell (a short-lived antibody-producing cell), produce one antibody in abundance. Another method of obtaining large amounts of a specific antibody is by fusing a B cell with a rapidly growing cancer cell. The resultant hybrid cell, known as a

hybridoma, multiplies rapidly in culture. Since the antibodies obtained from hybridomas are produced by clones derived from a single B cell, they are called monoclonal antibodies.

THE GENETICS OF CELLULAR IMMUNITY

Cellular immunity is mediated by T cells (a type of white blood cell) that can recognize infected body cells, cancer cells, and the cells of a foreign transplant. The control of cellular immune reactions is provided by a linked group of genes, known as the major histocompatibility complex (MHC). These genes code for the major histocompatibility antigens, which are found on the surface of almost all nucleated somatic cells. The major histocompatibility antigens were first discovered on the leukocytes (white blood cells) and are, therefore, usually referred to as the HLA (human leukocyte group A) antigens.

The advent of the transplantation of human organs in the 1950s made the question of tissue compatibility between donor and recipient of vital importance, and it was in this context that the HLA antigens and the MHC were elucidated. Investigators found that the MHC resides on the short arm of chromosome 6, on four closely associated sites designated HLA-A, HLA-B, HLA-C, and HLA-D. Each locus is highly polymorphic—i.e., each is represented by a great many alleles within the human gene pool. These alleles, like those of the ABO blood group system, are expressed in codominant fashion. Because of the large number of alleles at each HLA locus, there is an extremely low probability of any two individuals (other than siblings) having identical HLA genotypes. (Since a person inherits one chromosome 6 from each parent, siblings have a 25 percent probability of having received the

same paternal and maternal chromosomes 6 and thus of being HLA matched.)

Although HLA antigens are largely responsible for the rejection of organ transplants, it is obvious that the MHC did not evolve to prevent the transfer of organs from one person to another. Indeed, information obtained from the histocompatibility complex in the mouse (which is very similar in its genetic organization to that of the human) suggests that a primary function of the HLA antigens is to regulate the number of specific cytotoxic T killer cells, which have the ability to destroy virus-infected cells and cancer cells.

INFLUENCE OF THE ENVIRONMENT

Gene expression undergoes significant modification by factors in the environment. One example of how this occurs involves the recessively inherited disease called galactosemia, in which the enzyme necessary for the metabolism of galactose—a component of milk sugar—is defective. The sole source of galactose in the infant's diet is milk, which in this instance is toxic. The disease is treated via the removal of all natural forms of milk from the diet (environmental manipulation), which are replaced by a synthetic milk lacking galactose. The infant will then develop normally but will never be able to tolerate foods containing lactose. If milk were not a major part of the infant's diet, however, the mutant gene would never be able to express itself, and galactosemia would be unknown.

Another way of saying this is that no trait can exist or become actual without an environmental contribution. Thus, the old question of which is more important, heredity or environment, is without meaning. Both nature (heredity) and nurture (environment) are always important for every human attribute.

But this is not to say that the separate contributions of heredity and environment are equivalent for each characteristic. Dark pigmentation of the iris of the eye, for example, is under hereditary control in that one or more genes specify the synthesis and deposition in the iris of the pigment (melanin). This is one character that is relatively independent of environmental factors, such as diet or climate; thus, individual differences in eye colour tend to be largely attributable to hereditary factors rather than to ordinary environmental change.

On the other hand, it is unwarranted to assume that other traits (such as height, weight, or intelligence) are as little affected by environment as is eye colour. It is very easy to gather information that tall parents tend, on the average, to have tall children (and that short parents tend to produce short children), properly indicating a hereditary contribution to height. Nevertheless, it is equally manifest that growth can be stunted in the environmental absence of adequate nutrition. The dilemma arises that only the combined, final result of this nature–nurture interaction can be directly observed. There is no accurate way (in the case of a single individual) to gauge the separate contributions of heredity and environment to a characteristic such as height. An inferential way out of this dilemma is provided by studies of twins.

Fraternal Twins

Usually a fertile human female produces a single egg about once a month. Should fertilization occur (i.e., a zygote is formed), growth of the individual child normally proceeds after the fertilized egg has become implanted in the wall of the uterus (womb). In the unusual circumstance that two unfertilized eggs are simultaneously released by the ovaries, each egg may be fertilized by a different sperm

cell at about the same time, become implanted, and grow, to result in the birth of twins.

Twins formed from separate eggs and different sperm cells can be of the same or of either sex. No matter what their sex, they are designated as fraternal twins. This terminology is used to emphasize that fraternal twins are genetically no more alike than are siblings (brothers or sisters) born years apart. Basically they differ from ordinary siblings only in having grown side by side in the womb and in having been born at approximately the same time.

Identical Twins

In a major nonfraternal type of twinning, only one egg is fertilized; but during the cleavage of this single zygote into two cells, the resulting pair somehow become separated. Each of the two cells may implant in the uterus separately and grow into a complete, whole individual. In laboratory studies with the zygotes of many animal species, it has been found that in the two-cell stage (and later) a portion of the embryo, if separated under the microscope by the experimenter, may develop into a perfect, whole individual. Such splitting occurs spontaneously at the four-cell stage in some organisms (e.g., the armadillo) and has been accomplished experimentally with the embryos of salamanders, among others.

The net result of splitting at an early embryonic stage may be to produce so-called identical twins. Since such twins derive from the same fertilized egg, the hereditary material from which they originate is absolutely identical in every way, down to the last gene locus. While developmental and genetic differences between one "identical" twin and another still may arise through a number of processes (e.g., mutation), these twins are always found to be of the same sex. They are often breathtakingly similar in

appearance, frequently down to very fine anatomic and biochemical details (although their fingerprints are differentiable).

Diagnosis of Twin Types

Since the initial event in the mother's body (either splitting of a single egg or two separate fertilizations) is not observed directly, inferential means are employed for diagnosing a set of twins as fraternal or identical. The birth of fraternal twins is frequently characterized by the passage of two separate afterbirths. In many instances, identical twins are followed by only a single afterbirth, but exceptions to this phenomenon are so common that this is not a reliable method of diagnosis.

The most trustworthy method for inferring twin type is based on the determination of genetic similarity. By selecting those traits that display the least variation attributable to environmental influences (such as eye colour and blood types), it is feasible, if enough separate chromosome loci are considered, to make the diagnosis of twin type with high confidence. HLA antigens, which are very polymorphic, have become most useful in this regard.

Inferences from Twin Studies

Metric (Quantitative) Traits

By measuring the heights of a large number of ordinary siblings (brothers and sisters) and of twin pairs, it may be shown that the average difference between identical twins is less than half the difference for all other siblings. Any average differences between groups of identical twins are attributable with considerable confidence to the

environment. Thus, since the sample of identical twins who were reared apart (in different homes) differed little in height from identicals who were raised together, it appears that environmental-genetic influences on that trait tended to be similar for both groups.

Yet, the data for like-sexed fraternal twins reveal a much greater average difference in height (about the same as that found for ordinary siblings reared in the same home at different ages). Apparently the fraternal twins were more dissimilar than identicals (even though reared together) because the fraternals differed more among themselves in genotype. This emphasizes the great genetic similarity among identicals. Such studies can be particularly enlightening when the effects of individual genes are obscured or distorted by the influence of environmental factors on quantitative (measurable) traits (e.g., height, weight, and intelligence).

Any trait that can be objectively measured among identical and fraternal twins can be scrutinized for the particular combination of hereditary and environmental influences that impinge upon it. The effect of environment on identical twins reared apart is suggested by their relatively great average difference in body weight as compared with identical twins reared together. Weight appears to be more strongly modified by environmental variables than is height.

Study of comparable characteristics among farm animals and plants suggests that such quantitative human traits as height and weight are affected by allelic differences at a number of chromosome locations: that they are not simply affected by genes at a single locus. Investigation of these gene systems with multiple locations (polygenic systems) is carried out largely through selective-breeding experiments among large groups of plants and

lower animals. Humans select their mates in a much freer fashion, of course, and polygenic studies among people are thus severely limited.

Intelligence is a very complex human trait, the genetics of which has been a subject of controversy for some time. Much of the controversy arises from the fact that intelligence is so difficult to define. Information has been based almost entirely on scores on standardized IQ tests constructed by psychologists; in general such tests do not take into account cultural, environmental, and educational differences. As a result, the working definition of intelligence has been "the general factor common to a large number of diverse cognitive (IQ) tests." Even roughly measured as IQ, intelligence shows a strong contribution from the environment. Fraternal twins, however, show relatively great dissimilarity in IQ, suggesting an important contribution from heredity as well. In fact, it has been estimated that on the average between 60 and 80 percent of the variance in IQ test scores could be genetic. It is important to note that intelligence is polygenically inherited and that it has the highest degree of assortative mating of any trait; in other words, people tend to mate with people having similar IQ's. Moreover, twin studies involving psychological traits should be viewed with caution; for example, since identical twins tend to be singled out for special attention, their environment should not be considered equivalent even to that of other children raised in their own family.

Since the time of Galton, generalizations have been repeatedly made about racial differences in intelligence, with claims of genetic superiority of some races over others. These generalizations fail to recognize that races are composed of individuals, each of whom has a unique genotype made up by genes shared with other humans, and that the sources of intraracial variation are more numerous than those producing interracial differences.

Other Traits

For traits of a more qualitative (all-or-none) nature, the twin method can also be used in efforts to assess the degree of hereditary contribution. Such investigations are based on an examination of cases in which at least one member of the twin pair shows the trait. It was found in one study, for example, that in about 80 percent of all identical twin pairs in which one twin shows symptoms of the psychiatric disorder called schizophrenia, the other member of the pair also shows the symptoms (that is, the two are concordant for the schizophrenic trait). In the remaining 20 percent, the twins are discordant (that is, one lacks the trait). Since identical twins often have similar environments, this information by itself does not distinguish between the effects of heredity and environment. When pairs of like-sexed fraternal twins reared together are studied, however, the degree of concordance for schizophrenia is very much lower—only about 15 percent.

Schizophrenia thus clearly develops much more easily in some genotypes than among others; this indicates a strong hereditary predisposition to the development of the trait. Schizophrenia also serves as a good example of the influence of environmental factors since concordance for the condition does not appear in 100 percent of identical twins.

Studies of concordance and discordance between identical and fraternal twins have been carried out for many other human characteristics. For example, it has been known for many years that tuberculosis is a bacterial infection of environmental origin. Yet identical twins raised in the same home show concordance for the disease far more often than do fraternal twins. This finding seems to be explained by the high degree of genetic similarity among the identical twins. Although tuberculosis is not an

inherited disease, heredity does seem to make one more (or less) susceptible to this particular infection. Thus, the genes of one individual may provide the chemical basis for susceptibility to a disease, whereas the genes of another may fail to do so.

Indeed, there are genetic differences among infectious organisms themselves that result in differences in their virulence. Thus, whether a genetically susceptible person actually develops a disease also depends in part on the heredity of the particular strain of bacteria, virus, or parasite with which he or she must cope. Consequently, unless environmental factors such as these are adequately evaluated, the conclusions drawn from susceptibility studies can be unfortunately misleading.

The expression of the genotype can always be modified by the environment. It can be argued that all human illnesses have a genetic component and that the basis of all medical therapy is environmental modification. Specifically, this is the hope for the management of genetic diseases. The more that can be learned about the basic molecular and cellular dysfunctions associated with such diseases, the more amenable they will be to environmental manipulation.

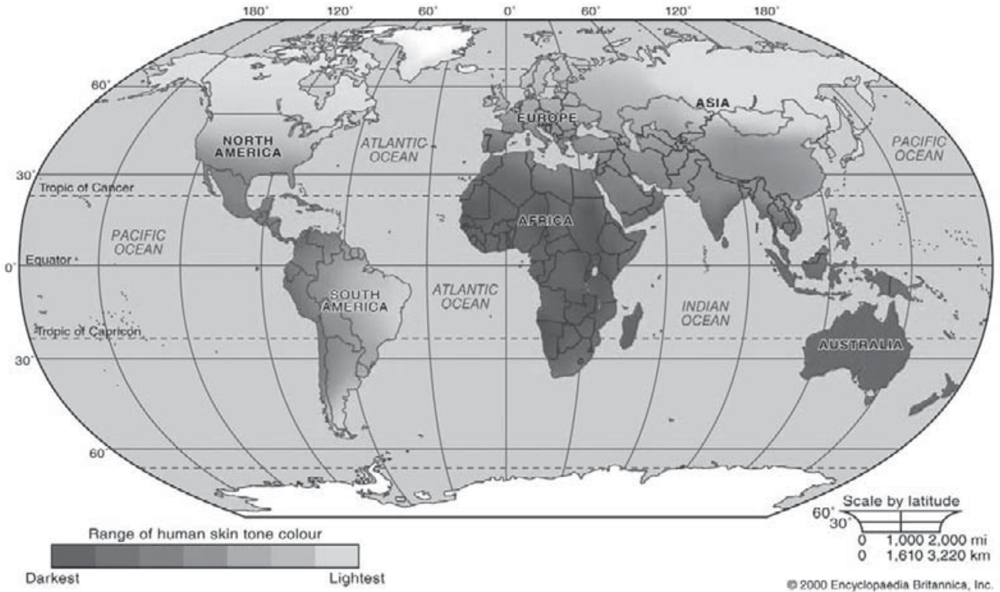
GENETICS AND THE CONCEPT OF RACE

Modern Scientific Explanations of Human Biological Variation

Human physical variations, especially in those traits that are normally used to classify people racially—skin colour, hair texture, facial features, and to some extent bodily structure—can be understood in terms of evolutionary processes and the long-range adaptation of human groups to differing environments. Variations between humans

that are not explained by these processes may simply reflect random mutations or functionally neutral changes in the genetic code.

In any given habitat, natural forces operate on all living forms, including human populations. The necessary



The distribution of skin colour variations of indigenous populations before colonization by Europeans. The map, compiled by Audrey Smedley, is a reconstruction of populations based on a number of sources. In some cases, areal characteristics have been estimated from descriptions (or drawings) of first contact by the earliest Europeans. In other cases, where there was little European contact or where there is scant information about native populations (as there is, for example, about the populations of inner Asia), skin colour was estimated from surrounding populations and geographic and climatological information. On a map of this scale, it is difficult to give more than a representation of current understanding. It must also be noted that many populations, even before the modern era, were quite heterogeneous for skin colour, and this heterogeneity is difficult to depict accurately on any scale. In areas of the world where the indigenous population was sparse and widely scattered (such as Australia), the map's colour density can be misleading. Another such problem is represented by the Tasmanians, who are virtually extinct, and the Maori, who have been widely mixed with Europeans, so only a few examples of "unmixed" individuals have been found in the historical records.

interaction with these forces will affect the survival and reproduction of the members of these societies. Such populations already have a wide and complex range of hereditary physical characteristics; indeed, human hereditary variability is a product of human sexual reproduction, whereby every individual receives half of his or her genetic endowment from each parent and no two individuals (except for identical twins) inherit the same combination of genetic features.

The global distribution of skin colour is the best example of adaptation, and the consequences of this process have long been well known. Skin colour clines (gradations) in indigenous populations worldwide correlate with latitude and amounts of sunlight. Indigenous populations within a broad band known as the tropics (the regions falling in latitude between the Tropics of Cancer and Capricorn) have darker skin colours than indigenous populations outside of these regions.

Within the tropics, skin colours vary from light tan to very dark brown or black, both among populations and among individuals within groups. The darkest skin colours are found in those populations long residing in regions where intense ultraviolet sunlight is greatest and there is little natural forest cover. The bluish black skins of some peoples—such as some of the Dravidians of South India, the peoples of Sri Lanka and Bangladesh, and peoples of the eastern Sudan zone, including the ancient region of Nubia, and the grasslands of Africa—are examples of the extremes of dark skin colour. Medium brown to dark brown peoples are found in the rest of tropical Africa and India and throughout Australia, Melanesia, and other parts of Southeast Asia.

Peoples with light skin colours evolved over thousands of years in northern temperate climates. Human groups intermittently migrating into Europe and the northern

parts of the Eurasian landmass over the past 25,000–50,000 years experienced a gradual loss of skin pigmentation. The changes were both physiological and genetic; that is, there were systemic changes in individuals and long-range genetic changes as a result of natural selection and, possibly, mutations. Those individuals with the lightest skin colours, with the lowest amounts of melanin, survived and reproduced in larger numbers and thus passed on their genes for lighter skin. Over time, entire populations living in northern climates evolved lighter skin tones than those individuals living in areas with higher levels of sunlight. Between populations with light skin and those with the darkest coloration are populations with various shades of light tan to brown. The cline in skin colours shows variation by infinite degrees; any attempts to place boundaries along this cline represent purely arbitrary decisions.

Scientists at the turn of the 21st century understood why these superficial visible differences developed. Melanin, a substance that makes the skin dark, has been shown to confer protection from sunburn and skin cancers in those very areas where ultraviolet sunlight is strongest. Dark skin, which tends to be thicker than light skin, may have other protective functions in tropical environments where biting insects and other vectors of disease are constant threats to human survival. But humans also need vitamin D, which is synthesized by sunlight from sterols (chemical compounds) present in the skin. Vitamin D affects bone growth, and, without a sufficient amount, the disease known as rickets would have been devastating to early human groups trying to survive in the cold, wintry weather of the north. As these groups adapted to northern climates with limited sunlight, natural selection brought about the gradual loss of melanin in favour of skin tones that enabled some individuals to better synthesize vitamin D.

Other physical characteristics indicate adaptations to cold or hot climates, to variations in elevation from sea level, to rain forests with high levels of rainfall, and to hot deserts. Body structure and the amount of body fat have also been explained by evolutionists in terms of human adaptation to differing environments. Long, linear body builds seems to be highly correlated with hot, dry climates. Such people inhabit the Sahara and the desiccated areas of the Sudan in Africa. Short, stocky body builds with stubby fingers and toes are correlated with cold, wet climates, such as are found in Arctic areas. People adapted to cold climates have acquired genetic traits that provide them extra layers of body fat, which accounts for the epicanthic fold over their eyes. People who live in areas of high elevation, as in the mountains of Peru, tend to have an adaptive feature not found among peoples who live at sea level; they have larger lungs and chest cavities. In an atmosphere where the oxygen supply is low, larger lungs are clearly adaptive.

Some adaptive variations are not obviously visible or measurable. Many peoples adapted to cold climates, for example, have protective physiological reactions in their blood supply. Their blood vessels either constrict the flow to extremities to keep the inner body warm while their surface skin may be very cold (vasoconstriction) or dilate to increase the blood flow to the hands, feet, and head to warm the outer surfaces (vasodilation).

The prevalence of diseases has been another major factor in the evolution of human diversity, and some of the most important of human genetic variations reflect differences in immunities to diseases. The sickle cell trait (hemoglobin S), for example, is found chiefly in those regions of the tropical world where malaria is endemic. Hemoglobin S in its heterozygous form (inherited from one parent only) confers some immunity to those people

who carry it, although it brings a deadly disease (sickle cell anemia) in its homozygous form (inherited from both parents).

In the last decades of the 20th century, scientists began to understand human physical variability in clinal terms and to recognize that it reflects much more complex gradation and combination than they had anticipated. To comprehend the full expression of a feature's genetic variability, it must be studied separately over geographic space and often in terms of its adaptive value. Many features are now known to relate to the environmental conditions of the populations that carry them.

The Scientific Debate Over "Race"

Although their numbers are dwindling, some scientists continue to believe that it is possible to divide *Homo sapiens* into discrete populations called races. They believe that the physical differences manifest in wide geographic regions are more than superficial—that they reflect innate intellectual, moral, emotional, and other behavioral differences between human groups. They deny that social circumstances and the cultural realities of racism have any effect on behaviour or the performance of children and adults on IQ tests. Those scientists who advocate the continued acceptance of race and racial differences have been labeled "splitters." Among the highly popularized reflections of this point of view was *The Bell Curve: Intelligence and Class Structure in American Life* (1994) by Richard Herrnstein and Charles Murray. This work is a representation of social Darwinism in that the authors argue not only that minority or low-status races have innate deficiencies but that poor people of all races, including whites, are genetically inferior.

Those who deny the biological salience of race or argue against the use of the term have been labeled "lumpers." The latter see their position as being buttressed and confirmed by ongoing genetic and other research. They emphasize the failure of science to establish exclusive boundaries around populations or lines of rigid distinctions that the term *race* conveys. They also point to the evidence demonstrating that all people regardless of their physical variations are capable of learning any kind of cultural behaviour. They argue that genes and cultural conditioning work in tandem and together contribute to the formation of individual personalities.

Today there is increasing evidence that the concept of race has outlived its usefulness. Social scientists, biologists, historians, and philosophers now point out that increasing migration and changes in attitudes toward human

Students take a break from classes to talk. Education often facilitates the formation of cultural identities, which are becoming increasingly prevalent as the significance of race is debated. © www.istockphoto.com/Chris Schmidt

differences have brought about extensive intermingling of peoples so that a growing number of people have ancestors originating in three or more continents. Such "mixed" people are not easily lumped into a single "racial" category. As a result, many scholars perceive that "race" is becoming more and more irrelevant and may eventually be eliminated as people increasingly are recognized in terms of their ethnic or cultural identities, occupations, education, and local affiliations.

A contradictory trend also seems to be occurring among some writers who find it difficult to relinquish some elements of race ideology. Instead, they "biologize" ethnic identity and interpret peoples' cultures and behaviour as if such features stem from genetic heredity. Should this trend expand, society may continue to manifest the broad elements of race ideology, though perhaps in diminished intensity or in a different form.

CHAPTER 8
Genetic Diseases of Humans

With the increasing ability to control infectious and nutritional diseases in developed countries, there has come the realization that genetic diseases are a major cause of disability, death, and human tragedy. Rare, indeed, is the family that is entirely free of any known genetic disorder.

Many thousands of different genetic disorders with defined clinical symptoms have been identified. Of the 3 to 6 percent of newborns with a recognized birth defect, at least half involve a predominantly genetic contribution. Furthermore, genetic defects are the major known cause of pregnancy loss in developed countries, and almost half of all spontaneous abortions (miscarriages) involve a chromosomally abnormal fetus. About 30 percent of all postnatal infant mortality in developed countries is due to genetic disease, and about 30 percent of pediatric and 10 percent of adult hospital admissions can be traced to a predominantly genetic cause. Finally, medical investigators estimate that genetic defects—however minor—are present in at least 10 percent of all adults. Thus, these are not rare events.

A congenital defect is any biochemical, functional, or structural abnormality that originates prior to or shortly after birth. It must be emphasized that birth defects do not all have the same basis, and it is even possible for apparently identical defects in different individuals to reflect different underlying causes. Though the genetic and biochemical bases for most recognized defects are still uncertain, it is evident that many of these disorders result from a combination of genetic and environmental factors.

CLASSES OF GENETIC DISEASE

Most human genetic defects can be categorized as resulting from either chromosomal, single-gene Mendelian, single-gene non-Mendelian, or multifactorial causes.

Diseases Caused by Chromosomal Aberrations

About 1 out of 150 live newborns has a detectable chromosomal abnormality. Yet even this high incidence represents only a small fraction of chromosome mutations since the vast majority are lethal and result in prenatal death or stillbirth. Indeed, 50 percent of all first-trimester miscarriages and 20 percent of all second-trimester miscarriages are estimated to involve a chromosomally abnormal fetus.

Diseases caused by chromosome defects often involve numerical or structural abnormalities. Numerical abnormalities, involving either the autosomes (one of the 22 pairs of body chromosomes) or sex chromosomes, are believed generally to result from meiotic nondisjunction—that is, the unequal division of chromosomes between daughter cells—that can occur during either maternal or paternal gamete formation. Meiotic nondisjunction leads to eggs or sperm with additional or missing chromosomes.

Although the biochemical basis of numerical chromosome abnormalities remains unknown, maternal age clearly has an effect, such that older women are at significantly increased risk to conceive and give birth to a chromosomally abnormal child. The risk increases with age in an almost exponential manner, especially after age 35, so that a pregnant woman age 45 or older has between a 1 in 20 and 1 in 50 chance that her child will have trisomy 21 (Down syndrome), while the risk is only 1 in 400 for a

A couple affected by Down syndrome. Shutterstock.com

35-year-old woman and less than 1 in 1,000 for a woman under the age of 30. There is no clear effect of paternal age on numerical chromosome abnormalities.

Although Down syndrome is probably the best-known and most commonly observed of the autosomal trisomies, being found in about 1 out of 800 live births, both trisomy 13 and trisomy 18 are also seen in the population, albeit at greatly reduced rates (1 out of 10,000 live births and 1 out of 6,000 live births, respectively). The vast majority of conceptions involving trisomy (when three chromosomes are present instead of a normal pair) for any of these three autosomes are nonetheless lost to miscarriage, as are all conceptions involving trisomy for any of the other autosomes. Similarly, monosomy (one chromosome instead of a pair) for any of the autosomes is lethal in utero and therefore is not seen in the population.

Structural abnormalities of the autosomes are even more common in the population than are numerical abnormalities and include translocations of large pieces of

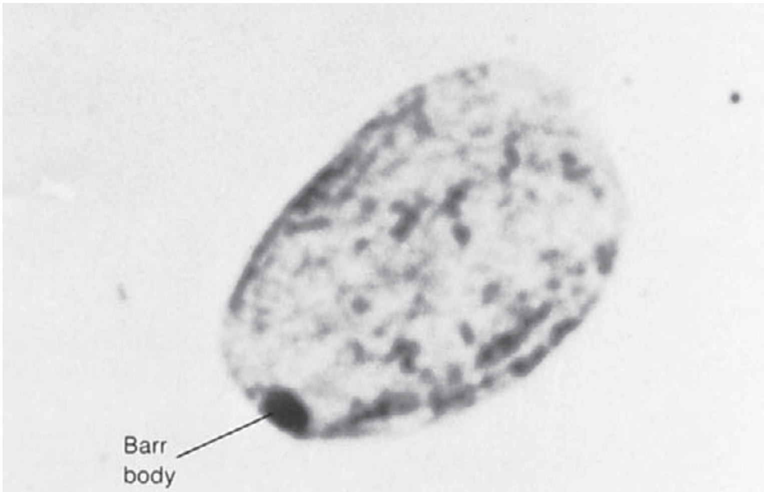
chromosomes, as well as smaller deletions, insertions, or rearrangements. Indeed, about 5 percent of all cases of Down syndrome result not from classic trisomy 21 but from the presence of excess chromosome 21 material attached to the end of another chromosome as the result of a translocation event. If balanced, structural chromosomal abnormalities may be compatible with a normal phenotype, although unbalanced chromosome structural abnormalities can be every bit as devastating as numerical abnormalities. Furthermore, because many structural defects are inherited from a parent who is a balanced carrier, couples who have one pregnancy with a structural chromosomal abnormality generally are at significantly increased risk above the general population to repeat the experience. Clearly, the likelihood of a recurrence would depend on whether a balanced form of the structural defect occurs in one of the parents.

Abnormalities of the Sex Chromosomes

About 1 in 400 male and 1 in 650 female live births demonstrate some form of sex chromosome abnormality, although the symptoms of these conditions are generally much less severe than are those associated with autosomal abnormalities. Turner syndrome is a condition of females who, in the classic form, carry only a single X chromosome (45,X). Turner syndrome is characterized by a collection of symptoms, including short stature, webbed neck, and incomplete or absent development of secondary sex characteristics, leading to infertility. Although Turner syndrome is seen in about 1 in 2,500 to 1 in 5,000 female live births, the 45,X karyotype accounts for 10 to 20 percent of the chromosomal abnormalities seen in spontaneously aborted fetuses, demonstrating that almost all 45,X conceptions are lost to miscarriage. Indeed, the majority of

liveborn females with Turner syndrome are diagnosed as mosaics, meaning that some proportion of their cells are 45,X while the rest are either 46,XX or 46,XY. The degree of clinical severity generally correlates inversely with the degree of mosaicism, so that females with a higher proportion of normal cells will tend to have a milder clinical outcome.

In contrast to Turner syndrome, which results from the absence of a sex chromosome, three alternative conditions result from the presence of an extra sex chromosome: Klinefelter syndrome, trisomy X, and 47,XYY syndrome. These conditions, each of which occurs in about 1 in 1,000 live births, are clinically mild, perhaps reflecting the fact that the Y chromosome carries relatively few genes, and, although the X chromosome is gene-rich, most of these genes become transcriptionally silent in all but one X



The Barr, or sex chromatin, body is an inactive X chromosome. It appears as a dense, dark-staining spot at the periphery of the nucleus of each somatic cell in the human female. From the Cytogenetics Laboratory of Dr. Arthur Robinson, National Jewish Hospital and Research Center/National Asthma Center, Denver, Colorado

chromosome in each somatic cell (i.e., all cells except eggs and sperm) via a process called X inactivation.

The phenomenon of X inactivation prevents a female who carries two copies of the X chromosome in every cell from expressing twice the amount of gene products encoded exclusively on the X chromosome, in comparison with males, who carry a single X. In brief, at some point in early development one X chromosome in each somatic cell of a female embryo undergoes chemical modification and is inactivated so that gene expression no longer occurs from that template. This process is apparently random in most embryonic tissues, so that roughly half of the cells in each somatic tissue will inactivate the maternal X while the other half will inactivate the paternal X. Cells destined to give rise to eggs do not undergo X inactivation, and cells of the extra-embryonic tissues preferentially inactivate the paternal X, although the rationale for this preference is unclear. The inactivated X chromosome typically replicates later than other chromosomes, and it physically condenses to form a Barr body, a small structure found at the rim of the nucleus in female somatic cells between divisions. The discovery of X inactivation is generally attributed to British geneticist Mary Lyon, and it is therefore often called "lyonization."

The result of X inactivation is that all normal females are mosaics with regard to this chromosome, meaning that they are composed of some cells that express genes only from the maternal X chromosome and others that express genes only from the paternal X chromosome. Although the process is apparently random, not every female has an exact 1:1 ratio of maternal to paternal X inactivation. Indeed, studies suggest that ratios of X inactivation can vary. Furthermore, not all genes on the X chromosome are inactivated; a small number escape

modification and remain actively expressed from both X chromosomes in the cell.

Klinefelter syndrome (47,XXY) occurs in males and is associated with increased stature and infertility. Gynecomastia (i.e., partial breast development in a male) is sometimes also seen. Males with Klinefelter syndrome, like normal females, inactivate one of their two X chromosomes in each cell, perhaps explaining, at least in part, the relatively mild clinical outcome.

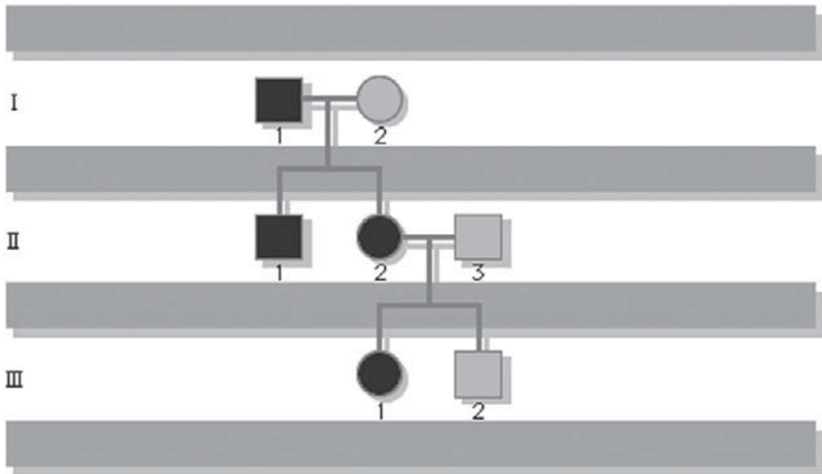
Trisomy X (47,XXX) is seen in females and is generally also considered clinically benign, although menstrual irregularities or sterility have been noted in some cases. Females with trisomy X inactivate two of the three X chromosomes in each of their cells, again perhaps explaining the clinically benign outcome.

47,XYY syndrome also occurs in males and is associated with tall stature but few, if any, other clinical manifestations. There is some evidence of mild learning disability associated with each of the sex chromosome trisomies, although there is no evidence of severe intellectual impairment in these persons.

Persons with karyotypes of 48,XXXY or 49,XXXXY have been reported but are extremely rare. These individuals show clinical outcomes similar to those seen in males with Klinefelter syndrome but with slightly increased severity. In these persons the "n - 1 rule" for X inactivation still holds, so that all but one of the X chromosomes present in each somatic cell is inactivated.

Diseases of Autosomal Dominant Inheritance

A disease trait that is inherited in an autosomal (non-sex-linked) dominant manner can occur in either sex and can be transmitted by either parent. It manifests itself in



Pedigree of a family with a history of achondroplasia, an autosomal dominantly inherited disease. The solid symbols signify affected individuals.

the heterozygote (designated *Aa*), who receives a mutant gene (designated *a*) from one parent and a normal ("wild-type") gene (designated *A*) from the other. In such a case the pedigree (i.e., a pictorial representation of family history) is vertical—that is, the disease passes from one generation to the next. In pedigrees of this sort, circles refer to females and squares to males; two symbols directly joined at the midpoint represent a mating, and those suspended from a common overhead line represent siblings, with descending birth order from left to right. Solid symbols represent affected individuals, and open symbols represent unaffected individuals. The Roman numerals denote generations, whereas the Arabic numerals identify individuals within each generation. Each person listed in a pedigree may therefore be specified uniquely by a combination of one Roman and one Arabic numeral, such as II-1.

An individual who carries one copy of a dominant mutation (*Aa*) is a heterozygote and thus has a 50 percent

chance of passing on the disease gene to each of his or her children. If an individual were to carry two copies of the dominant mutant gene (inherited from both parents), he or she would be homozygous (*AA*). The homozygote for a dominantly inherited abnormal gene may be equally affected with the heterozygote. Alternatively, he or she may be much more seriously affected; indeed, the homozygous condition may be lethal, sometimes even in utero or shortly after birth.

Although autosomal dominant traits are typically evident in multiple generations of a family, they can also arise from new mutations, so that two unaffected parents, neither of whom carries the mutant gene in their somatic cells, can conceive an affected child. Indeed, for some disorders the new mutation rate is quite high. Examples of autosomal dominant inheritance are common among human traits and diseases. More than 2,000 of these traits have been clearly identified.

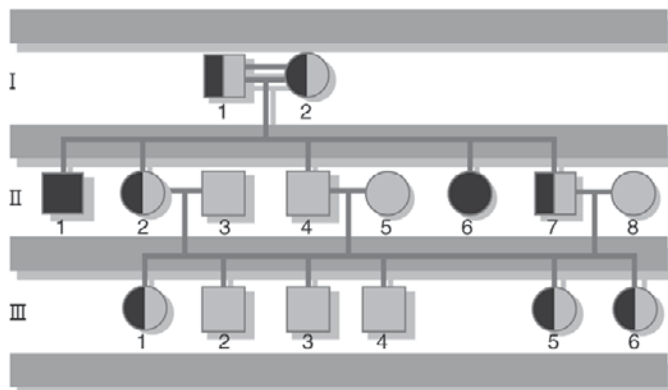
With regard to the physical manifestations (i.e., the phenotype) of some genetic disorders, a mutant gene may cause many different symptoms and may affect many different organ systems (pleiotropy). For example, along with the short-limbed dwarfism characteristic of achondroplasia, some individuals with this disorder also exhibit a long, narrow trunk, a large head with frontal bossing, and hyperextensibility of most joints, especially the knees. Similarly, for some genetic disorders, clinical severity may vary dramatically, even among affected members in the same family. These variations of phenotypic expression are called variable expressivity, and they are undoubtedly due to the modifying effects of other genes or environmental factors. Although for some disorders, such as achondroplasia, essentially all individuals carrying the mutant gene exhibit the disease phenotype, for other disorders some individuals who carry the mutant gene may

express no apparent phenotypic abnormalities at all. Such unaffected individuals are called "nonpenetrant," although they can pass on the mutant gene to their offspring, who could be affected.

Diseases of Autosomal Recessive Inheritance

Nearly 2,000 traits have been related to single genes that are recessive; that is, their effects are masked by normal ("wild-type") dominant alleles and manifest themselves only in individuals homozygous for the mutant gene. For example, sickle cell anemia, a severe hemoglobin disorder, results only when a mutant gene (*a*) is inherited from both parents. Each of the latter is a carrier, a heterozygote with one normal gene and one mutant gene (*Aa*) who is phenotypically unaffected. The chance of such a couple producing a child with sickle cell anemia is one out of four for each pregnancy. For couples consisting of one carrier (*Aa*) and one affected individual (*aa*), the chance of their having an affected child is one out of two for each pregnancy.

Many autosomal recessive traits reflect mutations in key metabolic enzymes and result in a wide variety of disorders classified as inborn errors of metabolism. One of the best-known examples of this class of disorders is phenylketonuria (PKU), which results from mutations in the gene encoding the enzyme phenylalanine hydroxylase (PAH). PAH normally catalyzes the conversion of phenylalanine, an amino acid prevalent in dietary proteins and in the artificial sweetener aspartame, to another amino acid called tyrosine. In persons with PKU, dietary phenylalanine either accumulates in the body or some of it is converted to phenylpyruvic acid, a substance that normally is produced only in small quantities. Individuals



The half-solid symbols represent heterozygous carriers of phenylketonuria; the double line between I–1 and I–2 signifies a consanguineous mating.

Pedigree of a family in which the gene for the disease phenylketonuria is segregating. In the first generation, both parents are heterozygous carriers of the disease. In further generations, solid black represents individuals affected by PKU, while solid gray represents individuals free from the disease. © 2002 Encyclopædia Britannica, Inc.

with PKU tend to excrete large quantities of this acid, along with phenylalanine, in their urine. When infants accumulate high concentrations of phenylpyruvic acid and unconverted phenylalanine in their blood and other tissues, neuronal damage in the brain may occur, which can lead to symptoms of intellectual disability. Fortunately, with early detection, strict dietary restriction of phenylalanine, and supplementation of tyrosine, this can be prevented.

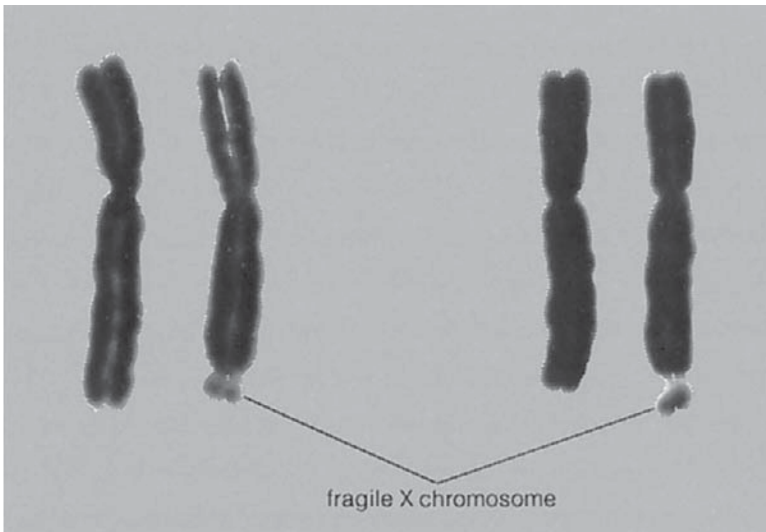
Since the recessive genes that cause inborn errors of metabolism are individually rare in the gene pool, it is not often that both parents are carriers; hence, the diseases are relatively uncommon. If the parents are related (consanguineous), however, they will be more likely to have inherited the same mutant gene from a common ancestor. For this reason, consanguinity is often more

common in the parents of those with rare, recessive inherited diseases.

Repeat Expansions

At least a dozen different disorders are known to result from triplet repeat expansions in the human genome, and these fall into two groups: (1) those that involve a polyglutamine tract within the encoded protein product that becomes longer upon expansion of a triplet repeat, an example of which is Huntington disease, and (2) those that have unstable triplet repeats in noncoding portions of the gene that, upon expansion, interfere with appropriate expression of the gene product, an example of which is fragile-X syndrome.

Both groups of disorders exhibit a distinctive pattern of non-Mendelian inheritance known as anticipation, in



The right-hand member in each of these two pairs of X chromosomes is a fragile X; the leader points to the fragile site at the tip of the long arm. Males hemizygous for this chromosome exhibit the fragile-X syndrome of mild to moderate intellectual disability..

which, following the initial appearance of the disorder in a given family, subsequent generations tend to show both increasing frequency and increasing severity of the disorder. This phenotypic anticipation is paralleled by increases in the relevant repeat length as it is passed from one generation to the next, with increasing size leading to increasing instability, until a "full expansion" mutation is achieved, generally several generations following the initial appearance of the disorder in the family. The full expansion mutation is then passed to subsequent generations in a standard Mendelian fashion—for example, autosomal dominant for Huntington disease and sex-linked for fragile-X syndrome.

Mitochondrial DNA Mutations

Disorders resulting from mutations in the mitochondrial genome demonstrate an alternative form of non-Mendelian inheritance, called maternal inheritance, in which the mutation and disorder are passed from mothers—never from fathers—to all of their children. The mutations generally affect the function of the mitochondrion, compromising, among other processes, the production of cellular adenosine triphosphate (ATP). Severity and even penetrance can vary widely for disorders resulting from mutations in the mitochondrial DNA, generally believed to reflect the combined effects of heteroplasmy (i.e., mixed populations of both normal and mutant mitochondrial DNA in a single cell) and other confounding genetic or environmental factors. There are close to 50 mitochondrial genetic diseases currently known.

Imprinted Gene Mutations

Some genetic disorders result from mutations in imprinted genes. Genetic imprinting involves a sex-specific process

of chemical modification to the imprinted genes, so that they are expressed unequally, depending on the sex of the parent of origin. So-called maternally imprinted genes are generally expressed only when inherited from the father, and so-called paternally imprinted genes are generally expressed only when inherited from the mother.

The disease gene associated with Prader-Willi syndrome is maternally imprinted, so that although every child inherits two copies of the gene (one maternal, one paternal), only the paternal copy is expressed. If the paternally inherited copy carries a mutation, the child will be left with no functional copies of the gene expressed, and the clinical traits of Prader-Willi syndrome will result. Similarly, the disease gene associated with Angelman syndrome is paternally imprinted, so that although every child inherits two copies of the gene, only the maternal copy is expressed. If the maternally inherited copy carries a mutation, the child again will be left with no functional copies of the gene expressed, and the clinical traits of Angelman syndrome will result. Individuals who carry the mutation but received it from the "wrong" parent can certainly pass it on to their children, although they will not exhibit clinical features of the disorder.

Upon rare occasion, persons are identified with an imprinted gene disorder who show no family history and do not appear to carry any mutation in the expected gene. These cases result from uniparental disomy, a phenomenon whereby a child is conceived who carries the normal complement of chromosomes but who has inherited both copies of a given chromosome from the same parent, rather than one from each parent, as is the normal fashion. If any key genes on that chromosome are imprinted in the parent of origin, the child may end up with no expressed copies, and a genetic disorder may result. Similarly, other genes may be overexpressed in cases of uniparental disomy,

perhaps also leading to clinical complications. Finally, uniparental disomy can account for very rare instances whereby two parents, only one of whom is a carrier of an autosomal recessive mutation, can nonetheless have an affected child, in the circumstance that the child inherits two mutant copies from the carrier parent.

Diseases Caused by Multifactorial Inheritance

Genetic disorders that are multifactorial in origin represent probably the single largest class of inherited disorders affecting the human population. By definition, these disorders involve the influence of multiple genes, generally acting in concert with environmental factors. Such common conditions as cancer, heart disease, and diabetes are multifactorial disorders. Indeed, improvements in the tools used to study this class of disorders have enabled the assignment of specific contributing gene loci to a number of common traits and disorders. Identification and characterization of these contributing genetic factors may not only enable improved diagnostic and prognostic indicators but may also identify potential targets for future therapeutic intervention.

Because the genetic and environmental factors that underlie multifactorial disorders are often unknown, the risks of recurrence are usually arrived at empirically. In general, it can be said that risks of recurrence are not as great for multifactorial conditions as for single-gene diseases and that the risks vary with the number of relatives affected and the closeness of their relationship. Moreover, close relatives of more severely affected individuals (e.g., those with bilateral cleft lip and cleft palate) are generally at greater risk than those related to persons with a less-severe form of the same condition (e.g., unilateral cleft lip).

VICTOR MCKUSICK

(b. Oct. 21, 1921, Parkman, Maine, U.S.—d. July 22, 2008, Baltimore, Md.)

American physician and genome researcher Victor McKusick pioneered the field of medical genetics.

McKusick was raised on a dairy farm in Maine. He attended Tufts University (1940–43) in Medford, Mass., before transferring to Johns Hopkins University School of Medicine (M.D., 1946) in Baltimore to train as a cardiologist. There he specialized in the study and treatment of heart murmurs. McKusick later published the influential textbook *Cardiovascular Sound in Health and Disease* (1958). An encounter with a heart patient whose malfunctioning aorta was symptomatic of Marfan syndrome, a rare inherited disease, triggered McKusick's switch to genetics. In 1957 he founded the first medical genetics clinic at Johns Hopkins, serving as its director until 1975. McKusick also chaired the department of medicine at Johns Hopkins (1973–85), where he remained as a professor of medical genetics (1985–2007).

McKusick's most significant research included identifying the gene that causes Marfan syndrome and pinpointing the genetic basis for a form of dwarfism known as McKusick-Kaufman syndrome, which is unusually common among Amish people. He was the founding president (1988–91) of the Human Genome Organisation (HUGO) and the creator of the multivolume reference work *Mendelian Inheritance in Man* (12 editions, 1966–98) and its Internet corollary (from 1987), the *Online Mendelian Inheritance in Man* (OMIM).

McKusick was elected to the National Academy of Sciences in 1973. He was the recipient of numerous honours, including Canada's Gairdner Award (1977), the Albert Lasker Award for Special Achievement in Medical Science (1997), the U.S. National Medal of Science (2001), and the Japan Prize in Medical Genomics and Genetics (2008).

GENETICS OF CANCER

Although at least 90 percent of all cancers are sporadic, meaning that they do not seem to run in families, nearly 10 percent of cancers are familial, and some are actually inherited in an apparently autosomal dominant manner.

Cancer may therefore be considered a multifactorial disease, resulting from the combined influence of many genetic factors acting in concert with environmental insults (e.g., ultraviolet radiation, cigarette smoke, and viruses).

Cancers, both familial and sporadic, generally arise from alterations in one or more of three classes of genes: oncogenes, tumour suppressor genes, and genes whose products participate in genome surveillance—for example, in DNA damage repair. For familial cancers, affected members inherit one mutant copy of a gene that falls into one of the latter two classes. That mutation alone is not sufficient to cause cancer, but it predisposes individuals to the disease because they are either more sensitive to spontaneous somatic mutations, as in the case of altered tumour suppressor genes, or are more prone to experience mutations, as in the case of impaired DNA repair enzymes. Of course, sporadic cancers can also arise from mutations in these same classes of genes, but because all of the mutations must arise in the individual *de novo*, as opposed to being inherited, they generally appear only later in life, and they do not run in families.

Retinoblastoma, an aggressive tumour of the eye that typically occurs in childhood, offers perhaps one of the clearest examples of the interplay between inherited and somatic mutations in the genesis of cancer. Data suggest that 60 to 70 percent of all cases of retinoblastoma are sporadic, whereas the rest are inherited. The relevant gene, *RB*, encodes a protein that normally functions as a suppressor of cell cycle progression and is considered a classic tumour suppressor gene. Children who inherit one mutant copy of the *RB* gene are at nearly 100 percent risk to develop retinoblastoma because the probability that their one remaining functional *RB* gene will sustain a mutation in at least one retinal cell is nearly assured. In

contrast, children who inherit two functional copies of the *RB* gene must experience two mutations at the *RB* locus in the same retinal cell in order to develop retinoblastoma; this is a very rare event. This "two-hit" hypothesis of retinoblastoma formation has provided a foundation upon which most subsequent theories of the genetic origins of familial cancer have been built.

Studies of both breast and colorectal cancers have revealed that, similar to retinoblastoma, these cancers are predominantly sporadic, although a small proportion are clearly familial. Sporadic breast cancer generally appears late in life, whereas the familial forms can present much earlier, often before age 40. For familial breast cancer, inherited mutations in one of two specific genes, *BRCA1* and *BRCA2* , account for at least half of the cases observed. The *BRCA1* and *BRCA2* genes both encode protein products believed to function in the pathways responsible for sensing and responding to DNA damage in cells. Whereas a woman in the general population has about a 10 percent lifetime risk of developing breast cancer, half of all women with *BRCA1* or *BRCA2* mutations will develop breast cancer by age 50, and close to 90 percent will develop the disease by age 80. Women with *BRCA1* mutations are also at increased risk to develop ovarian tumours. As with retinoblastoma, both men and women who carry *BRCA1* or *BRCA2* mutations, whether they are personally affected or not, can pass the mutated gene to their offspring, although carrier daughters are much more likely than carrier sons to develop breast cancer.

Two forms of familial colorectal cancer, hereditary nonpolyposis colorectal cancer (HNPCC) and familial adenomatous polyposis (FAP), have also been linked to predisposing mutations in specific genes. Persons with familial HNPCC have inherited mutations in one or more of their DNA mismatch repair genes, predominantly

MSH2 or *MLH1*. Similarly, persons with FAP carry inherited mutations in their *APC* genes, the protein product of which normally functions as a tumour suppressor. For individuals in both categories, the combination of inherited and somatic mutations results in a nearly 100 percent lifetime risk of developing colorectal cancer.

Although most cancer cases are not familial, all are undoubtedly diseases of the genetic material of somatic cells. Studies of large numbers of both familial and sporadic cancers have led to the conclusion that cancer is a disease of successive mutations, acting in concert to deregulate normal cell growth, provide appropriate blood supply to the growing tumour, and ultimately enable tumour cell movement beyond normal tissue boundaries to achieve metastasis (i.e., the dissemination of cancer cells to other parts of the body).

Many of the agents that cause cancer (e.g., X rays, certain chemicals) also cause mutations or chromosome abnormalities. For example, a large fraction of sporadic tumours have been found to carry oncogenes, altered forms of normal genes (proto-oncogenes) that have sustained a somatic "gain-of-function" mutation. An oncogene may be carried by a virus, or it can result from a chromosomal rearrangement, as is the case in chronic myelogenous leukemia, a cancer of the white blood cells characterized by the presence of the so-called Philadelphia chromosome in affected cells. The Philadelphia chromosome arises from a translocation in which one half of the long arm of chromosome 22 becomes attached to the end of the long arm of chromosome 9, creating the dominant oncogene *BCR/abl* at the junction point. The *BCR/abl* fusion protein stimulates the continuous proliferation of bone marrow stem cells that give rise to myelogenous cells. As the number of myelogenous cells in the body increases, symptoms such as fatigue and enlargement of

TELOMERES

Telomeres are segments of DNA occurring at the ends of chromosomes in eukaryotic cells. Telomeres are made up of repeated segments of DNA that consist of the sequence 5'-TTAGGG-3' (in which T, A, and G are the bases thymine, adenine, and guanine, respectively). Some human cells contain as many as 1,500 to 2,000 repeats of this sequence at each end of each chromosome. The number of repeats determines the maximum life span of a cell: each time a cell undergoes replication, multiple TTAGGG segments are lost. Once telomeres have been reduced to a certain size, the cell reaches a crisis point and is prevented from dividing further. As a consequence, the cell dies. Thus, the processes of cell aging and cell death are regulated in part by telomeres.

Telomeres are of special concern in the cellular mechanisms that underlie the development of some types of cancer. Telomeric control of cell life span appears to be inactivated by the expression of oncogenes (cancer-causing genes) or by the deactivation of tumour suppressor genes. In cells undergoing malignant transformation

Elizabeth Blackburn (left), *Jack Szostak, and Carol Greider are the three American scientists who shared the 2009 Nobel Prize for Physiology or Medicine for their work on telomeres. Blackburn made significant discoveries related to the genetic composition and function of telomeres and contributed to the discovery of the telomerase enzyme. Szostak's discoveries on the function of telomeres and Greider's discovery of telomerase in conjunction with Blackburn's work have substantial implications for the future of cancer research.* Justin Sullivan/Getty Images, Jodi Hilton/Getty Images, Chip Somodevilla/Getty Images

(progression to cancer), telomeres do shorten, but, as the crisis point nears, a formerly quiescent enzyme called telomerase becomes activated. This enzyme prevents the telomeres from shortening further and thereby prolongs the life of the cell.

Most malignant tumours—including breast cancer, colorectal cancer, prostate cancer, and ovarian cancer—exhibit telomerase activity. The more advanced the cancer, the greater the frequency of detectable telomerase in independent samples. Because cell immortality contributes to the growth of many cancers, telomerase is an attractive target for the development of new anticancer drugs. Three American scientists, Elizabeth Blackburn, Carol Greider, and Jack Szostak, won the 2009 Nobel Prize for Physiology or Medicine for their work on telomeres.

the spleen appear. Another example is Burkitt lymphoma, in which a rearrangement between chromosomes places the *myc* gene from chromosome 8 under the influence of regulatory sequences that normally control expression of immunoglobulin genes. Deregulation of the myc protein, which is involved in mediating cell cycle progression, is thought to be one of the major steps in the formation of Burkitt lymphoma.

GENETIC DAMAGE FROM ENVIRONMENTAL AGENTS

Exposure to certain chemical agents, both natural and human-made, can cause genetic damage. Among these agents are viruses; compounds produced by plants, fungi, and bacteria; industrial chemicals; products of combustion; alcohol; ultraviolet and ionizing radiation; and even oxygen in the air.

Certain viruses produce genetic damage because they integrate themselves into the genomes of host cells. In some cases, integration disrupts the regulation or function

of a gene that normally controls cell division, thereby causing unregulated cell proliferation. Alternatively, some viruses carry dominant oncogenes in their genomes, which can transform an infected cell and start it on the path toward cancer. Among the viruses that can cause cancer are Epstein-Barr virus, papilloma viruses, hepatitis B and C viruses, retroviruses (e.g., HIV), and herpes virus.

A number of human-made chemicals can also cause genetic disease in humans. Tens of thousands of different chemicals are routinely used in the production of plastics, fuels, food additives, and even medicines. Many of these chemicals are mutagens, and some have been found to be carcinogenic (cancer-producing).

Some products, such as the dibenzodioxins, which are produced as a result of incomplete combustion, are intensely mutagenic and have been demonstrated to cause cancer in laboratory rodents. Dioxins are associated with increased risk of a variety of human cancers.

Decomposition of ozone, which shields the Earth from ultraviolet radiation emitted from the Sun, has contributed to an increase in a variety of skin cancers (including melanoma) in certain areas of the world. Ultraviolet light, when acting on DNA, can lead to covalent linking of adjacent pyrimidine bases. Such pyrimidine dimerization is mutagenic. This damage can be repaired by an enzyme called photolyase, which utilizes the energy of longer wavelengths of light to cleave the dimers. However, people with a defect in the gene coding for photolyase develop xeroderma pigmentosum, a condition characterized by extreme sensitivity to sunlight. These individuals develop multiple skin cancers on all areas of exposed skin, such as the head, neck, and arms.

X rays and gamma rays, which are forms of ionizing radiation, can damage all biological macromolecules, including DNA, proteins, and polysaccharides. Ionizing

radiation has long been recognized as being mutagenic, carcinogenic, and lethal.

Likewise, molecular oxygen (O_2), which is broken down to form the reactive intermediates superoxide (O_{2-}), hydrogen peroxide (H_2O_2), and a hydroxyl radical (HO·), can initiate free-radical oxidation of important metabolites, inactivate certain enzymes, and cause release of iron from specific enzymes. This can subsequently lead to DNA damage, mutagenesis, and cell death.

A few genetic diseases are known to be related to oxygen radicals or to the enzymes that defend against them. One such disease is a familial form of amyotrophic lateral sclerosis (ALS), also known as Lou Gehrig disease, which is characterized by late-onset progressive paralysis due to the loss of motor neurons. Approximately 20 percent of cases of ALS have been shown to result from mutations affecting the enzyme superoxide dismutase, which normally quenches the activity of reactive intermediates and thereby controls the amount of oxidative damage sustained by DNA and other molecules. This form of ALS is genetically dominant, so that the mutant enzyme causes the disease even when half of the superoxide dismutase present in cells exists in the normal form.

MANAGEMENT OF GENETIC DISEASE

The management of genetic disease can be divided into counseling, diagnosis, and treatment. In brief, the fundamental purpose of genetic counseling is to help the individual or family understand their risks and options and to empower them to make informed decisions. Diagnosis of genetic disease is sometimes clinical, based on the presence of a given set of symptoms, and sometimes molecular, based on the presence of a recognized

gene mutation, whether clinical symptoms are present or not. The cooperation of family members may be required to achieve diagnosis for a given individual, and, once accurate diagnosis of that individual has been determined, there may be implications for the diagnoses of other family members. Balancing privacy issues within a family with the ethical need to inform individuals who are at risk for a particular genetic disease can become extremely complex.

Although effective treatments exist for some genetic diseases, for others there are none. It is perhaps this latter set of disorders that raises the most troubling questions with regard to presymptomatic testing, because phenotypically healthy individuals can be put in the position of hearing that they are going to become ill and potentially die and that there is nothing they or anyone else can do to stop it. Fortunately, with time and research, this set of disorders is slowly becoming smaller.

GENETIC COUNSELING

Genetic counseling represents the most direct medical application of the advances in understanding of basic genetic mechanisms. Its chief purpose is to help people make responsible and informed decisions concerning their own health or that of their children. Genetic counseling, at least in democratic societies, is nondirective; the counselor provides information, but decisions are left up to the individual or the family.

Calculating Risks of Known Carriers

Most couples who present themselves for preconceptional genetic counseling fall into one of two categories: those who have already had a child with genetically based

problems, and those who have one or more relatives with a disease they think might be inherited. The counselor must confirm the diagnosis in the affected person with meticulous accuracy, so as to rule out the possibility of alternative explanations for the clinical symptoms observed. A careful family history permits construction of a pedigree that may illuminate the nature of the inheritance (if any), may affect the calculation of risk figures, and may bring to light other genetic influences. The counselor, a certified health-care professional with special training in medical genetics, must then decide whether the disease in question has a strong genetic component and, if so, whether the heredity is single-gene, chromosomal, or multifactorial.

In the case of single-gene Mendelian inheritance, the disease may be passed on as an autosomal recessive, autosomal dominant, or sex-linked recessive trait. If the prospective parents already have a child with an autosomal recessive inherited disease, they both are considered by definition to be carriers, and there is a 25 percent risk that each future child will be affected. If one of the parents carries a mutation known to cause an autosomal dominant inherited disease, whether that parent is clinically affected or not, there is a 50 percent risk that each future child will inherit the mutation and therefore may be affected. If, however, the couple has borne a child with an autosomal dominant inherited disease though neither parent carries the mutation, then it will be presumed that a spontaneous mutation has occurred and that there is not a markedly increased risk for recurrence of the disease in future children. There is a caveat to this reasoning, however, because there is also the possibility that the new mutation might have occurred in a progenitor germ cell in one of the parents, so that some unknown proportion of

that individual's eggs or sperm may carry the mutation, even though it is absent from the somatic cells—including blood, which is generally the tissue sampled for testing. This scenario is called germline mosaicism. Finally, with regard to X-linked disorders, if the pedigree or carrier testing suggests that the mother carries a gene for a sex-linked disease, there is a 50 percent chance that each son will be affected and that each daughter will be a carrier.

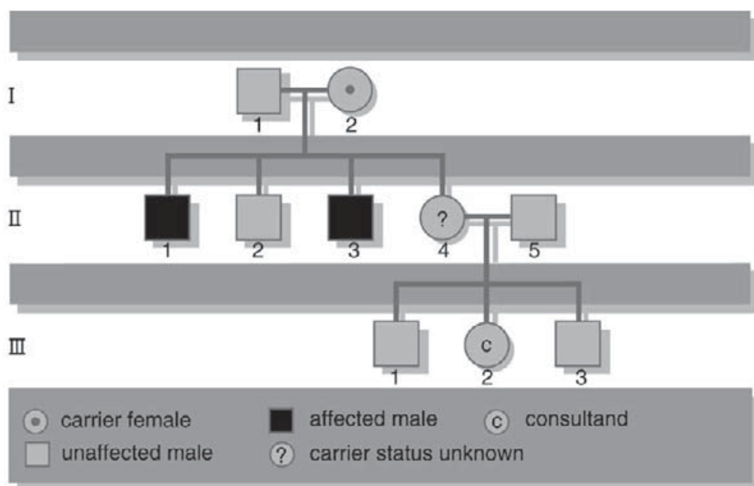
Counseling for chromosomal inheritance most frequently involves either an inquiring couple (consultands) who have had a child with a known chromosomal disorder, such as Down syndrome, or a couple who have experienced multiple miscarriages. To provide the most accurate recurrence risk values to such couples, both parents should be karyotyped to determine if one may be a balanced translocation carrier. Balanced translocations refer to genomic rearrangements in which there is an abnormal covalent arrangement of chromosome segments, although there is no net gain or loss of key genetic material. If both parents exhibit completely normal karyotypes, the recurrence risks cited are low and are strictly empirical.

Most of the common hereditary birth defects, however, are multifactorial. If the consulting couple have had one affected child, the empirical risk for each future child will be about 3 percent. If they have borne two affected children, the chance of recurrence will rise to about 10 percent. Clearly these are population estimates, so that the risks within individual families may vary.

Estimating Probability: Bayes's Theorem

The calculation of risks is relatively straightforward when the consultands are known carriers of diseases due to single genes of major effect that show regular Mendelian

inheritance. For a variety of reasons, however, the parental genotypes frequently are not clear and must be approximated from the available family data. Bayes's theorem, a statistical method first devised by the English clergyman-scientist Thomas Bayes in 1763, can be used to assess the relative probability of two or more alternative possibilities (e.g., whether a consultand is or is not a carrier). The likelihood derived from the appropriate Mendelian law (prior probability) is combined with any additional information that has been obtained from the consultand's family history or from any tests performed (conditional probability). A joint probability is then determined for each alternative



Pedigree of a family with a history of Duchenne muscular dystrophy, which is carried by females (circles) and affects half of a carrier's male children (squares). The consultand (III-2) wishes to know her risk of having an affected child. Since her grandmother (I-2) was a known carrier, the chances of her mother (II-4) having been a carrier are 1/5. Her own chances of being a carrier are therefore 1/5 × 1/2 = 1/10, and her chances of passing the syndrome to a male child are 1/10 × 1/2 × 1/2 = 1/40. Molecular testing would establish her status with certainty as either a carrier or a noncarrier, and consequently the chance of her male child having the disease would be either 1/2 or 0. Encyclopaedia Britannica, Inc.

outcome by multiplying the prior probability by all conditional probabilities.

By dividing the joint probability of each alternative by the sum of all joint probabilities, the posterior probability is arrived at. Posterior probability is the likelihood that the individual, whose genotype is uncertain, either carries the mutant gene or does not. One example application of this method is for the sex-linked recessive disease Duchenne muscular dystrophy (DMD). In this example, the consultand wishes to know her risk of having a child with DMD. It is known that the consultand's grandmother is a carrier, since she had two affected sons (spontaneous mutations occurring in both brothers would be extremely unlikely). What is uncertain is whether the consultand's mother is also a carrier. The Bayesian method for calculating the consultand's risk is as follows:

If the consultand's mother is a carrier (risk = 1/5), then there is a 1/2 chance that the consultand is also a carrier, so her total empirical risk is 1/5 × 1/2 = 1/10. If she becomes pregnant, there is a 1/2 chance that her child will be male and a 1/2 chance that the child, regardless of sex, will inherit the familial mutation. Hence, the total empirical risk for the consultand to have an affected child is 1/10 × 1/2 × 1/2 = 1/40. Of course, if the familial mutation is known, presumably from molecular testing of an affected family member, the carrier status of the consultand could be determined directly by molecular analysis, rather than estimated by Bayesian calculation. If the family is cooperative and an affected member is available for study, this is clearly the most informative route to follow because the risk for the consultand to carry the familial mutation would be either 1 or 0, and not 1/10. If her risk is 1, then each of her sons will have a 1/2 chance of being affected. If her risk is 0, none of her children will be affected (unless a new mutation occurs, which is very rare).

After determining the nature of the heredity, the counselor discusses with the consultand the likely risks and the available options to minimize impact of those risks on the individual and the family. In the case of a couple in which one member has a family history of a genetic disorder—for example, cystic fibrosis—typical options might include any of the following choices: (1) Accept the risks and take a chance that any future children may be affected. (2) Seek molecular testing for known mutations of cystic fibrosis in relevant family members to determine with greater accuracy whether either or both prospective parents are carriers for this recessive disorder. (3) If both members of the couple are determined to be carriers, utilize donor sperm for artificial insemination. This option is a good genetic solution only if the husband carries a dominant mutation, or if both parents are carriers of a recessive mutation. If the recessive trait is reasonably common, as are mutations for cystic fibrosis, however, it would be reasonable to ask that the sperm donor be checked for carrier status before pursuing this option. (4) Proceed with natural reproduction, but pursue prenatal diagnosis with the possibility of selective termination of an affected pregnancy, if desired by the parents. (5) Pursue in vitro fertilization with donor eggs, if the woman is the at-risk partner, or use both eggs and sperm from the couple but employ preimplantation diagnostics to select only unaffected embryos for implantation. (6) Decide against biological reproduction because the risks and available options are unacceptable; possibly pursue adoption.

PRENATAL DIAGNOSIS

Perhaps one of the most sensitive areas of medical genetics is prenatal diagnosis, the genetic testing of an unborn fetus, because of fears of eugenic misuse or because some

couples may choose to terminate a pregnancy depending on the outcome of the test. Nonetheless, prenatal testing in one form or another is now almost ubiquitous in most industrialized nations, and recent advances both in testing technologies and in the set of "risk factor" genes to be screened promise to make prenatal diagnosis even more widespread. Indeed, parents may soon be able to ascertain information not only about the sex and health status of their unborn child but also about his or her complexion, personality, and intellect. Whether parents should have access to all of this information and how they may choose to use it are matters of much debate.

Current forms of prenatal diagnosis can be divided into two classes, those that are apparently noninvasive and those that are more invasive. At present the noninvasive tests are generally offered to all pregnant women, while the more-invasive tests are generally recommended only if some risk factors exist. The noninvasive tests include ultrasound imaging and maternal serum tests. Serum tests include one for alpha-fetoprotein (AFP) or one for alphafetoprotein, estriol, and human chorionic gonadotropin (triple screen). These tests serve as screens for structural fetal malformations and for neural tube closure defects. The triple screen can also detect some cases of Down syndrome, although there is a significant false-positive and false-negative rate.

More-invasive tests include amniocentesis, chorionic villus sampling, percutaneous umbilical blood sampling, and, upon rare occasion, preimplantation testing of either a polar body or a dissected embryonic cell. Amniocentesis is a procedure in which a long, thin needle is inserted through the abdomen and uterus into the amniotic sac, enabling the removal of a small amount of the amniotic fluid bathing the fetus. This procedure is generally performed under ultrasound guidance between the 15th and

17th weeks of pregnancy, and, although it is generally regarded as safe, complications can occur, ranging from cramping to infection or loss of the fetus. The amniotic fluid obtained can be used in each of three ways: (1) living fetal cells recovered from this fluid can be induced to grow and can be analyzed to assess chromosome number, composition, or structure; (2) cells recovered from the fluid can be used for molecular studies; and (3) the amniotic fluid itself can be analyzed biochemically to determine the relative abundance of a variety of compounds associated with normal or abnormal fetal metabolism and development. Amniocentesis is typically offered to pregnant women over age 35 because of the significantly increased rate of chromosome disorders observed in the children of older mothers. A clear advantage of amniocentesis is the wealth of material obtained and the relative safety of the procedure. The disadvantage is timing: results may not be received until the pregnancy is already into the 19th week or beyond, at which point the possibility of termination may be much more physically and emotionally wrenching than if considered earlier.

Chorionic villus sampling (CVS) is a procedure in which either a needle is inserted through the abdomen or a thin tube is inserted into the vagina and cervix to obtain a small sample of placental tissue called chorionic villi. CVS has the advantage of being performed earlier in the pregnancy (generally 10–11 weeks), although the risk of complications is greater than that for amniocentesis. Risks associated with CVS include fetal loss and fetal limb reduction if the procedure is performed earlier than 10 weeks gestation. Another disadvantage of CVS reflects the tissue sampled: chorionic villi are not part of the embryo, and such a sample may not accurately represent the embryonic genetic constitution. In contrast, amniotic

cells are embryonic in origin, having been sloughed off into the fluid. Therefore, abnormalities, often chromosomal, may be seen in the chorionic villi but not in the fetus, or vice versa.

Both percutaneous umbilical blood sampling (PUBS) and preimplantation testing are rare, relatively high-risk, and performed only in very unusual cases. Preimplantation testing of embryos derived by in vitro fertilization is used only in cases of couples who are at high risk for having a fetus affected with a given familial genetic disorder and who find all other alternatives unacceptable. Preimplantation testing involves obtaining eggs and sperm from the couple, combining them in the laboratory, and allowing the resultant embryos to grow until they reach the early blastocyst stage of development, at which point a single cell is removed from the rest and harvested for fluorescence in situ hybridization (FISH) or other types of molecular analysis. The problem with this procedure is that one cell is scant material for diagnosis, so that a large array of tests cannot be performed. Similarly, if the test fails for any technical reason, it cannot be repeated. Finally, embryos determined to be normal and therefore selected for implantation into the mother are subject to other complications normally associated with in vitro fertilization—namely, that only a small fraction of the implanted embryos make it to term and that multiple, and therefore high-risk, pregnancies are common. Nonetheless, many at-risk couples find these complications easier to accept than the elective termination of the pregnancy.

The most common prenatal screening test is an assay of alpha-fetoprotein (AFP) in the maternal serum. Elevated levels are associated with neural tube defects in the fetus such as spina bifida (defective closure of the spine) and anencephaly (absence of brain tissue). When

alpha-fetoprotein levels are elevated, a more specific diagnosis is attempted using ultrasonography and amniocentesis to analyze the amniotic fluid for the presence of alpha-fetoprotein and acetylcholinesterase.

Other procedures for prenatal tests are being investigated. For example, fetal cells are known to circulate in the maternal blood, and procedures to isolate and analyze these cells are under development. These procedures, if successful, would offer a noninvasive alternative for prenatal testing.

GENETIC TESTING

In the case of genetic disease, options often exist for presymptomatic diagnosis—that is, diagnosis of individuals at risk for developing a given disorder, even though at the time of diagnosis they may be clinically healthy. Options also may exist for carrier testing, studies that determine whether an individual is at increased risk of having a child with a given disorder, even though he or she personally may never display symptoms. Accurate predictive information can enable early intervention, which often prevents the clinical onset of symptoms and the irreversible damage that may have already occurred by waiting for symptoms and then responding to them. In the case of carrier testing, accurate information can enable prospective parents to make more-informed family-planning decisions. Unfortunately, there can also be negative aspects to early detection, including such issues as privacy, individual responses to potentially negative information, discrimination in the workplace, or discrimination in access to or cost of health or life insurance. While some governments have outlawed the use of presymptomatic genetic testing information by insurance companies and

employers, others have embraced it as a way to bring spiraling health-care costs under control. Some communities have even considered instituting premarital carrier testing for common disorders in the populace.

Genetic testing procedures can be divided into two different groups: (1) testing of individuals considered at risk from phenotype or family history and (2) screening of entire populations, regardless of phenotype or personal family history, for evidence of genetic disorders common in that population. Both forms are currently pursued in many societies. Indeed, with the explosion of information about the human genome and the increasing identification of potential "risk genes" for common disorders, such as cancer, heart disease, or diabetes, the role of predictive genetic screening in general medical practice is increasing.

Adults are generally tested for evidence of genetic disease only if personal or family history suggests they are at increased risk for a given disorder. A typical example would be a young man whose father, paternal aunt, and older brother have all been diagnosed with early onset colorectal cancer. Although this person may appear perfectly healthy, he is at significantly increased risk to carry mutations associated with familial colorectal cancer, and accurate genetic testing could enable heightened surveillance (e.g., frequent colonoscopies) that might ultimately save his life.

Carrier testing for adults in most developed nations is generally offered only if family history or ethnic origins suggest an increased risk of having a particular disease. A typical example would be to offer carrier testing for cystic fibrosis to a couple including one member who has a sibling with the disorder. Another would be to offer carrier testing for Tay-Sachs disease to couples of Ashkenazic

Jewish origin, a population known to carry an increased frequency of Tay-Sachs mutations. The same would be true for couples of African or Mediterranean descent with regard to sickle cell anemia or thalassemia, respectively. Typically, in each of these cases a genetic counselor would be involved to help the individuals or couples understand their options and make informed decisions.

Genetic tests themselves can take many forms, and the choice of tests depends on a number of factors. For example, screening for evidence of sickle cell anemia, a hemoglobin disorder, is generally pursued at least initially by tests involving the hemoglobin proteins themselves, rather than DNA, because the relevant gene product (blood) is readily accessible, and because the protein test is currently cheaper to perform than the DNA test. In contrast, screening for cystic fibrosis, a disorder that predominantly affects the lungs and pancreas, is generally pursued in the at-risk newborn at the level of DNA because there is no cheap and accurate alternative. Older persons suspected of having cystic fibrosis, however, can also be diagnosed with a "sweat test" that measures sweat electrolytes.

Tests involving analysis of DNA are particularly powerful because they can be performed using very tiny samples; also, the DNA tested can originate from almost any tissue type, regardless of whether the gene of interest happens to be expressed in that tissue. Current technologies applied for mutation detection include traditional karyotyping and Southern blotting, as well as a multitude of new tests, including FISH with specific probes or the polymerase chain reaction (PCR). Which tests are applied depends on whether the genetic abnormalities are likely to be chromosomal (in which case karyotyping or FISH are appropriate), large deletions or other rearrangements

(best tested for by Southern blotting or PCR), or point mutations (best confirmed by PCR followed by oligonucleotide hybridization or restriction enzyme digestion). If a large number of different point mutations are sought, as is often the case, the most appropriate technology may be microarray hybridization analysis, which can test for tens to hundreds of thousands of different point mutations in the same sample simultaneously.

EUGENICS

Sir Francis Galton, detail of an oil painting by G. Graef, 1882; in the National Portrait Gallery, London. Courtesy of The National Portrait Gallery, London

Eugenics is the selection of desired heritable characteristics in order to improve future generations, typically in reference to humans. The term *eugenics* was coined in 1883 by the British explorer and natural scientist Francis Galton, who, influenced by Charles Darwin's theory of natural selection, advocated a system that would allow "the more suitable races or strains of blood a better chance of prevailing speedily over the less suitable." Social Darwinism, the popular theory in the late 19th century that life for humans in society was ruled by "survival of the fittest," helped advance eugenics into serious scientific study in the early 1900s. By World War I, many scientific authorities and political leaders supported eugenics. However, it ultimately failed as a science in the 1930s and 1940s, when the assumptions of eugenicists became heavily criticized and the Nazis used eugenics to support the extermination of entire races.

Early History

Although eugenics as understood today dates from the late 19th century, efforts to select matings in order to secure offspring with desirable traits date from ancient times. Plato's *Republic* (*c.* 378 BCE) depicts a society where efforts are undertaken to improve human beings through selective breeding. Later, Italian philosopher and poet Tomasso Campenella, in *City of the Sun* (1623), described a utopian community in which only the socially elite are allowed to procreate. Galton, in *Hereditary Genius* (1869), proposed that a system of arranged marriages between men of distinction and women of wealth would eventually produce a gifted race. In 1865, the basic laws of heredity were discovered by the father of modern genetics, Gregor Mendel. His experiments with peas demonstrated that each physical trait was the result of a combination of two units (now known as genes) and could be passed from one generation to another. However, his work was largely ignored until its rediscovery in 1900. This fundamental knowledge of heredity provided eugenicists—including Galton, who influenced his cousin Charles Darwin—with scientific evidence to support the improvement of humans through selective breeding.

The advancement of eugenics was concurrent with an increasing appreciation of Charles Darwin's account for change or evolution within society—what contemporaries referred to as social Darwinism. Darwin had concluded his explanations of evolution by arguing that the greatest step humans could make in their own history would occur when they realized that they were not completely guided by instinct. Rather, humans, through selective reproduction, had the ability to control their own future evolution. A language pertaining to reproduction and eugenics developed, leading to terms such as *positive eugenics*, defined as promoting the proliferation of "good stock," and *negative eugenics*, defined as prohibiting marriage and breeding between "defective stock." For eugenicists, nature was far more contributory than nurture in shaping humanity.

Popular Support for Eugenics

In 1910, the Eugenics Record Office (ERO), a clearinghouse for eugenics information, was opened at Cold Spring Harbor, N.Y. During the 1930s, eugenics gained considerable popular support across the United

States. Hygiene courses in public schools and eugenics courses in colleges spread eugenic-minded values to many. A eugenics exhibit titled "Pedigree-Study in Man" was featured at the Chicago World's Fair in 1933–34. Consistent with the fair's "Century of Progress" theme, stations were organized around efforts to show how favourable traits in the human population could best be perpetuated. Contrasts were drawn between the distinguished, presidential Roosevelt family and the degenerate "Ishmael" family (one of several pseudonymous family names used, the rationale for which was not given). Fairgoers were urged to adopt the view that responsible individuals should pursue marriage ever mindful of eugenics principles. Booths were set up at state fairs promoting "fitter families" contests, and medals were awarded to eugenically sound families. Popular eugenic advertisements claimed it was about time that humans received the same attention in the breeding of better babies that had been given to livestock and crops for centuries.

Antieugenics Sentiment

Antieugenics sentiment began to appear after 1910 and intensified during the 1930s. Most commonly it was based on religious grounds. For example, the 1930 papal encyclical *Casti connubii* condemned reproductive sterilization, though it did not specifically prohibit positive eugenic attempts to amplify the inheritance of beneficial traits. Many Protestant writings sought to reconcile age-old Christian warnings about the heritable sins of the father to pro-eugenic ideals. Indeed, most of the religion-based popular writings of the period supported positive means of improving the physical and moral makeup of humanity.

In the early 1930s, Nazi Germany adopted American measures to identify and selectively reduce the presence of those deemed to be "socially inferior" through involuntary sterilization. A rhetoric of positive eugenics in the building of a master race pervaded *Rassenhygiene* (racial hygiene) movements. When Germany extended its practices far beyond sterilization in efforts to eliminate the Jewish and other non-Aryan populations, the United States became increasingly concerned over its own support of eugenics. Many scientists, physicians, and political leaders began to denounce the work of the ERO publicly. After considerable reflection, the Carnegie Institution formally closed the ERO at the end of 1939.

During the aftermath of World War II, eugenics became stigmatized such that many individuals who had once hailed it as a science now spoke disparagingly of it as a failed pseudoscience. *Eugenics* was dropped from organization and publication names. In 1954, Britain's *Annals of Eugenics* was renamed *Annals of Human Genetics*. In 1972, the American Eugenics Society adopted the less-offensive name Society for the Study of Social Biology. Its publication, once popularly known as the *Eugenics Quarterly*, had already been renamed *Social Biology* in 1969.

U.S. Senate hearings in 1973, chaired by Edward Kennedy, revealed that thousands of U.S. citizens had been sterilized under federally supported programs. The U.S. Department of Health, Education, and Welfare proposed guidelines encouraging each state to repeal their respective sterilization laws. Other countries, most notably China, continue to support eugenics-directed programs openly in order to ensure the genetic makeup of their future.

The "New Eugenics"

Despite the dropping of the term *eugenics*, eugenic ideas remain prevalent in many issues surrounding human reproduction. Medical genetics, a post–World War II medical specialty, encompasses a wide range of health concerns, from genetic screening and counseling to fetal gene manipulation and the treatment of adults suffering from hereditary disorders. Because certain diseases (e.g., hemophilia and Tay-Sachs disease) are now known to be genetically transmitted, many couples choose to undergo genetic screening, in which they learn the chances that their offspring have of being affected by some combination of their hereditary backgrounds. Couples at risk of passing on genetic defects may opt to remain childless or to adopt children. Furthermore, it is now possible to diagnose certain genetic defects in the unborn. Many couples choose to terminate a pregnancy that involves a genetically disabled offspring. These developments have reinforced the eugenic aim of identifying and eliminating undesirable genetic material. Counterbalancing this trend, however, has been medical progress that enables victims of many genetic diseases to live fairly normal lives. Direct manipulation of harmful genes is also being studied. If perfected, it could obviate eugenic arguments for restricting reproduction among those who carry harmful genes. Such

conflicting innovations have complicated the controversy surrounding what many call the "new eugenics." Moreover, suggestions for expanding eugenics programs, which range from the creation of sperm banks for the genetically superior to the potential cloning of human beings, have met with vigorous resistance from the public, which often views such programs as unwarranted interference with nature or as opportunities for abuse by authoritarian regimes.

Applications of the Human Genome Project are often referred to as "Brave New World" genetics or the "new eugenics"; however, the ethical, legal, and social implications of this international project are monitored much more closely than were early 20th-century eugenics programs. Still, with or without the use of the term, many eugenics-related concerns are reemerging as a new group of individuals decide how to regulate the application of genetics science and technology. This gene-directed activity, in attempting to improve upon nature, may not be that distant from what Galton implied in 1909 when he described eugenics as the "study of agencies, under social control, which may improve or impair" future generations.

THE FUTURE OF GENETICS IN MEDICINE

Today genetics continues to advance at an extraordinary pace. The development of new technologies have enabled faster sequencing of DNA and have accelerated the discovery of genes and their functions. These advances, as well as the identification of genetic mutations that contribute to certain human diseases and improvements in the sensitivity of genetic testing methods, promise to speed the diagnosis of numerous conditions. Genetic diseases also are targets of personalized medicine. This approach to treatment is based on the concept that genetic screening for specific single nucleotide polymorphisms (or SNPs) in a person's genome can be used to select drugs most appropriate for that individual. Personalized medicine could be used to avoid potentially

dangerous drug responses that are the result of altered cellular metabolism caused by a specific SNP.

With the completion of sequencing of the human genome in 2003, scientists found at their disposal a practically endless volume of genetic information. Although many genes had already been identified by that time, scientists were only on the brink of genomics. The human genome has so far revealed valuable information about the evolutionary history of the human species and the relationship between humans and other primates. Further studies are expected to uncover details that help explain certain hereditary phenomena, such as inherited disease and epigenetics. Investigation of the human genome has been facilitated by the completion of genomic sequencing for a wide variety of other organisms. Since the mid-1990s, the genomes of more than 180 organisms have been sequenced in their entirety. Included among these organisms are: *Mus musculus*, the house mouse; *Anopheles gambiae*, a species of malaria-transmitting mosquito; *Bacillus anthracis*, the bacterium that causes anthrax; *Canis familiaris*, the domestic dog; *Oryza sativa*, the cultivated rice plant; *Gallus gallus*, the chicken; and *Deinococcus radiodurans*, a species of radiation-resistant bacteria.

Information culled from the DNA sequence of the human genome is also expected to provide insight into the genes that dictate human development. The interactions of genes, as demonstrated by cloning and stem cell research, plays a fundamental role in development. However, little is understood about how these interactions occur or what specific cell activities they function to control. Investigations aimed at answering these questions as they relate to humans are being aided by basic research into the genetic mechanisms underlying development and disease in organisms such as fruit flies and

nematodes. Improvements in rodent models of genetic disease also are proving to be of significance to these studies.

As genetics continues to race into the 21st century, scientists are anticipating an explosion of new information and knowledge about the human species. There is agreement too that this wealth of new information will reveal fascinating and previously unknown details about life on Earth as a whole.

Glossary

allele One of several forms of a gene that can be found at a particular location on a chromosome. Each allele produces a specific characteristic, such as eye colour or blood type. Usually one form of the allele will be expressed more than another in an individual.

amino acids Any of 20 different kinds of basic molecules that link into long chains to form proteins. They are called the building blocks of proteins.

anneal To bond.

autosomal dominance Heterozygous condition in which one inherited gene mutation on a non-sex chromosome is expressed as a trait or disorder in an individual.

autosome A chromosome that is not a sex chromosome. Autosomes control the inheritance of all characteristics except for gender.

cell The basic structural unit in any living organism. It is made up of a small, semipermeable, watery compartment filled with cytoplasm, a nucleus, various organelles, and a complete copy of the organism's genome in the form of DNA.

centromere The part of a chromosome at which both chromatids are joined. When a cell undergoes mitosis, the spindle fibre attaches to the centromere before the chromatids separate.

chromatin Genetic material consisting of DNA and proteins contained in a cell's nucleus. During cell division, chromatin condenses to form chromosomes.

chromosome Thread-like structure in the nucleus of eukaryotic cells consisting of tightly condensed DNA and protein that carries genes. Chromosomes occur in sets of two, one set inherited from the mother and

one from the father, and vary in number depending on the organism.

codons Set of three adjacent bases on a DNA or RNA chain that code for a special amino acid and allow the creation of a protein molecule.

deletion The loss of genetic material from a chromosome.

diploid Consisting of two sets of chromosomes, or double the haploid.

dominance Prevalence of one allele over another for a given trait. When an allele is dominant, the heterozygote will manifest the same trait as a homozygote with identical alleles.

DNA (deoxyribonucleic acid) A complex molecule, made up of a string of tiny molecules, that exists inside the nucleus of every cell and contains the genetic instructions for making living organisms.

enzyme A protein that promotes a biochemical reaction, most often causing it to accelerate.

eukaryote A complex, multicelled organism such as an animal, a plant, or a fungus, as opposed to a prokaryote.

gamete A mature reproductive cell that contains one set of unpaired chromosomes (i.e., the haploid number) and combines with another gamete from the opposite sex during the fertilization process.

gene Segment of DNA located on a chromosome that determines heredity by controlling protein synthesis.

genome The complete genetic content of an organism, including its full set of chromosomes in the nucleus; the term is sometimes extended to include the mitochondrial genome as well.

genotype An organism's genetic composition.

germ line The cells that produce gametes.

haploid Consisting of a single set of chromosomes, or half the diploid.

heterozygote An organism that has two different forms of a certain gene, one inherited from each parent.

homozygote An organism that has two identical forms of a certain gene, one inherited from each parent.

intron Non-coding DNA sequence that is transcribed to mRNA and then spliced prior to RNA maturation.

karyotype The composition and arrangement of chromosomes in a cell's nucleus.

meiosis Part of the process of sexual reproduction in which diploid (double) chromosomes are reduced to haploid (single) ones, which leads to the production of sex cells, including spores in plants and gametes in animals.

mitochondria Organelles contained in the cytoplasm of most eukaryotic cells. Mitochondria produce energy (adenosine triphosphate) for the cell.

mitochondrial DNA (mtDNA) A type of extranuclear DNA located in mitochondria. The genes contained in mtDNA facilitate the energy conversion of a cell.

mitosis Somatic cell division in which one cell nucleus divides and produces two "daughter" cells with the same genetic information as the parent cell.

mutation A permanent structural change in the genome of a cell.

nucleotides Molecules that make up the structure of RNA and DNA when they are joined together.

phenotype All traits and characteristics that can be observed and are produced by the interaction of the genotype with the environment.

polymer An organic or synthetic compound made up of large connected molecules.

prokaryote An organism, often single-celled, that lacks a clearly defined nucleus.

protein An organic compound composed of one or more chains of amino acids and arranged in a complex

structure. A protein can contain other components, such as sugar or fat, in addition to amino acids.

recessive Describing an allele that produces a trait observable only in the homozygous condition but hidden in the heterozygous condition.

RNA (ribonucleic acid) Nucleic acid structurally composed of ribose and phosphate units with uracil as a base rather than the thymine of DNA. RNA is critical to the process of protein synthesis as well as other cell functions.

somatic Related to any cell of the body with the exception of germ cells.

transposons Stretches of DNA that are unstable and able to move around, both on or between chromosomes. Also called jumping DNA or jumping genes.

vector A carrier of genetic material, often a segment of DNA.

For Further Reading

Barnes, Barry, and John Dupré. *Genomes and What to Make of Them*. Chicago, IL: University of Chicago Press, 2008.

Cavalli-Sforza, Luigi Luca. *Genes, Peoples, and Languages*. Berkeley, CA: University of California Press, 2001.

Conner, Jeffrey K., and Daniel L. Hartl. *A Primer of Ecological Genetics*. Sunderland, MA: Sinauer Associates, 2004.

Dawkins, Richard. *The Selfish Gene*. 30th Anniv. ed. New York, NY: Oxford University Press, 2006.

DeSalle, Rob, and Michael Yudell. *Welcome to the Genome: A User's Guide to the Genetic Past, Present, and Future*. Hoboken, NJ: John Wiley & Sons, 2005.

Keller, Evelyn Fox. *The Century of the Gene*. Cambridge, MA: Harvard University Press, 2002.

Lewontin, Richard. *The Triple Helix: Gene, Organism, and Environment*. Cambridge, MA: Harvard University Press, 2000.

Olson, Steve. *Mapping Human History: Discovering the Past Through Our Genes*. New York, NY: Houghton Mifflin, 2002.

Reilly, Philip R. *Abraham Lincoln's DNA and Other Adventures in Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Press, 2002.

Ridley, Mark. *The Cooperative Gene: How Mendel's Demon Explains the Evolution of Complex Beings*. New York, NY: Simon and Schuster, 2001.

Ridley, Matt. *Genome: Autobiography of a Species in 23 Chapters*. New York, NY: HarperCollins, 2006.

Ridley, Matt. *The Red Queen: Sex and Evolution of Human Nature*. New York, NY: HarperCollins, 2003.

Smolenyak, Megan Smolenyak, and Ann Turner. *Trace Your Roots with DNA: Using Genetic Tests to Explore Your Family Tree*. New York, NY: Rodale, 2004.

Stone, Linda, Paul F. Lurquin, and L. Luca Cavalli-Sforza. *Genes, Culture, and Human Evolution*. Malden, MA: Blackwell, 2007.

Sykes, Bryan. *The Seven Daughters of Eve: The Science That Reveals Our Genetic Ancestry*. New York, NY: W. W. Norton, 2001.

Tudge, Colin. *The Impact of the Gene: From Mendel's Peas to Designer Babies*. New York, NY: Farrar, Straus, & Giroux, 2000.

Watson, James D., and Andrew Berry. *DNA: The Secret of Life*. New York, NY: Random House, 2003.

Wells, Spencer. *Deep Ancestry: Inside the Genographic Project*. Washington, DC: National Geographic, 2007.

Wells, Spencer. *The Journey of Man: A Genetic Odyssey*. New York, NY: Random House, 2002.

Wilmut, Ian, and Roger Highfield. *After Dolly: The Uses and Misuses of Human Cloning*. New York, NY: W. W. Norton, 2007.

Index

A

achondroplasia, 228
alleles, about, 26, 39–40
alpha-fetoprotein (AFP), screening for in maternal serum, 249, 251–252
amniocentesis, 249–250, 252
animals
  breeding of, 22, 101, 137, 158, 174–189
  and breeding systems, 187–189
  cloning of, 137, 147, 148–153, 155, 158
  evaluation of in breeding, 183–184
  heritability and genetic correlations in breeding, 179–181
  methods of selection in breeding, 181–187
  and progeny testing, 186–187
antibody formation, genetics of, 202–204
Arber, Werner, 124
Aristotle, 23, 24
autosomal dominant inheritance, diseases of, 226–229, 235
autosomal recessive inheritance, diseases of, 229–231
autosomes, 43, 44, 86
Avery, Oswald, 31, 34, 58

B

Bakewell, Robert, 175
Barr body, 45, 225
Bateson, William, 29
Bayes's theorem, 245–247
Beadle, George, 30
behaviour genetics, 104–108
  methods of study, 105–108
Benzer, Seymour, 31
Blackburn, Elizabeth, 240
blood, genetics of human, 198–202
blood types, 199–200
bone marrow transplants, 117
Boyer, Herbert W., 124
breast cancer, 237
breeding
  animal, 101, 158, 174–189
  plant, 158–174
Brenner, Sydney, 31–32
Bridges, Calvin Blackman, 50, 51
Briggs, Robert W., 147

C

cancer, genetics of, 235–240, 241
cell theory of fertilization, 27–28
cellular immunity, genetics of, 204–205
Chargaff, Erwin, 59–60
Chase, Martha, 58
chorionic villus sampling (CVS), 249, 250–251
chromosomal aberrations, 52–56
  aneuploidy, 55–56
  deletions, 52–53, 54
  diseases caused by, 221–223
  duplications, 53, 54
  inversions, 53–54

polyploidy, 54–55
translocations, 54
chromosomes
about, 41–56
about human, 190–191
definition of, 30
classical genetics, 85–86
cloning, 123, 137, 146–157, 260
and ethical controversies, 147, 150, 151, 155–157, 259
mammalian, 151–152, 155–156, 158
reproductive, 137, 149–151, 155, 158
therapeutic, 153–155, 156–157
Cohen, Stanley N., 124
colorectal cancer, 237–238, 253
congenital defects, 220
Correns, Carl Erich, 28–29
Creighton, Harriet, 83
Crick, Francis, 31–32, 33–34, 35, 36, 37, 60
crossbreeding of animals, 187–188
crossing over, 47, 51, 53, 54, 96
cross-pollinated species of plants, breeding, 171–174
cross-pollination, experiments with, 25–28
cystic fibrosis, 248, 253, 254
cytogenetics, 86–88, 190

D

Darwin, Charles, 24, 29, 48–49, 51, 105, 217, 255, 256
de Vries, Hugo, 28–29
diseases, genetic, 220–261
classes of, 221–234
and genetic counseling, 243–248
and genetic testing, 252–255
management of, 242–243
and prenatal diagnosis, 248–252
DNA, 57–84
early discoveries, 31–33, 57–58
and expression of genetic code, 64–71
and genetic engineering, 123–155
in humans, 191–193, 195, 196, 197–198
and PCR, 90–93
repetitive, 81–84
replication of, 60–63
structure/composition of, 31, 33–34, 35, 38–39, 59–60, 85, 89–90
DNA cloning, 125–134
DNA fingerprinting, 93–94, 136
DNA library, 129–130
DNA sequencing, 132–133
Dolly (cloned sheep), 137, 148, 149, 151–153
Down syndrome, 55, 87, 221–223, 245, 249
Drosophila, experiments with, 48, 49–51, 109, 191

E

embryology, experiments in, 48–49
environment
influence on genetics of, 205–208, 240–242
physical characteristics and adaptation to, 196–197, 214–217
environmental agents, genetic damage from, 240–242
epigenetics, 110–112, 260
eugenics, 156, 248, 255–259

evolution, theory of, 24, 29, 48–49, 52, 256, 260
experimental breeding, 101

F

Fire, Andrew Z., 80, 120
Fisher, Ronald Aylmer, 105
fluorescence in situ hybridization (FISH), 87–88, 194, 251, 254
47,XYY syndrome, 224–225, 226
founder effects, 197
founder principle, 104
fragile-X syndrome, 231, 232
Franklin, Rosalind, 33, 37, 60

G

Galton, Francis, 104–105, 210, 255, 256, 259
Garrod, Archibald, 30
gene cloning, 123
gene expression, regulation of, 76–79
gene mutation, 71–76
genes
about, 38–41
naming and discovery of, 22, 29–30, 50
gene silencing, 79–81, 120–122
gene therapy, 90, 135, 144
genetically modified organisms (GMOs), 101, 137–146
in agriculture, 139–142
and controversy, 139, 141, 145–146
in medicine and research, 143–144, 146
role of in environmental management, 144–145
sociopolitical relevance of, 145–146
genetic code, expression of and DNA, 64–71
genetic counseling, 243–248
genetic damage from environmental agents, 240–242
genetic diseases in humans, 190, 193, 198, 211–212, 220–261
classes of, 221–234
and genetic counseling, 243–248
and genetic testing, 252–255
management of, 242–243
and prenatal diagnosis, 248–252
genetic drift, 103–104
genetic engineering, 89, 123–157
explained, 123–125
genetics
and antibody formation, 202–204
applied to animals, 158, 174–189
applied to plants, 158–174
behaviour, 104–108
and cellular immunity, 204–205
classical, 85–86
and concept of race, 212–219
cytogenetics, 86–88
early experiments in, 25–30
epigenetics, 110–112, 260
future of in medicine, 259–261
genomics, 95–100
history of, 21–37
and human blood, 198–202
immunogenetics, 95

influence of environment on, 205–208, 240–242
microbial, 88–89
molecular, 30–31, 89–94
population, 100–104
reverse, 135–136
study of, 85–122
study of human, 108–109, 190–219
genetic testing, 252–255
genomic library, 90, 129–130, 131
genomics, 95–100, 260
comparative, 99–100
functional, 98
genotype, explained, 40, 41
Gilbert, Walter, 133
Greider, Carol, 240
Griffith, Frederick, 57

H

Hardy, Godfrey Harold, 102
Hardy-Weinberg law, 100, 102–103
hemoglobin, 201–202
heredity
ancient beliefs and ideas about, 21, 22–23, 24
DNA as agent of, 57–84
early experiments in, 25–30
physical basis of, 38–56
Hershey, Alfred D., 58
Heston, Leonard, 108
Hippocrates, 23
human blood, genetics of, 198–202
human diseases, genetic, 220–261
human embryos, cloning of, 150, 154–157
human genetics, study of, 108–109, 190–219
human genome, 191–198, 260
origins of, 195–197
role of in research, 194–195, 198
social impact of research on, 197–198
Human Genome Project, 36, 259
human migration, tracking, 196–197
Hwang Woo Suk, 154–155

I

immunogenetics, 95
imprinted gene mutations, 232–234
inbreeding of animals, 188–189
independent assortment, law of, 47
in vitro mutagenesis, 133–134

J

Jacob, François, 32–33
Jeffreys, Alec, 93
Johannsen, Wilhelm, 22, 29, 50

K

Khorana, Har Gobind, 32
King, Thomas J., 147
Klinefelter syndrome, 56, 87, 224–225, 226

L

Lamarck, Jean-Baptiste, 24
Leighton, E.A., 184–185
Levan, Albert, 190

linkage groups, 47–51
Lyon, Mary, 225

M

MacLeod, Colin M., 31
macular degeneration, 121
Maxam, Allan M., 133
McCarty, Maclyn, 31
McClintock, Barbara, 82–83
McKusick, Victor, 235
medical genetics, 190
meiosis, 39–40, 44, 46–47, 51, 55, 96, 147
Mello, Craig C., 80, 120
Mendel, Gregor, 21–22, 23, 24, 25–29, 30, 38, 47, 49, 50, 51, 85, 86, 105, 246, 256
Meselson, Matthew, 31, 60–61
messenger RNA (mRNA), 67, 69–71, 74, 76, 79, 129
Miescher, Johann Friedrich, 57
mitochondrial DNA mutations, 232
mitosis, 43, 45–46, 147, 190
molecular genetics, 30–31
Monod, Jacques, 32–33
Morgan, Thomas Hunt, 30, 48–51
Muller, Hermann Joseph, 50, 51
Mullis, Kary B., 91
multifactorial inheritance, disease caused by, 234, 236

N

Nägeli, Karl Wilhelm von, 28
Nathans, Daniel, 124
natural selection, 24, 49, 215, 255
“nature-nurture” argument, 104, 205–206
Nirenberg, Marshall, 32
nuclear transfer, 147, 148, 157
nucleotides, 31–32, 39

O

Okazaki fragments, 63

P

pangenesis, theory of, 23, 24, 49
Parkinson disease, 114, 118, 154
percutaneous umbilical blood sampling, 249, 251
phenotype, explained, 40–41
phenylketonuria (PKU), 229–230
physical characteristics of humans and adaptation to environment, 196–197, 214–217
plants/crops
  breeding/selecting of, 22, 101, 137, 158–174
  evaluation of and breeding, 161–163
  and GMOs, 139–142
  methods of breeding, 163–174
polymerase chain reactions, 90–93, 147, 194, 254–255
preformation, theory of, 24, 49
preimplantation testing of embryos, 249, 251
prenatal diagnosis, 248–252
progeny testing, and animal breeding, 186–187

R

race, genetics and the concept of, 212–219

modern scientific explanations of, 212–217
scientific debate over, 217–219
recombinant DNA technology, 89, 90, 91, 95, 123–137
and animal breeding, 158
and antibodies, 202–203
and diagnostics, 136
and DNA cloning, 125–132
and DNA sequencing, 132–133
and gene therapy, 135
and GMOs, 137–139, 144
and in vitro mutagenesis, 133–134
and protein manufacture, 136–137
and reproductive cloning, 148
and reverse genetics, 135–136
repeat expansions, disease caused by, 231–232
retinoblastoma, 236–237
retroviruses, 66, 119, 120
reverse genetics, 135–136
RNA (ribonucleic acid), 33, 37, 41, 58, 78, 87, 99, 119, 120–122
and gene mutation, 74
structure/composition of, 59, 66
and transcription, 66, 67–69, 79
and translation, 66, 69–71
RNA interference (RNAi), 79–81, 120–122

S

Sanger, Frederick, 133
segregation, law of, 28, 29, 47
selective breeding, 106, 174–176, 181–187, 256
self-pollinated species of plants, breeding, 165–171
serum proteins, 200–201
sex chromosomes, about, 43, 44–45
disease caused by abnormalities of, 223–226
sex linkage, 49, 51, 86
sickle cell anemia, 144, 216–217, 229, 254
skin colour, explanation for differences in, 196–197, 214–215
Smith, Hamilton O., 124
social Darwinism, 217, 255, 256
somatic cell nuclear transfer (SCNT), 148–149, 150–151, 155
Speman, Hans, 147
Stahl, Franklin W., 31, 60–61
stem cells, 111, 112–120, 148, 260
adult, 115–120
embryonic, 114–115, 118–119, 150, 155, 157
ethical issues surrounding use of, 115, 118
and GMOs, 144
induced pluripotent, 118–120
and therapeutic cloning, 153–154
Sturtevant, Alfred Henry, 50, 51
Szostak, Jack, 240

T

Tatum, Edward, 30
telomeres, about, 239–240
Tjio, Joe Hin, 190
transcription, 66, 67–69, 79

TITLES IN THIS SERIES

NEW THINKING ABOUT
EVOLUTION

NEW THINKING ABOUT
GENETICS

NEW THINKING ABOUT
POLLUTION

Britannica
Educational Publishing