Let’s take a closer look at the four evolutionary processes by examining a null hypothesis—what happens to allele frequencies when the evolutionary mechanisms are not operating.

25.1 Analyzing Change in Allele Frequencies: The Hardy-Weinberg Principle

To study how the four evolutionary processes affect populations, biologists take a three-pronged approach. First they create mathematical models that track the fate of alleles over time. Then they collect data to test predictions made by the models’ equations. Finally, they apply the results to solve problems in human genetics, conservation of endangered species, or other fields.

This research strategy began in 1908, when both G. H. Hardy and Wilhelm Weinberg published a major result independently. At the time, it was commonly believed that changes in allele frequency occur simply as a result of sexual reproduction—meiosis followed by the random fusion of gametes (eggs and sperm) to form offspring. Some biologists claimed that dominant alleles inevitably increase in frequency. Others predicted that two alleles of the same gene inevitably reach a frequency of 0.5.

To test these hypotheses, Hardy and Weinberg analyzed what happens to the frequencies of alleles when many individuals in a population mate and produce offspring. Instead of thinking about the consequences of a mating between two parents with a specific pair of genotypes, as we did with Punnett squares in Chapter 13, Hardy and Weinberg wanted to know what happened in an entire population, when all of the individuals—and thus all possible genotypes—bred. Like Darwin, Hardy and Weinberg were engaged in population thinking.

To analyze the consequences of matings among all of the individuals in a population, Hardy and Weinberg invented a novel approach: They imagined that all of the gametes produced in each generation go into a single group called the gene pool and then combine at random to form offspring. Something very much like this happens in species like clams and sea stars and sea urchins, which release their gametes into the water, where they mix randomly with gametes from other individuals to symbolize the frequency of \( A_1 \) alleles in the gene pool and \( q \) symbolize the frequency of \( A_2 \) alleles in the same gene pool. Because there are only two alleles, the two frequencies must add up to 1; that is, \( p + q = 1 \). Although \( p \) and \( q \) can have any value between 0 and 1, let’s suppose that the initial frequency of \( A_1 \) is 0.7 and that of \( A_2 \) is 0.3 (Figure 25.1, step 1). In the case, 70 percent of the gametes in the gene pool carry \( A_1 \) and 30 percent carry \( A_2 \) (Figure 25.1, step 2).

Because only two alleles are present, three genotypes are possible: \( A_1A_1, A_1A_2, \) and \( A_2A_2 \) (Figure 25.1, step 3). What will the frequency of these three genotypes be in the next generation? Figure 25.1, step 4, explains the logic of Hardy’s and Weinberg’s result:

- The frequency of the \( A_1A_1 \) genotype is \( p^2 \).
- The frequency of the \( A_1A_2 \) genotype is \( 2pq \).
- The frequency of the \( A_2A_2 \) genotype is \( q^2 \).

The genotype frequencies in the offspring generation must add up to 1, which means that \( p^2 + 2pq + q^2 = 1 \). In our numerical example, \( 0.49 + 0.42 + 0.09 = 1 \). Figure 25.1, step 5, shows how the frequencies of alleles \( A_1 \) and \( A_2 \) are calculated from these genotype frequencies. In our example, the frequency of allele \( A_1 \) is still 0.7 and the frequency of allele \( A_2 \) is still 0.3. Thus, the frequency of allele \( A_1 \) in the next generation is still 0.7, and the frequency of allele \( A_2 \) is still 0.3. No allele frequency change occurred. Even if \( A_1 \) is dominant to \( A_2 \), it does not increase in frequency (Figure 25.1, step 6). And there is no trend toward both alleles reaching a frequency of 0.5. Figure 25.1 illustrates the same result a little differently. The figure uses a Punnett square in a novel way: to predict the outcome of random mating—meaning, random combinations of all gametes—in a population. The outcome is the same as in Figure 25.1.

This result is called the Hardy-Weinberg principle. It makes two fundamental claims:

1. If the frequencies of alleles \( A_1 \) and \( A_2 \) in a population are given by \( p \) and \( q \), then the frequencies of genotypes \( A_1A_1, A_1A_2, \) and \( A_2A_2 \) will be given by \( p^2, 2pq, \) and \( q^2 \) for generation after generation.
2. When alleles are transmitted according to the rules of Mendelian inheritance, their frequencies do not change.
models' equations. Finally, they apply the results to solve problems in human genetics, conservation of endangered species, or other fields.

This research strategy began in 1908, when both G. H. Hardy and Wilhelm Weinberg published a major result independently. At the time, it was commonly believed that changes in allele frequency occur simply as a result of sexual reproduction—meiosis followed by the random fusion of gametes (eggs and sperm) to form offspring. Some biologists claimed that dominant alleles inevitably increase in frequency. Others predicted that two alleles of the same gene inevitably reach a frequency of 0.5.

To test these hypotheses, Hardy and Weinberg analyzed what happens to the frequencies of alleles when many individuals in a population mate and produce offspring. Instead of thinking about the consequences of a mating between two parents with a specific pair of genotypes, as we did with Punnett squares in Chapter 13, Hardy and Weinberg wanted to know what happened in an entire population, when all of the individuals—and thus all possible genotypes—bred. Like Darwin, Hardy and Weinberg were engaged in population thinking.

To analyze the consequences of matings among all of the individuals in a population, Hardy and Weinberg invented a novel approach: They imagined that all of the gametes produced in each generation go into a single group called the *gene pool* and then combine at random to form offspring. Something very much like this happens in species like clams and sea stars and sea urchins, which release their gametes into the water, where they mix randomly with gametes from other individuals in the population and combine to form zygotes.

To determine which genotypes would be present in the next generation and in what frequency, Hardy and Weinberg simply had to calculate what happened when two gametes were plucked at random out of the gene pool, many times, and each of these gamete pairs was then combined to form offspring. These calculations would predict the genotypes of the offspring that would be produced, as well as the frequency of each genotype.

The researchers began by analyzing the simplest situation possible—that just two alleles of a particular gene exist in a population. Let’s call these alleles $A_1$ and $A_2$. We’ll use $p$ to

- The frequency of the $A_1A_1$ genotype is $p^2$.
- The frequency of the $A_1A_2$ genotype is $2pq$.
- The frequency of the $A_2A_2$ genotype is $q^2$.

The genotype frequencies in the offspring generation must add up to 1, which means that $p^2 + 2pq + q^2 = 1$. In our numerical example, $0.49 + 0.42 + 0.09 = 1$. Figure 25.1, step 3, shows how the frequencies of alleles $A_1$ and $A_2$ are calculated from these genotype frequencies. In our example, the frequency of allele $A_1$ is still 0.7 and the frequency of allele $A_2$ is still 0.3. Thus, the frequency of allele $A_1$ in the next generation is still $p$ and the frequency of allele $A_2$ is still $q$. No allele frequency change occurred. Even if $A_1$ is dominant to $A_2$, it does not increase in frequency (Figure 25.1, step 6). And there is no trend toward both alleles reaching a frequency of 0.5. Figure 25.2 illustrates the same result a little differently. The figure uses a Punnett square in a novel way: to predict the outcome of random mating—meaning, random combinations of all gametes in a population. The outcome is the same as in Figure 25.1.

This result is called the *Hardy-Weinberg principle*. It makes two fundamental claims:

1. If the frequencies of alleles $A_1$ and $A_2$ in a population are given by $p$ and $q$, then the frequencies of genotypes $A_1A_1$, $A_1A_2$, and $A_2A_2$ will be given by $p^2$, $2pq$, and $q^2$ for generation after generation.

2. When alleles are transmitted according to the rules of Mendelian inheritance, their frequencies do not change over time. For evolution to occur, some other factor or factors must come into play.

What are these other factors?

**The Hardy-Weinberg Model Makes Important Assumptions**

The Hardy-Weinberg model is based on important assumptions about how populations and alleles behave. Specifically, for a population to conform to the Hardy-Weinberg principle, none of the four mechanisms of evolution can be acting on the population. In addition, the model assumes that mating is random with respect
A NUMERICAL EXAMPLE OF THE HARDY-WEINBERG PRINCIPLE

Allele frequencies in parental generation: 
- Allele $A_1$: $p = 0.7$
- Allele $A_2$: $q = 0.3$

1. Suppose allele frequencies in the parental generation were 0.7 and 0.3.

2. 70% of gametes in the gene pool carry allele $A_1$, and 30% carry allele $A_2$.

3. Pick two gametes at random from the gene pool to form offspring. You have a 70% chance of picking allele $A_1$ and a 30% chance of picking allele $A_2$.

4. Three genotypes are possible. Calculate the frequencies of these three combinations of alleles.

5. When the offspring breed, imagine their gametes entering a gene pool. Calculate the frequencies of the two alleles in this gene pool.

6. The frequencies of $A_1$ and $A_2$ have not changed from parental to offspring generation. Evolution has not occurred.

A Numerical Example of the Hardy-Weinberg Principle. To understand the logic behind calculating the frequency of $A_1A_1$ genotypes in step 4, see BioSkills 9.

In the gene in question. Thus, here are the five conditions that must be met:

- No natural selection at the gene in question. In step 2 of Figure 25.1, the model assumed that all members of the parental generation survived and contributed equal numbers of gametes to the gene pool, no matter what their genotype.

- No genetic drift, or random allele frequency changes, affecting the gene in question. We avoided this type of allele frequency change in step 4 of Figure 25.1 by assuming that we
A Numerical Example of the Hardy-Weinberg Principle. To understand the logic

in question. Thus, here are the five conditions that

natural selection at the gene in question. In step 2 of the model assumed that all members of the parental generation survived and contributed equal numbers to the gene pool, no matter what their genotype.

Genetic drift, or random allele frequency changes, affects the gene in question. We avoided this type of allele frequency change in step 4 of Figure 25.1 by assuming that we used alleles in their exact frequencies \( p \) and \( q \), and not at different values caused by chance. For example, allele could not "get lucky" and get drawn more than 70 percent of the time. No random changes due to luck occurred.

Gene flow. No new alleles were added by immigration through emigration anywhere in Figure 25.1. As a result, all of the alleles in the offspring population came from the original population's gene pool.

Mutation. We didn't consider that new \( A_1 \) or \( A_2 \) or new alleles might be introduced into the gene pool in step 3 or step 5 of Figure 25.1.

### Allele frequencies in parental generation:

\[
\begin{align*}
A_1 & = p = 0.7 \\
A_2 & = q = 0.3
\end{align*}
\]

### All eggs in gene pool

<table>
<thead>
<tr>
<th></th>
<th>0.7 ( A_1 )</th>
<th>0.3 ( A_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7 ( A_1 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3 ( A_2 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Genotype frequencies in offspring generation:

\[
\begin{align*}
A_1 A_1 & = p^2 = 0.49 \\
A_1 A_2 & = 2pq = 0.42 \\
A_2 A_2 & = q^2 = 0.09
\end{align*}
\]

### Allele frequencies in offspring generation:

\[
\begin{align*}
A_1 & = p = 0.49 + \frac{1}{2}(0.42) = 0.70 \\
A_2 & = q = \frac{1}{2}(0.42) + 0.09 = 0.30
\end{align*}
\]

**FIGURE 25.2** A Punnett Square Illustrates the Hardy-Weinberg Principle.
5. Random mating with respect to the gene in question. We enforced this condition by picking gametes from the gene pool at random in step 3 of Figure 25.1. We did not allow individuals to choose a mate based on their genotype.

The Hardy-Weinberg principle tells us what to expect if selection, genetic drift, gene flow, and mutation are not affecting a gene, and if mating is random with respect to that gene. Under these conditions, the genotypes \( A_1A_1, A_1A_2, \) and \( A_2A_2 \) should be in the Hardy-Weinberg proportions \( p^2, 2pq, \) and \( q^2 \), and no evolution will occur.

**How Does the Hardy-Weinberg Principle Serve as a Null Hypothesis?**

Recall from Chapter 1 that a null hypothesis predicts there are no differences among the treatment groups in an experiment. Biologists often want to test whether natural selection is acting on a particular gene, nonrandom mating is occurring, or one of the other evolutionary mechanisms is at work. In addressing questions like these, the Hardy-Weinberg principle functions as a null hypothesis. Given a set of allele frequencies, it predicts what genotype frequencies will be when natural selection, mutation, genetic drift, and gene flow are not affecting the gene; and when mating is random with respect to that gene. If biologists observe genotype frequencies that do not conform to the Hardy-Weinberg prediction, it means that something interesting is going on: Either nonrandom mating is occurring, or allele frequencies are changing for some reason. Further research is needed to determine which of the five Hardy-Weinberg conditions is being violated.

Let’s consider two examples to illustrate how the Hardy-Weinberg principle is used as a null hypothesis: MN blood types and HLA genes, both in humans.

**Are MN Blood Types in Humans in Hardy-Weinberg Equilibrium?** One of the first genes that geneticists could study in natural populations was the MN blood group of humans. Most human populations have two alleles, designated \( M \) and \( N \), at this gene. Because the gene codes for a protein found on the surface of red blood cells, researchers could determine whether individuals are \( MM, MN, \) or \( NN \) by treating blood samples with antibodies to each protein (this technique was first introduced in Chapter 8). To estimate the frequency of each genotype in the population, geneticists obtain data from a large number of individuals and then divide the number of individuals with a specific genotype by the total number of individuals in the sample.

Table 25.1 shows MN genotype frequencies for populations from throughout the world and illustrates how observed genotype frequencies are compared with the genotype frequencies expected if the Hardy-Weinberg principle holds. The analysis is based on the following steps:

1. Estimate genotype frequencies by observation—in this case, by testing many blood samples for the \( M \) and \( N \) alleles. These frequencies are given in the rows labeled “observed” in Table 25.1.

2. Calculate observed allele frequencies from the observed genotype frequencies. In this case, the frequency of the \( M \) allele is the frequency of \( MM \) homozygotes plus half the frequency of \( MN \) heterozygotes; the frequency of the \( N \) allele is the frequency of \( NN \) homozygotes plus half the frequency of \( MN \) heterozygotes. (You can review the logic behind this calculation in steps 5 and 6 of Figure 25.1.)

3. Use the observed allele frequencies to calculate the genotype expected according to the Hardy-Weinberg principle. Under the null hypothesis of no evolution and random mating, the expected genotype frequencies are \( p^2 : 2pq : q^2 \).

<table>
<thead>
<tr>
<th>Population and Location</th>
<th>MM</th>
<th>MN</th>
<th>NN</th>
<th>M</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inuit (Greenland)</td>
<td>0.835</td>
<td>0.156</td>
<td>0.009</td>
<td>0.913</td>
<td>0.087</td>
</tr>
<tr>
<td>Native Americans (U.S.)</td>
<td>0.600</td>
<td>0.351</td>
<td>0.049</td>
<td>0.776</td>
<td>0.224</td>
</tr>
<tr>
<td>Caucasians (U.S.)</td>
<td>0.292</td>
<td>0.494</td>
<td>0.213</td>
<td>0.540</td>
<td>0.460</td>
</tr>
<tr>
<td>Aborigines (Australia)</td>
<td>0.025</td>
<td>0.304</td>
<td>0.672</td>
<td>0.178</td>
<td>0.825</td>
</tr>
<tr>
<td>Ainu (Japan)</td>
<td>0.179</td>
<td>0.302</td>
<td>0.319</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**EXERCISE** Fill in the values for allele frequencies and expected genotype frequencies for the Ainu people of Japan.