CHAPTER OUTLINE

The African cheetah. *This species has a relatively low level of genetic variation because the population was reduced to a small size approximately 10,000 to 12,000 years ago.*

24 POPULATION GENETICS

Until now, we have primarily focused our attention on genes within individuals and their related family members. In this chapter and Chapters 25 and 26, we turn to the study of genes in a population or species. The field of **population genetics** is concerned with changes in genetic variation within a group of individuals over time. Population geneticists want to know the extent of genetic variation within populations, why it exists, and how it changes over the course of many generations. The field of population genetics emerged as a branch of genetics in the 1920s and 1930s. Its mathematical foundations were developed by theoreticians who extended the principles of Gregor Mendel and Charles Darwin by deriving formulas to explain the occurrence of genotypes within populations. These foundations can be largely attributed to three scientists: Sir Ronald Fisher, Sewall Wright, and J. B. S. Haldane. As we will see, support for their mathematical theories was provided by several researchers who analyzed the genetic composition of natural and experimental populations. More recently, population geneticists have used techniques to probe genetic variation at the molecular level. In addition, staggering advances in computer technology have aided population geneticists in the analysis of their genetic theories and data. In this chapter, we will explore the genetic variation that occurs in populations and consider the reasons why the genetic composition of populations may change over the course of several generations.

24.1 GENES IN POPULATIONS AND THE HARDY-WEINBERG EQUATION

Population genetics may seem like a significant departure from other topics in this textbook, but it is a direct extension of our understanding of Mendel's laws of inheritance, molecular genetics, and the ideas of Darwin. In the field of population genetics, the focus shifts away from the individual and toward the population of which the individual is a member. Conceptually, all of the alleles of every gene in a population make up the **gene pool.** In this regard, each member of the population is viewed as receiving its genes from its parents, which, in turn, are members of the

gene pool. Furthermore, individuals that reproduce contribute to the gene pool of the next generation. Population geneticists study the genetic variation within the gene pool and how such variation changes from one generation to the next. The emphasis is often on allelic variation. In this introductory section, we will examine some of the general features of populations and gene pools.

A Population Is a Group of Interbreeding Individuals That Share a Gene Pool

In genetics, the term population has a very specific meaning. With regard to sexually reproducing species, a **population** is a group of individuals of the same species that occupy the same region and can interbreed with one another. Many species occupy a wide geographic range and are divided into discrete populations. For example, distinct populations of a given species may be located on different continents, or populations on the same continent could be divided by a geographical feature such as a large mountain range.

A large population usually is composed of smaller groups called **local populations,** or **demes.** The members of a local population are far more likely to breed among themselves than with other members of the general population. Local populations are often separated from each other by moderate geographic barriers. As shown in **Figure 24.1**, the large ground finch (*Geospiza magnirostris*) is found on a small volcanic island called Daphne Major, which is one of the Galápagos Islands. Daphne Major is located northwest of the much larger Santa Cruz Island. The population of large ground finch on Daphne Major constitutes a local population of this species. Breeding is much more apt to occur among members of a local population than between members of neighboring populations. On relatively rare occasions, however, a bird may fly from Daphne Major to Santa Cruz Island, which would make breeding between the two different local populations possible.

Populations typically are dynamic units that change from one generation to the next. A population may change its size, geographic location, and genetic composition. With regard to size, natural populations commonly go through cycles of "feast or famine," during which environmental factors cause the population to swell or shrink. In addition, natural predators or disease may periodically decrease the size of a population to significantly lower levels; the population later may rebound to its original size. Populations or individuals within populations may migrate to a new site and establish a distinct population in this location. The environment at this new geographic location may differ from the original site. What are the consequences of such changes? As population sizes and locations change, their genetic composition generally changes as well. As described later, population geneticists have developed mathematical models that predict how the gene pool will change in response to fluctuations in size, migration, and new environments.

(a) Large ground finch (b) Daphne major

FIGURE 24.1 A local population of the large ground finch.
(a) The large ground finch (*Geospiza magnirostris*) on Daphne Major. **(b)** Daphne Major, one of the Galápagos Islands.

At the Population Level, Some Genes May Be Monomorphic, But Most Are Polymorphic

In population genetics, the term **polymorphism** (meaning many forms) refers to the observation that many traits display variation within a population. Historically, polymorphism first referred to the variation in traits that are observable with the naked eye. Polymorphisms in color and pattern have long attracted the attention of population geneticists. These include studies involving yellow and red varieties of the elder-flowered orchid, and brown, pink, and yellow land snails, which are discussed later in this chapter. **Figure 24.2** illustrates a striking example of polymorphism in the Hawaiian happy-face spider (*Theridion grallator*). The three individuals shown in this figure are from the same species, but they differ in alleles that affect color and pattern.

What is the underlying cause of polymorphism? At the DNA level, polymorphism may be due to two or more alleles that influence the phenotype of the individual that inherits them. In other words, it is due to genetic variation. Geneticists also use the term **polymorphic** to describe a gene that commonly exists as two or more alleles in a population. By comparison, a **monomorphic** gene exists predominantly as a single allele in a population. By convention, when a single allele is found in at least 99% of all cases, the gene is considered monomorphic. (Some geneticists view an allele frequency of 95% or greater to be monomorphic.)

At the level of a particular gene, a polymorphism may involve various types of changes such as a deletion of a significant region of the gene, a duplication of a region, or a change in a single nucleotide. This last phenomenon is called a **single-nucleotide polymorphism (SNP).** SNPs are the smallest type of genetic change that can occur within a given gene and are also the most common. In humans, for example, SNPs represent 90% of all the variation in DNA sequences that occurs among different people. SNPs are found very frequently in genes. In the human population, a gene that is 2000 to 3000 bp in length, on average, contains 10 different sites that are polymorphic. The high frequency of SNPs indicates that polymorphism is the norm for most human genes. Likewise, relatively large, healthy populations of nearly all species exhibit a high level of genetic variation as evidenced by the occurrence of SNPs within most genes.

Within a population, the alleles of a given gene may arise by different types of genetic changes. **Figure 24.3** considers a gene that exists in multiple forms in humans. This example is a

FIGURE 24.2 Polymorphism in the Hawaiian happy-face spider.
Genes → Traits These three spiders are members of the same species and carry the same genes. However, several genes that affect pigmentation patterns are polymorphic, meaning that more than one allele occurs in each gene within the population. This polymorphism within the Hawaiian happy-face spider population produces members that look quite different from each other.

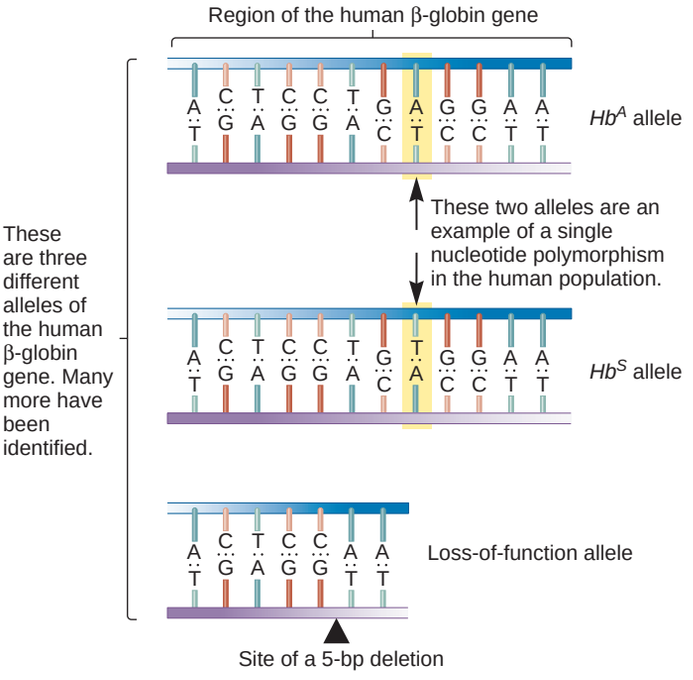


FIGURE 24.3 The relationship between alleles and various types of mutations. The DNA sequence shown here is a small portion of the β-globin gene in humans. Mutations have altered the gene to create the three different alleles in this figure. The top two alleles differ by a single base pair and are referred to as a single-nucleotide polymorphism (SNP). The bottom allele has a 5-bp deletion that begins right after the arrowhead. The deletion results in a nonfunctional polypeptide. It is a loss-of-function allele.

short segment of DNA found within the human β-globin gene. The top sequence is an allele designated Hb^A, whereas the middle sequence is called Hb^S. These alleles differ from each other by a single nucleotide, so they are an example of a single-nucleotide polymorphism. As discussed in Chapter 4, the Hb^S allele causes sickle cell disease in a homozygote. The bottom sequence contains a short, 5-bp deletion compared with the other two alleles. This deletion results in a nonfunctional β-globin polypeptide. Therefore, the bottom sequence is an example of a loss-of-function allele.

Population Genetics Is Concerned with Allele and Genotype Frequencies

As we have seen, population geneticists want to understand the prevalence of polymorphic genes within populations. Their goal is to identify the causative factors that govern changes in genetic variation. Much of their work evaluates the frequency of alleles in a quantitative way. Two fundamental calculations are central to population genetics: **allele frequencies** and **genotype frequencies.** The allele and genotype frequencies are defined as

$$\text{Allele frequency} = \frac{\text{Number of copies of an allele in a population}}{\text{Total number of all alleles for that gene in a population}}$$

$$\text{Genotype frequency} = \frac{\text{Number of individuals with a particular genotype in a population}}{\text{Total number of individuals in a population}}$$

Though these two frequencies are related, a clear distinction between them must be kept in mind. As an example, let's consider a hypothetical population of 100 frogs with the following genotypes:

64 dark green frogs with the genotype *GG*

32 medium green frogs with the genotype *Gg*

4 light green frogs with the genotype *gg*

When calculating an allele frequency, homozygous individuals have two copies of an allele, whereas heterozygotes have only one. For example, in tallying the *g* allele, each of the 32 heterozygotes has one copy of the *g* allele, and each light green frog has two copies. The allele frequency for *g* equals

$$g = \frac{32 + 2(4)}{2(64) + 2(32) + 2(4)}$$

$$= \frac{40}{200} = 0.2, \text{ or } 20\%$$

This result tells us that the allele frequency of *g* is 20%. In other words, 20% of the alleles for this gene in the population are the *g* allele.

Let's now calculate the genotype frequency of *gg* (light green) frogs:

$$gg = \frac{4}{64 + 32 + 4}$$

$$= \frac{4}{100} = 0.04, \text{ or } 4\%$$

We see that 4% of the individuals in this population are light green frogs.

Allele and genotype frequencies are always less than or equal to 1 (i.e., ≤100%). If a gene is monomorphic, the allele frequency for the single allele will equal or be close to a value of 1.0. For polymorphic genes, if we add up the frequencies for all of the alleles in the population, we should obtain a value of 1.0. In our frog example, the allele frequency of *g* equals 0.2. The frequency of the other allele, *G*, equals 0.8. If we add the two together, we obtain a value of 0.2 + 0.8 = 1.0.

The Hardy-Weinberg Equation Can Be Used to Calculate Genotype Frequencies Based on Allele Frequencies

Now that we have a general understanding of genes in populations, we can begin to relate these concepts to mathematical expressions as a way to examine whether allele and genotype frequencies are changing over the course of many generations. In 1908, a British mathematician, Godfrey Harold Hardy, and a German physician, Wilhelm Weinberg, independently derived a simple mathematical expression that predicts stability of allele and genotype frequencies from one generation to the next. It is called the **Hardy-Weinberg equilibrium,** because (under a given set of conditions, described later) the allele and genotype frequencies do not change over the course of many generations.

Why is the Hardy-Weinberg equilibrium a useful concept? An equilibrium is a null hypothesis, which suggests that evolutionary change is not occurring. In reality, however, populations rarely achieve an equilibrium. Therefore, the main usefulness of the Hardy-Weinberg equilibrium is that it provides a framework on which to understand changes in allele frequencies within a population when such an equilibrium is violated.

To appreciate the Hardy-Weinberg equilibrium, let's return to our hypothetical frog example in which a gene is polymorphic and exists as two different alleles: *G* and *g*. If the allele frequency of *G* is denoted by the variable *p*, and the allele frequency of *g* by *q*, then

$$p + q = 1$$

For example, if $p = 0.8$, then q must be 0.2. In other words, if the allele frequency of *G* equals 80%, the remaining 20% of alleles must be *g*, because together they equal 100%.

The Hardy-Weinberg equilibrium relates allele frequencies and genotype frequencies. For a diploid species, each individual inherits two copies of most genes. The Hardy-Weinberg equilibrium assumes that the alleles for the next generation for any given individual are chosen randomly and independently of each other. Therefore, we can use the product rule and multiply the sum, $p + q$, together. Because $p + q = 1$, we also know that their product also equals 1:

$$(p + q)(p + q) = 1$$

$$p^2 + 2pq + q^2 = 1 \textbf{ (Hardy-Weinberg equation)}$$

The Hardy-Weinberg equation applies to a gene in a diploid species that is found in only two alleles. If this equation is related to our hypothetical frog population in which a gene exists in alleles designated *G* and *g*, then

p^2 equals the genotype frequency of *GG*

$2pq$ equals the genotype frequency of *Gg*

q^2 equals the genotype frequency of *gg*

If $p = 0.8$ and $q = 0.2$ and if the population is in Hardy-Weinberg equilibrium, then

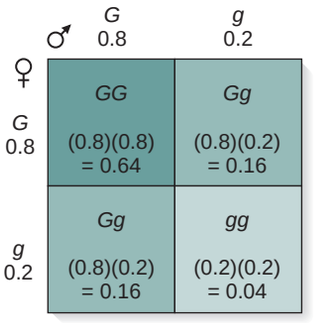
$$GG = p^2 = (0.8)^2 = 0.64$$

$$Gg = 2pq = 2(0.8)(0.2) = 0.32$$

$$gg = q^2 = (0.2)^2 = 0.04$$

In other words, if the allele frequency of *G* is 80% and the allele frequency of *g* is 20%, the genotype frequency of *GG* is 64%, *Gg* is 32%, and *gg* is 4%.

To illustrate the relationship between allele frequencies and genotypes, **Figure 24.4** compares the Hardy-Weinberg equation with the way that gametes combine randomly with each other to produce offspring. In a population, the frequency of a gamete carrying a particular allele is equal to the allele frequency in that population. In this example, the frequency of a gamete carrying the *G* allele equals 0.8.



GG genotype = 0.64 = 64%
Gg genotype = 0.16 + 0.16 = 0.32 = 32%
gg genotype = 0.04 = 4%

FIGURE 24.4 A comparison between allele frequencies and the union of alleles in a Punnett square. In a population in Hardy-Weinberg equilibrium, the frequency of gametes carrying a particular allele is equal to the allele frequency in the population.

We can use the product rule to determine the frequency of genotypes. For example, the frequency of producing a *GG* homozygote is $0.8 \times 0.8 = 0.64$, or 64%. Likewise, the probability of inheriting both *g* alleles is $0.2 \times 0.2 = 0.04$, or 4%. As seen in Figure 24.4, heterozygotes can be produced in two different ways. An offspring could inherit the *G* allele from its father and *g* from its mother, or *G* from its mother and *g* from its father. Therefore, the frequency of heterozygotes is $pq + pq$, which equals $2pq$; in our example, this is $2(0.8)(0.2) = 0.32$, or 32%.

One particularly useful feature of the Hardy-Weinberg equation is that it allows us to estimate the frequency of heterozygotes for recessive genetic diseases. As an example, let's consider cystic fibrosis, which involves a gene that encodes a chloride transporter. Persons afflicted with this disorder have an irregularity in salt and water balance. One of the symptoms is thick mucus in the lungs that can contribute to repeated lung infections. In Caucasian populations, the frequency of affected individuals is approximately 1 in 2500. Because this is a recessive disorder, affected individuals are homozygotes. If *q* represents the allele frequency of the disease-causing allele, then

$$q^2 = 1/2500$$

$$q^2 = 0.0004$$

We take the square root to determine *q*:

$$q = \sqrt{0.0004}$$

$$q = 0.02$$

If *p* represents the normal allele,

$$p = 1 - q$$

$$p = 1 - 0.02 = 0.98$$

The frequency of heterozygous carriers is

$$2pq = 2(0.98)(0.02) = 0.0392, \text{ or } 3.92\%$$

In the absence of evolutionary changes, the Hardy-Weinberg equation predicts an equilibrium—unchanging allele and genotype frequencies from generation to generation—if certain conditions are met in a population. With regard to the gene of interest, these conditions are as follows:

1. No new mutations: The gene of interest does not incur any new mutations.
2. No genetic drift: The population is so large that allele frequencies do not change due to random sampling effects.
3. No migration: Individuals do not travel between different populations.
4. No natural selection: All of the genotypes have the same reproductive success.
5. Random mating: With respect to the gene of interest, the members of the population mate with each other without regard to their phenotypes and genotypes.

The Hardy-Weinberg equation provides a quantitative relationship between allele and genotype frequencies in a population. **Figure 24.5** describes this relationship for different allele frequencies of *g* and *G*. As expected, when the allele frequency of *g* is very low, the *GG* genotype predominates; when the *g* allele frequency is high, the *gg* homozygote is most prevalent in the population. When the allele frequencies of *g* and *G* are intermediate in value, the heterozygote predominates.

In reality, no population satisfies the Hardy-Weinberg equilibrium completely. Nevertheless, in large natural populations with little migration and negligible natural selection, the Hardy-Weinberg equilibrium may be nearly approximated for certain genes. In addition, the Hardy-Weinberg equilibrium can be extended to situations in which a single gene exists in three or more alleles, as described in solved problem S4 at the end of the chapter.

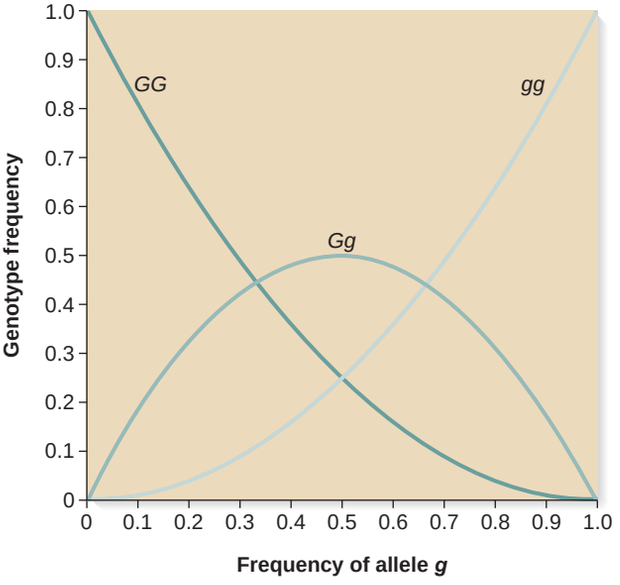


FIGURE 24.5 The relationship between allele frequencies and genotype frequencies according to the Hardy-Weinberg equilibrium. This graph assumes that *g* and *G* are the only two alleles for this gene.

As discussed in Chapter 2, the chi square test can determine if observed and expected data are in agreement. Therefore, we can use a chi square test to determine whether a population exhibits Hardy-Weinberg equilibrium for a particular gene. To do so, it is necessary to distinguish between the homozygotes and heterozygotes, either phenotypically or at the molecular level. This is necessary so that we can determine both the allele and genotype frequencies. As an example, let's consider a human blood type called the MN type. In this case, the blood type is determined by two codominant alleles, *M* and *N*. In an Inuit population in East Greenland, it was found that among 200 people, 168 were *MM*, 30 were *MN*, and 2 were *NN*. We can use these observed data to calculate the expected number of each genotype based on the Hardy-Weinberg equation.

$$\text{Allele frequency of } M = \frac{2(168) + 30}{400} = 0.915$$

$$\text{Allele frequency of } N = \frac{2(2) + 30}{400} = 0.085$$

Expected frequency of $MM = p^2 = (0.915)^2 = 0.837$

Expected number of *MM* individuals $= 0.837 \times 200 = 167.4$ (or 167 if rounded to the nearest individual)

Expected frequency of $NN = q^2 = (0.085)^2 = 0.007$

Expected number of *NN* individuals $= 0.007 \times 200 = 1.4$ (or 1 if rounded to the nearest individual)

Expected frequency of $MN = 2pq = 2(0.915)(0.085) = 0.156$

Expected number of *MN* individuals $= 0.156 \times 200 = 31.2$ (or 31 if rounded to the nearest individual)

$$\chi^2 = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} + \frac{(O_3 - E_3)^2}{E_3}$$

$$\chi^2 = \frac{(168 - 167)^2}{167} + \frac{(30 - 31)^2}{31} + \frac{(2 - 1)^2}{1}$$

$$= 1.04$$

Let's refer back to Table 2.1 in Chapter 2 to evaluate the calculated chi square value. Because the gene exists in two alleles, the degrees of freedom equal 1. For a chi square of 1.04, the *P* value is between 0.5 and 0.2, which is well within the acceptable range. Therefore, we fail to reject the null hypothesis. In this case, the alleles for this gene appear to be in Hardy-Weinberg equilibrium.

When researchers have investigated other genes in various populations, a high chi square value may be obtained, and the hypothesis is rejected that the allele and genotype frequencies are in Hardy-Weinberg equilibrium. In these cases, we would say that such a population is in **disequilibrium.** Deviation from a Hardy-Weinberg equilibrium indicates evolutionary change. As discussed next, factors such as genetic drift, migration, natural selection, and inbreeding may disrupt Hardy-Weinberg equilibrium. Therefore, when population geneticists discover that a population is not in equilibrium, they try to determine which factors are at work.

24.2 FACTORS THAT CHANGE ALLELE AND GENOTYPE FREQUENCIES IN POPULATIONS

The genetic variation in natural populations typically changes over the course of many generations. The term **microevolution** describes changes in a population's gene pool from generation to generation. Such change is rooted in two related phenomena (**Table 24.1**). First, the introduction of new genetic variation into a population is one essential aspect of microevolution. As discussed later in this chapter, gene variation can originate by a variety of molecular mechanisms. For example, new alleles of preexisting genes can arise by random mutations. Such events provide a continuous source of new variation to populations. However, due to their low rate of occurrence, mutations do not act as a major factor in promoting widespread changes in a population. If mutations were the only type of change occurring in a

TABLE 24.1
Factors That Govern Microevolution

Source of New Allelic Variation*	
Mutation	In this section, we consider allelic variation. Random mutations within preexisting genes introduce new alleles into populations, but at a very low rate. New mutations may be beneficial, neutral, or deleterious. For new alleles to rise to a significant percentage in a population, evolutionary mechanisms (i.e., random genetic drift, migration, natural selection) must operate on them.
Mechanisms That Alter Existing Genetic Variation	
Random genetic drift	This is a change in genetic variation from generation to generation due to random sampling error. Allele frequencies may change as a matter of chance from one generation to the next. This tends to have a greater effect in a small population.
Migration	Migration can occur between two different populations. The introduction of migrants into a recipient population may change the allele frequencies of that population.
Natural selection	This is the phenomenon in which the environment selects for individuals that possess certain traits that favor reproductive success. For example, natural selection may be related to the survival of members to reproductive age.
Nonrandom mating	This is the phenomenon in which individuals select mates based on their phenotypes or genetic lineage. This can alter the relative proportion of homozygotes and heterozygotes predicted by the Hardy-Weinberg equation but does not change allele frequencies.

*Allelic variation is just one source of new genetic variation. Section 24.3 considers a variety of mechanisms through which new genetic variation can occur.

population, that population would not evolve at a significant rate because mutations are so rare.

Microevolution also involves the action of evolutionary mechanisms that alter the prevalence of a given allele or genotype in a population. These mechanisms are random genetic drift, migration, natural selection, and nonrandom mating (see Table 24.1). The collective contributions of these evolutionary mechanisms over the course of many generations have the potential to promote widespread genetic changes in a population. In this section, we will examine how random genetic drift, migration, natural selection, and nonrandom mating can affect the type of genetic variation that occurs when a gene exists in two or more alleles in a population. As you will learn, these mechanisms may cause a particular allele to be favored, or they may create a balance where two or more alleles are maintained in a population.

Mutations Provide the Source of Genetic Variation

As discussed in Chapters 8 and 16, mutations involve changes in gene sequences, chromosome structure, and/or chromosome number. Mutations are random events that occur spontaneously at a low rate or are caused by mutagens at a higher rate. In 1926, the Russian geneticist Sergei Chetverikov was the first to suggest that mutational variability provides the raw material for evolution but does not constitute evolution itself. In other words, mutation can provide new alleles to a population but does not substantially alter allele frequencies. Chetverikov proposed that populations in nature absorb mutations like a sponge and retain them in a heterozygous condition, thereby providing a source of variability for future change.

Population geneticists often consider how new mutations affect the survival and reproductive potential of the individual that inherits them. A new mutation may be deleterious, neutral, or beneficial. For genes that encode proteins, the effects of new mutations depend on their influence on protein function. Deleterious and neutral mutations are far likelier to occur than beneficial ones. For example, alleles can be altered in many different ways that render an encoded protein defective. As discussed in Chapter 16, deletions and point mutations such as frameshift mutations, missense mutations, and nonsense mutations all may cause a gene to express a protein that is nonfunctional or less functional than the wild-type protein. Also, mutations in noncoding regions can alter gene expression (refer back to Table 16.3). Neutral mutations can also occur in several different ways. For example, a neutral mutation can change the wobble base without affecting the amino acid sequence of the encoded protein, or it can be a missense mutation that has no effect on protein function. Such point mutations occur at specific sites within the coding sequence. Neutral mutations can also occur within introns, the noncoding sequences of genes. By comparison, beneficial mutations are relatively uncommon. To be advantageous, a new mutation could alter the amino acid sequence of a protein to yield a better-functioning product. Although such mutations do occur, they are expected to be very rare for a population in a stable environment.

The **mutation rate** is defined as the probability that a gene will be altered by a new mutation. The rate is typically expressed as the number of new mutations in a given gene per generation. A common value for the mutation rate is in the range of 1 in 100,000 to 1 in 1,000,000, or 10^{-5} to 10^{-6} per generation. However, mutation rates vary depending on species, cell types, chromosomal location, and gene size. Furthermore, in experimental studies, the mutation rate is usually measured by following the change of a normal (functional) gene to a deleterious (nonfunctional) allele. The mutation rate producing beneficial alleles is expected to be substantially less.

It is clear that new mutations provide genetic variability, but population geneticists also want to know how much the mutation rate affects the allele frequencies in a population. Can random mutations have a large effect on allele frequencies over time? To answer this question, let's take this simple case: a gene exists in an allele, *A*; the allele frequency of *A* is denoted by the variable p. A mutation can convert the *A* allele into a different allele called *a*. The allele frequency of *a* is designated by q. The conversion of the *A* allele into the *a* allele by mutation occurs at a rate that is designated μ. If we assume that the rate of the reverse mutation (*a* to *A*) is negligible, the increase in the frequency of the *a* allele after one generation is

$$\Delta q = \mu p$$

For example, let's consider the following conditions:

$p = 0.8$ (i.e., frequency of *A* is 80%)

$q = 0.2$ (i.e., frequency of *a* is 20%)

$\mu = 10^{-5}$ (i.e., the mutation rate of converting *A* to *a*)

$\Delta q = (10^{-5})(0.8) = (0.00001)(0.8) = 0.000008$

Therefore, in the next generation (designated $n + 1$),

$q_{n+1} = 0.2 + 0.000008 = 0.200008$

$p_{n+1} = 0.8 - 0.000008 = 0.799992$

As we can see from this calculation, new mutations do not significantly alter the allele frequencies in a single generation.

We can use the following equation to calculate the change in allele frequency after any number of generations:

$$(1 - \mu)^t = \frac{p_t}{p_0}$$

where

μ is the mutation rate of the conversion of *A* to *a*

t is the number of generations

p_0 is the allele frequency of *A* in the starting generation

p_t is the allele frequency of *A* after t generations

As an example, let's suppose that the allele frequency of *A* is 0.8, $\mu = 10^{-5}$, and we want to know what the allele frequency will be after 1000 generations ($t = 1000$). Plugging these values into the preceding equation and solving for p_t,

$$(1 - 0.00001)^{1000} = \frac{p_t}{0.8}$$

$$p_t = 0.792$$

Therefore, after 1000 generations, the frequency of *A* has dropped only from 0.8 to 0.792. Again, these results point to how slowly the occurrence of new mutations changes allele frequencies. In natural populations, the rate of new mutation is rarely a significant catalyst in shaping allele frequencies. Instead, other processes such as genetic drift, migration, and natural selection have far greater effects on allele frequencies. Next, we will examine how these other factors work.

In Small Populations, Allele Frequencies Can Be Altered by Random Genetic Drift

In the 1930s, geneticist Sewall Wright played a key role in developing the concept of **random genetic drift,** or simply, **genetic drift,** which refers to changes in allele frequencies in a population due to random fluctuations. As a matter of chance, the frequencies of alleles found in gametes that unite to form zygotes vary from generation to generation. Over the long run, genetic drift usually results in either the loss of an allele or its fixation at 100% in the population. The rate at which this occurs depends on the population size and on the initial allele frequencies. **Figure 24.6** illustrates the potential consequences of genetic drift in one large ($N = 1000$) and five small ($N = 20$) populations. At the beginning of this hypothetical simulation, all of these populations have identical allele frequencies: $A = 0.5$ and $a = 0.5$. In the five small populations, this allele frequency fluctuates substantially from generation to generation. Eventually, one of the alleles is eliminated and the other is fixed at 100%. At this point, the allele has become monomorphic and cannot fluctuate any further. By comparison, the allele frequencies in the large population fluctuate much less, because random sampling error is expected to have a smaller effect. Nevertheless, genetic drift leads to homozygosity even in large populations, but this takes many more generations to occur.

Now let's ask two questions:

1. How many new mutations do we expect in a natural population?
2. How likely is it that any new mutation will be either fixed in, or eliminated from, a population due to random genetic drift?

With regard to the first question, the average number of new mutations depends on the mutation rate (μ) and the number of individuals in a population (*N*). If each individual has two copies of the gene of interest, the expected number of new mutations in this gene is

$$\text{Expected number of new mutations} = 2N\mu$$

From this, we see that a new mutation is more likely to occur in a large population than in a small one. This makes sense, because the larger population has more copies of the gene to be mutated.

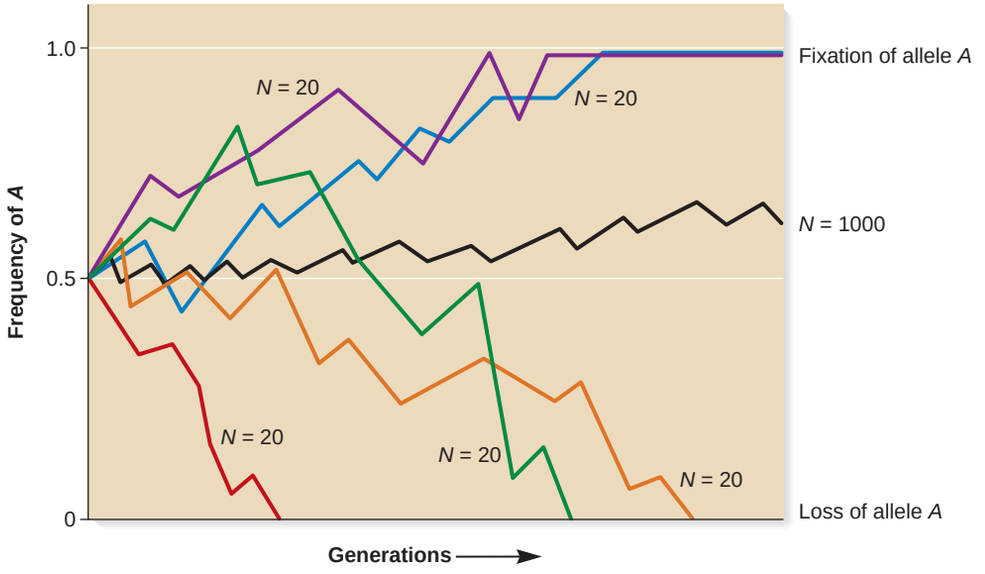
With regard to the second question, the probability of fixation of a newly arising allele due to genetic drift is

$$\text{Probability of fixation} = \frac{1}{2N}$$ (assuming equal numbers of males and females contribute to the next generation)

In other words, the probability of fixation is the same as the initial allele frequency in the population. For example, if $N = 20$, the probability of fixation of a new allele equals $1/(2 \times 20)$, or 2.5%. Conversely, a new allele may be lost from the population.

$$\text{Probability of elimination} = 1 - \text{probability of fixation}$$
$$= 1 - \frac{1}{2N}$$

If $N = 20$, the probability of elimination equals $1 - 1/(2 \times 20)$, or 97.5%. As you may have noticed, the value of *N* has opposing effects with regard to new mutations and their eventual fixation in a population. When *N* is very large, new mutations are much more likely to occur. Each new mutation, however, has a greater chance



ONLINE ANIMATION

FIGURE 24.6 **A hypothetical simulation of random genetic drift.** In all cases, the starting allele frequencies are $A = 0.5$ and $a = 0.5$. The colored lines illustrate five populations in which $N = 20$; the black line shows a population in which $N = 1000$.

of being eliminated from the population due to random genetic drift. On the other hand, when N is small, the probability of new mutations is also small, but if they occur, the likelihood of fixation is relatively large.

Now that we have an appreciation for the phenomenon of genetic drift, we can ask a third question:

3. If fixation of a new allele does occur, how many generations is it likely to take?

The formula for calculating this also depends on the number of individuals in the population:

$$\bar{t} = 4N$$

where

$\bar{t}$ equals the average number of generations to achieve fixation

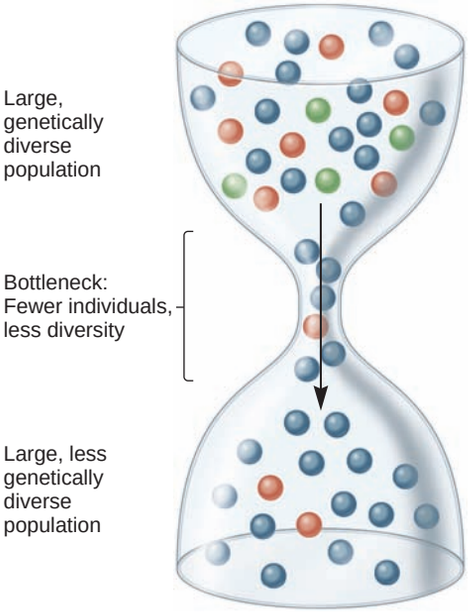
N equals the number of individuals in the population, assuming that males and females contribute equally to each succeeding generation

As you may have expected, allele fixation takes much longer in large populations. If a population has 1 million breeding members, it takes, on average, 4 million generations, perhaps an insurmountable period of time, to reach fixation. In a small group of 100 individuals, however, fixation takes only 400 generations, on average. As discussed in Chapter 26, the drifting of neutral alleles among different populations and species provides a way to measure the rate of evolution and can be used to determine evolutionary relationships.

The preceding discussion of random genetic drift has emphasized two important points. First, genetic drift ultimately operates in a random manner with regard to allele frequency and, over the long run, leads to either allele fixation or elimination. The process is random with regard to particular alleles. Genetic drift can lead to the fixation of deleterious, neutral, or beneficial alleles. A second important feature of genetic drift is that its influence is greatly affected by population size. Genetic drift may lead more quickly to allele loss or fixation in a small population.

In nature, there are different ways that geography and population size influence how genetic drift affects the genetic composition of a species. Some species occupy wide ranges in which small, local populations become geographically isolated from the rest of the species. The allele frequencies within these small populations are more susceptible to genetic drift. Because this is a random process, small isolated populations tend to be more genetically disparate in relation to other populations. This may occur in the following two ways.

Bottleneck Effect Changes in population size may influence genetic drift via the **bottleneck effect.** In nature, a population can be reduced dramatically in size by events such as earthquakes, floods, drought, or human destruction of habitat. Such events may randomly eliminate most of the members of the population without regard to genetic composition. The initial bottleneck may be greatly influenced by genetic drift because the surviving members may have allele frequencies that differ from those of the original population. In addition, allele frequencies are expected to drift substantially during the generations when the population size is small. In extreme cases, alleles may even be eliminated. Eventually, the population with the bottleneck may regain its original size (Figure 24.7). However, the new population has less genetic variation than the original large population. As an example, the African cheetah population lost a substantial amount of its genetic variation due to a bottleneck effect. DNA analysis by population geneticists has suggested that a severe bottleneck occurred approximately 10,000 to 12,000 years ago when the population size was dramatically reduced. The population eventually rebounded, but the bottleneck significantly decreased the genetic variation.



(a) Bottleneck effect

(b) An African cheetah

FIGURE 24.7 The bottleneck effect, an example of genetic drift. **(a)** A representation of the bottleneck effect. Note that the genetic variation denoted by the green balls has been lost. **(b)** The African cheetah. The modern species has low genetic variation due to a genetic bottleneck that is thought to have occurred about 10,000 to 12,000 years ago.

Founder Effect Geography and population size may also influence genetic drift via the **founder effect.** Compared to the

bottleneck effect, the key difference is that the founder effect involves migration; a small group of individuals separates from a larger population and establishes a colony in a new location. For example, a few individuals may migrate from a large continental population and become the founders of an island population. The founder effect has two important consequences. First, the founding population is expected to have less genetic variation than the original population from which it was derived. Second, as a matter of chance, the allele frequencies in the founding population may differ markedly from those of the original population.

Population geneticists have studied many examples of isolated populations that were started from a few members of another population. In the 1960s, Victor McKusick studied allele frequencies in the Old Order Amish of Lancaster County, Pennsylvania. At that time, this was a group of about 8000 people, descended from just three couples that immigrated to the United States in 1770. Among this population of 8000, a genetic disease known as the Ellis-van Creveld syndrome (a recessive form of dwarfism) was found at a frequency of 0.07, or 7%. By comparison, this disorder is extremely rare in other human populations, even the population from which the founding members had originated. The high frequency of dwarfism in the Lancaster County population is a chance occurrence due to the founder effect. Perhaps one of the original founders carried the recessive allele for dwarfism, or a new mutation may have occurred in one of their early descendants.

Migrations Between Two Populations Can Alter Allele Frequencies

We have just seen how migration to a new location by a relatively small group can result in a population with an altered genetic composition. In addition, migration between two different established populations can alter allele frequencies. For example, a species of birds may occupy two geographic regions that are separated by a large body of water. On rare occasions, the prevailing winds may allow birds from the western population to fly over this body of water and become members of the eastern population. If the two populations have different allele frequencies and if migration occurs in sufficient numbers, this migration may alter the allele frequencies in the eastern population.

After migration has occurred, the new (eastern) population is called a **conglomerate.** To calculate the allele frequencies in the conglomerate, we need two kinds of information. First, we must know the original allele frequencies in the donor and recipient populations. Second, we must know the proportion of the conglomerate population that is due to migrants. With these data, we can calculate the change in allele frequency in the conglomerate population using the following equation:

$$\Delta p_C = m(p_D - p_R)$$

where

- Δp_C is the change in allele frequency in the conglomerate population
- p_D is the allele frequency in the donor population
- p_R is the allele frequency in the original recipient population
- m is the proportion of migrants that make up the conglomerate population

$$m = \frac{\text{Number of migrants in the conglomerate population}}{\text{Total number of individuals in the conglomerate population}}$$

As an example, let's suppose the allele frequency of *A* is 0.7 in the donor population and 0.3 in the recipient population. A group of 20 individuals migrates and joins the recipient population, which originally had 80 members. Thus,

$$m = \frac{20}{20 + 80}$$
$$= 0.2$$
$$\Delta p_C = m(p_D - p_R)$$
$$= 0.2(0.7 - 0.3)$$
$$= 0.08$$

We can now calculate the allele frequency in the conglomerate:

$$p_C = p_R + \Delta p_C$$
$$= 0.3 + 0.08 = 0.38$$

Therefore, in the conglomerate population, the allele frequency of *A* has changed from 0.3 (its value before migration) to 0.38. This increase in allele frequency arises from the higher allele frequency of *A* in the donor population. **Gene flow** occurs whenever individuals migrate between populations and the migrants are able to breed successfully with the members of the recipient population. Gene flow depends not only on migration, but also on the ability of the migrants' alleles to be passed to subsequent generations.

In our previous example, we considered the consequences of a unidirectional migration from a donor to a recipient population. In nature, it is common for individuals to migrate in both directions. What are the main consequences of bidirectional migration? Depending on its rate, migration tends to reduce differences in allele frequencies between neighboring populations. In fact, population geneticists can analyze allele frequencies in two different populations to evaluate the rate of migration between them. Populations that frequently mix their gene pools via migration tend to have similar allele frequencies, whereas isolated populations are expected to be more disparate. By comparison, genetic drift, which we considered earlier, tends to make local populations more disparate from each other.

Migration can also enhance genetic diversity within a population. As mentioned, new mutations are relatively rare events. Therefore, a particular mutation may arise only in one population. Migration may then introduce this new allele into neighboring populations.

Natural Selection Is Based on the Relative Reproductive Success of Genotypes

In the 1850s, Charles Darwin and Alfred Russel Wallace independently proposed the theory of evolution by **natural selection.** According to this theory, the conditions found in nature

result in the selective survival and reproduction of individuals whose characteristics make them better adapted to their environment. These surviving individuals are more likely to reproduce and contribute offspring to the next generation. Natural selection can be related not only to differential survival but also to mating efficiency and fertility.

A modern restatement of the principles of natural selection can relate our knowledge of molecular genetics to the phenotypes of individuals.

1. Within a population, allelic variation arises in various ways, such as through random mutations that cause differences in DNA sequences. A mutation that creates a new allele may alter the amino acid sequence of the encoded protein, which, in turn, may alter the function of the protein.
2. Some alleles may encode proteins that enhance an individual's survival or reproductive capability compared with that of other members of the population. For example, an allele may produce a protein that is more efficient at a higher temperature, conferring on the individual a greater probability of survival in a hot climate.
3. Individuals with beneficial alleles are more likely to survive and contribute to the gene pool of the next generation.
4. Over the course of many generations, allele frequencies of many different genes may change through this process, thereby significantly altering the characteristics of a species. The net result of natural selection is a population that is better adapted to its environment and more successful at reproduction. Even so, it should be emphasized that species are not perfectly adapted to their environments, because mutations are random events and because the environment tends to change from generation to generation.

As mentioned at the beginning of the chapter, Fisher, Wright, and Haldane developed mathematical relationships to explain the theory of natural selection. As our knowledge of the process of natural selection has increased, it has become apparent that it operates in many different ways. In this chapter, we will consider a few examples of natural selection involving a single trait or a single gene that exists in two alleles. In reality, however, natural selection acts on populations of individuals in which many genes are polymorphic and each individual contains thousands or tens of thousands of different genes.

To begin our quantitative discussion of natural selection, we must examine the concept of **Darwinian fitness**—the relative likelihood that one genotype will contribute to the gene pool of the next generation rather than other genotypes. Natural selection acts on phenotypes that are derived from individuals' genotypes. Although Darwinian fitness often correlates with physical fitness, the two concepts are not identical. Darwinian fitness is a measure of reproductive success. An extremely fertile genotype may have a higher Darwinian fitness than a less fertile genotype that appears more physically fit.

To consider Darwinian fitness, let's use our example of a gene existing in the *A* and *a* alleles. If the three genotypes have the same level of mating success and fertility, we can assign fitness values to each genotype class based on their likelihood of surviving to reproductive age. For example, let's suppose that the relative survival to adulthood of each of the three genotype classes is as follows: For every five *AA* individuals that survive, four *Aa* individuals survive, and one *aa* individual survives. By convention, the genotype with the highest reproductive ability is given a fitness value of 1.0. Relative fitness values are denoted by the variable *W*. The fitness values of the other genotypes are assigned values relative to this 1.0 value:

$$\text{Fitness of } AA\text{: } W_{AA} = 1.0$$

$$\text{Fitness of } Aa\text{: } W_{Aa} = 4/5 = 0.8$$

$$\text{Fitness of } aa\text{: } W_{aa} = 1/5 = 0.2$$

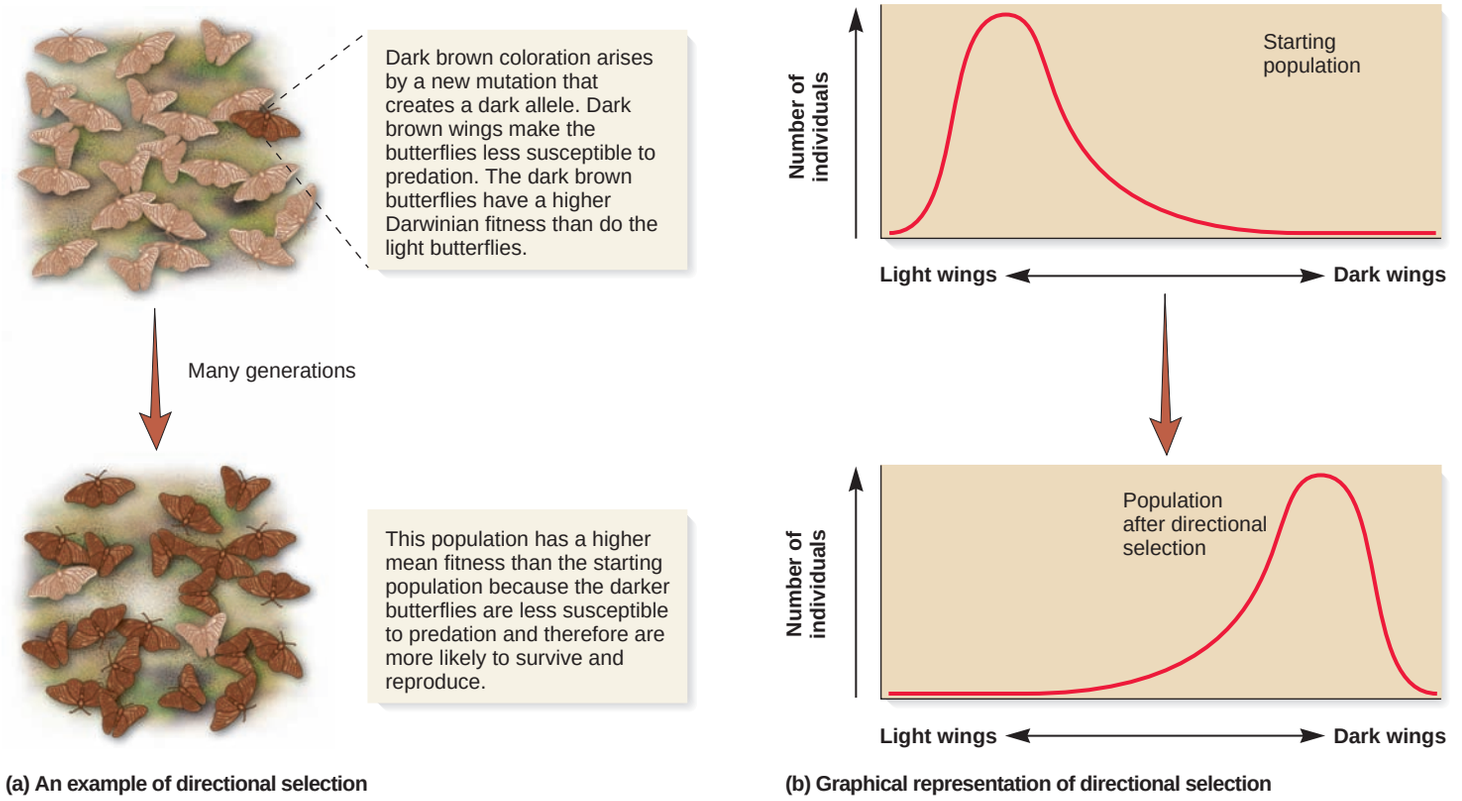
Keep in mind that differences in reproductive success among genotypes may stem from various reasons. In this case, the fittest genotype is more likely to survive to reproductive age. In other situations, the most fit genotype is more likely to mate. For example, a bird with brightly colored feathers may have an easier time attracting a mate than a bird with duller plumage. Finally, a third possibility is that the fittest genotype may be more fertile. It may produce a higher number of gametes or gametes that are more successful at fertilization.

By studying species in their native environments, population geneticists have discovered that natural selection can occur in several ways. The patterns of natural selection depend on the relative fitness values of the different genotypes and on the variation of environmental effects. The four patterns of natural selection that we will consider are called directional, stabilizing, disruptive, and balancing selection. In most of the examples described next, natural selection leads to adaptation so that a species is better able to survive to reproductive age.

Directional Selection Favors the Extreme Phenotype

Directional selection favors individuals at one extreme of a phenotypic distribution that are more likely to survive and reproduce in a particular environment. Different phenomena may initiate the process of directional selection. One way that directional selection may arise is that a new allele may be introduced into a population by mutation, and the new allele may promote a higher fitness in individuals that carry it (**Figure 24.8**). If the homozygote carrying the favored allele has the highest fitness value, directional selection may cause this favored allele to eventually become the predominant allele in the population, perhaps even becoming a monomorphic allele.

Another possibility is that a population may be exposed to a prolonged change in its living environment. Under the new environmental conditions, the relative fitness values may change to favor one genotype, which will promote the elimination of other genotypes. As an example, let's suppose a population of finches on the mainland already has genetic variation in beak size. A small number of birds migrate to an island where the seeds are generally larger than they are on the mainland. In this new environment, birds with larger beaks have a higher fitness because they are better able to crack open the larger seeds and thereby survive to reproductive age. Over the course of many



(a) An example of directional selection

(b) Graphical representation of directional selection

FIGURE 24.8 **Directional selection.** **(a)** A new mutation arises in a population that confers higher Darwinian fitness. In this example, butterflies with dark wings are more likely to survive and reproduce. Over many generations, directional selection favors the prevalence of darker individuals. **(b)** A graphical representation of directional selection.

generations, directional selection produces a population of birds carrying alleles that promote larger beak size.

In the case of directional selection, allele frequencies may change in a step-by-step, generation-per-generation way. To appreciate how this occurs, let's take a look at how fitness affects the Hardy-Weinberg equilibrium and allele frequencies. Again, let's suppose a gene exists in two alleles: *A* and *a*. The three fitness values, which are based on relative survival levels, are

$$W_{AA} = 1.0$$
$$W_{Aa} = 0.8$$
$$W_{aa} = 0.2$$

In the next generation, we expect that the Hardy-Weinberg equilibrium is modified in the following way due to directional selection:

Frequency of *AA*: p^2W_{AA}

Frequency of *Aa*: $2pqW_{Aa}$

Frequency of *aa*: q^2W_{aa}

In a population that is changing due to natural selection, these three terms may not add up to 1.0, as they do in the Hardy-Weinberg equilibrium. Instead, the three terms sum to a value known as the **mean fitness of the population ($\overline{W}$):**

$$p^2W_{AA} + 2pqW_{Aa} + q^2W_{aa} = \overline{W}$$

Dividing both sides of the equation by the mean fitness of the population,

$$\frac{p^2W_{AA}}{\overline{W}} + \frac{2pqW_{Aa}}{\overline{W}} + \frac{q^2W_{aa}}{\overline{W}} = 1$$

Using this equation, we can calculate the expected genotype and allele frequencies after one generation of directional selection:

Frequency of *AA* genotype: $\frac{p^2W_{AA}}{\overline{W}}$

Frequency of *Aa* genotype: $\frac{2pqW_{Aa}}{\overline{W}}$

Frequency of *aa* genotype: $\frac{q^2W_{aa}}{\overline{W}}$

Allele frequency of *A*: $p_A = \frac{p^2W_{AA}}{\overline{W}} + \frac{pqW_{Aa}}{\overline{W}}$

Allele frequency of *a*: $q_a = \frac{q^2W_{aa}}{\overline{W}} + \frac{pqW_{Aa}}{\overline{W}}$

As an example, let's suppose that the starting allele frequencies are $A = 0.5$ and $a = 0.5$, and use fitness values of 1.0, 0.8, and 0.2 for the three genotypes, *AA*, *Aa*, and *aa*, respectively. We begin by calculating the mean fitness of the population:

$$p^2W_{AA} + 2pqW_{Aa} + q^2W_{aa} = \overline{W}$$

$$\overline{W} = (0.5)^2(1) + 2(0.5)(0.5)(0.8) + (0.5)^2(0.2)$$

$$\overline{W} = 0.25 + 0.4 + 0.05 = 0.7$$

After one generation of directional selection,

Frequency of *AA* genotype: $\frac{p^2W_{AA}}{\overline{W}} = \frac{(0.5)^2(1)}{0.7} = 0.36$

Frequency of *Aa* genotype: $\frac{2pqW_{Aa}}{\overline{W}} = \frac{2(0.5)(0.5)(0.8)}{0.7} = 0.57$

Frequency of *aa* genotype: $\frac{q^2W_{aa}}{\overline{W}} = \frac{(0.5)^2(0.2)}{0.7} = 0.07$

Allele frequency of *A*: $p_A = \frac{p^2W_{AA}}{\overline{W}} + \frac{pqW_{Aa}}{\overline{W}}$

$$= \frac{(0.5)^2(1)}{0.7} + \frac{(0.5)(0.5)(0.8)}{0.7} = 0.64$$

Allele frequency of *a*: $q_a = \frac{q^2W_{aa}}{\overline{W}} + \frac{pqW_{Aa}}{\overline{W}}$

$$= \frac{(0.5)^2(0.2)}{0.7} + \frac{(0.5)(0.5)(0.8)}{0.7} = 0.36$$

After one generation, the allele frequency of *A* has increased from 0.5 to 0.64, and the frequency of *a* has decreased from 0.5 to 0.36. This has occurred because the *AA* genotype has the highest fitness, whereas the *Aa* and *aa* genotypes have lower fitness values. Another interesting feature of natural selection is that it raises the mean fitness of the population. If we assume the individual fitness values are constant, the mean fitness of this next generation is

$$\begin{aligned}\overline{W} &= p^2W_{AA} + 2pqW_{Aa} + q^2W_{aa} \\ &= (0.64)^2(1) + 2(0.64)(0.36)(0.8) + (0.36)^2(0.2) \\ &= 0.80\end{aligned}$$

The mean fitness of the population has increased from 0.7 to 0.8.

What are the consequences of natural selection at the population level? This population is better adapted to its environment than the previous one. Another way of viewing this calculation is that the subsequent population has a greater reproductive potential than the previous one. We could perform the same types of calculations to find the allele frequencies and mean fitness value in the next generation. If we assume the individual fitness values remain constant, the frequencies of *A* and *a* in the next generation are 0.85 and 0.15, respectively, and the mean fitness increases to 0.931. As we can see, the general trend is to increase *A*, decrease *a*, and increase the mean fitness of the population.

In the previous example, we considered the effects of natural selection by beginning with allele frequencies at intermediate levels (namely, $A = 0.5$ and $a = 0.5$). **Figure 24.9** illustrates what would happen if a new mutation introduced the *A* allele into a population that was originally monomorphic for the *a* allele. As

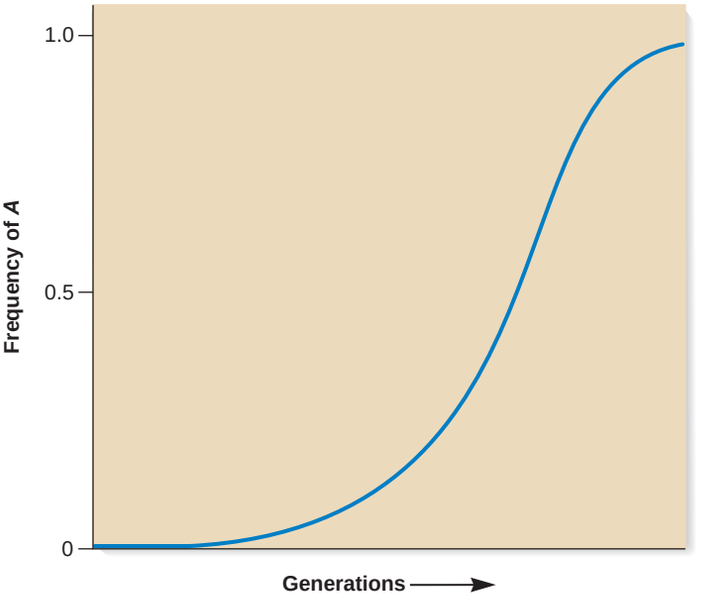


FIGURE 24.9 The fate of a beneficial allele that is introduced as a new mutation into a population. A new allele (*A*) is beneficial in the homozygous condition: $W_{AA} = 1.0$. The heterozygote, *Aa* ($W_{Aa} = 0.8$), and homozygote, *aa* ($W_{aa} = 0.2$), have lower fitness values.

before, the *AA* homozygote has a fitness of 1.0, the *Aa* heterozygote 0.8, and the recessive *aa* homozygote 0.2. Initially, the *A* allele is at a very low frequency in the population. If it is not lost initially due to genetic drift, its frequency slowly begins to rise and then, at intermediate values, rises much more rapidly.

Eventually, this type of natural selection may lead to fixation of a beneficial allele. However, a new beneficial allele is in a precarious situation when its frequency is very low. As we have seen, random genetic drift is likely to eliminate new mutations, even beneficial ones, due to chance fluctuations.

Researchers have identified many examples of directional selection in nature. As mentioned in Chapter 7, resistance to antibiotics is a growing concern in the treatment of infection. The selection of bacterial strains resistant to one or more antibiotics typically occurs in a directional manner. Similarly, the resistance of insects to pesticides, such as DDT (dichlorodiphenyltrichloroethane), occurs in a directional manner. DDT usage began in the 1940s as a way to decrease the populations of mosquitoes and other insects. However, certain insect species can become resistant to DDT by a dominant mutation in a single enzyme-encoding gene. The mutant enzyme detoxifies DDT, making it harmless to the insect. **Figure 24.10** shows the results of an experiment in which mosquito larvae (*Aedes aegypti*) were exposed to DDT over the course of seven generations. The starting population showed a low level of DDT resistance, as evidenced by the low percentage of survivors after exposure to DDT. By comparison, in seven generations, nearly 100% of the population was DDT-resistant. These results illustrate the power of directional selection in promoting change in a population. Since the 1950s, resistance to nearly every known insecticide has evolved within 10 years of its commercial introduction!

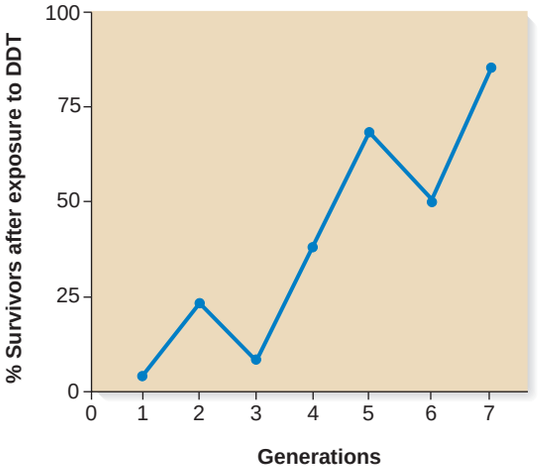


FIGURE 24.10 **Directional selection for DDT resistance in a mosquito population.** In this experiment, mosquito larvae (*Aedes aegypti*) were exposed to 10 mg/L of DDT. The percentage of survivors was recorded, and then the survivors of each generation were used as parents for the next generation.

Stabilizing Selection Favors Individuals with Intermediate Phenotypes

In **stabilizing selection,** the extreme phenotypes for a trait are selected against, and those individuals with the intermediate phenotypes have the highest fitness values. Stabilizing selection tends to decrease genetic diversity for a particular gene because it eliminates alleles that cause a greater variation in phenotypes. An example of stabilizing selection involves clutch size in birds, which was first proposed by British biologist David Lack in 1947. Under stabilizing selection, birds that lay too many or too few eggs have lower fitness values than those that lay an intermediate value (**Figure 24.11**). Laying too many eggs may cause many offspring to die due to inadequate parental care and food. In addition, the strain on the parents themselves may decrease their likelihood of survival and therefore their ability to produce more offspring. Having too few offspring, on the other hand, does not contribute many individuals to the next generation. Therefore, the most successful parents are those that produce an intermediate clutch size. In the 1980s, Swedish evolutionary biologist Lars Gustafsson and colleagues examined the phenomenon of stabilizing selection in the collared flycatcher, *Ficedula albicollis*, on the island of Gotland, which is southeast of the mainland of Sweden. They discovered that Lack's hypothesis that clutch size is subject to the action of stabilizing selection also appears to be true for this species.

Disruptive Selection Favors Multiple Phenotypes

Disruptive selection, also known as diversifying selection, favors the survival of two or more different genotypes that produce different phenotypes (**Figure 24.12**). In disruptive selection, the fitness values of a particular genotype are higher in one environment and lower in a different one. Disruptive selection is likely to occur in populations that occupy diverse environments so that some members of the species survive in each type of environmental condition.

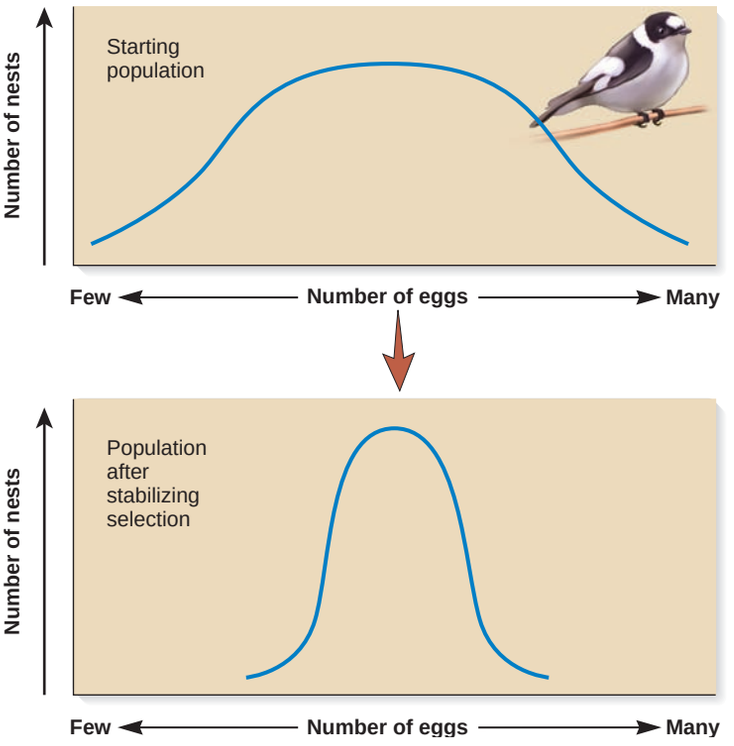


FIGURE 24.11 **Stabilizing selection.** In this pattern of natural selection, the extremes of a phenotypic distribution are selected against. Those individuals with intermediate traits have the highest fitness. This results in a population with less diversity and more uniform traits.

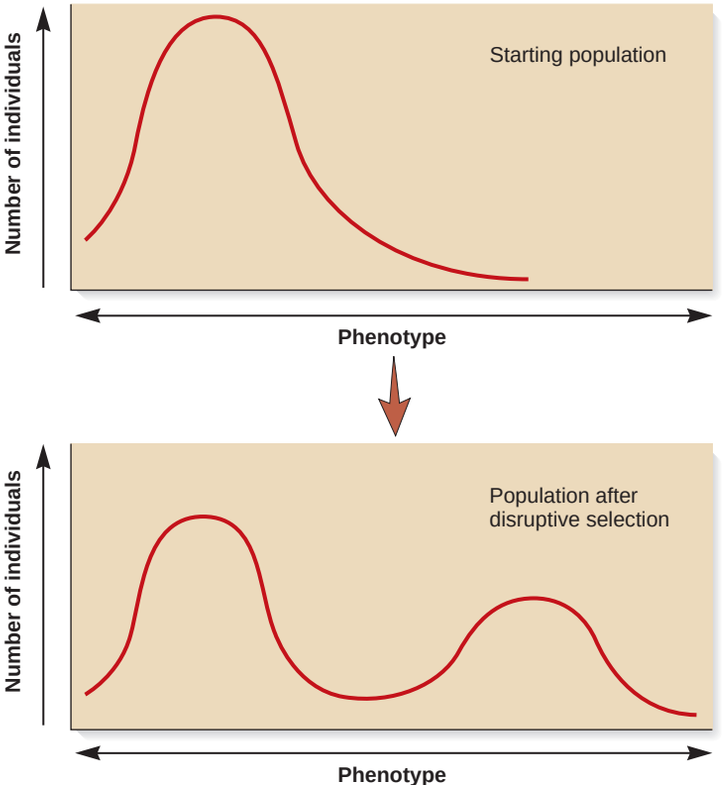


FIGURE 24.12 **Disruptive selection.** Over time, this form of selection favors two or more phenotypes due to heterogeneous environments.

(a) Land snails

Habitat	*Brown*	*Pink*	*Yellow*
Beechwoods	0.23	0.61	0.16
Deciduous woods	0.05	0.68	0.27
Hedgerows	0.05	0.31	0.64
Rough herbage	0.004	0.22	0.78

(b) Frequency of snail color

FIGURE 24.13 Polymorphism in the land snail, *Cepaea nemoralis*. **(a)** This species of snail can exist in several different colors and banding patterns. **(b)** Coloration of the snails is correlated with the specific environments where they are located.

Genes → Traits Snail coloration is an example of genetic polymorphism due to heterogeneous environments; the genes governing shell coloration are polymorphic. The predation of snails is correlated with their ability to be camouflaged in their natural environment. Snails with brown shells are most prevalent in beechwoods, where the soil is dark. Pink snails are most abundant in the leaf litter of beechwoods and deciduous woods. Yellow snails are most prevalent in sunnier locations, such as hedgerows and rough herbage.

As an example, **Figure 24.13a** shows a photograph of land snails, *Cepaea nemoralis*, that live in woods and open fields. This snail is polymorphic in color and banding patterns. In 1954, Arthur Cain and Philip Sheppard found that snail color was correlated with the environment. As shown in **Figure 24.13b**, the highest frequency of brown shell color was found in snails in the beechwoods, where there are wide expanses of dark soil. Their frequency was substantially less in other environments. By comparison, pink snails are most common in the leaf litter of forest floors, and the yellow snails are most abundant in the sunny, grassy areas of hedgerows and rough herbage. Researchers have suggested that this disruptive selection can be explained by different levels of predation by thrushes. Depending on the environment, certain snail phenotypes may be more easily seen by their predators than others. Migration can occasionally occur between the snail populations, which keeps the polymorphism in balance among these different environments.

Balanced Polymorphisms May Occur Due to Heterozygote Advantage or Negative Frequency-Dependent Selection

As we have just seen, polymorphisms may occur when a species occupies a diverse environment. Researchers have discovered other patterns of natural selection that favor the maintenance of two or more alleles in a more homogeneous environment. This pattern, called **balancing selection,** results in a genetic polymorphism in a population.

For genetic variation involving a single gene, balancing selection may arise when the heterozygote has a higher fitness than either corresponding homozygote, a situation called **heterozygote advantage.** In this case, an equilibrium is reached in which both alleles are maintained in the population. If the fitness values are known for each of the genotypes, the allele frequencies at equilibrium can be calculated. To do so, we must consider the **selection coefficient (*s*),** which measures the degree to which a genotype is selected against.

$$s = 1 - W$$

By convention, the genotype with the highest fitness has an s value of zero. Genotypes at a selective disadvantage have s values that are greater than 0 but less than or equal to 1.0. An extreme case is a recessive lethal allele. It would have an s value of 1.0 in the homozygote, while the s value in the heterozygote could be 0.

Let's consider the following case of relative fitness, where

$$W_{AA} = 0.7$$
$$W_{Aa} = 1.0$$
$$W_{aa} = 0.4$$

The selection coefficients are

$$s_{AA} = 1 - 0.7 = 0.3$$
$$s_{Aa} = 1 - 1.0 = 0$$
$$s_{aa} = 1 - 0.4 = 0.6$$

The population reaches an equilibrium when

$$s_{AA}p = s_{aa}q$$

If we take this equation, let $q = 1 - p$, and then solve for p:

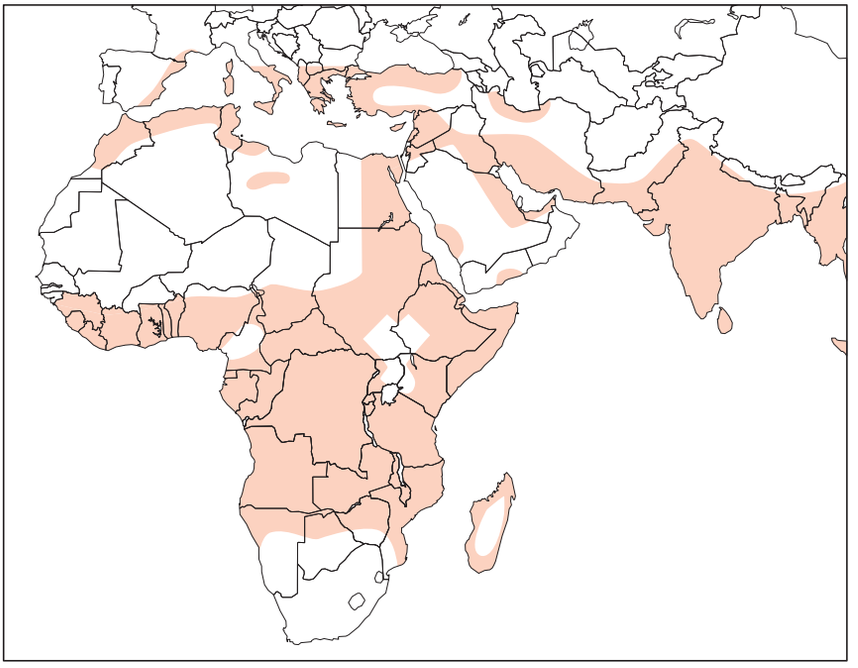
$$p = \text{Allele frequency of } A = \frac{s_{aa}}{s_{AA} + s_{aa}}$$
$$= \frac{0.6}{0.3 + 0.6} = 0.67$$

If we let $p = 1 - q$ and then solve for q:

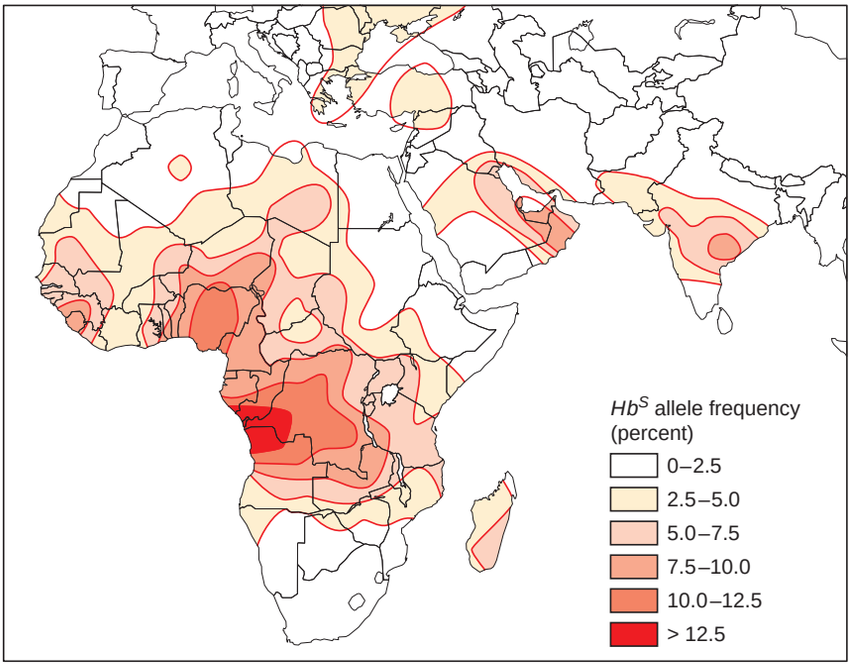
$$q = \text{Allele frequency of } a = \frac{s_{AA}}{s_{AA} + s_{aa}}$$
$$= \frac{0.3}{0.3 + 0.6} = 0.33$$

In this example, balancing selection maintains the two alleles in the population at frequencies in which A equals 0.67 and a equals 0.33.

Heterozygote advantage can sometimes explain the high frequency of alleles that are deleterious in a homozygous condition. A classic example is the Hb^S allele of the human β-globin gene. A homozygous Hb^SHb^S individual displays sickle cell disease, a disorder that leads to the sickling of the red blood cells. The Hb^SHb^S homozygote has a lower fitness than a homozygote with two copies of the more common β-globin allele, Hb^AHb^A. However, the heterozygote, Hb^AHb^S, has a higher level of fitness than either homozygote in areas where malaria is endemic (**Figure 24.14**).



(a) Malaria prevalence



(b) Hb^S allele frequency

FIGURE 24.14 The geographic relationship between malaria and the frequency of the sickle cell allele in human populations.
(a) The geographic prevalence of malaria in Africa and surrounding areas. **(b)** The frequency of the Hb^S allele in the same areas.

Genes → Traits The sickle cell allele of the β-globin gene is maintained in human populations as a balanced polymorphism. In areas where malaria is prevalent, the heterozygote carrying one copy of the Hb^S allele has a greater fitness than either of the corresponding homozygotes (Hb^AHb^A and Hb^SHb^S). Therefore, even though the Hb^SHb^S homozygotes suffer the detrimental consequences of sickle cell disease, this negative aspect is balanced by the beneficial effects of malarial resistance in the heterozygotes.

Compared with $Hb^A Hb^A$ homozygotes, heterozygotes have a 10 to 15% better chance of survival if infected by the malarial parasite, *Plasmodium falciparum*. Therefore, the Hb^S allele is maintained in populations where malaria is prevalent, even though the allele is detrimental in the homozygous state.

In addition to sickle cell disease, other gene mutations that cause human disease in the homozygous state are thought to be prevalent because of heterozygote advantage. For example, the high prevalence of the allele causing cystic fibrosis may be related to this phenomenon, but the advantage that a heterozygote may possess is not understood.

Negative frequency-dependent selection is a second mechanism of balancing selection. In this pattern of natural selection, the fitness of a genotype decreases when its frequency becomes higher. In other words, rare individuals have a higher fitness than more common individuals. Therefore, rare individuals are more likely to reproduce, whereas common individuals are less likely, thereby producing a balanced polymorphism in which no genotype becomes too rare or too common.

An interesting example of negative frequency-dependent selection involves the elder-flowered orchid, *Dactylorhiza sambucina* (**Figure 24.15**). Throughout its range, both yellow- and red-flowered individuals are prevalent. The explanation for this polymorphism is related to its pollinators, which are mainly bumblebees such as *Bombus lapidarius* and *B. terrestris*. The pollinators increase their visits to the flower color of *D. sambucina* as it becomes less common in a given area. One reason why this may occur is because *D. sambucina* is a rewardless flower; that is, it does not provide its pollinators with any reward for visiting, such as sweet nectar. Pollinators learn that the more common color of *D. sambucina* in a given area does not offer a reward, and they increase their visits to the less-common flower. Thus, the relative fitness of the less-common flower increases.

FIGURE 24.15 The two color variations found in the elder-flowered orchid, *Dactylorhiza sambucina*. The two colors are maintained in the population due to negative frequency-dependent selection.

EXPERIMENT 24A

The Grants Have Observed Natural Selection in Galápagos Finches

Let's now turn to a study that demonstrates natural selection in action. Since 1973, Peter Grant, Rosemary Grant, and their colleagues have studied the process of natural selection in finches found on the Galápagos Islands. For over 30 years, the Grants have focused much of their research on one of the Galápagos Islands known as Daphne Major. This small island (0.34 km^2) has a moderate degree of isolation (8 km from the nearest island), an undisturbed habitat, and a resident population of finches, including the medium ground finch, *Geospiza fortis* (**Figure 24.16**).

To study natural selection, the Grants observed various traits in the medium ground finch, including beak size, over the course of many years. The medium ground finch has a relatively small crushing beak, suitable for breaking open small, tender seeds. The Grants quantified beak size among the medium ground finches of Daphne Major by carefully measuring beak depth (a measurement of the beak from top to bottom, at its base) on individual birds. During the course of their studies, they compared the beak sizes of parents and offspring by examining many broods over several years. The depth of the beak was transmitted from parents to offspring, regardless of environmental conditions, indicating that differences in beak sizes are due to genetic differences in the population. In other words, they found that beak depth is a heritable trait.

FIGURE 24.16 The medium ground finch (*Geospiza fortis*), which is found on Daphne Major.

By measuring many birds every year, the Grants were able to assemble a detailed portrait of natural selection from generation to generation. In the study shown in **Figure 24.17**, they measured beak depth in 1976 and 1978.

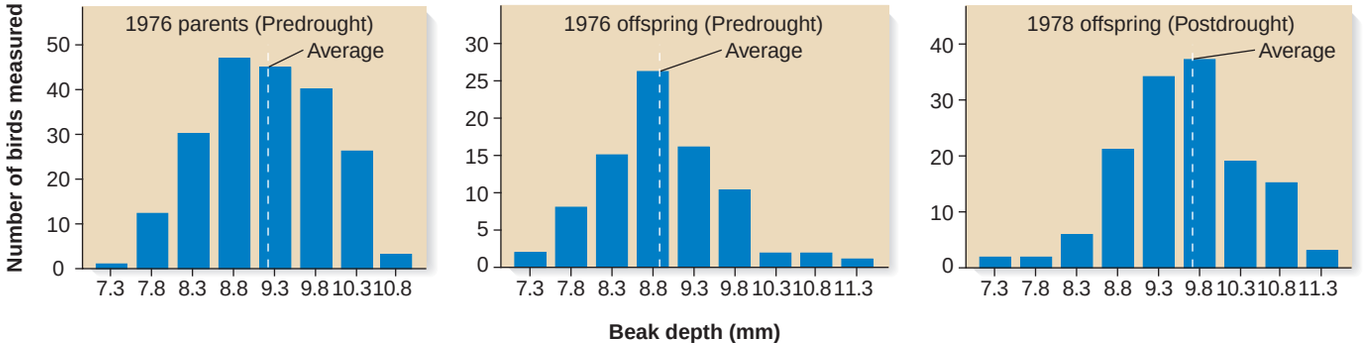
THE HYPOTHESIS

Beak size will be influenced by natural selection. Environments that produce larger seeds will select for birds with large beaks.

TESTING THE HYPOTHESIS — FIGURE 24.17 Natural selection in medium ground finches of Daphne Major.

	Experimental level	Conceptual level
1. In 1976, measure beak depth in parents and offspring of the species *G. fortis*.	Capture birds and measure beak depth. 	This is a way to measure a trait that may be subject to natural selection.
2. Repeat the procedure on offspring that were born in 1978 and had reached mature size. A drought had occurred in 1977 that caused plants on the island to produce mostly larger seeds and relatively few small seeds.	Capture birds and measure beak depth. 	This is a way to measure a trait that may be subject to natural selection.

THE DATA



Data from: B. Rosemary Grant and Peter R. Grant (2003) What Darwin's finches can teach us about the evolutionary origin and regulation of biodiversity. *Bioscience 53,* 965–975.

INTERPRETING THE DATA

In the wet year of 1976, the plants of Daphne Major produced the small seeds that these finches were able to eat in abundance. However, a drought occurred in 1977. During this year, the plants on Daphne Major tended to produce few of the smaller seeds, which the finches rapidly consumed. To survive, the finches resorted to eating larger, drier seeds, which were harder to crush. As a result, the birds that survived tended to have larger beaks, because they were better able to break open these large seeds. In the year after the drought, the average beak depth of offspring in the population increased to approximately 9.8 mm because the surviving birds with larger beaks passed this trait on to their offspring. This is likely to be due to directional selection (see Figure 24.8), although genetic drift could also contribute to these data. Overall, these results illustrate the power of natural selection to alter the nature of a trait, in this case, beak depth, in a given population.

A self-help quiz involving this experiment can be found at www.mhhe.com/brookergenetics4e.

Nonrandom Mating May Occur in Populations

As mentioned earlier, one of the conditions required to establish the Hardy-Weinberg equilibrium is random mating. This means that individuals choose their mates irrespective of their genotypes and phenotypes. In many cases, particularly in human populations, this condition is violated frequently.

When mating is nonrandom in a population, the process is called **assortative mating.** Positive assortative mating occurs when individuals with similar phenotypes choose each other as mates. The opposite situation, where dissimilar phenotypes mate preferentially, is called negative assortative mating. In addition, individuals may choose a mate that is part of the same genetic lineage. The mating of two genetically related individuals, such as cousins, is called **inbreeding.** This is also termed consanguinity. Inbreeding sometimes occurs in human societies and is more likely to take place in nature when population size becomes

very limited. In Chapter 25, we will examine how inbreeding is a useful strategy for developing agricultural breeds or strains with desirable characteristics. Conversely, **outbreeding,** which involves mating between unrelated individuals, can create hybrids that are heterozygous for many genes.

In the absence of other evolutionary processes, inbreeding and outbreeding do not affect allele frequencies in a population. However, these patterns of mating do disrupt the balance of genotypes that is predicted by the Hardy-Weinberg equation. Let's first consider inbreeding in a family pedigree. **Figure 24.18** illustrates a human pedigree involving a mating between cousins. Individuals III-2 and III-3 are cousins and have produced the daughter labeled IV-1. She is said to be inbred, because her parents are genetically related to each other.

During inbreeding, the gene pool is smaller, because the parents are related genetically. In the 1940s, Gustave Malécot developed methods to quantify the degree of inbreeding. The **inbreeding coefficient** is the probability that two alleles in a particular individual will be identical for a given gene because both copies are due to descent from a common ancestor. An inbreeding coefficient (F) can be computed by analyzing the degree of relatedness within a pedigree.

As an example, let's determine the inbreeding coefficient for individual IV-1. To begin this problem, we must first identify all of this individual's common ancestors. A common ancestor is anyone who is an ancestor to both of an individual's parents. In Figure 24.18, IV-1 has one common ancestor, I-2, her great-grandfather. I-2 is the grandfather of III-2 and III-3.

Our next step is to determine the inbreeding paths. An inbreeding path for an individual is the shortest path through the pedigree that includes both parents and the common ancestor. In a pedigree, there is an inbreeding path for each common ancestor. The length of each inbreeding path is calculated by adding together all of the individuals in the path except the individual of interest. In this case, there is only one path because IV-1 has only one common ancestor. To add the members of the path, we begin with individual IV-1, but we do not count her. We then move to her father (III-2); to her grandfather (II-2); to I-2, her great-grandfather (the common ancestor); back down to her other grandmother (II-3); and finally to her mother (III-3). This path has five members. Finally, to calculate the inbreeding coefficient, we use the following formula:

$$F = \Sigma(1/2)^n(1 + F_A)$$

where

- F is the inbreeding coefficient of the individual of interest
- n is the number of individuals in the inbreeding path, excluding the inbred offspring
- F_A is the inbreeding coefficient of the common ancestor
- Σ indicates that we add together $(1/2)^n(1 + F_A)$ for each inbreeding path

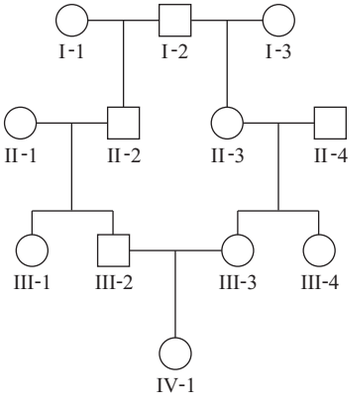


FIGURE 24.18 A human pedigree containing inbreeding. Individual IV-1 is the result of inbreeding because her parents are related.

In this case, there is only one common ancestor and, therefore, only one inbreeding path. Also, we do not know anything about the heritage of the common ancestor, so we assume that F_A is zero. Thus, in our example of Figure 24.18,

$$F = \Sigma(1/2)^n(1 + 0)$$
$$= (1/2)^5 = 1/32 = 3.125\%$$

What does this value mean? Our inbreeding coefficient, 3.125%, tells us the probability that a gene in the inbred individual (IV-1) is homozygous due to its inheritance from a common ancestor (I-2). In this case, therefore, each gene in individual IV-1 has a 3.125% chance of being homozygous because she has inherited the same allele twice from her great-grandfather (I-2), once through each parent.

As an example, let's suppose that the common ancestor (I-2) is heterozygous for the gene involved with cystic fibrosis. His genotype would be *Cc*, where *c* is the recessive allele that causes cystic fibrosis. There is a 3.125% probability that the inbred individual (IV-1) is homozygous (*CC* or *cc*) for this gene because she has inherited both copies from her great-grandfather. She has a 1.56% probability of inheriting both normal alleles (*CC*) and a 1.56% probability of inheriting both mutant alleles (*cc*). The inbreeding coefficient is denoted by the letter F (for Fixation) because it is the probability that an allele will be fixed in the homozygous condition. The term fixation signifies that the homozygous individual can pass only one type of allele to their offspring.

In other pedigrees, an individual may have two or more common ancestors. In this case, the inbreeding coefficient F is calculated as the sum of the inbreeding paths. Such an example is described in solved problem S2 at the end of the chapter.

In addition to pedigree analysis, the effects of inbreeding and outbreeding can also be considered within a population. For example, let's consider the situation in which the frequency of $A = p$ and the frequency of $a = q$. In a given population, the genotype frequencies are determined in the following way:

- $p^2 + fpq$ equals the frequency of *AA*
- $2pq(1 - f)$ equals the frequency of *Aa*
- $q^2 + fpq$ equals the frequency of *aa*

where f is a measure of how much the genotype frequencies deviate from Hardy-Weinberg equilibrium due to nonrandom mating. The value of f ranges from -1 to $+1$. When inbreeding occurs, the value is greater than zero. When outbreeding occurs, the value is less than zero.

As an example, let's suppose that $p = 0.8$, $q = 0.2$, and $f = 0.25$. We can calculate the frequencies of the *AA*, *Aa*, and *aa* genotypes under these conditions as follows:

$$AA = p^2 + fpq = (0.8)^2 + (0.25)(0.8)(0.2) = 0.68$$

$$Aa = 2pq(1 - f) = 2(0.8)(0.2)(1 - 0.25) = 0.24$$

$$aa = q^2 + fpq = (0.2)^2 + (0.25)(0.8)(0.2) = 0.08$$

There will be 68% *AA* homozygotes, 24% heterozygotes, and 8% *aa* homozygotes. If mating had been random (i.e., $f = 0$), the genotype frequencies of *AA* would be p^2, which equals 64%, and *aa* would be q^2, which equals 4%. The frequency of heterozygotes would be $2pq$, which equals 32%. When comparing these numbers, we see that inbreeding raises the proportions of homozygotes and decreases the proportion of heterozygotes. In natural populations, the value of f tends to become larger as a population becomes smaller, because each individual has a more limited choice in mate selection.

What are the consequences of inbreeding in a population? From an agricultural viewpoint, it results in a higher proportion of homozygotes, which may exhibit a desirable trait. For example, an animal breeder may use inbreeding to produce animals that are larger because they have become homozygous for alleles promoting larger size. On the negative side, many genetic diseases are inherited in a recessive manner (see Chapter 22). For these disorders, inbreeding increases the likelihood that an individual will be homozygous and therefore afflicted with the disease. Also, in natural populations, inbreeding lowers the mean fitness of the population if homozygous offspring have a lower fitness value. This can be a serious problem as natural populations become smaller due to human habitat destruction. As the population shrinks, inbreeding becomes more likely because individuals have fewer potential mates from which to choose. The inbreeding, in turn, produces homozygotes that are less fit, thereby decreasing the reproductive success of the population. This phenomenon is called **inbreeding depression.** Conservation biologists sometimes try to circumvent this problem by introducing individuals from one population into another. For example, the endangered Florida panther (*Felis concolor coryi*) suffers from inbreeding-related defects, which include poor sperm quality and quantity and morphological abnormalities. To help alleviate these effects, panthers of the same species from Texas have been introduced into the Florida population.

24.3 SOURCES OF NEW GENETIC VARIATION

In the previous section, we primarily focused on genetic variation in which a single gene exists in two or more alleles. This simplified scenario allows us to appreciate the general principles behind evolutionary mechanisms. As researchers have analyzed genetic variation at the molecular, cellular, and population level, however, they have come to understand that new genetic variation occurs in many ways (**Table 24.2**). Among eukaryotic species, sexual reproduction is an important way that new genetic variation occurs among offspring. In Chapters 3 and 6, we considered how independent assortment and crossing over during sexual reproduction may produce new combinations of alleles among different genes, thereby producing new genetic variation in the resulting offspring. Similarly, in Chapter 26, we will consider

TABLE 24.2
Sources of New Genetic Variation That Occur in Populations

Type	Description
Independent assortment	The independent segregation of different homologous chromosomes may give rise to new combinations of alleles in offspring (see Chapter 3).
Crossing over	Recombination (crossing over) between homologous chromosomes can also produce new combinations of alleles that are located on the same chromosome (see Chapter 6).
Interspecies crosses	On occasion, members of different species may breed with each other to produce hybrid offspring. This topic is discussed in Chapter 26.
Prokaryotic gene transfer	Prokaryotic species possess mechanisms of genetic transfer such as conjugation, transduction, and transformation (see Chapter 7).
New alleles	Point mutations can occur within a gene to create single-nucleotide polymorphisms (SNPs). In addition, genes can be altered by small deletions and additions. Gene mutations are also discussed in Chapter 16.
Gene duplications	Events, such as misaligned crossovers, can add additional copies of a gene into a genome and lead to the formation of gene families. This topic is discussed in Chapter 8.
Chromosome structure and number	Chromosome structure may be changed by deletions, duplications, inversions, and translocations. Changes in chromosome number result in aneuploid, polyploid, and alloploid offspring. These mechanisms are discussed in Chapters 8 and 26.
Exon shuffling	New genes can be created when exons of preexisting genes are rearranged to make a gene that encodes a protein with a new combination of protein domains.
Horizontal gene transfer	Genes from one species can be introduced into another species and become incorporated into that species' genome.
Changes in repetitive sequences	Short repetitive sequences are common in genomes due to the occurrence of transposable elements and due to tandem arrays. The number and lengths of repetitive sequences tend to show considerable variation in natural populations.

how breeding between members of different species may produce hybrid offspring that harbor new combinations of genetic material. Such hybridization events have been important in the evolution of new species, particularly those in the plant kingdom. Though prokaryotic species reproduce asexually, they also possess mechanisms for gene transfer, such as conjugation, transduction, and transformation (see Chapter 7). These mechanisms are important for fostering genetic variation among bacterial and archaeal populations.

Rare mutations in DNA may also give rise to new types of variation (see Table 24.2). As discussed earlier in this chapter (see Figure 24.3) and in Chapter 16, mutations may occur within a particular gene to create new alleles of that gene. Such allelic variation is common in natural populations. Also, as described in Chapter 8, gene duplications may create a gene family; each family member acquires independent mutations and often evolves more specialized functions. An example is the globin gene family (see Figure 8.7). Changes in chromosome structure and number are also important in the evolution of new species (see Chapter 26). In this section, we examine some additional mechanisms through which an organism can acquire new genetic variation. These include exon shuffling, horizontal gene transfer, and changes in repetitive sequences. The myriad of mechanisms that foster genetic variation underscores its profound importance in the evolution of species that are both well adapted to their native environments and successful at reproduction.

New Genes Are Produced in Eukaryotes via Exon Shuffling

Sources of new genetic variation are revealed when the parts of genes that encode protein domains are compared within a single species. Many proteins, particularly those found in eukaryotic species, have a modular structure composed of two or more domains with different functions. For example, certain transcription factors have discrete domains involved with hormone binding, dimerization, and DNA binding. As described in Chapter 15, the glucocorticoid receptor has a domain that binds the hormone, a second domain that facilitates protein dimerization, and a third that allows the glucocorticoid receptor to bind to glucocorticoid response elements (GREs) next to genes (see Figure 15.7). By comparing the modular structure of eukaryotic proteins with the genes that encode them, geneticists have discovered that each domain tends to be encoded by one coding sequence, or exon, or by a series of two or more adjacent exons.

During the evolution of eukaryotic species, many new genes have been created by a process known as **exon shuffling,** in which an exon and its flanking introns are inserted into a gene, thereby producing a new gene that encodes a protein with an additional domain (**Figure 24.19**). This process may also involve the duplication and rearrangement of exons. Exon shuffling results in novel genes that express proteins with diverse functional modules. Such proteins can then alter traits in the organism and may be acted on by evolutionary processes such as genetic drift and natural selection.

Exon shuffling may occur by more than one mechanism. As described in Chapter 17, transposable elements may promote the insertion of exons into the coding sequences of other genes. Alternatively, a double crossover event could promote the insertion of an exon into another gene (this is the case in Figure 24.19). This is called nonhomologous recombination because the two regions involved in the crossover are not homologous to each other.

New Genes Are Acquired via Horizontal Gene Transfer

Species also accumulate genetic changes by a process called **horizontal gene transfer,** in which an organism incorporates genetic material from another organism without being the offspring of that organism. It often involves the exchange of genetic material between different species. **Figure 24.20** illustrates one possible mechanism for horizontal gene transfer. In this example, a eukaryotic cell has engulfed a bacterium by endocytosis. During the degradation of the bacterium, a bacterial gene escapes to the nucleus of the cell, where it is inserted into one of the chromosomes. In this way, a gene has been transferred from a bacterial species to a eukaryotic species. By analyzing gene sequences among many different species, researchers have discovered that horizontal gene transfer is a common phenomenon. This process can occur from prokaryotes to eukaryotes, from eukaryotes to prokaryotes, between different species of prokaryotes, and between different species of eukaryotes.

Gene transfer among bacterial species is relatively widespread. As discussed in Chapter 7, bacterial species may carry out three natural mechanisms of gene transfer known as conjugation, transduction, and transformation. By analyzing the genomes of bacterial species, scientists have determined that many genes within a given bacterial genome are derived from horizontal gene transfer. Genome studies have suggested that as much as 20 to 30% of the variation in the genetic composition of modern prokaryotic species can be attributed to this process. For example, roughly 17% of the genes of *E. coli* and *Salmonella typhimurium* have been acquired via horizontal gene transfer during the past 100 million years. The roles of these acquired genes are quite varied, though they commonly involve functions that are readily acted on by natural selection. These include genes that confer antibiotic resistance, the ability to degrade toxic compounds, and pathogenicity (the ability to cause disease).

Genetic Variation Is Produced via Changes in Repetitive Sequences

Another source of genetic variation comes from changes in **repetitive sequences**—short sequences typically a few base pairs to a few thousand base pairs long that are repeated many times within a species' genome. Repetitive sequences usually come from two types of sources. First, transposable elements are genetic sequences that can move from place to place in a species' genome (see Chapter 17). The prevalence and movement of transposable elements provides a great deal of genetic variation between species and within a single species. In certain eukaryotic species, transposable elements have become fairly abundant (see Table 17.3).

A second type of repetitive sequence is nonmobile and involves short sequences that are tandemly repeated. In a **microsatellite** (also called short tandem repeats, STRs), the repeat unit is usually 1 to 6 bp long, and the whole tandem repeat is less than a couple hundred base pairs in length. For example, the most common microsatellite encountered in humans is a sequence

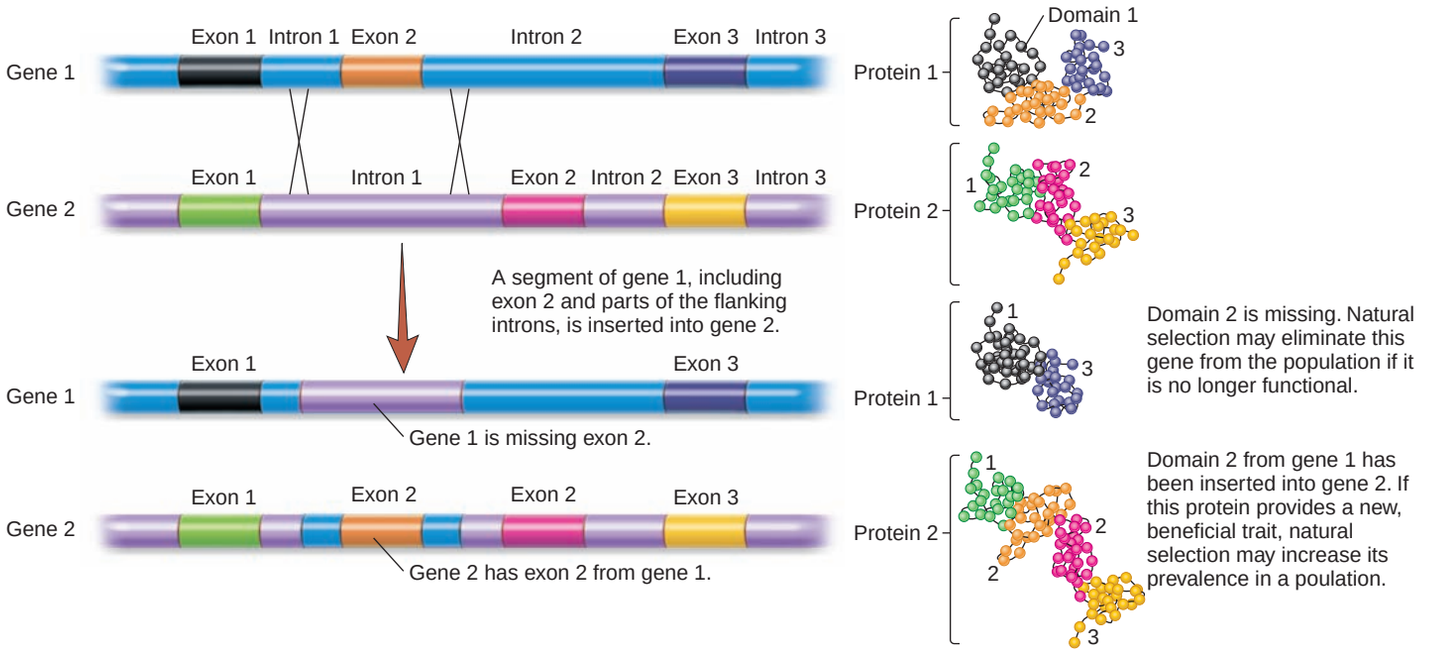
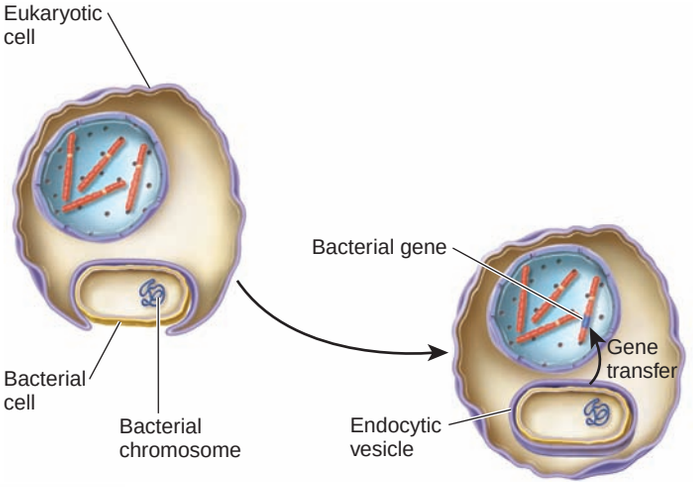


FIGURE 24.19 **The process of exon shuffling.** In this example, a segment of one gene containing an exon and its flanking introns has been inserted into another gene. A rare, abnormal crossing-over event called nonhomologous recombination may cause this to happen. This results in proteins that have new combinations of domains and possibly new functions.



ONLINE ANIMATION

FIGURE 24.20 **Horizontal gene transfer from a bacterium to a eukaryote.** In this example, a bacterium is engulfed by a eukaryotic cell, and a bacterial gene is transferred to one of the eukaryotic chromosomes.

$(CA)_N$, where N may range from 5 to more than 50. In other words, this dinucleotide sequence can be tandemly repeated 5 to 50 or more times. The $(CA)_N$ microsatellite is found, on average, about every 10,000 bases in the human genome. In a **minisatellite,** the repeat unit is typically 6 to 80 bp in length, and the size of the minisatellite ranges from 1 kbp to 20 kbp. An example of a minisatellite in humans is telomeric DNA. In a human sperm cell, for example, the repeat unit is 6 bp and the size of a telomere is about 15 kbp. (Note: Tandem repeat sequences are called satellites because they sediment away from the rest of the chromosomal DNA during equilibrium gradient centrifugation.)

Tandem repetitive sequences such as microsatellites and minisatellites tend to undergo mutation in which the number of tandem repeats changes. For example, a microsatellite with a 4-bp repeat unit and a length of 64 bp may undergo a mutation that adds three more repeat units and become 76 bp long. Micro- and minisatellites may change in length by different mechanisms. One common mechanism to explain this phenomenon is DNA strand slippage that occurs during DNA replication. As discussed in Chapter 16, this mechanism can explain trinucleotide repeat expansion (see Figure 16.12).

Because repetitive sequences tend to vary within a population, they have become a common tool that geneticists use in a variety of ways. For example, as described in Chapters 20 and 22, microsatellites can be used as molecular markers to map the locations of genes (see Figure 22.7). Likewise, population geneticists analyze microsatellites or minisatellites to study variation at the population level and to determine the relationships among individuals and neighboring populations. The sizes of microsatellites and minisatellites found in closely related individuals tend to be more similar than they are in unrelated individuals. As described next, this phenomenon is the basis for DNA fingerprinting.

DNA Fingerprinting Is Used for Identification and Relationship Testing

The technique of **DNA fingerprinting,** also known as **DNA profiling,** analyzes individuals based on the occurrence of repetitive sequences in their genome. When subjected to traditional DNA fingerprinting, the chromosomal DNA gives rise to a series of bands on a gel (**Figure 24.21**). The sizes and order of bands is an individual's DNA fingerprint. Like the human fingerprint, the DNA of each individual has a distinctive pattern. It is the unique pattern of these bands that makes it possible to distinguish individuals.

A comparison of the DNA fingerprints among different individuals has found two applications. First, DNA fingerprinting can be used as a method of identification. In forensics, DNA fingerprinting can identify a crime suspect. In medicine, the technique can identify the type of bacterium that is causing an infection in a particular patient. A second use of DNA fingerprinting is relationship testing. Closely related individuals have more similar fingerprints than do distantly related ones (see solved problem S6). In humans, this can be used in paternity testing. In population genetics, DNA fingerprinting can provide evidence regarding the degree of relatedness among members of a population. Such information may help geneticists determine if a population is likely to be suffering from inbreeding depression.

The development of DNA fingerprinting has relied on the identification of DNA sites that vary greatly in length among members of a population. This naturally occurring variation causes each individual to have a unique DNA fingerprint. In the 1980s, Alec Jeffreys and his colleagues found that certain minisatellites within human chromosomes are particularly variable in their lengths. As discussed earlier, minisatellites tend to vary within populations due to changes in the number of tandem repeats at each site.

In the past decade, the technique of DNA fingerprinting has become automated, much like the automation that changed the procedure of DNA sequencing described in Chapter 18. DNA fingerprinting is now done using the technique of polymerase chain reaction (PCR), which amplifies microsatellites. Like minisatellites, microsatellites are found in multiple sites in the genome of humans and other species and vary in length among different individuals. In this procedure, the microsatellites from a sample of DNA are amplified by PCR using primers that flank the repetitive region and then separated by gel electrophoresis according to their molecular masses. As in automated DNA sequencing, the amplified microsatellite fragments are fluorescently labeled. A laser excites the fluorescent molecule within a microsatellite, and a detector records the amount of fluorescence emission for each microsatellite. As shown in **Figure 24.22**, this type of DNA fingerprint yields a series of peaks, each peak having a characteristic molecular mass. In this automated approach, the pattern of peaks rather than bands constitutes an individual's DNA fingerprint.

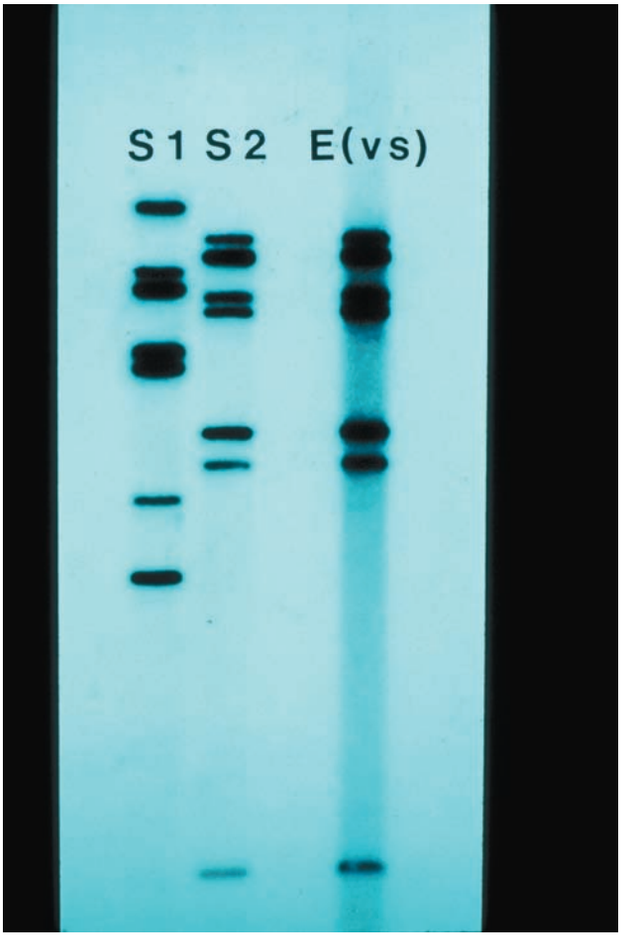


FIGURE 24.21 **A comparison of two DNA fingerprints.** The chromosomal DNA from two different individuals (Suspect 1—S1, and Suspect 2—S2) was subjected to DNA fingerprinting. The DNA evidence at a crime scene, E(vs), was also subjected to DNA fingerprinting. Following the hybridization of a radiolabeled probe, the DNA appears as a series of bands on a gel. The dissimilarity in the pattern of these bands distinguishes different individuals, much as the differences in physical fingerprint patterns can be used for identification. As seen here, S2 matches the DNA found at the crime scene.

ONLINE ANIMATION

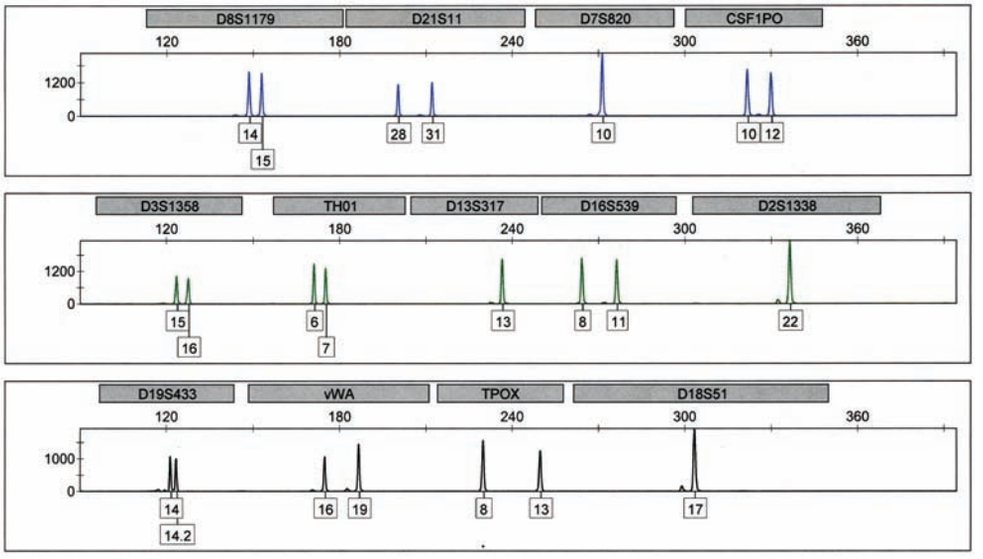


FIGURE 24.22 **Automated DNA fingerprinting.** In automated DNA fingerprinting, a sample of DNA is amplified, using primers that recognize the ends of microsatellites. The microsatellite fragments are fluorescently labeled and then separated by gel electrophoresis. The fluorescent molecules within each microsatellite are excited with a laser, and the amount of fluorescence is measured via a fluorescence detector. A printout from the detector is shown here. The gray boxes indicate the names of specific microsatellites. The peaks show the relative amounts of each microsatellite. The boxes beneath each peak indicate the number of tandem repeats in a given microsatellite. In this example, the individual is heterozygous for certain microsatellites (e.g., D8S1179) and homozygous for others (e.g., D7S820).

KEY TERMS

Page 670. population genetics, gene pool
Page 671. population, local populations, demes, polymorphism, polymorphic, monomorphic, single-nucleotide polymorphism (SNP)
Page 672. allele frequencies, genotype frequencies
Page 673. Hardy-Weinberg equilibrium, Hardy-Weinberg equation
Page 675. disequilibrium, microevolution
Page 676. mutation rate
Page 677. random genetic drift, genetic drift
Page 678. bottleneck effect, founder effect
Page 679. conglomerate, gene flow, natural selection
Page 680. Darwinian fitness, directional selection
Page 681. mean fitness of the population ($\bar{W}$)
Page 683. stabilizing selection, disruptive selection
Page 684. balancing selection, heterozygote advantage, selection coefficient (s)
Page 686. negative frequency-dependent selection
Page 687. assortative mating, inbreeding
Page 688. outbreeding, inbreeding coefficient
Page 689. inbreeding depression
Page 690. exon shuffling, horizontal gene transfer, repetitive sequences, microsatellite
Page 691. minisatellite, DNA fingerprinting, DNA profiling

CHAPTER SUMMARY

- Population genetics is concerned with changes in genetic variation within a population of individuals over time.

24.1 Genes in Populations and the Hardy-Weinberg Equation

- All of the alleles of every gene in a population constitute the gene pool.
- For sexually reproducing organisms, a population is a group of individuals of the same species that occupy the same region and can interbreed with one another (see Figure 24.1).
- Polymorphism refers to traits or genes that exhibit variation in a population (see Figure 24.2).
- Single-nucleotide polymorphisms are the most common type of variation in genes (see Figure 24.3).
- Geneticists analyze genetic variation by determining allele and genotype frequencies.
- The Hardy-Weinberg equation can be used to calculate genotype frequencies based on allele frequencies (see Figures 24.4, 24.5).
- Deviation from a Hardy-Weinberg equilibrium indicates that evolutionary change is occurring.

24.2 Factors That Change Allele and Genotype Frequencies in Populations

- Mutations are the source of new genetic variation. However, the occurrence of new mutations does not greatly change allele frequencies because they are relatively rare (see Table 24.1).
- Random genetic drift, migration, natural selection, and nonrandom mating are mechanisms that may alter allele or genotype frequencies in a population over time.
- Genetic drift is a change in allele frequency due to chance fluctuations. Over the long run, it often results in allele fixation or loss. The effect of genetic drift is greater in small populations (see Figure 24.6).
- Two examples of genetic drift are the bottleneck effect and the founder effect (see Figure 24.7).
- Migration can alter allele frequencies and tends to make the allele frequencies in neighboring populations more similar.
- Natural selection is the process that changes allele frequencies from one generation to the next based on fitness, which is the relative reproductive successes of different genotypes.
- Directional selection favors the extreme phenotype (see Figures 24.8–24.10).
- Stabilizing selection favors individuals with intermediate phenotypes (see Figure 24.11).
- Disruptive selection favors multiple phenotypes. This may occur due to heterogeneous environments (see Figures 24.12, 24.13).
- Balancing selection results in a stable polymorphism. Examples include heterozygote advantage and negative-frequency dependent selection (see Figures 24.14, 24.15).
- The Grants observed natural selection in a finch population. The selection involved a change in beak size due to drought conditions (see Figures 24.16, 24.17).
- Nonrandom mating may alter the genotype frequencies that would be based on a Hardy-Weinberg equilibrium. Inbreeding promotes homozygosity (see Figure 24.18).

24.3 Sources of New Genetic Variation

- A variety of different mechanisms can bring about genetic variation (see Table 24.2).
- New genes in eukaryotes are produced by exon shuffling (see Figure 24.19).
- A species may acquire a new gene from another species via horizontal gene transfer (see Figure 24.20).
- A common source of genetic variation in populations involves changes in repetitive sequences, such as microsatellites.
- DNA fingerprinting is a technique that relies on variation in repetitive sequences within a population. It is used as a means of identification and in relationship testing (see Figures 24.21, 24.22).

PROBLEM SETS & INSIGHTS

Solved Problems

S1. The phenotypic frequency of people who cannot taste phenylthiocarbamide (PTC) is approximately 0.3. The inability to taste this bitter substance is due to a recessive allele. If we assume there are only two alleles in the population (namely, tasters, *T*, and nontasters, *t*) and that the population is in Hardy-Weinberg equilibrium, calculate the frequencies of these two alleles.

Answer: Let p = allele frequency of the taster allele and q = the allele frequency of the nontaster allele. The frequency of nontasters is 0.3.

This is the frequency of the genotype *tt*, which in this case is equal to q^2:

$$q^2 = 0.3$$

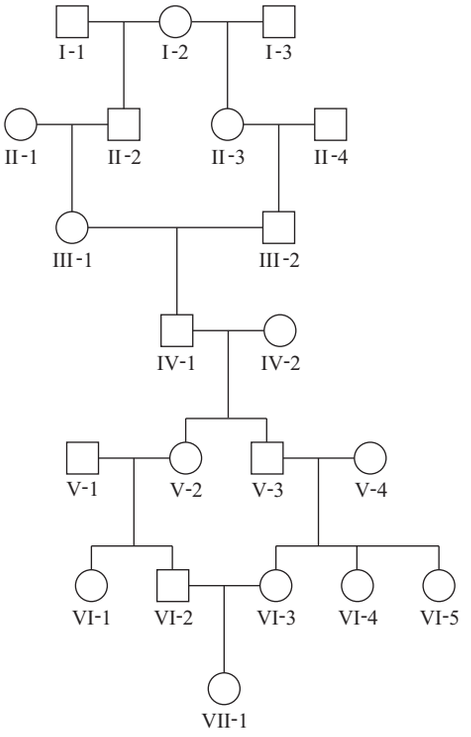
To determine the frequency q of the nontaster allele, we take the square root of both sides of this equation:

$$q = 0.55$$

With this value, we can calculate the frequency p of the taster allele:

$$p = 1 - q$$
$$= 1 - 0.55 = 0.45$$

S2. In the pedigree shown here, answer the following questions with regard to individual VII-1:



A. Who are the common ancestors of her parents?

B. What is the inbreeding coefficient?

Answer:

A. The common ancestors are IV-1 and IV-2. They are the grandparents of VI-2 and VI-3, who are the parents of VII-1.

B. The inbreeding coefficient is calculated using the formula

$$F = \Sigma(1/2)^n(1 + F_A)$$

In this case, there are two common ancestors, IV-1 and IV-2. Also, IV-1 is inbred, because I-2 is a common ancestor to both of IV-1's parents. The first step is to calculate F_A, the inbreeding coefficient for this common ancestor. The inbreeding path for IV-1 contains five people: III-1, II-2, I-2, II-3, and III-2. Therefore,

$$n = 5$$

$$F_A = (1/2)^5 = 0.03$$

Now we can calculate the inbreeding coefficient for VII-1. Each inbreeding path contains five people: VI-2, V-2, IV-1, V-3, and VI-3; and VI-2, V-2, IV-2, V-3, and VI-3. Thus,

$$F = (1/2)^5\,(1 + 0.03) + (1/2)^5\,(1 + 0)$$

$$= 0.032 + 0.031 = 0.063$$

S3. The Hardy-Weinberg equation provides a way to predict genotype frequency based on allele frequency. In the case of mammals, males are hemizygous for X-linked genes, whereas females have two copies. Among males, the frequency of any X-linked trait equals the frequency of males with the trait. For example, if an allele frequency for an X-linked disease-causing allele was 5%, then 5% of all males would be affected with the disorder. Female genotype frequencies are computed using the Hardy-Weinberg equation.

As a specific example, let's consider the human X-linked trait known as hemophilia A (see Chapter 22 for a description of this disorder). In human populations, the allele frequency of the hemophilia A allele is approximately 1 in 10,000, or 0.0001. The other allele for this gene is the normal allele. Males can be affected or unaffected, whereas females can be affected, unaffected carriers, or unaffected noncarriers.

A. What are the allele frequencies for the mutant and normal allele in the human population?

B. Among males, what is the frequency of affected individuals?

C. Among females, what is the frequency of affected individuals and heterozygous carriers?

D. Within a population of 100,000 people, what is the expected number of affected males? In this same population, what is the expected number of carrier females?

Answer: Let p represent the normal allele and q represent the allele that causes hemophilia.

A. X^H normal allele, frequency = 0.9999 = p

X^h hemophilia allele, frequency = 0.0001 = q

B. X^hY genotype frequency of affected males $= q = 0.0001$

C. X^hX^h genotype frequency of affected females $= q^2 = (0.0001)^2 = 0.00000001$

X^HX^h genotype frequency of carrier females $= 2pq = 2(0.9999)(0.0001) = 0.0002$

D. We assume this population is composed of 50% males and 50% females.

Number of affected males $= 50{,}000 \times 0.0001 = 5$

Number of carrier females $= 50{,}000 \times 0.0002 = 10$

S4. The Hardy-Weinberg equation can be modified to include situations of three or more alleles. In its standard (two-allele) form, the Hardy-Weinberg equation reflects the Mendelian notion that each individual inherits two copies of each allele, one from each parent. For a two-allele situation, it can also be written as

$(p + q)^2 = 1$ (Note: The number 2 in this equation reflects the idea that the genotype is due to the inheritance of two alleles, one from each parent.)

This equation can be expanded to include three or more alleles. For example, let's consider a situation in which a gene exists as three alleles: *A1*, *A2*, and *A3*. The allele frequency of *A1* is designated by the letter *p*, *A2* by the letter *q*, and *A3* by the letter *r*. Under these circumstances, the Hardy-Weinberg equation becomes

$$(p + q + r)^2 = 1$$

$$p^2 + q^2 + r^2 + 2pq + 2pr + 2qr = 1$$

where

p^2 is the genotype frequency of *A1A1*

q^2 is the genotype frequency of *A2A2*

r^2 is the genotype frequency of *A3A3*

$2pq$ is the genotype frequency of *A1A2*

$2pr$ is the genotype frequency of *A1A3*

$2qr$ is the genotype frequency of *A2A3*

Now here is the question. As discussed in Chapter 4, the gene that affects human blood type can exist in three alleles. In a Japanese population, the allele frequencies are

$I^A = 0.28$

$I^B = 0.17$

$i = 0.55$

Based on these allele frequencies, calculate the different possible genotype frequencies and blood type frequencies.

Answer: If we let p represent I^A, q represent I^B, and r represent i, then

p^2 is the genotype frequency of I^AI^A,
which is type A blood $= (0.28)^2 = 0.08$

q^2 is the genotype frequency of I^BI^B,
which is type B blood $= (0.17)^2 = 0.03$

r^2 is the genotype frequency of ii,
which is type O blood $= (0.55)^2 = 0.30$

$2pq$ is the genotype frequency of I^AI^B,
which is type AB blood $= 2(0.28)(0.17) = 0.09$

$2pr$ is the genotype frequency of I^Ai,
which is type A blood $= 2(0.28)(0.55) = 0.31$

$2qr$ is the genotype frequency of I^Bi,
which is type B blood $= 2(0.17)(0.55) = 0.19$

Type A $= 0.08 + 0.31 = 0.39$, or 39%

Type B $= 0.03 + 0.19 = 0.22$, or 22%

Type O $= 0.30$, or 30%

Type AB $= 0.09$, or 9%

S5. Let's suppose that pigmentation in a species of insect is controlled by a single gene existing in two alleles, *D* for dark and *d* for light. The heterozygote *Dd* is intermediate in color. In a heterogeneous environment, the allele frequencies are $D = 0.7$ and $d = 0.3$. This polymorphism is maintained because the environment contains some dimly lit forested areas and some sunny fields. During a hurricane, a group of 1000 insects is blown to a completely sunny area. In this environment, the fitness values are $DD = 0.3$, $Dd = 0.7$, and $dd = 1.0$. Calculate the allele frequencies in the next generation.

Answer: The first step is to calculate the mean fitness of the population:

$$p^2W_{DD} + 2pqW_{Dd} + q^2W_{dd} = \overline{W}$$

$$\overline{W} = (0.7)^2(0.3) + 2(0.7)(0.3)(0.7) + (0.3)^2(1.0)$$

$$= 0.15 + 0.29 + 0.09 = 0.53$$

After one generation of selection, we get

$$\text{Allele frequency of } D\text{: } p_D = \frac{p^2W_{DD}}{\overline{W}} + \frac{pqW_{Dd}}{\overline{W}}$$

$$= \frac{(0.7)^2(0.3)}{0.53} + \frac{(0.7)(0.3)(0.7)}{0.53}$$

$$= 0.55$$

$$\text{Allele frequency of } d\text{: } q_d = \frac{q^2W_{dd}}{\overline{W}} + \frac{pqW_{Dd}}{\overline{W}}$$

$$= \frac{(0.3)^2(1.0)}{0.53} + \frac{(0.7)(0.3)(0.7)}{0.53}$$

$$= 0.45$$

After one generation, the allele frequency of *D* has decreased from 0.7 to 0.55, while the frequency of *d* has increased from 0.3 to 0.45.

S6. An important application of DNA fingerprinting is relationship testing. Persons who are related genetically have some bands or peaks in common. The number they share depends on the closeness of their genetic relationship. For example, an offspring is expected to receive half of his or her minisatellites from one parent and the rest from the other. The diagram on the next page schematically shows a traditional DNA fingerprint of an offspring, mother, and two potential fathers.

In paternity testing, the offspring's DNA fingerprint is first compared with that of the mother. The bands that the offspring have in common with the mother are depicted in purple. The bands that are not similar between the offspring and the mother must have been inherited from the father. These bands are depicted in red. Which male could be the father?

Answer: Male 2 does not have many of the paternal bands. Therefore, he can be excluded as being the father of this child. However, male 1 has all of the paternal bands. He is very likely to be the father.

Geneticists can calculate the likelihood that the matching bands between the offspring and a prospective father could occur just as a matter of random chance. To do so, they must analyze the frequency of each band in a reference population (e.g., Caucasians living in the United States). For example, let's suppose that DNA fingerprinting analyzed 40 bands. Of these, 20 bands matched with the mother and 20 bands matched with a prospective father. If the probability of each of these bands in a reference population was 1/4, the likelihood of such a match occurring by random chance would be $(1/4)^{20}$, or roughly 1 in 1 trillion. Therefore, a match between two samples is rarely a matter of random chance.

Conceptual Questions

C1. What is the gene pool? How is a gene pool described in a quantitative way?

C2. In genetics, what does the term population mean? Pick any species you like and describe how its population might change over the course of many generations.

C3. What is a genetic polymorphism? What is the source of genetic variation?

C4. State for each of the following whether it is an example of an allele, genotype, and/or phenotype frequency:

A. Approximately 1 in 2500 Caucasians is born with cystic fibrosis.

B. The percentage of carriers of the sickle cell allele in West Africa is approximately 13%.

C. The number of new mutations for achondroplasia, a genetic disorder, is approximately 5×10^{-5}.

C5. The term polymorphism can refer to both genes and traits. Explain the meaning of a polymorphic gene and a polymorphic trait. If a gene is polymorphic, does the trait that the gene affects also have to be polymorphic? Explain why or why not.

C6. Cystic fibrosis (CF) is a recessive autosomal trait. In certain Caucasian populations, the number of people born with this disorder is about 1 in 2500. Assuming Hardy-Weinberg equilibrium for this trait,

A. What are the frequencies for the normal and CF alleles?

B. What are the genotype frequencies of homozygous normal, heterozygous, and homozygous affected individuals?

C. Assuming random mating, what is the probability that two phenotypically unaffected heterozygous carriers will choose each other as mates?

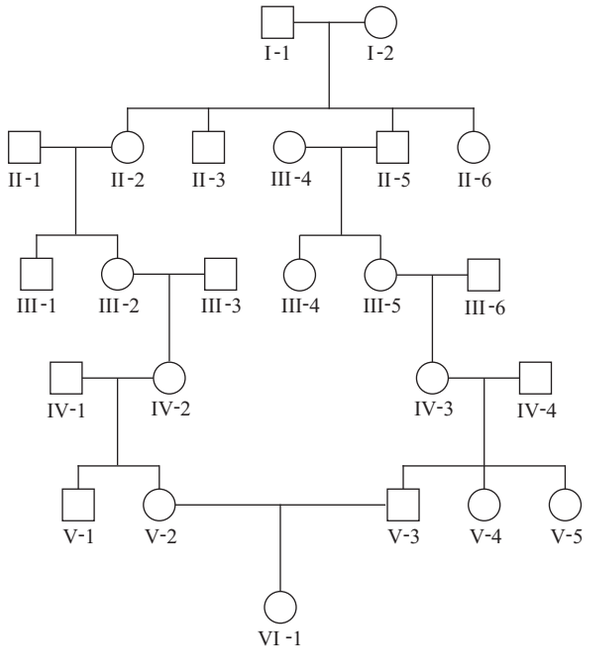
C7. Does inbreeding affect allele frequencies? Why or why not? How does it affect genotype frequencies? With regard to rare recessive diseases, what are the consequences of inbreeding in human populations?

C8. For a gene existing in two alleles, what are the allele frequencies when the heterozygote frequency is at its maximum value? What if there are three alleles?

C9. In a population, the frequencies of two alleles are $B = 0.67$ and $b = 0.33$. The genotype frequencies are $BB = 0.50$, $Bb = 0.37$, and $bb = 0.13$. Do these numbers suggest inbreeding? Explain why or why not.

C10. The ability to roll your tongue is inherited as a recessive trait. The frequency of the rolling allele is approximately 0.6, and the dominant (nonrolling) allele is 0.4. What is the frequency of individuals who can roll their tongues?

C11. Using the pedigree shown here, answer the following questions for individual VI-1:



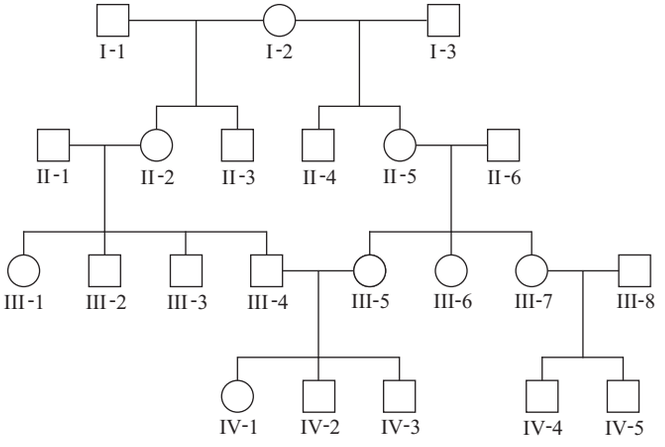
A. Is this individual inbred?

B. If so, who are her common ancestor(s)?

C. Calculate the inbreeding coefficient for VI-1.

D. Are the parents of VI-1 inbred?

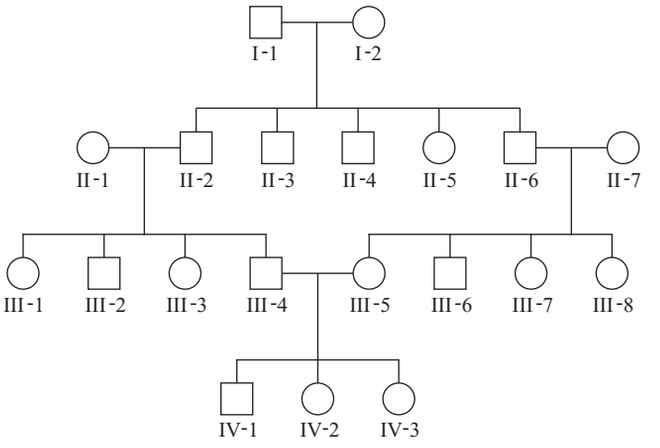
C12. A family pedigree is shown here.



A. What is the inbreeding coefficient for individual IV-3?

B. Based on the data shown in this pedigree, is individual IV-4 inbred?

C13. A family pedigree is shown here.



A. What is the inbreeding coefficient for individual IV-2? Who is/are her common ancestors?

B. Based on the data shown in this pedigree, is individual III-4 inbred?

C14. What evolutionary factors can cause allele frequencies to change and possibly lead to a genetic polymorphism? Discuss the relative importance of each type of process.

C15. In the term genetic drift, what is drifting? Why is this an appropriate term to describe this phenomenon?

C16. What is the difference between a random and an adaptive evolutionary process? Describe two or more examples of each. At the molecular level, explain how mutations can be random or adaptive.

C17. Let's suppose the mutation rate for converting a *B* allele into a *b* allele is 10^{-4}. The current allele frequencies are $B = 0.6$ and $b = 0.4$. How long will it take for the allele frequencies to equal each other, assuming that no genetic drift is taking place?

C18. Why is genetic drift more significant in small populations? Why does it take longer for genetic drift to cause allele fixation in large populations than in small ones?

C19. A group of four birds flies to a new location and initiates the formation of a new colony. Three of the birds are homozygous *AA*, and one bird is heterozygous *Aa*.

A. What is the probability that the *a* allele will become fixed in the population?

B. If fixation of the *a* allele occurs, how long will it take?

C. How will the growth of the population, from generation to generation, affect the answers to parts A and B? Explain.

C20. Describe what happens to allele frequencies as a result of the bottleneck effect. Discuss the relevance of this effect with regard to species that are approaching extinction.

C21. When two populations frequently intermix due to migration, what are the long-term consequences with regard to allele frequencies and genetic variation?

C22. Discuss the similarities and differences among directional, disruptive, balancing, and stabilizing selection.

C23. What is Darwinian fitness? What types of characteristics can promote high fitness values? Give several examples.

C24. What is the intuitive meaning of the mean fitness of a population? How does its value change in response to natural selection?

C25. Antibiotics are commonly used to combat bacterial and fungal infections. During the past several decades, however, antibiotic-resistant strains of microorganisms have become alarmingly prevalent. This has undermined the effectiveness of antibiotics in treating many types of infectious disease. Discuss how the following processes that alter allele frequencies may have contributed to the emergence of antibiotic-resistant strains:

A. Random mutation

B. Genetic drift

C. Natural selection

C26. With regard to genetic drift, are the following statements true or false? If a statement is false, explain why.

A. Over the long run, genetic drift leads to allele fixation or loss.

B. When a new mutation occurs within a population, genetic drift is more likely to cause the loss of the new allele rather than the fixation of the new allele.

C. Genetic drift promotes genetic diversity in large populations.

D. Genetic drift is more significant in small populations.

C27. Two populations of antelope are separated by a mountain range. The antelope are known to occasionally migrate from one population to the other. Migration can occur in either direction. Explain how migration affects the following phenomena:

A. Genetic diversity in the two populations

B. Allele frequencies in the two populations

C. Genetic drift in the two populations

C28. Do the following examples describe directional, disruptive, balancing, or stabilizing selection?

A. Polymorphisms in snail color and banding pattern as described in Figure 24.13

B. Thick fur among mammals exposed to cold climates

C. Birth weight in humans

D. Sturdy stems and leaves among plants exposed to windy climates

Experimental Questions

E1. You will need to be familiar with the techniques described in Chapter 18 to answer this question. Gene polymorphisms can be detected using a variety of cellular and molecular techniques. Which techniques would you use to detect gene polymorphisms at the following levels?

A. DNA level

B. RNA level

C. Polypeptide level

E2. You will need to understand solved problem S4 to answer this question. The gene for coat color in rabbits can exist in four alleles termed C (full coat color), c^{ch} (chinchilla), c^{h} (Himalayan), and c (albino). In a population of rabbits in Hardy-Weinberg equilibrium, the allele frequencies are

$C = 0.34$

$c^{ch} = 0.17$

$c^{h} = 0.44$

$c = 0.05$

Assume that C is dominant to the other three alleles. c^{ch} is dominant to c^{h} and c, and c^{h} is dominant to c.

A. What is the frequency of albino rabbits?

B. Among 1000 rabbits, how many would you expect to have a Himalayan coat color?

C. Among 1000 rabbits, how many would be heterozygotes with a chinchilla coat color?

E3. In a large herd of 5468 sheep, 76 animals have yellow fat, compared with the rest of the members of the herd, which have white fat. Yellow fat is inherited as a recessive trait. This herd is assumed to be in Hardy-Weinberg equilibrium.

A. What are the frequencies of the white and yellow fat alleles in this population?

B. Approximately how many sheep with white fat are heterozygous carriers of the yellow allele?

E4. The human MN blood group is determined by two codominant alleles, *M* and *N*. The following data were obtained from various human populations:

		Percentages		
Population	**Place**	***MM***	***MN***	***NN***
Inuit	East Greenland	83.5	15.6	0.9
Navajo Indians	New Mexico	84.5	14.4	1.1
Finns	Karajala	45.7	43.1	11.2
Russians	Moscow	39.9	44.0	16.1
Aborigines	Queensland	2.4	30.4	67.2

(Data from E.B. Speiss (1990). *Genes in Populations*, 2d ed. Wiley-Liss, New York.)

A. Calculate the allele frequencies in these five populations.

B. Which populations appear to be in Hardy-Weinberg equilibrium?

C. Which populations do you think have had significant intermixing due to migration?

E5. You will need to understand solved problem S4 before answering this question. In an island population, the following data were obtained for the numbers of people with each of the four blood types:

Type O	721
Type A	932
Type B	235
Type AB	112

Is this population in Hardy-Weinberg equilibrium? Explain your answer.

E6. In a donor population, the allele frequencies for the common (Hb^{A}) and sickle cell alleles (Hb^{S}) are 0.9 and 0.1, respectively. A

group of 550 individuals migrates to a new population containing 10,000 individuals; in the recipient population, the allele frequencies are $Hb^A = 0.99$ and $Hb^S = 0.01$.

A. Calculate the allele frequencies in the conglomerate population.

B. Assuming the donor and recipient populations are each in Hardy-Weinberg equilibrium, calculate the genotype frequencies in the conglomerate population prior to further mating between the donor and recipient populations.

C. What will be the genotype frequencies of the conglomerate population in the next generation, assuming it achieves Hardy-Weinberg equilibrium in one generation?

E7. A recessive lethal allele has achieved a frequency of 0.22 due to genetic drift in a very small population. Based on natural selection, how would you expect the allele frequencies to change in the next three generations? Note: Your calculation can assume that genetic drift is not altering allele frequencies in either direction.

E8. Among a large population of 2 million gray mosquitoes, one mosquito is heterozygous for a body color gene; this mosquito has one gray allele and one blue allele. There is no selective advantage or disadvantage between gray and blue body color. All of the other mosquitoes carry the gray allele.

A. What is the probability of fixation of the blue allele?

B. If fixation happens to occur, how many generations is it likely to take?

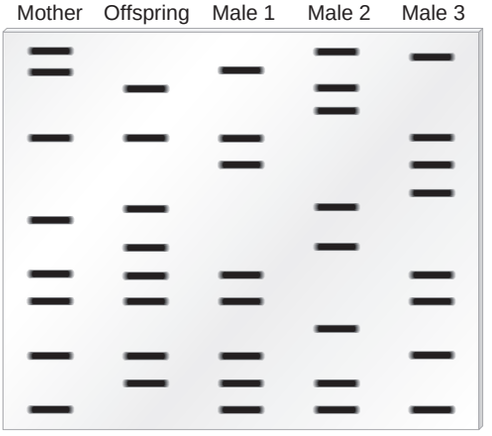
C. Qualitatively, how would the answers to parts A and B be affected if the blue allele conferred a slight survival advantage?

E9. Resistance to the poison warfarin is a genetically determined trait in rats. Homozygotes carrying the resistance allele (*WW*) have a lower fitness because they suffer from vitamin K deficiency, but heterozygotes (*Ww*) do not. However, the heterozygotes are still resistant to warfarin. In an area where warfarin is applied, the heterozygote has a survival advantage. Due to warfarin resistance, the heterozygote is also more fit than the normal homozygote (*ww*). If the relative fitness values for *Ww*, *WW*, and *ww* individuals are 1.0, 0.37, and 0.19, respectively, in areas where warfarin is applied, calculate the allele frequencies at equilibrium. How would this equilibrium be affected if the rats were no longer exposed to warfarin?

E10. Describe, in as much experimental detail as possible, how you would test the hypothesis that snail color distribution is due to predation.

E11. In the Grants' study of the medium ground finch, do you think the pattern of natural selection was directional, stabilizing, disruptive, or balancing? Explain your answer. If the environment remained dry indefinitely (for many years), what do you think would be the long-term outcome?

E12. Here are traditional DNA fingerprints of five people: a child, mother, and three potential fathers:



Which males can be ruled out as being the father? Explain your answer. If one of the males could be the father, explain the general strategy for calculating the likelihood that he could match the offspring's DNA fingerprint by chance alone. (See solved problem S6 before answering this question.)

E13. What is DNA fingerprinting? How can it be used in human identification?

E14. When analyzing the automated DNA fingerprints of a father and his biological daughter, a technician examined 50 peaks and found that 30 of them were a perfect match. In other words, 30 out of 50 peaks, or 60%, were a perfect match. Is this percentage too high, or would you expect a value of only 50%? Explain why or why not.

E15. What would you expect to be the minimum percentage of matching peaks in an automated DNA fingerprint for the following pairs of individuals?

A. Mother and son

B. Sister and brother

C. Uncle and niece

D. Grandfather and grandson

Questions for Student Discussion/Collaboration

1. Discuss examples of positive and negative assortive mating in natural populations, human populations, and agriculturally important species.
2. Discuss the role of mutation in the origin of genetic polymorphisms. Suppose that a genetic polymorphism has two alleles at frequencies of 0.45 and 0.55. Describe three different scenarios to explain these observed allele frequencies. You can propose that the alleles are neutral, beneficial, or deleterious.
3. Most new mutations are detrimental, yet rare beneficial mutations can be adaptive. With regard to the fate of new mutations, discuss whether you think it is more important for natural selection to select against detrimental alleles or to select in favor of beneficial ones. Which do you think is more significant in human populations?

Note: All answers appear at the website for this textbook; the answers to even-numbered questions are in the back of the textbook.