PART

3

Evolution

Mechanisms of Evolution

15

On November 11, 1918, an armistice agreement signed in France signaled the end of World War I. But the death toll from four years of war was soon surpassed by the casualties of a massive influenza epidemic that began in the spring of 1918 among soldiers in a U.S. Army barracks. Over the next 18 months, this particular strain of flu virus spread across the globe, killing more than 50 million people worldwide—more than twice the number of World War I-related combat deaths.

The 1918–1919 pandemic was noteworthy because the death rate among young adults—who are usually less likely to die from influenza than are the elderly or the very young—was 20 times higher than in flu epidemics before or since. Why was that particular virus so deadly, especially to typically hardy individuals? The 1918 flu strain triggered an especially intense reaction in the human immune system. This overreaction meant that people with strong immune systems were likely to be more severely affected.

In most cases, however, our immune system helps us fight viruses; this response is the basis of *vaccination.* Since 1945, programs to administer flu vaccines have helped keep the number and severity of influenza outbreaks in check. Last year's vaccine, however, will probably not be effective against this year's virus. New strains of flu virus are evolving continuously, ensuring genetic variation in the population. If these viruses did not evolve, we would become resistant to them and annual vaccination would become unnecessary. But because they do evolve, biologists must develop a new and different flu vaccine each year.

Flu victims are treated at a U.S. Army hospital in 1918.

Vertebrate immune systems recognize proteins on the viral surface, and changes in these proteins mean that the virus can escape immune detection. Virus strains with the greatest number of changes to their surface proteins are most likely to avoid detection and infect their hosts, and thus have an advantage over other strains. Biologists can observe evolution in action by following changes in influenza virus proteins from year to year.

We learn a great deal about the processes of evolution by examining rapidly evolving organisms such as viruses, and these studies contribute to the development of evolutionary theory. Evolutionary theory, in turn, is put to practical uses, such as the development of better strategies for combating deadly diseases.

QUESTION

How do biologists use evolutionary theory to develop better flu vaccines?

*You will find the answer to this question on page 312.

KEY CONCEPTS

15.1 Evolution Is Both Factual and the Basis of Broader Theory

15.2 Mutation, Selection, Gene Flow, Genetic Drift, and Nonrandom Mating Result in Evolution

15.3 Evolution Can Be Measured by Changes in Allele Frequencies

15.4 Selection Can Be Stabilizing, Directional, or Disruptive

15.5 Genomes Reveal Both Neutral and Selective Processes of Evolution

15.6 Recombination, Lateral Gene Transfer, and Gene Duplication Can Result in New Features

15.7 Evolutionary Theory Has Practical Applications

concept 15.1 Evolution Is Both Factual and the Basis of Broader Theory

That biological populations change over time, or **evolve,** is a fact that is not disputed by scientists. We can, and do, observe evolutionary change on a regular basis, both in laboratory experiments and in natural populations. We measure the rate at which new mutations arise, observe the spread of new genetic variants through a population, and see the effects of genetic change on the form and function of organisms. In the fossil record, we observe the long-term morphological changes (which are the result of underlying genetic changes) that have occurred among living organisms. These underlying changes in the genetic makeup of populations (sometimes referred to as *microevolution*) drive the origin and extinction of species and fuel the diversification of life (*macroevolution*).

In addition to observing and recording physical changes over evolutionary time, biologists have accumulated a large body of evidence about *how* these changes occur, and about *what* evolutionary changes have occurred in the past. The resulting understanding and application of the mechanisms of evolutionary change to biological problems is known as **evolutionary theory**.

Evolutionary theory has many useful applications. We constantly apply it to the study and treatment of diseases; to the development of better agricultural crops and practices; and to the development of industrial processes that produce new molecules with useful properties. At a more basic level, knowledge of evolutionary theory allows biologists to understand how life diversified and has provided insight into how species interact. It also helps us to make predictions about the biological world.

In everyday speech, people tend to use the word "theory" to mean an untested hypothesis, or even a guess. But *evolutionary theory* does not refer to any single hypothesis, and it certainly is not guesswork. The concept of evolutionary change among living organisms was present among a few scientists even before Charles Darwin so clearly described his observations, presented his conclusions, and articulated the premise of natural selection in *The Origin of Species.* The rediscovery of Mendel's experiments and the subsequent establishment of the principles of genetic inheritance early in the 1900s set the stage for vast amounts of research. By the end of the twentieth century, findings from many fields of biology firmly upheld Darwin's basic premises about the common ancestry of life and the role of natural selection as an important mechanism of evolution. Today a vast and rich array of geological, morphological, and molecular data all support and expand the factual basis of evolution.

When we refer to evolutionary theory, we are referring to our understanding of the mechanisms that result in genetic changes in populations over time and to our use of that understanding to interpret changes in and interactions among living organisms. We can directly observe the evolution of influenza viruses, but it is evolutionary theory that allows us to apply our observations to the task of developing more effective vaccines. Several mechanisms of evolutionary change are recognized, and the scientific community is continually using evolutionary theory to expand its understanding of how and when these mechanisms apply to particular biological problems.

Darwin and Wallace introduced the idea of evolution by natural selection

In the early 1800s, it was not yet evident to many people that life evolves. But several biologists had suggested that the species living on Earth had changed over time—that is, that evolution had taken place. Jean-Baptiste Lamarck, for one, presented strong evidence for the fact of evolution in 1809, but his ideas about *how* it occurred were not convincing. At that time, no one had yet envisioned a viable mechanism for evolution.

Charles Robert Darwin

In the 1820s, a young Charles Darwin became passionately interested in the subjects of geology (with its new sense of Earth's great age) and *natural history* (the scientific study of how different organisms function and carry out their lives in nature). Despite these interests, he planned, at his father's behest, to become a doctor. But surgery conducted without anesthesia nauseated Darwin, and he gave up medicine to study at Cambridge University for a career as a clergyman in the Church of England. Always more interested in science than in theology, he gravitated toward scientists on the faculty, especially the botanist John Henslow. In 1831, Henslow recommended Darwin for a position on HMS *Beagle*, a Royal Navy vessel that was preparing for a survey voyage around the world.

HMS *Beagle*

Whenever possible during the 5-year voyage (**FIGURE 15.1**), Darwin went ashore to study rocks and to observe and collect plants and animals. He noticed striking differences between the species he saw in South America and those of Europe. He

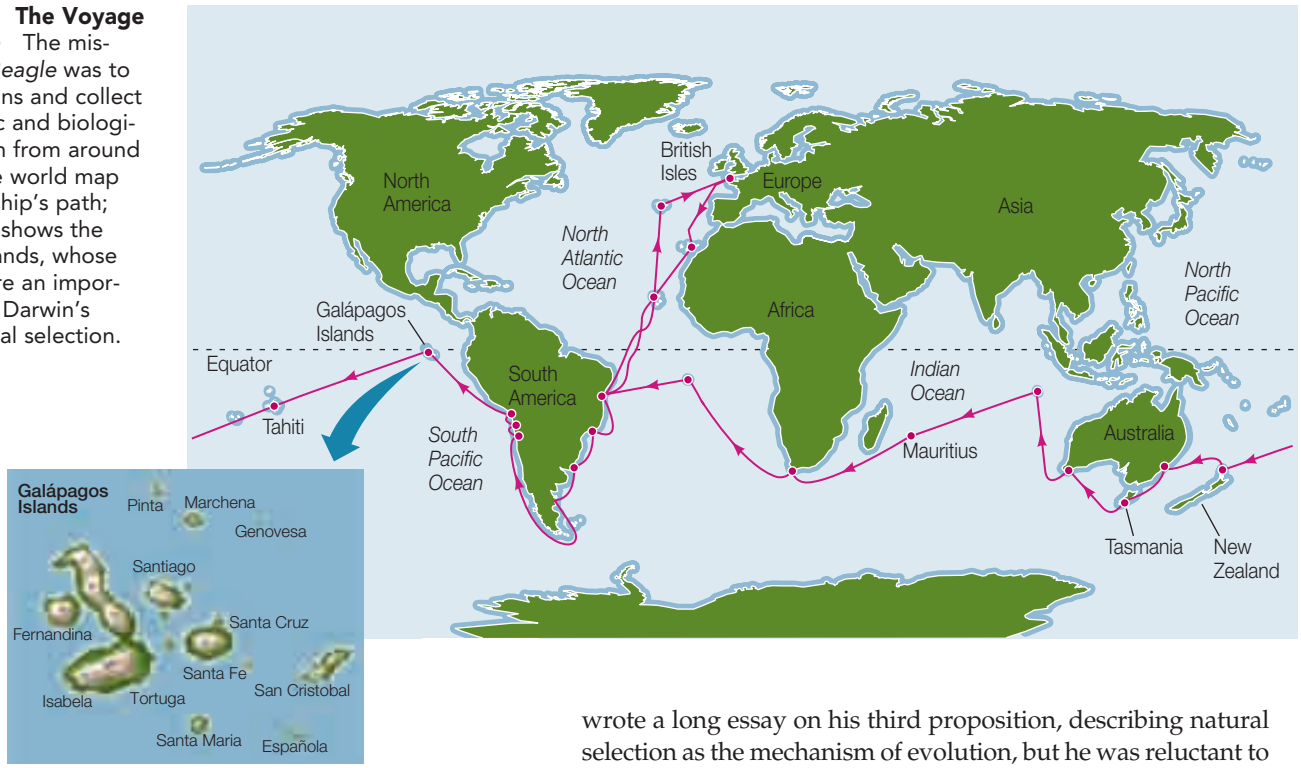


FIGURE 15.1 The Voyage of the *Beagle* The mission of HMS *Beagle* was to chart the oceans and collect oceanographic and biological information from around the world. The world map indicates the ship's path; the inset map shows the Galápagos Islands, whose organisms were an important source of Darwin's ideas on natural selection.

observed that the species of the temperate regions of South America (Argentina and Chile) were more similar to those of tropical South America (Brazil) than they were to temperate European species. When he explored the islands of the Galápagos archipelago west of Ecuador, he noted that most of the animals were *endemic* to the islands (that is, unique and found nowhere else), although they were similar to animals found on the mainland of South America. Darwin also observed that the fauna of the Galápagos differed from island to island. He postulated that some animals had come to the archipelago from mainland South America and had subsequently undergone different changes on each of the islands. He wondered what might account for these changes.

When he returned to England in 1836, Darwin continued to ponder his observations. His ruminations were strongly influenced by the geologist Charles Lyell, who had recently popularized the idea that Earth had been shaped by slow-acting forces that are still at work today. Darwin reasoned that similar thinking could be applied to the living world. Within a decade, he had developed the framework of an explanatory theory for evolutionary change based on three major propositions:

- Species are not immutable; they change over time.
- Divergent species share a common ancestor.
- The mechanism that produces changes in species is **natural selection**: the differential survival and reproduction of individuals in a population based on variation in their traits.

The first of these propositions was not unique to Darwin; several earlier authors had argued for the fact of evolution. A more revolutionary idea was his second proposition, that *divergent species are related to one another through common descent*. In 1844, Darwin wrote a long essay on his third proposition, describing natural selection as the mechanism of evolution, but he was reluctant to publish it, preferring to assemble more evidence first.

Darwin's hand was forced in 1858, when he received a letter and manuscript from another traveling English naturalist, Alfred Russel Wallace, who was studying the biota of the Malay Archipelago. Wallace asked Darwin to evaluate his manuscript, which included an explanation of natural selection almost identical to Darwin's. Darwin was at first dismayed, believing Wallace to have preempted his idea. Parts of Darwin's 1844 essay, together with Wallace's manuscript, were presented to the Linnaean Society of London on July 1, 1858, thereby crediting both men for the idea of natural selection. Darwin then worked quickly to finish his own book, *The Origin of Species*, which was published the following year.

yourBioPortal.com

Go to ANIMATED TUTORIAL 15.1
Natural Selection

Although Darwin and Wallace independently articulated the concept of natural selection, Darwin developed his ideas first. Furthermore, *The Origin of Species* proved to be a stunning work of scholarship that provided exhaustive evidence from many fields supporting both the premise of evolution itself and the notion of natural selection as a mechanism of evolution. Thus both concepts are more closely associated with Darwin than with Wallace.

The publication of *The Origin of Species* in 1859 stirred considerable interest (and controversy) among scientists and the public alike. Scientists spent much of the rest of the nineteenth century amassing biological and paleontological data to test evolutionary ideas and document the history of life on Earth. By 1900, the fact of biological evolution (by then defined as change in the physical characteristics of populations over time)

was established beyond any reasonable doubt. But the *genetic* basis of evolutionary change was not yet understood.

Evolutionary theory has continued to develop over the past century

In 1900, several individuals rediscovered the work of Gregor Mendel (which had been published in 1866 but rarely read or cited), and the basic mechanisms of genetic inheritance began to be unraveled. In the first decades of the twentieth century, Thomas Hunt Morgan's studies on fruit flies led to his discovery of the role of chromosomes in inheritance. In the 1920s and early 1930s, the major principles of population genetics were established, the genetic basis of new variation (i.e., mutations) began to be understood, and mechanisms of evolution such as genetic drift were described (see Concept 15.2). This work set the stage for a "modern synthesis" of genetics and evolution that took place over the period 1936–1947. Some of the major contributors to this synthesis and a few of their books are listed in **FIGURE 15.2**.

Although chromosomes were now understood to be the basis of genetic transmission in eukaryotes, their molecular structure remained a mystery until soon after the modern synthesis. Then, in 1953, Watson and Crick published their paper on the structure of DNA, opening the door to our current detailed understanding of molecular evolutionary mechanisms. By the 1960s, biologists could study and document changes in allele frequencies in populations over time (see Concept 15.3). Most of this early work necessarily focused on variants of proteins that differed within and between populations and species; even

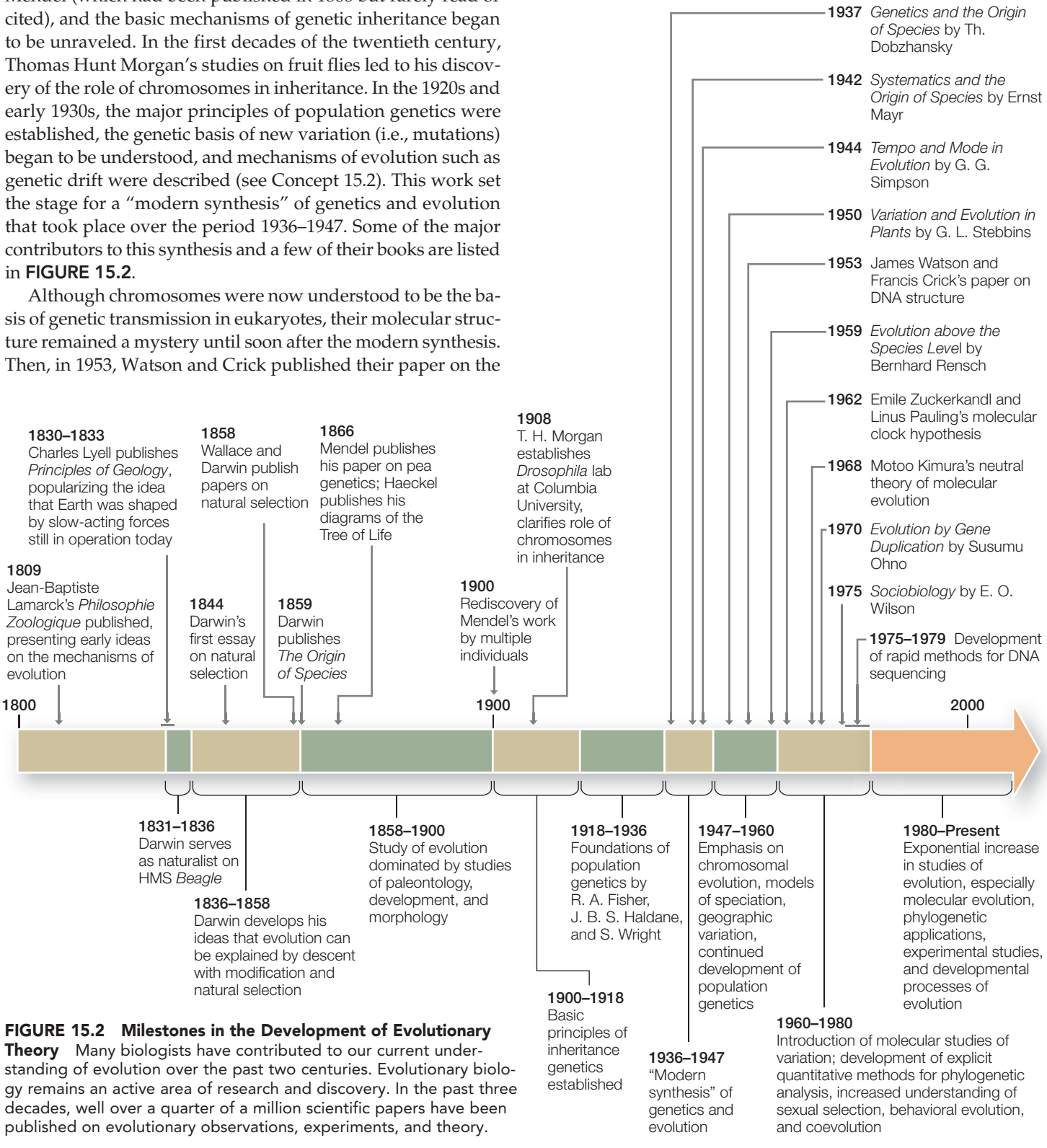


FIGURE 15.2 Milestones in the Development of Evolutionary Theory Many biologists have contributed to our current understanding of evolution over the past two centuries. Evolutionary biology remains an active area of research and discovery. In the past three decades, well over a quarter of a million scientific papers have been published on evolutionary observations, experiments, and theory.

though the molecular structure of DNA was known, it was not yet practical to sequence long stretches of DNA. Nonetheless, many important advances occurred in evolutionary theory during this time (see Figure 15.2), and these advances were not focused solely on a genetic understanding of evolution. E. O. Wilson's 1975 book *Sociobiology*, for example, invigorated studies of the evolution of behavior (a subject that had fascinated Darwin).

In the late 1970s, several techniques were developed that allowed the rapid sequencing of long stretches of DNA, which in turn allowed researchers to ascertain the amino acid sequences of proteins. This ability opened a new door for evolutionary biologists, who can now explore the structure of genes and proteins and document evolutionary changes within and between species in ways never before possible.

Do You Understand Concept 15.1?

- Why do biologists speak of "evolutionary theory" if the facts of evolution are not in doubt?
- Why do you think Darwin and Wallace formulated their ideas on natural selection at about the same time?
- Discuss the significance of each of the following scientific advances for evolutionary theory:
  a. Elucidation of the principles of chromosomal inheritance
  b. The discovery of DNA, its structure, and the universal genetic code
  c. Technology that allows us to sequence long segments of DNA

  Keep your discussion in mind as you continue reading this chapter.

Darwin had important insights into the mechanisms of evolution, even though he had a poor understanding of genetic transmission. Next we'll consider the primary mechanisms of evolution in light of our current understanding of genetics.

concept 15.2 Mutation, Selection, Gene Flow, Genetic Drift, and Nonrandom Mating Result in Evolution

Although the word "evolution" is often used in a general sense to mean simply "change," in a biological context **evolution** refers specifically to changes in the genetic makeup of populations over time. Developmental changes that occur in a single individual over the course of the life cycle are not the result of evolutionary change. Evolution is genetic change occurring in a **population**—a group of individuals of a single species that live and interbreed in a particular geographic area at the same time. It is important to remember that *individuals do not evolve; populations do*.

The premise of natural selection was one of Darwin's principal insights and has been demonstrated to be an important mechanism of evolution, but natural selection does not act alone. Three additional processes—gene flow, genetic drift, and nonrandom mating—affect the genetic makeup of populations over time. Before we consider how these processes change the frequencies of gene variants in a population, however, we need to understand how mutation brings such variants into existence.

Mutation generates genetic variation

The origin of genetic variation is mutation. As described in Concept 9.3, a *mutation* is any change in the nucleotide sequences of an organism's DNA. The process of DNA replication is not perfect, and some changes appear almost every time a genome is replicated. Mutations occur randomly with respect to an organism's needs; it is natural selection acting on this random variation that results in adaptation. Most mutations are either harmful to their bearers (*deleterious mutations*) or have no effect (*neutral mutations*). But a few mutations are *beneficial*, and even previously deleterious or neutral alleles may become advantageous if environmental conditions change. In addition, mutation can restore genetic variation that other evolutionary processes have removed. Thus mutation both creates and helps maintain genetic variation in populations.

Mutation rates can be high, as we saw in the case of the influenza viruses described at the opening of this chapter, but in many organisms the mutation rate is very low (on the order of 10^{-8} to 10^{-9} changes per base pair of DNA per generation). Even low overall mutation rates, however, create considerable genetic variation, because each of a large number of genes may change, and populations often contain

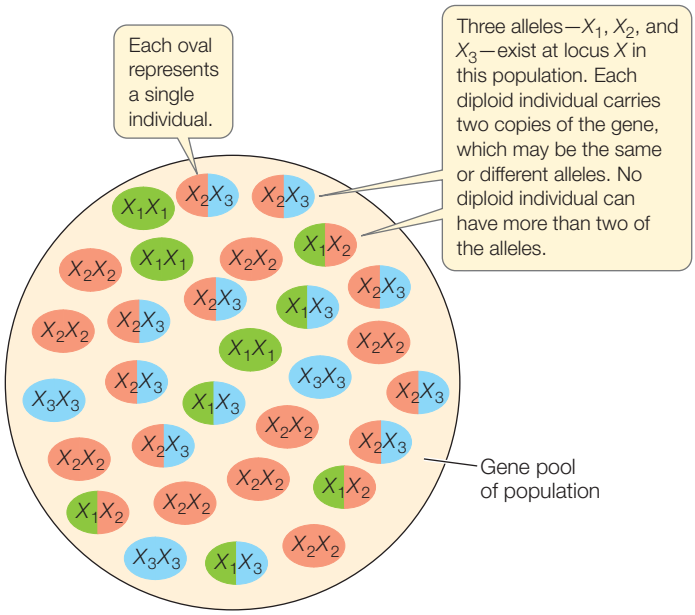


FIGURE 15.3 A Gene Pool A gene pool is the sum of all the alleles found in a population or at a particular locus. This figure shows the gene pool for one locus, *X*. The allele frequencies in this case are 0.20 for X_1, 0.50 for X_2, and 0.30 for X_3 (see Figure 15.10).

large numbers of individuals. For example, if the probability of a point mutation (an addition, deletion, or substitution of a single base) were 10^{-9} per base pair per generation, then each human gamete—the DNA of which contains 3×10^9 base pairs—would average three new point mutations ($3 \times 10^9 \times 10^{-9} = 3$), and each zygote would carry an average of six new mutations. The current human population of about 7 billion people would thus be expected to carry about 42 billion new mutations (i.e., changes in the nucleotide sequences of their DNA that were not present one generation earlier). So even though the mutation rate in humans is low, human populations still contain enormous genetic variation on which other evolutionary mechanisms can act.

As a result of mutation, different forms of a gene, known as *alleles*, may exist at a particular chromosomal locus. At any particular locus, a single diploid individual has no more than two of the alleles found in the population to which it belongs. The sum of all copies of all alleles at all loci found in a population constitutes its **gene pool** (**FIGURE 15.3**). (We can also refer to the gene pool for a particular chromosomal locus or loci.) The gene pool is the sum of the genetic variation in the population. The proportion of each allele in the gene pool is the **allele frequency.** Likewise, the proportion of each genotype among individuals in the population is the **genotype frequency**.

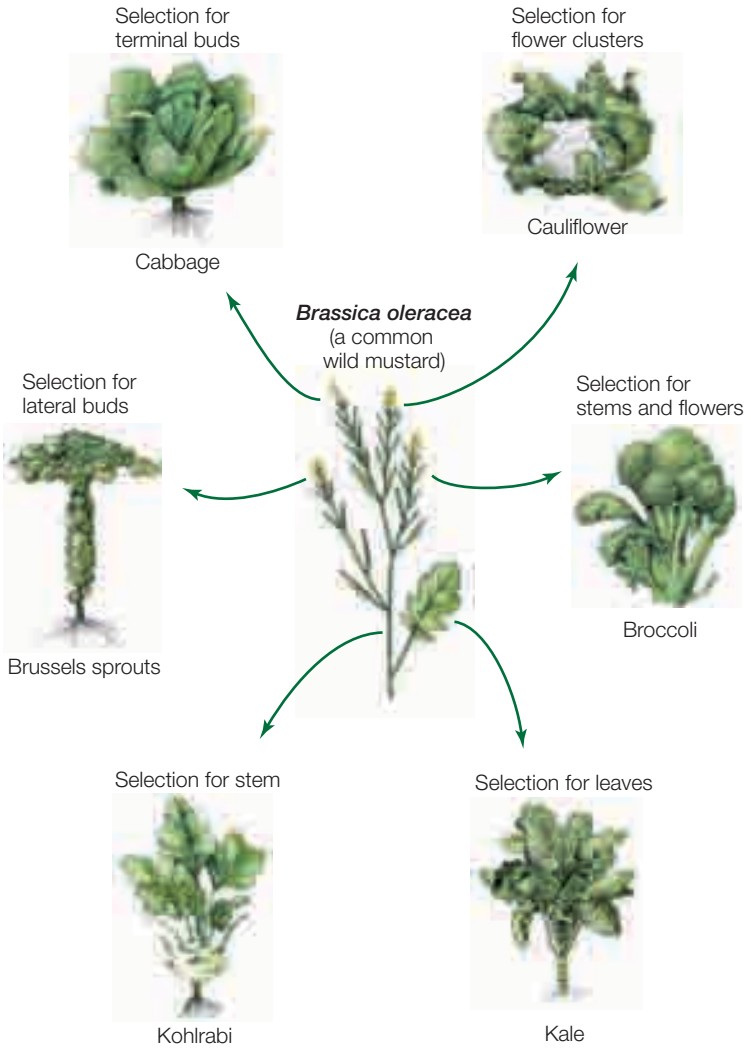


FIGURE 15.4 Many Vegetables from One Species All of the crop plants shown here derive from a single wild mustard species. European agriculturalists produced these crop species by selecting and breeding plants with unusually large buds, stems, leaves, or flowers. The results substantiate the vast amount of variation present in a gene pool.

LINK Review the nature of alleles and genetic inheritance in Concepts 8.1 and 8.2

Selection on genetic variation leads to new phenotypes

As a result of mutation, the gene pools of nearly all populations contain variation for many characters. Selection on different characters in a single European species of wild mustard produced many important crop plants (**FIGURE 15.4**). Agriculturalists were able to achieve these results because the original mustard population had genetic variation for the characters of interest (such as stem thickness or number of leaves).

Darwin compared this **artificial selection** by animal and plant breeders with natural selection. Many of Darwin's observations on the nature of variation and selection came from domesticated plants and animals. Darwin bred pigeons and thus knew firsthand the astonishing diversity in color, size, form, and behavior that breeders could achieve (**FIGURE 15.5**). He recognized close parallels between selection by breeders and selection in nature. Natural selection resulted in traits that helped organisms survive and reproduce more effectively; artificial selection resulted in traits that were preferred by the human breeders, for whatever reason.

Laboratory experiments also demonstrate the existence of considerable genetic variation in populations. In one such experiment, investigators bred populations of the fruit fly *Drosophila melanogaster* with high or low numbers of bristles on their abdomens from an initial population with intermediate numbers of bristles. After 35 generations, all flies in both the high- and low-bristle lineages had bristle numbers that fell well outside the range found in the original population (**FIGURE 15.6**). Selection for high and low bristle numbers resulted in new combinations of the many different genes that were present in the original population, so that the phenotypic variation seen in subsequent generations fell outside the phenotypic variation seen in the original population.

Natural selection increases the frequency of beneficial mutations in populations

Darwin knew that far more individuals of most species are born than survive to reproduce. He also knew that, although offspring tend to resemble their parents, the offspring of most organisms are not identical either to their parents or to one another. He suggested that slight differences among individuals affect the chance that a given individual will survive and

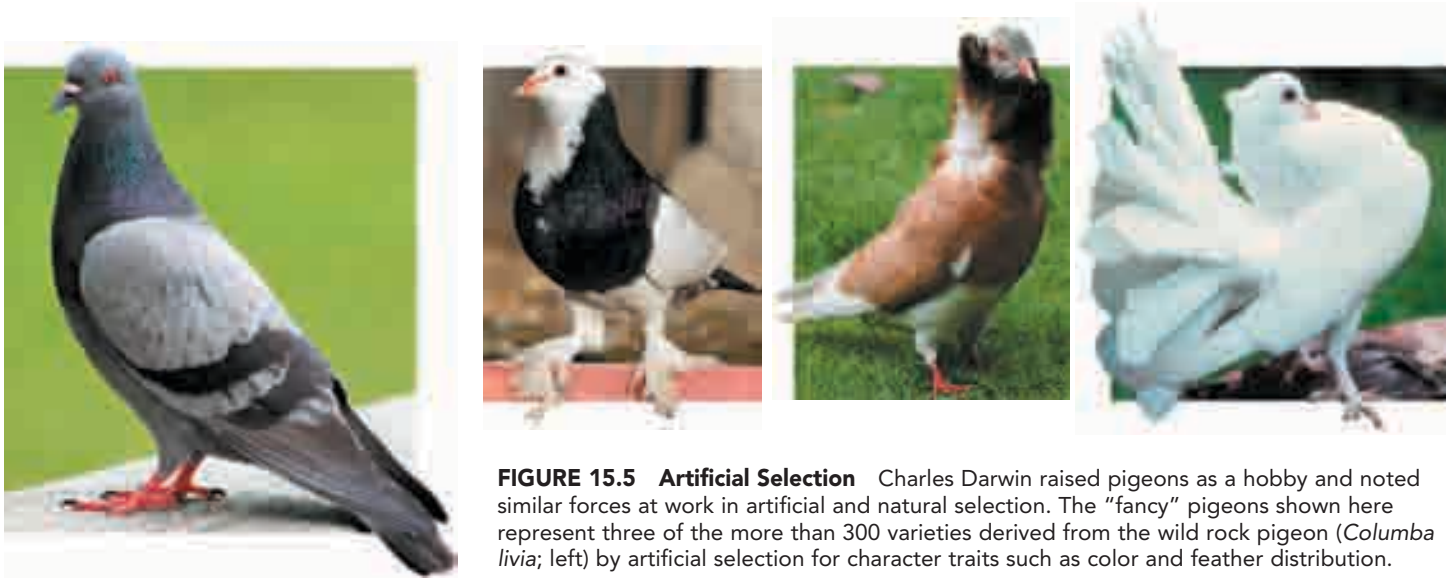


FIGURE 15.5 Artificial Selection Charles Darwin raised pigeons as a hobby and noted similar forces at work in artificial and natural selection. The "fancy" pigeons shown here represent three of the more than 300 varieties derived from the wild rock pigeon (*Columba livia*; left) by artificial selection for character traits such as color and feather distribution.

reproduce, which increases the frequency of the favored trait in the next generation. A favored trait that evolves through natural selection is known as an **adaptation**; this word is used to describe both the trait itself and the process that produces the trait.

Biologists regard an organism as being *adapted* to a particular environment when they can demonstrate that a slightly different organism reproduces and survives less well in that environment. To understand adaptation, biologists compare the performances of individuals that differ in their traits.

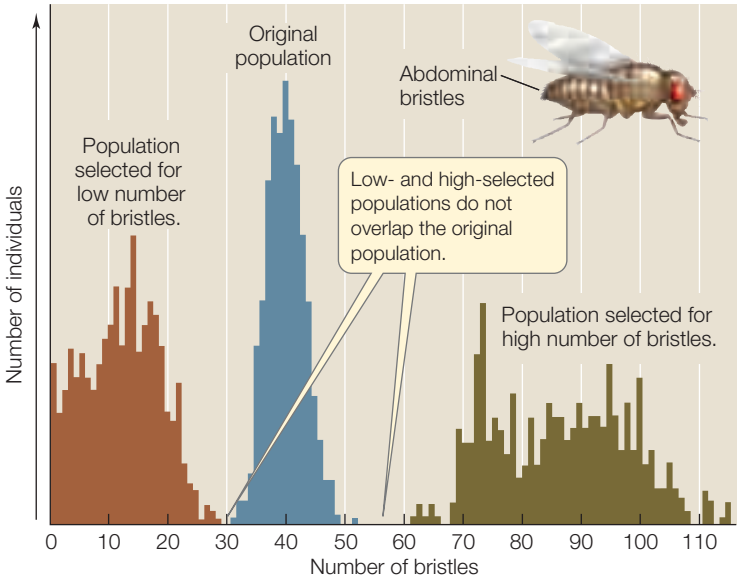


FIGURE 15.6 Artificial Selection Reveals Genetic Variation When investigators subjected *Drosophila melanogaster* to artificial selection for abdominal bristle number, that trait evolved rapidly. The graph shows the number of flies with different numbers of bristles in the original population and after 35 generations of artificial selection. The bristle numbers of the selected lineages clearly diverged from those of the original population.

Natural selection also acts to remove deleterious mutations from populations. Individuals with deleterious mutations are less likely to survive and reproduce, so they are less likely to pass their alleles on to the next generation.

Gene flow may change allele frequencies

Few populations are completely isolated from other populations of the same species. Migration of individuals and movements of gametes between populations—a phenomenon called **gene flow**—can change allele frequencies in a population. If the arriving individuals survive and reproduce in their new location, they may add new alleles to the population's gene pool, or they may change the frequencies of alleles present in the original population.

LINK If gene flow between two populations stops, those populations may diverge and become different species; see Concept 17.2

Genetic drift may cause large changes in small populations

In small populations, **genetic drift**—random changes in allele frequencies from one generation to the next—may produce large changes in allele frequencies over time. Harmful alleles may increase in frequency, and rare advantageous alleles may be lost. Even in large populations, genetic drift can influence the frequencies of neutral alleles (which do not affect the survival and reproductive rates of their bearers).

As an example, suppose there are only two females in a small population of mice, and one of these females carries a newly arisen dominant allele that produces black fur. Even in the absence of any selection, it is unlikely that the two females will produce exactly the same number of offspring. Even if they do produce identical litter sizes and identical numbers of litters, chance events that have nothing to do with genetic

characteristics are likely to result in differential mortality among their offspring. If each female produces one litter, but a flood envelops the black female's nest and kills all of her offspring, the novel allele could be lost from the population in just one generation. In contrast, if the wild-type female's litter is lost, then the frequency of the newly arisen allele (and phenotype) for dark fur will rise dramatically in just one generation.

Genetic drift also operates when a population is reduced dramatically in size. Even populations that are normally large may occasionally pass through environmental events that only a small number of individuals survive, a situation known as a **population bottleneck**. The effect of genetic drift in such a situation is illustrated in **FIGURE 15.7**, in which red and yellow beans represent two alleles of a gene. Most of the beans in the small sample of the "population" that "survives" the bottleneck event are, just by chance, red, so the new population has a much higher frequency of red beans than the previous generation had. In a real population, the red and yellow allele frequencies would be described as having "drifted."

A population forced through a bottleneck is likely to lose much of its genetic variation. For example, when Europeans first arrived in North America, millions of greater prairie-chickens (*Tympanuchus cupido*) inhabited the midwestern prairies. As a result of hunting and habitat destruction by the new settlers, the Illinois population of this species plummeted from about 100 million birds in 1900 to fewer than 50 individuals in the 1990s. A comparison of DNA from birds collected in Illinois during the middle of the twentieth century with DNA from the surviving population in the 1990s showed that Illinois prairie-chickens have lost most of their genetic diversity. Loss of genetic variation in small populations is one of the problems facing biologists who attempt to protect endangered species.

Genetic drift can have similar effects when a few pioneering individuals colonize a new region. Because of its small size, the colonizing population is unlikely to possess all of the alleles found in the gene pool of its source population. The resulting change in genetic variation, called a **founder effect**, is equivalent to that in a large population reduced by a bottleneck.

Nonrandom mating can change genotype or allele frequencies

Mating patterns often alter genotype frequencies because the individuals in a population do not choose mates at random. For example, self-fertilization (*selfing*) is common in many groups of organisms, especially plants. Any time individuals mate preferentially with other individuals of the same genotype (including themselves), homozygous genotypes will increase in frequency and heterozygous genotypes will decrease in frequency over time. The opposite effect (more heterozygotes, fewer homozygotes) is expected when individuals mate primarily or exclusively with individuals of different genotypes.

Sexual selection results from a specific type of nonrandom mating in which an organism's phenotype influences its ability to attract mates. For example, female peacocks may choose their male mates on the basis of his bright tail feathers and associated mating display. Males with brighter feathers are more likely to attract females. The higher reproductive success of colorful males results in an increase in the frequency of the alleles associated with colorful tail feathers in the next generation.

In *The Origin of Species*, Darwin devoted a few pages to sexual selection, but in 1871 he wrote an entire book about it: *The Descent of Man, and Selection in Relation to Sex*. Sexual selection was Darwin's explanation for the evolution of conspicuous characters that would appear to inhibit survival, such as bright colors, long tails, and elaborate courtship displays in males of many species. He hypothesized that these features either improved the ability of their bearers to compete for access to mates (*intrasexual selection*) or made their bearers more attractive to members of the opposite sex (*intersexual selection*). The concept of sexual selection was either ignored or questioned for many decades, but recent investigations have demonstrated its importance.

Whereas Darwin associated natural selection with traits that enhance the survival of their bearers or their bearers' descendants, sexual selection is primarily about successful reproduction. Of course, an animal must survive long enough to reproduce, but if it survives and fails to reproduce, it makes no contribution to the next generation. Thus sexual selection may favor traits that enhance an individual's chances of reproduction even when these traits reduce its chances of survival. For example, females may be more likely to see or hear males with a given trait (and thus be more likely to mate with those males), even though the favored trait also increases the chances that the male will be seen or heard by a predator.

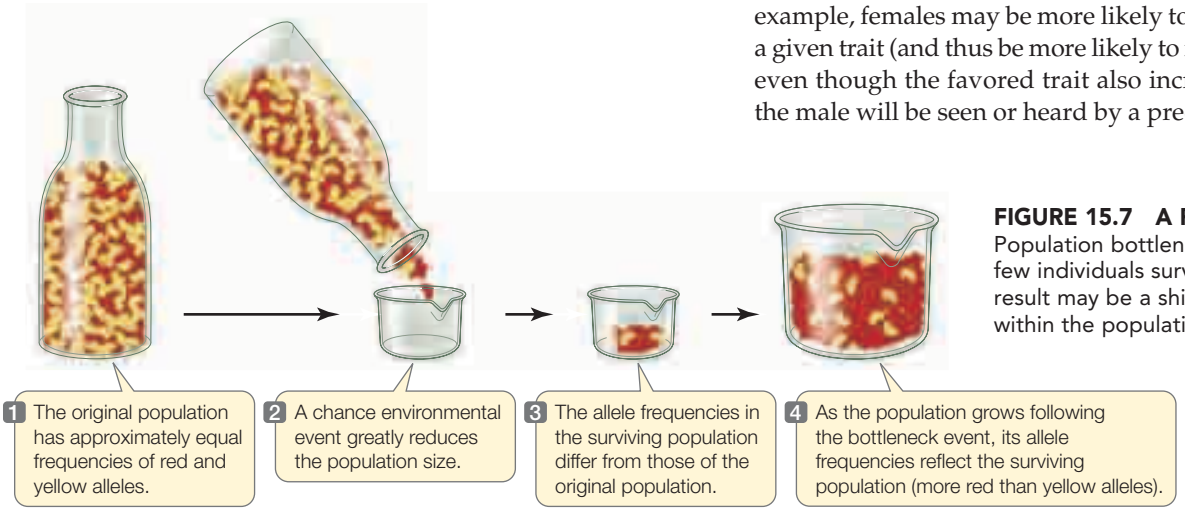


FIGURE 15.7 A Population Bottleneck Population bottlenecks occur when only a few individuals survive a random event. The result may be a shift in allele frequencies within the population.

In other cases, a male's sexual signal directly indicates a successful genotype. In many species of frogs, for example, females prefer males with low-frequency calls. Males' calls vary with body size, and a low-frequency call is indicative of a large-bodied frog. Frogs exhibit *indeterminate growth*—that is, they continue to grow throughout their lives—so a large frog is a long-lived frog, and size is an indication of survivorship. In this case, the sexual signal represents what is known as an *honest signal* of the male's ability to survive in the local environment.

LINK Some of the animal behaviors that have evolved in response to sexual selection are described in Concepts 41.5 and 41.6

One example of a trait that Darwin attributed to sexual selection is the remarkable tail of the male African long-tailed widowbird (*Euplectes progne*), which is longer than the bird's head and body combined (**FIGURE 15.8**). Male widowbirds normally select, and defend from other males, a territory where they perform courtship displays to attract females. To investigate whether sexual selection drove the evolution of widowbird tails, Malte Andersson, a behavioral ecologist at Gothenburg University in Sweden, clipped the tails of some captured male widowbirds and lengthened the tails of others by gluing on additional feathers. He then cut and reglued the tail feathers of still other males, which served as controls. Both short- and long-tailed males successfully defended their display territories, indicating that

FIGURE 15.8 What Is the Advantage? The extensive tail of the male African long-tailed widowbird actually inhibits its ability to fly. Darwin attributed the evolution of this seemingly nonadaptive trait to sexual selection.

INVESTIGATION

FIGURE 15.9 Sexual Selection in Action Behavioral ecologist Malte Andersson tested Darwin's hypothesis that excessively long tails evolved in male widowbirds because female preference for longer-tailed males increased their mating and reproductive success.

HYPOTHESIS

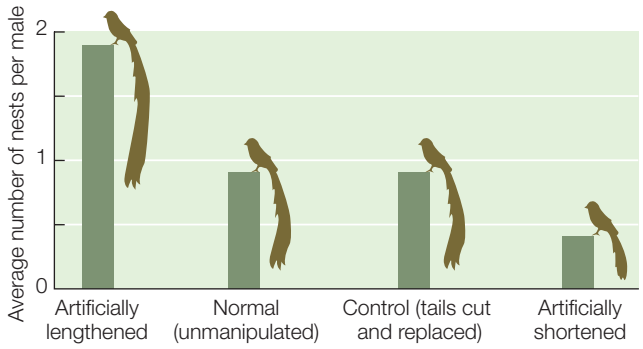
Female widowbirds prefer to mate with the male that displays the longest tail; longer-tailed males thus are favored by sexual selection because they will father more offspring.

METHOD

1. Capture males and artificially lengthen or shorten tails by cutting or gluing on feathers. In a control group, cut and replace tails to their normal length (to control for the effects of tail-cutting).
2. Release the males to establish their territories and mate.
3. Count the nests with eggs or young on each male's territory.

RESULTS

Male widowbirds with artificially shortened tails established and defended display sites sucessfully but fathered fewer offspring than did control or unmanipulated males. Males with artificially lengthened tales fathered the most offspring.



CONCLUSION

Sexual selection in *Euplectes progne* has favored the evolution of long tails in the male.

ANALYZE THE DATA

Are the differences plotted above significantly different? See Working with Data 15.1 at **yourBioPortal.com** for a simple method to test the statistical significance of the differences using the following data.

Group	Number of nests per male		
	Shortened tail	Control	Elongated tail
1	0	0	2
2	0	0	2
3	2	3	5
4	1	2	4
5	0	1	2
6	0	1	2
7	0	1	0
8	0	0	0
9	1	0	0

Go to **yourBioPortal.com** for original citations, discussions, and relevant links for all INVESTIGATION figures.

a long tail does not confer an advantage in male–male competition. However, males with artificially elongated tails attracted about four times more females than did males with shortened tails (**FIGURE 15.9**).

Why do female widowbirds prefer males with long tails? One possibility is that the ability to grow and maintain a costly feature such as a long tail may indicate that the male bearing it is vigorous and healthy, even though the tail impairs his ability to fly. If so, then females that are attracted to long tails are indirectly attracted to vigorous, healthy males, which are likely to carry beneficial genes that will lead to higher survivorship of their offspring.

Do You Understand Concept 15.2?

- How do deleterious, neutral, and beneficial mutations differ?
- Can you explain how natural selection results in an increase in the frequency of beneficial alleles in a population over time, and a decrease in the frequency of deleterious alleles?
- How can genetic drift cause large changes in small populations?
- How do selfing and sexual selection differ in their expected effects on genotype and allele frequencies over time?

The mechanisms of mutation, selection, gene flow, genetic drift, and nonrandom mating can all result in evolutionary change. We will consider next how evolutionary change that results from these mechanisms is measured.

concept 15.3 Evolution Can Be Measured by Changes in Allele Frequencies

Much of evolution occurs through gradual changes in the relative frequencies of different alleles in a population from one generation to the next. Major genetic changes can also be sudden, as happens when two formerly separated populations merge and hybridize, or when genes within a population are duplicated within the genome (see Concept 15.6). We can measure evolution by looking at changes in allele frequencies in populations.

Allele frequencies are usually estimated in locally interbreeding populations. To measure allele frequencies in a population precisely, we would need to count every allele at every locus in every individual in the population. Fortunately, we do not need to make such complete measurements because we can reliably estimate allele frequencies for a given locus by counting alleles in a sample of individuals from the population. The sum of all allele frequencies at a locus is equal to 1, so measures of allele frequency range from 0 to 1.

yourBioPortal.com

Go to INTERACTIVE TUTORIAL 15.1
Genetic Drift

An allele's frequency is calculated using the following formula:

$$\frac{\text{number of copies of the allele in the population}}{\text{total number of copies of all alleles in the population}}$$

If only two alleles (we'll call them *A* and *a*) for a given locus are found among the members of a diploid population, those alleles can combine to form three different genotypes: *AA*, *Aa*, and *aa* (see Figure 15.3). A population with more than one allele at a locus is said to be *polymorphic* ("many forms") at that locus. Applying the formula above as shown in **FIGURE 15.10**, we can calculate the relative frequencies of alleles *A* and *a* in a population of *N* individuals as follows:

- Let N_{AA} be the number of individuals that are homozygous for the *A* allele (*AA*).
- Let N_{Aa} be the number that are heterozygous (*Aa*).
- Let N_{aa} be the number that are homozygous for the *a* allele (*aa*).

Note that $N_{AA} + N_{Aa} + N_{aa} = N$, the total number of individuals in the population, and that the total number of copies of both

RESEARCH TOOLS

FIGURE 15.10 Calculating Allele and Genotype Frequencies Allele and genotype frequencies for a gene locus with two alleles in the population can be calculated using the equations in panel 1. When the equations are applied to two populations (panel 2), we find that the *frequencies of alleles A and a in the two populations are the same*, but the alleles are distributed differently between heterozygous and homozygous genotypes.

1 In any population, where *N* is the total number of individuals in the population:

$$\text{Frequency of allele } A = p = \frac{2N_{AA} + N_{Aa}}{2N} \qquad \text{Frequency of allele } a = q = \frac{2N_{aa} + N_{Aa}}{2N}$$

Frequency of genotype $AA = N_{AA}/N$
Frequency of genotype $Aa = N_{Aa}/N$
Frequency of genotype $aa = N_{aa}/N$

2 Compute the allele and genotype frequencies for two separate populations of $N = 200$:

Population 1 (mostly homozygotes)	Population 2 (mostly heterozygotes)
$N_{AA} = 90$, $N_{Aa} = 40$, and $N_{aa} = 70$	$N_{AA} = 45$, $N_{Aa} = 130$, and $N_{aa} = 25$
$p = \frac{180 + 40}{400} = 0.55$	$p = \frac{90 + 130}{400} = 0.55$
$q = \frac{140 + 40}{400} = 0.45$	$q = \frac{50 + 130}{400} = 0.45$
Freq. *AA* = 90/200 = 0.45	Freq. *AA* = 45/200 = 0.225
Freq. *Aa* = 40/200 = 0.20	Freq. *Aa* = 130/200 = 0.65
Freq. *aa* = 70/200 = 0.35	Freq. *aa* = 25/200 = 0.125

alleles present in the population is $2N$, because each individual is diploid. Each *AA* individual has two copies of the *A* allele, and each *Aa* individual has one copy of the *A* allele. Therefore, the total number of *A* alleles in the population is $2N_{AA} + N_{Aa}$. Similarly, the total number of *a* alleles in the population is $2N_{aa} + N_{Aa}$. If p represents the frequency of *A*, and q represents the frequency of *a*, then

$$p = \frac{2N_{AA} + N_{Aa}}{2N}$$

and

$$q = \frac{2N_{aa} + N_{Aa}}{2N}$$

Figure 15.10 applies these formulas to calculate the allele and genotype frequencies in two hypothetical populations, each containing 200 diploid individuals. The calculations in Figure 15.10 demonstrate two important points. First, notice that for each population, $p + q = 1$, which means that $q = 1 - p$. So when there are only two alleles at a given locus in a population, we can calculate the frequency of one allele and obtain the second allele's frequency by subtraction. If there is only one allele at a given locus in a population, its frequency is 1: the population is then *monomorphic* at that locus, and the allele is said to be *fixed*.

The second thing to notice is that population 1 (consisting mostly of homozygotes) and population 2 (consisting mostly of heterozygotes) have the same allele frequencies for *A* and *a*. Thus they have the same gene pool for this locus. Because the alleles in the gene pool are distributed differently among individuals, however, the *genotype frequencies* of the two populations differ.

The frequencies of the different alleles at each locus and the frequencies of the different genotypes in a population describe that population's **genetic structure**. Allele frequencies measure the amount of genetic variation in a population; genotype frequencies show how a population's genetic variation is distributed among its members. Other measures, such as the proportion of loci that are polymorphic, are also used to measure variation in populations. With these measurements, it becomes possible to consider how the genetic structure of a population changes or remains the same over generations—that is, to measure evolutionary change.

yourBioPortal.com

Go to ANIMATED TUTORIAL 15.2
Hardy–Weinberg Equilibrium

Evolution will occur unless certain restrictive conditions exist

In 1908, the British mathematician Godfrey Hardy and the German physician Wilhelm Weinberg independently deduced the conditions that must prevail if the genetic structure of a population is to remain the same over time. If the conditions they identified do not exist, then evolution will occur. The resulting principle is known as **Hardy–Weinberg equilibrium**. Hardy–Weinberg equilibrium describes a model in which allele frequencies do not change across generations and genotype frequencies can be predicted from allele frequencies (**FIGURE 15.11**). The principles of Hardy–Weinberg equilibrium apply only to sexually reproducing organisms. Several conditions must be met for a population to be at Hardy–Weinberg equilibrium (which, you should notice, correspond precisely to the five principal mechanisms of evolution discussed in Concept 15.2):

- *There is no mutation*. The alleles present in the population do not change, and no new alleles are added to the gene pool.

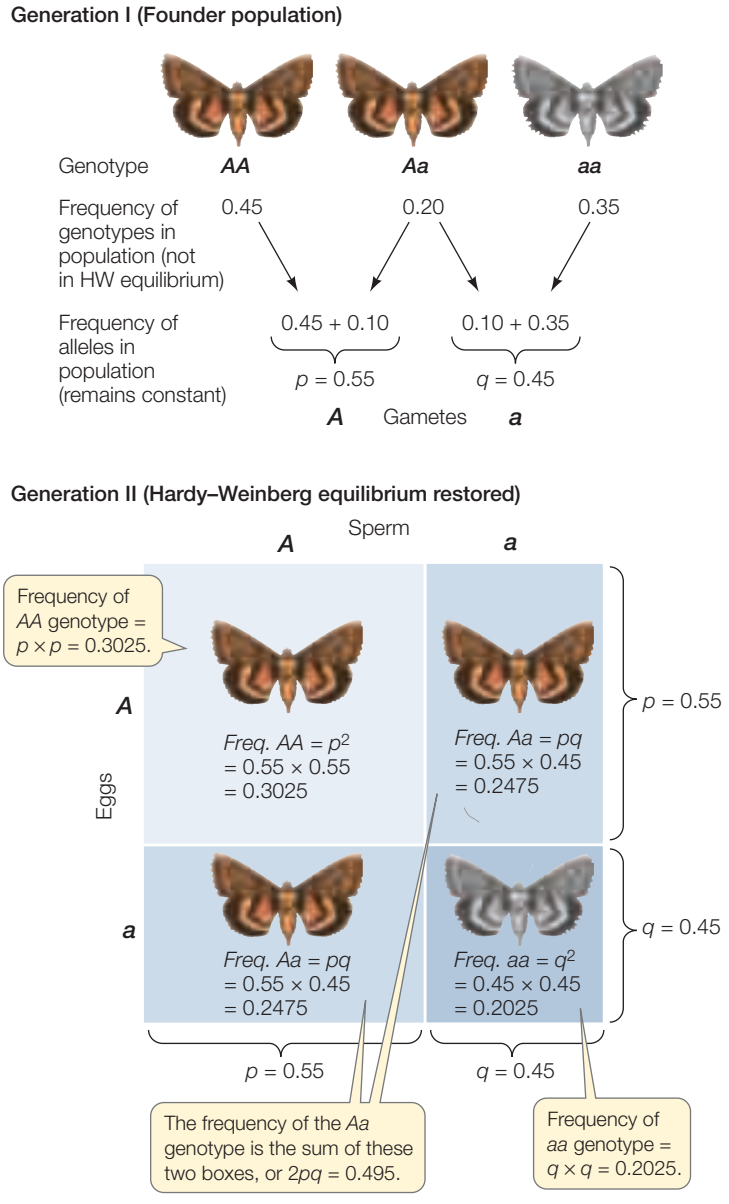


FIGURE 15.11 One Generation of Random Mating Restores Hardy–Weinberg Equilibrium Generation I of this population is made up of migrants from several source populations, and so is not initially in Hardy–Weinberg equilibrium. After one generation of random mating, the allele frequencies are unchanged, and the genotype frequencies return to Hardy–Weinberg expectations. The lengths of the sides of each rectangle are proportional to the allele frequencies in the population; the areas of the rectangles are proportional to the genotype frequencies.

- *There is no differential selection among genotypes*. Individuals with different genotypes have equal probabilities of survival and equal rates of reproduction.
- *There is no gene flow.* There is no movement of individuals into or out of the population or reproductive contact with other populations.
- *Population size is infinite*. The larger a population, the smaller will be the effect of genetic drift.
- *Mating is random*. Individuals do not preferentially choose mates with certain genotypes.

If these "ideal" conditions hold, two major consequences follow. First, the frequencies of alleles at a locus remain constant from generation to generation. Second, following one generation of random mating, the genotype frequencies occur in the following proportions:

Genotype	*AA*	*Aa*	*aa*
Frequency	p^2	$2pq$	q^2

To understand why these consequences are important, start by considering a population that is *not* in Hardy–Weinberg equilibrium, such as generation I in Figure 15.11. This could occur, for example, if the initial population is founded by migrants from several other populations, thus violating the Hardy–Weinberg assumption of no gene flow. In this example, generation I has more homozygous individuals and fewer heterozygous individuals than would be expected under Hardy–Weinberg equilibrium (a condition known as *heterozygote deficiency*).

Even with a starting population that is not in Hardy–Weinberg equilibrium, we can predict that after a single generation of random mating, and if the other Hardy–Weinberg assumptions are not violated, the *allele frequencies* will remain unchanged, but the *genotype frequencies* will return to Hardy–Weinberg expectations. Let's explore why this is true.

In generation I of Figure 15.11, the frequency of the *A* allele (p) is 0.55. Because we assume that individuals select mates at random, without regard to their genotype, gametes carrying *A* or *a* combine at random—that is, as predicted by the allele frequencies p and q. Thus, in this example, the probability that a particular sperm or egg will bear an *A* allele is 0.55. In other words, 55 out of 100 randomly sampled sperm or eggs will bear an *A* allele. Because $q = 1 - p$, the probability that a sperm or egg will bear an *a* allele is $1 - 0.55 = 0.45$.

LINK You may wish to review the discussion of probability and inheritance in Concept 8.1

To obtain the probability of two *A*-bearing gametes coming together at fertilization, we multiply the two independent probabilities of their occurrence:

$$p \times p = p^2 = (0.55)^2 = 0.3025$$

Therefore, 0.3025, or 30.25 percent, of the offspring in generation II will have homozygous genotype *AA*. Similarly, the probability of two *a*-bearing gametes coming together is

$$q \times q = q^2 = (0.45)^2 = 0.2025$$

Thus 20.25 percent of generation II will have the *aa* genotype.

There are two ways of producing a heterozygote: an *A* sperm may combine with an *a* egg, the probability of which is $p \times q$; or an *a* sperm may combine with an *A* egg, the probability of which is $q \times p$. Consequently, the overall probability

APPLY THE CONCEPT

Evolution can be measured by changes in allele frequencies

Imagine you have discovered a new population of curly-tailed lizards established on an island after immigrants have arrived from several different source populations during a hurricane. You collect and tabulate genotype data (right) for the lactate dehydrogenase gene (*Ldh*) for each of the individual lizards. Use the table to answer the following questions.

1. Calculate the allele and genotype frequencies of *Ldh* in this newly founded population.
2. Is the population in Hardy–Weinberg equilibrium? If not, which genotypes are over- or underrepresented? Given the population's history, what is a likely explanation of your answer?
3. Under Hardy–Weinberg assumptions, what allele and genotype frequencies do you predict for the next generation?
4. Imagine that you are able to continue studying this population and determine the next generation's actual allele and genotype frequencies. What are some of the principal reasons you might expect the observed allele and genotype frequencies to differ from the Hardy–Weinberg expectations you calculated in question 3?

INDIVIDUAL NUMBER	SEX	INDIVIDUAL GENOTYPE FOR *Ldh*
1	Male	*Aa*
2	Male	*AA*
3	Female	*AA*
4	Male	*aa*
5	Female	*aa*
6	Female	*AA*
7	Male	*aa*
8	Male	*aa*
9	Female	*Aa*
10	Male	*AA*

of obtaining a heterozygote is $2pq$, or 0.495. The frequencies of the *AA*, *Aa*, and *aa* genotypes in generation II of Figure 15.11 now meet Hardy–Weinberg expectations, and the frequencies of the two alleles (p and q) have not changed from generation I.

Under the assumptions of Hardy–Weinberg equilibrium, allele frequencies p and q remain constant from generation to generation. If Hardy–Weinberg assumptions are violated and the genotype frequencies in the parental generation are altered (say, by the loss of a large number of *AA* individuals from the population), then the allele frequencies in the next generation will be altered. However, based on the new allele frequencies, another generation of random mating will be sufficient to restore the genotype frequencies to Hardy–Weinberg equilibrium.

Deviations from Hardy–Weinberg equilibrium show that evolution is occurring

You probably have realized that populations in nature never meet the stringent conditions necessary to be at Hardy–Weinberg equilibrium—which explains why all biological populations evolve. Why, then, is this model considered so important for the study of evolution? There are two reasons. First, the equation is useful for predicting the approximate genotype frequencies of a population from its allele frequencies. Second—and crucially—the model allows biologists to evaluate which mechanisms are acting on the evolution of a particular population. The specific patterns of deviation from Hardy–Weinberg equilibrium can help us identify the various mechanisms of evolutionary change.

Do You Understand Concept 15.3?

- Why is the concept of Hardy–Weinberg equilibrium important even though the assumptions on which it is based are never completely met in nature?
- Although the stringent assumptions of Hardy–Weinberg equilibrium are never met completely in real populations, the genotype frequencies of many populations do not deviate significantly from Hardy–Weinberg expectations. Can you explain why?
- Suppose you examine a population of toads breeding in a single pond and find that heterozygous genotypes at several different loci are present at significantly lower frequencies than predicted by Hardy–Weinberg equilibrium. What are some possible explanations?

Our discussion so far has focused on changes in allele frequencies at a single gene locus. Genes do not exist in isolation, however, but interact with one another (and with the environment) to produce an organism's phenotype. What effects can these interactions have on selection?

concept 15.4 Selection Can Be Stabilizing, Directional, or Disruptive

Until now, we have only discussed traits influenced by alleles at a single locus. Such traits are often distinguished by discrete qualities (black versus white, or smooth versus wrinkled), and so are called *qualitative traits*. Many traits, however, are influenced by alleles at more than one locus. Such traits are likely to show continuous quantitative variation rather than discrete qualitative variation, and so are known as *quantitative traits*. For example, the distribution of body sizes of individuals in a population, a trait that is influenced by genes at many loci as well as by the environment, is likely to resemble a continuous bell-shaped curve.

Natural selection can act on characters with quantitative variation in any one of several different ways, producing quite different results (**FIGURE 15.12**):

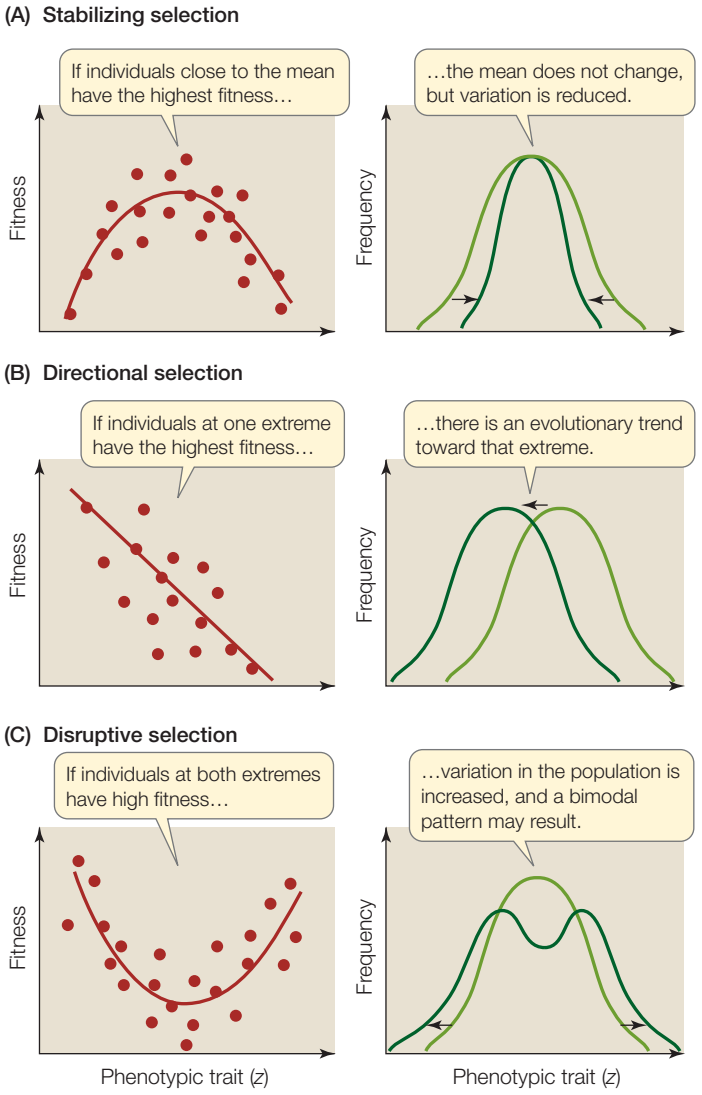


FIGURE 15.12 Natural Selection Can Operate in Several Ways The graphs in the left-hand column show the fitness of individuals with different phenotypes of the same trait. The graphs on the right show the distribution of the phenotypes in the population before (light green) and after (dark green) the influence of selection.

- **Stabilizing selection** preserves the average characteristics of a population by favoring average individuals.
- **Directional selection** changes the characteristics of a population by favoring individuals that vary in one direction from the mean of the population.
- **Disruptive selection** changes the characteristics of a population by favoring individuals that vary in both directions from the mean of the population.

Stabilizing selection reduces variation in populations

If the smallest and largest individuals in a population contribute fewer offspring to the next generation than do individuals closer to the average size, then stabilizing selection is operating on size (see Figure 15.12A). Stabilizing selection reduces variation in populations, but it does not change the mean. Natural selection frequently acts in this way, countering increases in variation brought about by sexual recombination, mutation, or gene flow. Rates of phenotypic change in many species are slow because natural selection is often stabilizing. Stabilizing selection operates, for example, on human birth weight. Babies who are lighter or heavier at birth than the population mean die at higher rates than babies whose weights are close to the mean (**FIGURE 15.13**). In discussions of specific genes, stabilizing selection is often called *purifying selection* because there is selection against any deleterious mutations to the usual gene sequence.

Directional selection favors one extreme

Directional selection is operating when individuals at one extreme of a character distribution contribute more offspring to the next generation than other individuals do, shifting the average value of that character in the population toward that extreme. In the case of a single gene locus, directional selection may result in favoring a particular genetic variant—referred to as *positive selection* for that variant. By favoring one phenotype over another, directional selection results in an increase of the frequencies of alleles that produce the favored phenotype (as with the surface proteins of influenza discussed in the opening of this chapter).

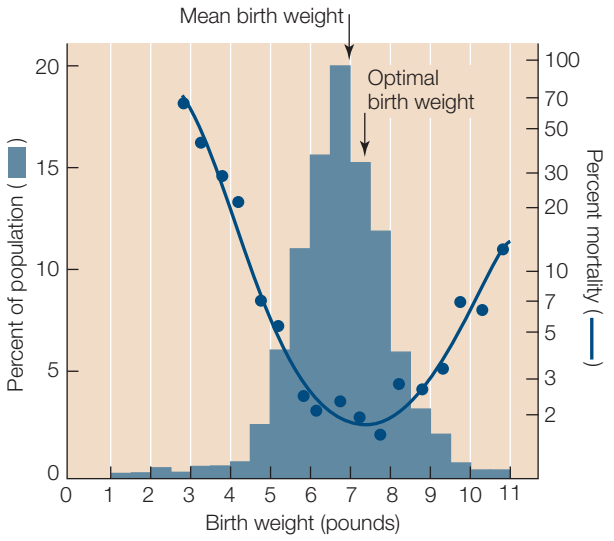


FIGURE 15.13 Human Birth Weight Is Influenced by Stabilizing Selection Babies that weigh more or less than average are more likely to die soon after birth than babies with weights close to the population mean.

FIGURE 15.14 Long Horns Are the Result of Directional Selection Long horns were advantageous for defending young calves from attacks by predators, so horn length increased in feral herds of Spanish cattle in the American Southwest between the early 1500s and the 1860s. The result was the familiar Texas Longhorn breed. This evolutionary trend has been maintained in modern times by ranchers practicing artificial selection.

If directional selection operates over many generations, an *evolutionary trend* is seen in the population (see Figure 15.12B). Evolutionary trends often continue for many generations, but they can be reversed if the environment changes and different phenotypes are favored, or halted when an optimal phenotype is reached or trade-offs between different adaptational advantages oppose further change. The character then undergoes stabilizing selection.

Many cases of directional selection have been observed directly, and long-term examples abound in the fossil record. The long horns of Texas Longhorn cattle (**FIGURE 15.14**) are an example of a trait that has evolved through directional selection. Texas Longhorns are descendants of cattle brought to the New World by Christopher Columbus, who picked up a few cattle in the Canary Islands and brought them to the island of Hispaniola in 1493. The cattle multiplied, and their descendants were taken to the mainland of Mexico. Spaniards exploring what would become Texas and the southwestern United States brought these cattle with them, some of which escaped and formed feral herds. Populations of feral cattle increased greatly over the next few hundred years, but there was heavy predation from bears, mountain lions, and wolves, especially on the young calves. Cows with longer horns were more successful in protecting their calves against attacks, and over a few hundred years the average horn length in the feral herds increased considerably. In addition, the cattle evolved resistance to endemic diseases of the Southwest, as well as higher fecundity and longevity. Texas Longhorns often live and produce calves

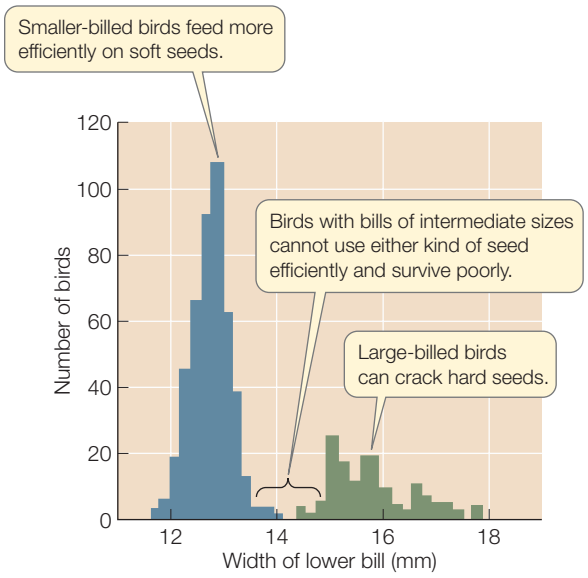


FIGURE 15.15 Disruptive Selection Results in a Bimodal Character Distribution The bimodal distribution of bill sizes in the black-bellied seedcracker of West Africa is a result of disruptive selection, which favors individuals with larger and smaller bill sizes over individuals with intermediate-sized bills.

well into their twenties—about twice as long as many breeds of cattle that have been artificially selected by humans for traits such as high fat content or high milk production (which are examples of artificial directional selection).

Disruptive selection favors extremes over the mean

When disruptive selection operates, individuals at opposite extremes of a character distribution contribute more offspring to the next generation than do individuals close to the mean, which increases variation in the population (see Figure 15.12C).

The strikingly *bimodal* (two-peaked) distribution of bill sizes in the black-bellied seedcracker (*Pyrenestes ostrinus*), a West African finch (**FIGURE 15.15**), illustrates how disruptive selection can influence populations in nature. The seeds of two types of sedges (marsh plants) are the most abundant food source for these finches during part of the year. Birds with large bills can readily crack the hard seeds of the sedge *Scleria verrucosa*. Birds with small bills can crack *S. verrucosa* seeds only with difficulty; however, they feed more efficiently on the soft seeds of *S. goossensii* than do birds with larger bills. Young finches whose bills deviate markedly from the two predominant bill sizes do not survive as well as finches whose bills are close to one of the two sizes represented by the distribution peaks. Because there are few abundant food sources in the finches' environment, and because the seeds of the two sedges do not overlap in hardness, birds with intermediate-sized bills are less efficient in using either one of the species' principal food sources. Disruptive selection therefore maintains a bimodal bill size distribution.

Do You Understand Concept 15.4?

- What are the different expected outcomes of stabilizing, directional, and disruptive selection?
- Why would you expect selection on human birth weight to be stabilizing rather than directional?
- Can you think of examples of extreme phenotypes in animal or plant populations that could be explained by directional selection?

Our discussion so far has largely focused on the evolution of phenotypes (what organisms look like and how they behave). We will now consider the specific mechanistic processes that operate at the level of genes and genomes.

concept 15.5 Genomes Reveal Both Neutral and Selective Processes of Evolution

Most natural populations harbor far more genetic variation than we would expect to find if genetic variation were influenced by natural selection alone. This discovery, combined with the knowledge that many mutations do not change molecular function, provided a major stimulus to the development of the field of *molecular evolution*.

To discuss the evolution of genes, we need to consider the specific types of mutations that are possible. A *nucleotide substitution* is a change in a single nucleotide in a DNA sequence (a type of point mutation). Many nucleotide substitutions have no effect on phenotype, even if the change occurs in a gene that encodes a protein, because most amino acids are specified by more than one codon. A substitution that does not change the encoded amino acid is known as a *silent substitution* or **synonymous substitution** (**FIGURE 15.16A**). Synonymous substitutions do not affect the functioning of a protein (although they may have other effects, such as changes in mRNA stability or translation rates) and are therefore less likely to be influenced by natural selection.

A nucleotide substitution that *does* change the amino acid sequence encoded by a gene is known as a *missense substitution* or **nonsynonymous substitution** (**FIGURE 15.16B**). In general, nonsynonymous substitutions are likely to be deleterious to the organism. But not every amino acid replacement alters a protein's shape and charge (and hence its functional properties). Therefore, some nonsynonymous substitutions are selectively neutral, or nearly so. A third possibility is that a nonsynonymous substitution alters a protein in a way that confers an advantage to the organism, and is therefore favored by natural selection.

LINK The genetic code determines the amino acid that is encoded by each codon; see Figure 10.11

The rate of synonymous substitutions in most protein-coding genes is much higher than the rate of nonsynonymous substitutions. In other words, *substitution rates are highest at nucleotide*

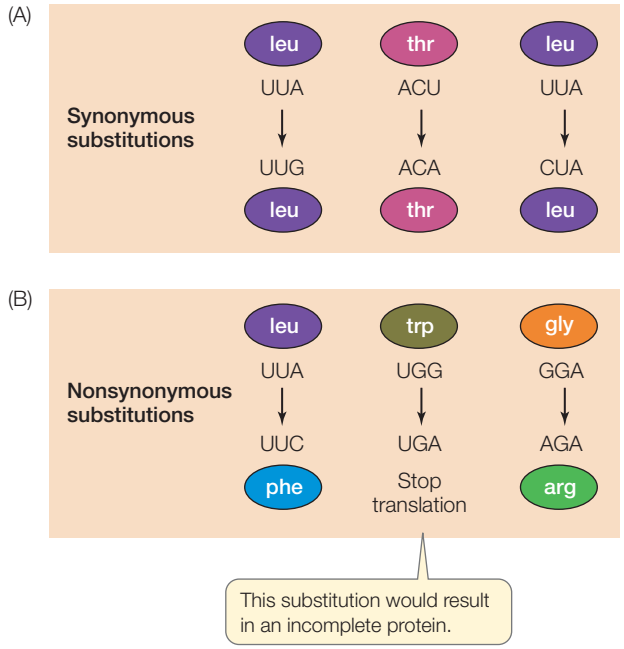


FIGURE 15.16 When One Nucleotide Changes (A) Synonymous substitutions do not change the amino acid specified and do not affect protein function. Such substitutions are less likely to be subject to natural selection, although they contribute greatly to the buildup of neutral genetic variation in a population. (B) Nonsynonymous substitutions do change the amino acid sequence and are likely to have an effect (often deleterious, but sometimes beneficial) on protein function. Such nucleotide substitutions are targets for natural selection.

positions that do not change the amino acid being expressed (**FIGURE 15.17**). The rate of substitution is even higher in **pseudogenes**, which are copies of genes that are no longer functional.

Insertions, deletions, and rearrangements of DNA sequences are all mutations that may affect a larger portion of the gene or genome than do point mutations (see Concept 9.3). Insertions and deletions of nucleotides in a protein-coding sequence interrupt its reading frame, unless they occur in multiples of three nucleotides (the length of one codon). Rearrangements may merely change the order of whole genes along chromosomes, or they may rearrange functional domains among individual genes.

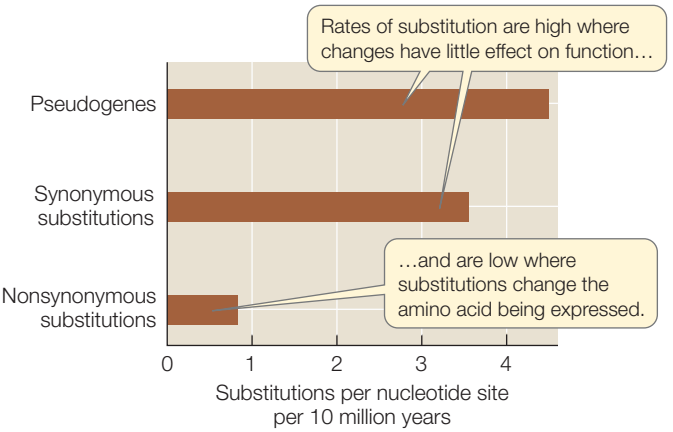


FIGURE 15.17 Rates of Substitution Differ Rates of nonsynonymous substitution are typically much lower than rates of synonymous substitution, and much lower than substitution rates in pseudogenes. This pattern reflects stronger stabilizing selection in functional genes than in pseudogenes.

When biologists began to examine the details of genetic variation of populations, they soon discovered many gene variants that had little or no effect on function. This gave rise to new ideas about how these *neutral variants* arise and spread in populations.

Much of molecular evolution is neutral

Motoo Kimura proposed the **neutral theory** in 1968. He suggested that, at the molecular level, the majority of variants found in most populations are selectively neutral. That is, most gene variants confer neither an advantage nor a disadvantage on their bearers. Therefore, these neutral variants must accumulate through genetic drift rather than through positive selection.

We saw in Concept 15.2 that genetic drift of existing gene variants tends to be greatest in small populations. However, the *rate of fixation* of neutral mutations by genetic drift is independent of population size. To see why this is so, consider a population of size N and a neutral mutation rate of μ (mu) per gamete per generation at a particular locus. The number of new mutations would be, on average, $\mu \times 2N$, because $2N$ gene copies are available to mutate in a population of diploid organisms. The probability that a given mutation will be fixed by drift alone is its frequency, which equals $1/(2N)$ for a newly arisen mutation. We can multiply these two terms to get the rate of fixation of neutral mutations in a given population of N individuals:

$$2N\mu \frac{1}{2N} = \mu$$

Therefore, the rate of fixation of neutral mutations depends only on the neutral mutation rate μ and is independent of population size. Any given mutation is more likely to appear in a large population than in a small one, but any mutation that does appear is more likely to become fixed in a small population. These two influences of population size cancel each other out. Therefore, the rate of fixation of neutral mutations is equal to the mutation rate.

As long as the underlying mutation rate is constant, macromolecules evolving in different populations should diverge from one another in neutral changes at a constant rate. The rate of evolution of particular genes and proteins is indeed often relatively constant over time, and therefore can be used as a "molecular clock" to calculate evolutionary divergence times between species (see Concept 16.3).

Although much of the genetic variation present in a population is the result of neutral evolution, the neutral theory does not imply that most mutations have no effect on the individual organism. Many mutations are never observed in populations because they are lethal or strongly detrimental, and the individuals that carry them are quickly removed from the population through natural selection. Similarly, because mutations that confer a selective advantage tend to be quickly fixed in

populations, they also do not result in significant variation at the population level. Nonetheless, if we compare homologous proteins from different populations or species, some amino acid positions will remain constant under purifying selection, others will vary through neutral genetic drift, and still others will differ between species as a result of positive selection for change. How can these evolutionary processes be distinguished?

Positive and purifying selection can be detected in the genome

Positive and purifying selection are defined with respect to the **fitness** of the genotype, or change in the relative frequency of the genotype in the population from one generation to the next. Genotypes of higher fitness increase in frequency over time; those of lower fitness decrease over time.

Relative rates of synonymous and nonsynonymous substitution differ among codons of protein-coding genes as a function of selection. Each codon specifies an amino acid residue in the encoded protein. Changes in some amino acid residues have a large effect on protein function, whereas other changes have little or no effect on function. The nature and rates of substitutions in the corresponding genes can identify codons and genes that are evolving under neutral or selective processes:

- If the rates of synonymous and nonsynonymous substitution at a codon position are very similar (that is, the ratio of the two rates is close to 1), then the corresponding amino acid residue is likely drifting neutrally among states.
- If the rate of nonsynonymous substitution exceeds the rate of synonymous substitution at a codon position, then positive selection likely accounts for change in the corresponding amino acid residue.
- If the rate of synonymous substitution exceeds the rate of nonsynonymous substitution at a codon position, then purifying selection is likely resisting change in the corresponding amino acid residue.

FRONTIERS Biologists have compared the complete genomes of humans and our closest living relatives, chimpanzees. Analysis of ratios of rates of nonsynonymous to synonymous substitution reveals hundreds of genes that are evolving under positive selection in one or both lineages. Further analysis of these genes is expected to provide insights into the major selective changes that have occurred in humans and chimpanzees since our most recent common ancestor.

The evolution of lysozyme illustrates how and why particular codons in a gene sequence might be under different modes of selection. Lysozyme is an enzyme that is found in almost all animals; it is produced in the tears, saliva, and milk of mammals and in the albumen (whites) of bird eggs. Lysozyme digests the cell walls of bacteria, rupturing and killing them. Most animals defend themselves against bacteria by digesting them, which is probably why most animals have lysozyme. Some animals, however, also use lysozyme to digest their food.

Among mammals, a mode of digestion called *foregut fermentation* has evolved twice. In mammals with this mode of digestion, the foregut—consisting of the posterior esophagus and/or the stomach—has been converted into a chamber in which bacteria break down ingested plant matter by fermentation. Foregut fermenters can obtain nutrients from the otherwise indigestible cellulose that makes up a large proportion of plant tissue. Foregut fermentation evolved independently in *ruminants* (a group of hoofed mammals that includes cattle) and in certain leaf-eating monkeys, such as langurs. We know that these evolutionary events were independent because both langurs and ruminants have close relatives that are not foregut fermenters.

In both mammalian foregut-fermenting lineages, lysozyme has been modified to play a new, nondefensive role. The modified lysozyme enzyme ruptures some of the bacteria that live in the foregut, releasing nutrients metabolized by the bacteria, which the mammal then absorbs. How many changes in the lysozyme molecule were needed to allow it to perform this function amid the digestive enzymes and acidic conditions of the mammalian foregut? To answer this question, biologists compared the lysozyme-coding sequences in foregut fermenters with those in several of their nonfermenting relatives. They determined which amino acids differed and which were shared among the species (**FIGURE 15.18A**), as well as the rates of synonymous and nonsynonymous substitution in lysozyme genes across the evolutionary history of the sampled species.

The researchers found that the rate of synonymous substitution within the gene that codes for lysozyme was much higher than the rate of nonsynonymous substitution. This observation indicates that many of the amino acids that make up lysozyme are evolving under purifying selection. In other words, there is selection against change in the lysozyme protein at these positions, and the encoded amino acids must therefore be critical for lysozyme function. At other positions, several different amino acids function equally well, and the corresponding codons have similar rates of synonymous and nonsynonymous substitution.

The most striking finding was that amino acid replacements in lysozyme happened at a much higher rate in the lineage leading to langurs than in any other primates. The high rate of nonsynonymous substitution in the langur lysozyme gene shows that lysozyme went through a period of rapid change in adapting to the stomachs of langurs. Moreover, the lysozymes of langurs and cattle share five convergent amino acid replacements, all of which lie on the surface of the lysozyme molecule, well away from the enzyme's active site. Several of these shared replacements are changes from arginine to lysine, which make the protein more resistant to degradation by the stomach enzyme pepsin. By understanding the functional significance of amino acid replacements, biologists can explain the observed changes in amino acid sequences in terms of changes in the functioning of the protein.

A large body of fossil, morphological, and molecular evidence shows that langurs and cattle do not share a recent common ancestor. However, langur and ruminant *lysozymes* share several amino acids that neither mammal shares with the

(A) *Semnopithecus* sp.

Bos taurus

(B) *Opisthocomus hoazin*

The lysozymes of langurs and cattle are convergent for 5 amino acid residues, indicative of the independent evolution of foregut fermentation in these two species.

	Langur	Baboon	Human	Rat	Cattle	Horse
Langur		14	18	38	32	65
Baboon	0		14	33	39	65
Human	0	1		37	41	64
Rat	0	0	0		55	64
Cattle	5	0	0	0		71
Horse	0	0	0	0	1	

FIGURE 15.18 Convergent Molecular Evolution of Lysozyme
(A) The numbers of amino acid differences in the lysozymes of several pairs of mammals are shown above the diagonal line; the numbers of similarities that arose from convergence between species are shown below the diagonal. The two foregut-fermenting species (cattle and langur) share five convergent amino acid replacements related to this digestive adaptation. (B) The hoatzin—the only known foregut-fermenting bird species—has been evolving independently from mammals for hundreds of millions of years but has independently evolved modifications to lysozyme similar to those found in cattle and langurs.

lysozymes of its own closer relatives. The lysozymes of these two mammals have undergone *convergent evolution* of some amino acid residues despite their very different ancestry. The amino acids they share give these lysozymes the ability to lyse the bacteria that ferment plant material in the foregut.

The hoatzin, an unusual leaf-eating South American bird (**FIGURE 15.18B**) and the only known avian foregut fermenter, offers another remarkable example of the convergent evolution of lysozyme. Many birds have an enlarged esophageal chamber called a *crop*. The hoatzin crop contains lysozyme and bacteria and acts as a fermentation chamber. Many of the amino acid replacements that occurred in the adaptation of hoatzin lysozyme are identical to those that evolved in ruminants and langurs. Thus, even though the hoatzin and foregut-fermenting mammals have not shared a common ancestor in hundreds of millions of years, similar adaptations have evolved in their lysozyme enzymes, enabling both groups to recover nutrients from fermenting bacteria.

APPLY THE CONCEPT

Genomes reveal both neutral and selective processes of evolution

Analysis of synonymous and nonsynonymous substitutions in protein-coding genes can be used to detect neutral evolution, positive selection, and purifying selection. An investigator compared many gene sequences that encode the protein hemagglutinin (a surface protein of influenza virus) sampled over time, and collected the data at right. Use the table to answer the following questions.

1. Which codon positions encode amino acids that have probably changed as a result of positive selection? Why?
2. Which codon position is most likely to encode an amino acid that drifts neutrally among states? Why?
3. Which codon positions encode amino acids that have probably changed as a result of purifying selection? Why?

CODON POSITION	NUMBER OF SYNONYMOUS SUBSTITUTIONS IN CODON	NUMBER OF NONSYNONYMOUS SUBSTITUTIONS IN CODON
12	0	7
15	1	9
61	0	12
80	7	0
137	12	1
156	24	2
165	3	4
226	38	3

Heterozygote advantage maintains polymorphic loci

In many cases, different alleles of a particular gene are advantageous under different environmental conditions. Most organisms, however, experience a wide diversity of environments. A night is dramatically different from the preceding day. A cold, cloudy day differs from a clear, hot one. Day length and temperature change seasonally. For many genes, a single allele is unlikely to perform well under all these conditions. In such situations, a heterozygous individual (with two different alleles) is likely to outperform individuals that are homozygous for either one of the alleles.

Colias butterflies of the Rocky Mountains live in environments where dawn temperatures often are too cold, and afternoon temperatures too hot, for the butterflies to fly. Populations of these butterflies are polymorphic for the gene that encodes phosphoglucose isomerase (PGI), an enzyme that influences how well an individual flies at different temperatures. Butterflies with certain PGI genotypes can fly better during the cold hours of early morning; those with other genotypes perform better during midday heat. The optimal body temperature for flight is 35°C–39°C, but some butterflies can fly with body temperatures as low as 29°C or as high as 40°C. Heat-tolerant genotypes are favored during spells of unusually hot weather; during spells of unusually cool weather, cold-tolerant genotypes are favored.

Heterozygous *Colias* butterflies can fly over a greater temperature range than homozygous individuals because they produce two different forms of PGI. This greater range of activity should give them an advantage in foraging and finding mates. A test of this prediction did find a mating advantage in heterozygous males, and further found that this mating advantage maintains the polymorphism in the population (**FIGURE 15.19**). The heterozygous condition can never become fixed in the population, however, because the offspring of two heterozygotes will always include both classes of homozygotes in addition to heterozygotes.

Genome size and organization also evolve

We know that genome size varies tremendously among organisms. Across broad taxonomic categories, there is some correlation between genome size and organismal complexity. The genome of the tiny bacterium *Mycoplasma genitalium* has only 470 genes. *Rickettsia prowazekii*, the bacterium that causes typhus, has 634 genes. *Homo sapiens*, by

INVESTIGATION

FIGURE 15.19 A Heterozygote Mating Advantage Among butterflies of the genus *Colias*, males that are heterozygous for two alleles of the PGI enzyme can fly farther under a broader range of temperatures than males that are homozygous for either allele. Does this ability give heterozygous males a mating advantage?

HYPOTHESIS

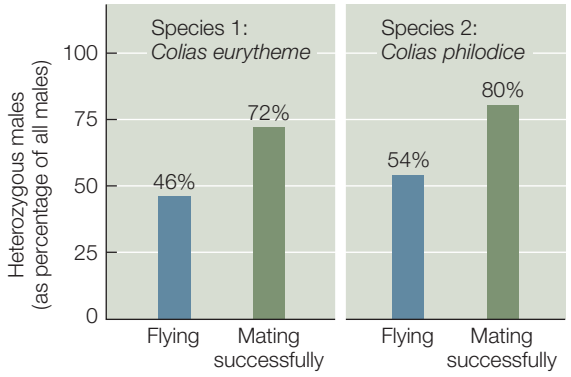
Heterozygous male *Colias* will have proportionally greater mating success than homozygous males.

METHOD

1. For each of two *Colias* species, capture butterflies in the field. In the laboratory, determine their genotypes and allow them to mate.
2. Determine the genotypes of the offspring, thus revealing paternity and mating success of the males.

RESULTS

For both species, the proportion of heterozygous males that mated successfully was higher than the proportion of all males seeking females ("flying").



CONCLUSION

Heterozygous *Colias* males have a mating advantage over homozygous males.

ANALYZE THE DATA

Analyze this sampling data collected during the experiment (only one of several samples is shown for each species).

Species	All viable males*		Mating males	
	Heterozygous/ total	% heterozygous	Heterozygous/ total	% heterozygous
C. philodice	32/74	43.2	31/50	62.0
C. eurytheme	44/92	47.8	45/59	76.3

*"Viable males" are all males captured flying with females (hence with the potential to mate)

A. Under the assumption that the proportions of each genotype should be the same as the proportions seen among all viable males, calculate the number of *mating males* expected to be heterozygous.

B. Use a chi-square test (see Appendix B) to evaluate the significance of the difference in your expected numbers in (A) and the observed percentages of heterozygous mating males. The critical value ($P = 0.05$) of the chi-square distribution with one degree of freedom is 3.841. Are the observed and expected numbers of heterozygotes among mating males significantly different in these samples?

For more, go to Working with Data 15.2 at **yourBioPortal.com**.

Go to **yourBioPortal.com** for original citations, discussions, and relevant links for all INVESTIGATION figures.

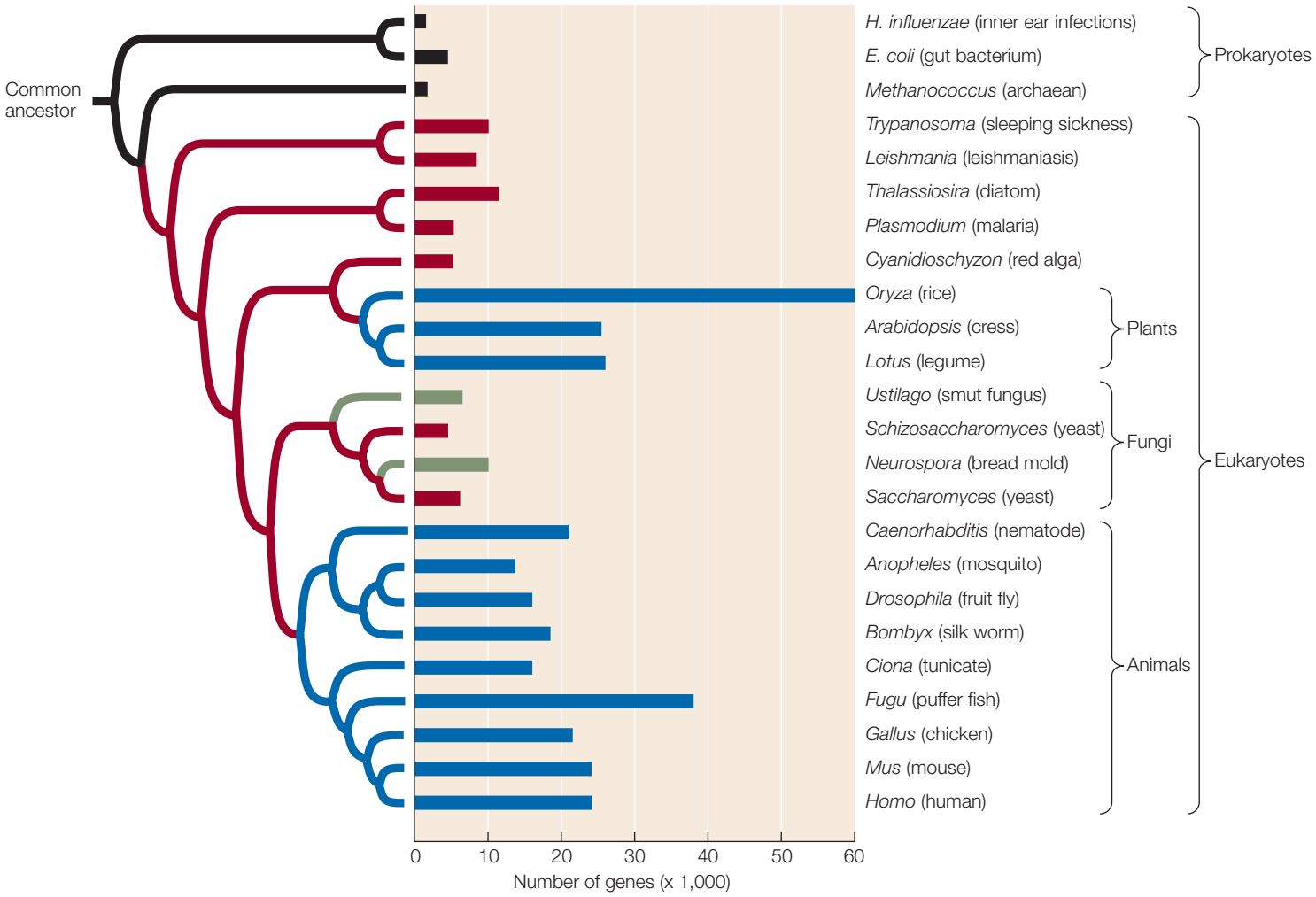


FIGURE 15.20 Genome Size Varies Widely This figure shows the number of genes from a sample of organisms whose genomes have been fully sequenced, arranged by their evolutionary relationships. Bacteria and archaea (black branches) typically have fewer genes than most eukaryotes. Among eukaryotes, multicellular organisms with tissue organization (plants and animals; blue branches) have more genes than single-celled organisms (red branches) or multicellular organisms that lack pronounced tissue organization (green branches).

contrast, has about 23,000 protein-coding genes. **FIGURE 15.20** shows the number of genes from a sample of organisms whose genomes have been fully sequenced, arranged by their evolutionary relationships. As this figure reveals, however, a larger genome does not always indicate greater complexity (compare rice with the other plants, for example). It is not surprising that more complex genetic instructions are needed for building and maintaining a large, multicellular organism than a small, single-celled bacterium. What is surprising is that some organisms, such as lungfishes, some salamanders, and lilies, have about 40 times as much DNA as humans do (**FIGURE 15.21**). Structurally, a lungfish or a lily is not 40 times more complex than a human. So why does genome size vary so much?

Differences in genome size are not so great if we take into account only the portion of DNA that actually encodes RNAs or proteins. The organisms with the largest total amounts of nuclear DNA (some ferns and flowering plants) have 80,000 times as much DNA as do the bacteria with the smallest genomes, but no species has more than about 100 times as many protein-coding genes as a bacterium. Therefore, much of the variation in genome size lies not in the number of functional genes, but in the amount of noncoding DNA (see Figure 15.21).

Why do the cells of most eukaryotic organisms have so much noncoding DNA? Does this noncoding DNA have a function? Although some of this DNA does not appear to have a direct function, it can alter the expression of the genes surrounding it. The degree or timing of gene expression can vary dramatically depending on the gene's position relative to noncoding sequences. Other regions of noncoding DNA consist of pseudogenes that are simply carried in the genome because the cost of doing so is very small. These pseudogenes may become the raw material for the evolution of new genes with novel functions. Some noncoding sequences function in maintaining chromosomal structure. Still others consist of parasitic transposable elements that spread through populations because they reproduce faster than the host genome.

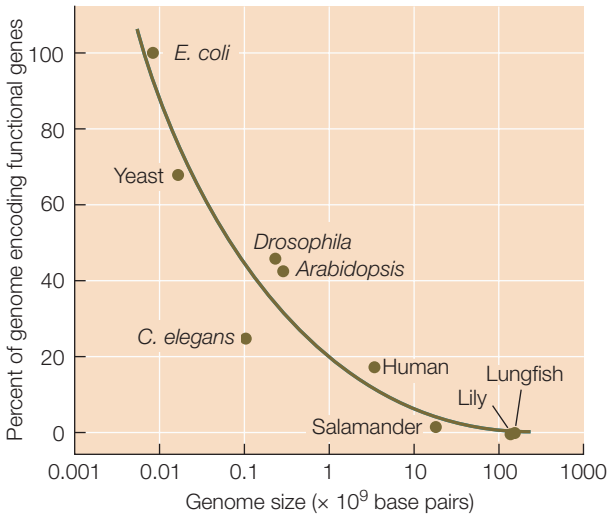


FIGURE 15.21 A Large Proportion of DNA Is Noncoding
Most of the DNA of bacteria and yeasts encodes RNAs or proteins, but a large percentage of the DNA of multicellular species is noncoding.

Another hypothesis is that the proportion of noncoding DNA is related primarily to population size. Noncoding sequences that are only slightly deleterious to the organism are likely to be purged by selection most efficiently in species with large population sizes. In species with small populations, the effects of genetic drift can overwhelm selection against noncoding sequences that have small deleterious consequences. Therefore, selection against the accumulation of noncoding sequences is most effective in species with large populations, so such species (such as bacteria or yeasts) have relatively little noncoding DNA compared with species with small populations (see Figure 15.21).

Do You Understand Concept 15.5?

- How can the ratio of synonymous to nonsynonymous substitutions be used to determine whether a particular gene is evolving neutrally, under positive selection, or under stabilizing selection?
- Why is the rate of fixation of neutral mutations independent of population size?
- Why do heterozygous individuals sometimes have an advantage over homozygous individuals?
- Why can a mutation that results in the replacement of one amino acid by another be a neutral event in some cases and in other cases be detrimental or beneficial? (Hint: Review the information about amino acids in Table 3.2 and the details of protein structure in Concept 3.2.)
- Postulate and contrast two hypotheses for the wide diversity of genome sizes among different organisms.

Most of our discussion so far has centered on changes in existing genes and phenotypes. Next we consider how new genes with novel functions arise in populations in the first place.

concept 15.6 Recombination, Lateral Gene Transfer, and Gene Duplication Can Result in New Features

Several evolutionary processes can result in the acquisition of major new characteristics in populations. Each of these processes results in larger and more rapid evolutionary changes than do single point mutations.

Sexual recombination amplifies the number of possible genotypes

In asexually reproducing organisms, each new individual is genetically identical to its parent unless there has been a mutation. When organisms reproduce sexually, however, offspring differ from their parents because of crossing over and independent assortment of chromosomes during meiosis, as well as the combination of genetic material from two different gametes, as described in Concept 7.4. Sexual recombination generates an endless variety of genotype combinations that increase the evolutionary potential of populations—a long-term advantage of sex. Although some species may reproduce asexually most of the time, most asexual species have some means of achieving genetic recombination.

The evolution of meiosis and sexual recombination was a crucial event in the history of life. Exactly how these mechanisms arose is puzzling, however, because in the short term, sex has at least three striking disadvantages:

- Recombination breaks up adaptive combinations of genes.
- Sex reduces the rate at which females pass genes on to their offspring.
- Dividing offspring into separate genders greatly reduces the overall reproductive rate.

To see why this last disadvantage exists, consider an asexual female that produces the same number of offspring as a sexual female. Assume that both females produce two offspring, but that half of the sexual female's offspring are males. In the next (F_1) generation, then, each of the two asexual F_1 females will produce two more offspring—but there is only one sexual F_1 female to produce offspring. Thus, the effective reproductive rate of the asexual lineage is twice that of the sexual lineage. The evolutionary problem is to identify the advantages of sex that can overcome such short-term disadvantages.

A number of hypotheses have been proposed to explain the existence of sex, none of which are mutually exclusive. One is that sexual recombination facilitates repair of damaged DNA, because breaks and other errors in DNA on one chromosome can be repaired by copying the intact sequence from the homologous chromosome.

Another advantage of sexual reproduction is that it permits the elimination of deleterious mutations through recombination followed by selection. As Concept 9.2 described, DNA replication is not perfect, and many replication errors result in lower fitness. Meiotic recombination distributes these deleterious mutations unequally among gametes. Sexual reproduction

then produces some individuals with more deleterious mutations and some with fewer. The individuals with fewer deleterious mutations are more likely to survive. Therefore, sexual reproduction allows natural selection to eliminate particular deleterious mutations from the population over time.

In asexual reproduction, deleterious mutations can be eliminated only by the death of the lineage or by a rare back mutation (that is, when a subsequent mutation returns a mutated sequence to its original DNA sequence). Hermann J. Muller noted that deleterious mutations in a non-recombining genome accumulate— "ratchet up" —at each replication. Mutations occur and are passed on each time a genome replicates, and these mutations accumulate with each subsequent generation. This accumulation of deleterious mutations in lineages that lack genetic recombination is known as *Muller's ratchet*.

Another explanation for the existence of sex is that the great variety of genetic combinations created in each generation can itself be advantageous. For example, genetic variation can be a defense against pathogens and parasites. Most pathogens and parasites have much shorter life cycles than their hosts and can rapidly evolve counteradaptations to host defenses. Sexual recombination might give the host's defenses a chance to keep up.

Sexual recombination does not directly influence the frequencies of alleles. Rather, *it generates new combinations of alleles on which natural selection can act*. It expands variation in quantitative characters by creating new genotypes. That is why artificial selection for bristle number in *Drosophila* (see Figure 15.6) resulted in flies that had either more or fewer bristles than the flies in the initial population had.

Lateral gene transfer can result in the gain of new functions

The tree of life is usually visualized as a branching diagram, with each lineage diverging into two (or more) lineages over time, from one common ancestor to the millions of species that are alive today. Ancestral lineages divide into descendant lineages, and it is those speciation events that the tree of life captures. However, there are also processes of **lateral gene transfer**, which allow individual genes, organelles, or fragments of genomes to move horizontally from one lineage to another. Some species may pick up fragments of DNA directly from the environment. A virus may pick up some genes from one host and transfer them to a new host when the virus becomes integrated into the new host's genome. Hybridization between species also results in the lateral transfer of large numbers of genes.

Lateral gene transfer can be highly advantageous to the species that incorporates novel genes from a distant relative. Genes that confer antibiotic resistance, for example, are commonly transferred among different species of bacteria. Lateral gene transfer is another way, in addition to mutation and recombination, that species can increase their genetic variation.

The degree to which lateral gene transfer events occur in various parts of the tree of life is a matter of considerable current investigation and debate. Lateral gene transfer appears to be relatively uncommon among most eukaryote lineages, although the two major endosymbioses that gave rise to mitochondria and chloroplasts involved lateral transfers of entire bacterial genomes to the eukaryote lineage. Some groups of eukaryotes, most notably some plants, are subject to relatively high levels of hybridization among closely related species. Hybridization leads to the exchange of many genes among recently separated lineages of plants. The greatest degree of lateral transfer, however, appears to occur among bacteria. Many genes have been transferred repeatedly among bacteria, to the point that relationships and boundaries among species of bacteria are sometimes hard to decipher.

Many new functions arise following gene duplication

Gene duplication is yet another way in which genomes can acquire new functions. When a gene is duplicated, one copy of that gene is potentially freed from having to perform its original function. The identical copies of a duplicated gene can have any one of four different fates:

- Both copies of the gene may retain their original function (which can result in a change in the amount of gene product that is produced by the organism).
- Both copies of the gene may retain the ability to produce the original gene product, but the expression of the genes may diverge in different tissues or at different times in development.
- One copy of the gene may be incapacitated by the accumulation of deleterious mutations and become a functionless pseudogene.
- One copy of the gene may retain its original function while the second copy changes and evolves a new function.

How often do gene duplications arise, and which of these four outcomes is most likely? Investigators have found that rates of gene duplication are fast enough for a yeast or *Drosophila* population to acquire several hundred duplicate genes over the course of a million years. They have also found that most of the duplicated genes that are still present in these organisms are very young. Many duplicated genes are lost from a genome within 10 million years—an eyeblink on an evolutionary time scale.

Many gene duplications affect only one or a few genes at a time, but in some cases entire genomes may be duplicated. When all the genes are duplicated, there are massive opportunities for new functions to evolve. That is exactly what seems to have happened during the course of vertebrate evolution. The genomes of the jawed vertebrates have four diploid sets of many major genes, which leads biologists to conclude that two genome-wide duplication events occurred in the ancestor of these species. These duplications allowed considerable specialization of individual vertebrate genes, many of which are now highly tissue-specific in their expression.

LINK See Concept 14.4 for a discussion of the role of duplicated Hox genes in vertebrate evolution

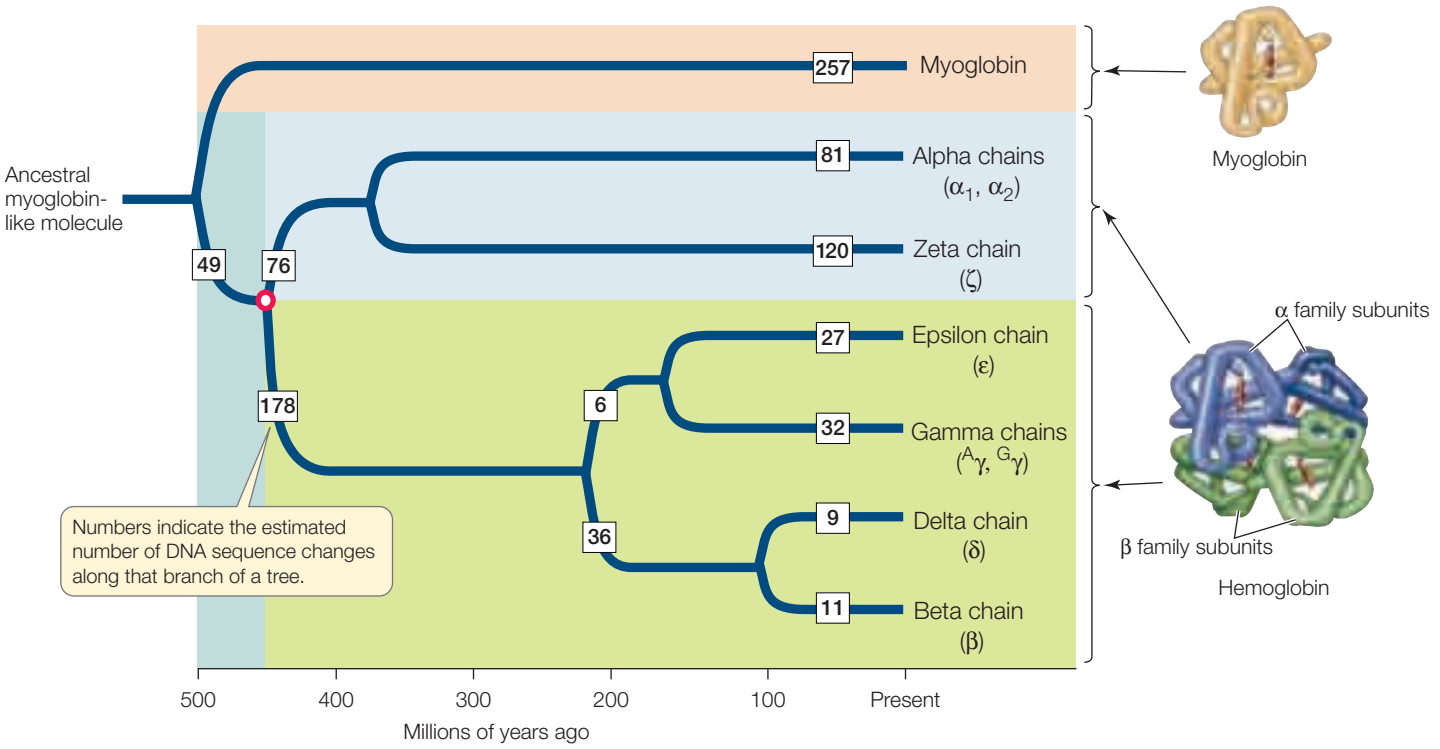


FIGURE 15.22 A Globin Family Gene Tree This gene tree suggests that the α-globin and β-globin gene clusters diverged about 450 million years ago (open circle), soon after the origin of the vertebrates.

yourBioPortal.com

Go to WEB ACTIVITY 15.1
Gene Tree Construction

Several successive rounds of duplication and sequence evolution may result in a *gene family*, a group of homologous genes with related functions, often arrayed in tandem along a chromosome. An example of this process is provided by the globin gene family (**FIGURE 15.22**). Comparisons of the amino acid sequences among globins strongly suggest that this family of proteins arose via gene duplications.

Hemoglobin is a tetramer (four-subunit molecule) consisting of two α-globin and two β-globin polypeptide chains. It carries oxygen in the blood. Myoglobin, a monomer, is the primary O_2 storage protein in muscle. Myoglobin's affinity for O_2 is much higher than that of hemoglobin, but hemoglobin has evolved to be more diversified in its role. Hemoglobin binds O_2 in the lungs or gills, where the O_2 concentration is relatively high, transports it to deep body tissues, where the O_2 concentration is low, and releases it in those tissues. With its more complex tetrameric structure, hemoglobin is able to carry four molecules of O_2, as well as hydrogen ions and carbon dioxide, in the blood. Hemoglobin and myoglobin are estimated to have arisen through gene duplication about 500 million years ago.

Do You Understand Concept 15.6?

- What are some of the potential advantages of lateral gene transfer to the organisms that gain new genes by this mechanism?
- Why is gene duplication considered important for long-term evolutionary change?
- Why is sexual reproduction so prevalent in nature, despite its having at least three short-term evolutionary disadvantages?

The development of evolutionary theory has helped reveal how biological molecules function, how genetic diversity is created and maintained, and how organisms develop new features. Next we will see how biologists put this theory into practice.

concept 15.7 Evolutionary Theory Has Practical Applications

Evolutionary theory has many practical applications across biology, and new ones are being developed every day. Here we'll discuss a few of these applications to fields such as agriculture, industry, and medicine.

Knowledge of gene evolution is used to study protein function

Earlier in this chapter we discussed the ways in which biologists can detect codons or genes that are under positive selection for change. These methods have greatly increased our

understanding of the functions of many genes. Consider, for example, the gated sodium channel genes. Sodium channels have many functions, including the control of nerve impulses in the nervous system (see Concept 34.2). Sodium channels can become blocked when they bind certain toxins, one of which is the tetrodotoxin (TTX) present in puffer fishes and many other animals. A human who eats puffer fish tissues that contain TTX can become paralyzed and die because the toxin-blocked sodium channels prevent nerves and muscles from functioning properly.

But puffer fish themselves have sodium channels, so why doesn't the TTX in their system paralyze them? Nucleotide substitutions in the puffer fish genome have resulted in structural changes in the proteins that form the sodium channels, and those changes prevent TTX from binding to the channel pore. Several different substitutions that result in such resistance have evolved in the various duplicated sodium channel genes of the many species of puffer fish. Many other changes that have nothing to do with the evolution of tetrodotoxin resistance have occurred in these genes as well.

So how does what we have learned about the evolution of TTX-resistant sodium channels affect our lives? Mutations in human sodium channel genes are responsible for a number of neurological pathologies. By studying the function of sodium channels and understanding which changes have produced tetrodotoxin resistance, we are learning a great deal about how these crucial channels work and how various mutations affect them. Biologists do this by comparing rates of synonymous and nonsynonymous substitutions across sodium channel genes in various animals that have evolved TTX resistance. In a similar manner, molecular evolutionary principles are used to understand function and diversification of function in many other proteins.

In vitro evolution produces new molecules

Living organisms produce thousands of compounds that humans have found useful. The search for naturally occurring compounds that can be used for pharmaceutical, agricultural, or industrial purposes has been termed *bioprospecting*. These compounds are the result of millions of years of molecular evolution across millions of species of living organisms. Yet biologists can imagine molecules that could have evolved but have not, in the absence of the right combination of selection pressures and opportunities.

For instance, we might want to find a molecule that binds a particular environmental contaminant so that the contaminant can be isolated and extracted from the environment. But if the contaminant is synthetic (not produced naturally), then it is unlikely that any living organism would have evolved a molecule with the function we desire. This problem was the inspiration for the field of **in vitro evolution**, in which new molecules are produced in the laboratory to perform novel and useful functions.

The principles of in vitro evolution are based on principles of molecular evolution that we have learned from the natural world. Consider a new RNA molecule that was produced in the laboratory using the principles of mutation and selection. The new molecule's intended function was to join two other RNA molecules (acting as a ribozyme with a function similar to that of the naturally occurring DNA ligase described in Concept 9.2, but for RNA molecules). The process started with a large pool of random RNA sequences (10^{15} different sequences, each about 300 nucleotides long), which were then selected for displaying any ligase activity (**FIGURE 15.23**). None were very effective ligases, but some were slightly better than others. The most functional of the ribozymes were selected and reverse-transcribed into cDNA (using the enzyme reverse transcriptase). The cDNA molecules were then amplified using the polymerase chain reaction (PCR; see Figure 9.15). PCR amplification is not perfect, and it introduced many new mutations

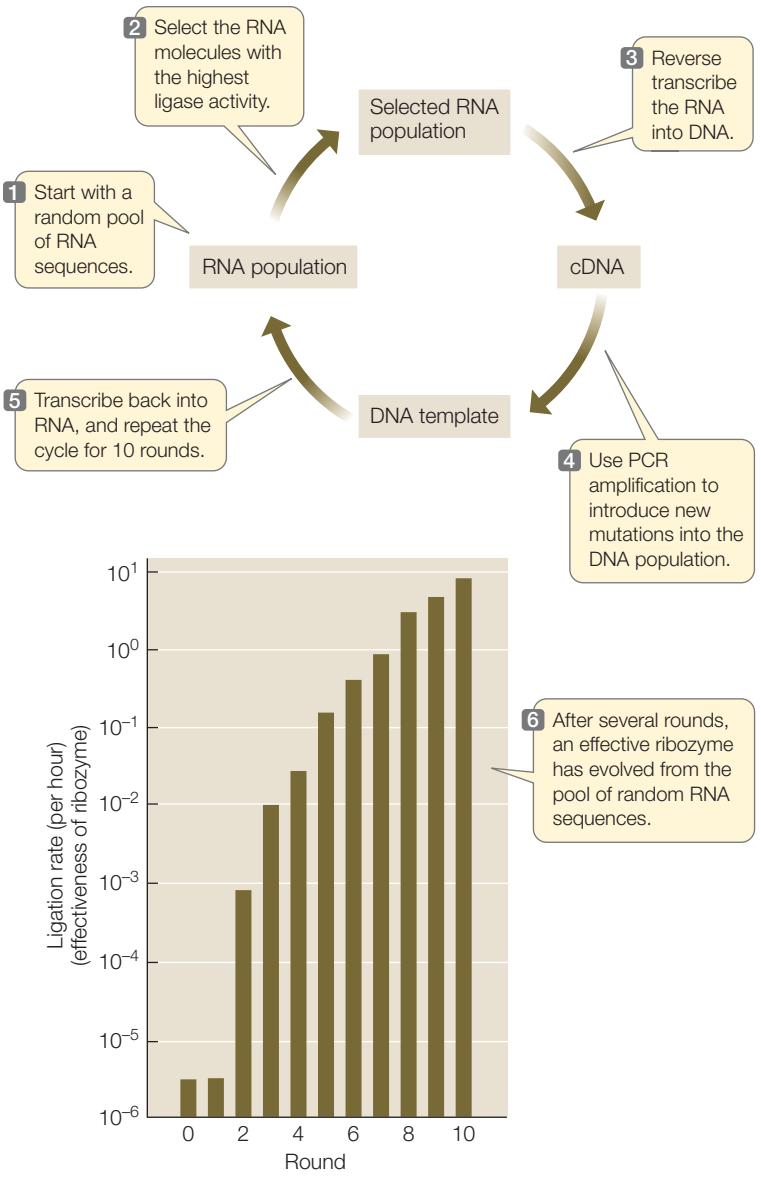


FIGURE 15.23 In Vitro Evolution Starting with a large pool of random RNA sequences, David Bartel and Jack Szostak of Massachusetts General Hospital produced a new ribozyme through rounds of mutation and selection for the ability to ligate RNA sequences.

into the pool of RNA sequences. These sequences were then transcribed back into RNA molecules using RNA polymerase, and the process was repeated.

The ligase activity of the RNAs evolved quickly; after 10 rounds of in vitro evolution, it had increased by about 7 million times. Similar techniques have been used to create a wide variety of molecules with novel enzymatic and binding functions.

FRONTIERS In vitro evolution of a bacterial enzyme has identified molecular changes that give rise to antibiotic resistance in bacteria. This research may lead to the design of new antibiotics that are harder for bacteria to evolve resistance against. Developing new, effective antibiotics is critical to human health, since many of the existing antibiotic drugs are rapidly losing their effectiveness against bacterial pathogens.

Evolutionary theory provides multiple benefits to agriculture

Well before humans had a clear understanding of evolution, they were selecting beneficial traits in the plants and animals they used for food. Modern agricultural practices have benefitted from a clearer understanding of evolutionary principles. Agriculturists have also used knowledge of evolutionary relationships and principles to incorporate beneficial genes into our food crops from many wild species.

Evolutionary theory has also proved important for understanding how to reduce the threats of pesticide and herbicide resistance. When farmers use the same pesticide over many seasons, the pests they are trying to kill gradually evolve resistance to the pesticide. Each year, a few pest individuals are slightly better at surviving in the presence of the pesticide, and those individuals produce most of the next generation of crop pests. Because their genes allow them to survive at a higher rate, and because they pass these resistant genes on to their offspring, pesticide resistance quickly evolves in the entire population. To combat this problem, evolutionary biologists have devised pesticide application and rotation schemes to reduce the rate of evolution of pesticide resistance, thus allowing farmers to use pesticides more effectively for longer periods of time.

Knowledge of molecular evolution is used to combat diseases

Many of the most problematic human diseases are caused by living, evolving organisms that present a moving target for modern medicine, as we described for influenza at the start of this chapter. The control of these and many other human diseases depends on techniques that can track the evolution of pathogenic organisms over time.

During the past century, transportation advances have allowed humans to move around the world with unprecedented speed and increasing frequency. Unfortunately, this mobility has increased the rate at which pathogens are transmitted among human populations, leading to the global emergence of many "new" diseases. Most of these emerging diseases are caused by viruses, and virtually all new viral diseases have been identified by evolutionary comparison of their genomes with those of known viruses. In recent years, rodent-borne hantaviruses have been identified as the source of widespread respiratory illnesses, and the virus that causes sudden acute respiratory syndrome (SARS) has been identified, as has its host, using evolutionary comparisons of genes. Studies of the origins, timing of emergence, and global diversity of many human pathogens (including HIV, the human immunodeficiency virus) depend on evolutionary principles and methods, as do efforts to develop effective vaccines against these pathogens.

At present, it is difficult to identify many common infections (the viral strains that cause "colds," for instance). As genomic databases increase, however, automated methods of sequencing and making evolutionary comparisons of sequences will allow us to identify and treat a much wider array of human (and other) diseases. Once biologists have collected genome data for enough infectious organisms, it will be possible to identify an infection by sequencing a portion of the pathogen's genome and comparing this sequence with other sequences on an evolutionary tree.

Do You Understand Concept 15.7?

- How can gene evolution be used to study protein function?
- How are principles of evolutionary biology used to identify emerging diseases?
- What are the key elements of in vitro evolution, and how do these elements correspond to natural evolutionary processes?

The mechanisms of evolution have produced a remarkable variety of organisms, some of which are adapted to most environments on Earth. In the next chapter, we will describe how biologists study the evolutionary relationships across the great diversity of life.

How do biologists use evolutionary theory to develop better flu vaccines?

ANSWER Many different strains of influenza virus circulate among human populations and other vertebrate hosts each year, but only a few of those strains survive to leave descendants. Selection among these circulating influenza strains results in rapid evolution of the viral genome. One of the ways that influenza strains differ is in the configuration of proteins on their surface. These surface proteins are the targets of recognition by the host immune system (**FIGURE 15.24**).

When changes occur in the surface proteins of an influenza virus, the host immune system may no longer detect the invading virus, so the virus is more likely to replicate

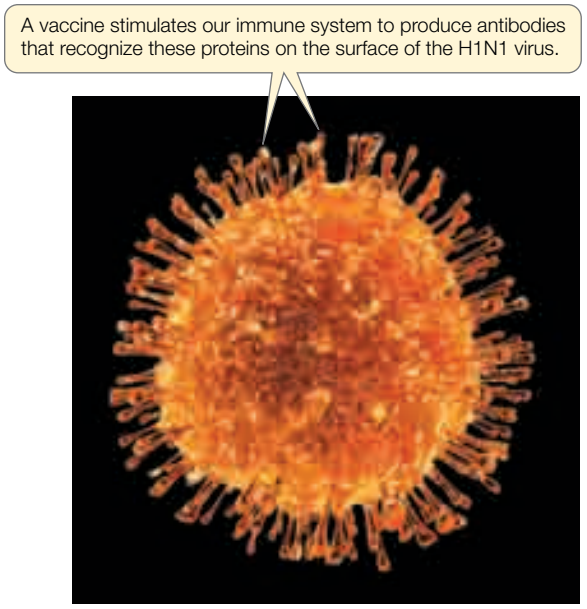


FIGURE 15.24 Evolutionary Analysis of Surface Proteins Leads to Improved Flu Vaccines This computer-generated image is of the H1N1 virus that was the target of a 2009–2010 flu vaccine. Rapidly evolving surface proteins ("spikes" in this illustration) allow flu viruses to escape detection by the host's immune system. Analyzing the surface proteins among current strains of the virus can help biologists anticipate which strains are most likely to be the cause of future epidemics.

successfully. The viral strains with the greatest number of changes to their surface proteins are most likely to escape detection by the host immune system, and are therefore most likely to spread among the host population and result in future flu epidemics. In other words, there is positive selection for change in the surface proteins of influenza.

By comparing the survival and proliferation rates of virus strains that have different gene sequences coding for their surface proteins, biologists can study adaptation of the viruses over time (Concept 15.2). If biologists can predict which of the currently circulating flu virus strains are most likely to escape host immune detection, then they can identify the strains of that are most likely to be involved in upcoming influenza epidemics and can target those strains for vaccine production.

How can biologists make such predictions? By examining the ratio of synonymous to nonsynonymous substitutions in genes that encode viral surface proteins, biologists can detect which codon changes (i.e., mutations) are under positive selection (Concept 15.5). They can then assess which of the currently circulating flu strains show the greatest number of changes in these positively selected codons. It is these flu strains that are most likely to survive and lead to the flu epidemics of the future, so they are the best targets for new vaccines. This practical application of evolutionary theory leads to more effective flu vaccines—and thus fewer illnesses and influenza-related deaths each year.

15 SUMMARY

concept 15.1 Evolution Is Both Factual and the Basis of Broader Theory

- **Evolution** is genetic change in populations over time. Evolution can be observed directly in living populations as well as in the fossil record of life.
- **Evolutionary theory** refers to our understanding and application of the mechanisms of evolutionary change.
- Charles Darwin in best known for his ideas on the common ancestry of divergent species and on **natural selection** as a mechanism of evolution. See ANIMATED TUTORIAL 15.1
- Since Darwin's time, many biologists have contributed to the development of evolutionary theory, and rapid progress in our understanding continues today. Review Figure 15.2

concept 15.2 Mutation, Selection, Gene Flow, Genetic Drift, and Nonrandom Mating Result in Evolution

- Mutation produces new genetic variants (**alleles**).
- Within populations, natural selection acts to increase the frequency of beneficial alleles and decrease the frequency of deleterious alleles.
- **Adaptation** refers both to a trait that evolves through natural selection and to the process that produces such traits.
- Migration or mating of individuals between populations results in **gene flow**.
- In small populations, **genetic drift**—the random loss of individuals and the alleles they possess—may produce large changes in allele frequencies from one generation to the next and greatly reduce genetic variation.
- **Population bottlenecks** occur when only a few individuals survive a random event, resulting in a drastic shift in allele frequencies within the population and the loss of variation. Similarly, a population established by a small number of individuals colonizing a new region may lose variation via a **founder effect**. Review Figure 15.7
- **Nonrandom mating** may result in changes in genotype frequencies in a population.
- **Sexual selection** results from differential mating success of individuals based on their phenotype. Review Figure 15.9 and WORKING WITH DATA 15.1

concept 15.3 Evolution Can Be Measured by Changes in Allele Frequencies

- Allele frequencies measure the amount of genetic variation in a population. Genotype frequencies show how a population's genetic variation is distributed among its members. Together, allele and genotype frequencies describe a population's **genetic structure**. Review Figure 15.10 and INTERACTIVE TUTORIAL 15.1
- **Hardy–Weinberg equilibrium** predicts genotype frequencies from allele frequencies in the absence of evolution. Deviation from these frequencies indicates that evolutionary mechanisms are at work. Review Figure 15.11 and ANIMATED TUTORIAL 15.2

concept 15.4 Selection Can Be Stabilizing, Directional, or Disruptive

- Natural selection can act on characters with quantitative variation in three different ways. Review Figure 15.12
- **Stabilizing selection** acts to reduce variation without changing the mean value of a trait.
- **Directional selection** acts to shift the mean value of a trait toward one extreme.
- **Disruptive selection** favors both extremes of trait values, resulting in a bimodal character distribution.

concept 15.5 Genomes Reveal Both Neutral and Selective Processes of Evolution

- **Nonsynonymous substitutions** of nucleotides result in amino acid replacements in proteins, but **synonymous substitutions** do not. Review Figure 15.16
- Rates of synonymous substitution are typically higher than rates of nonsynonymous substitution in protein-coding genes (a result of stabilizing selection). Review Figure 15.17
- Much of the change in nucleotide sequences over time is a result of neutral evolution. The rate of fixation of neutral mutations is independent of population size and is equal to the mutation rate.
- Positive selection for change in a protein-coding gene may be detected by a higher rate of nonsynonymous than synonymous substitution.
- Specific codons within a given gene sequence can be under different modes of selection. Review Figure 15.19 and WORKING WITH DATA 15.2
- The total size of genomes varies much more widely across multicellular organisms than does the number of functional genes. Review Figures 15.20 and 15.21
- Even though many noncoding regions of the genome may not have direct functions, these regions can affect the phenotype of an organism by influencing gene expression.
- Functionless **pseudogenes** can serve as the raw material for the evolution of new genes.

concept 15.6 Recombination, Lateral Gene Transfer, and Gene Duplication Can Result in New Features

- Despite its short-term disadvantages, sexual reproduction generates countless genotype combinations that increase genetic variation in populations.
- **Lateral gene transfer** can result in the rapid acquisition of new functions from distantly related species.
- **Gene duplications** can result in increased production of the gene's product, in divergence of the duplicated genes' expression, in pseudogenes, or in new gene functions. Several rounds of gene duplication can give rise to multiple genes with related functions, known as a gene family. Review Figure 15.22 and WEB ACTIVITY 15.1

concept 15.7 Evolutionary Theory Has Practical Applications

- Protein function can be studied by examining gene evolution. Detection of positive selection can be used to identify molecular changes that have resulted in functional changes.
- Agricultural applications of evolution include the development of new crop plants and domesticated animals, as well as a reduction in the rate of evolution of pesticide resistance.
- **In vitro evolution** is used to produce synthetic molecules with particular desired functions. Review Figure 15.23
- Many diseases are identified, studied, and combated through molecular evolutionary investigations.