19

POPULATION GENETICS

The Hardy-Weinberg Equilibrium and Mating Systems

STUDY OBJECTIVES

1. To understand the concept of population-level genetic processes 553
2. To learn the assumptions and nature of the Hardy-Weinberg equilibrium and its extensions 554
3. To test whether a population is in Hardy-Weinberg equilibrium 557
4. To analyze the process and consequences of nonrandom mating in diploid populations 560

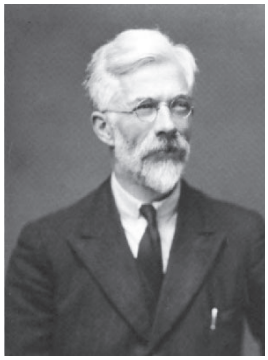
STUDY OUTLINE

The cheetah (*Acinonyx jubatus*) is in peril of extinction; it has very low genetic variability.

(Gregory G. Dimijian, MD/Photo Researchers, Inc.)

Evolution is a process that takes place in populations of organisms. To study evolution, we need to shift our focus to population genetics, the algebraic description of the genetic makeup of a population and the changes in allelic frequencies in populations over time. This chapter is the first of three that looks at what population genetics can tell us about the way evolution proceeds.

Almost all of the mathematical foundations of genetic changes in populations were developed in a short period of time during the 1920s and 1930s by three men: R. A. Fisher, J. B. S. Haldane, and S. Wright. Some measure of disagreement emerged among these men, but they disagreed on which evolutionary processes were more important, not on how the processes worked. Since the 1960s, excitement has arisen in the field of population genetics, primarily on three fronts. First, the high-speed computer has made it possible to do a large amount of arithmetic in a very short period of time; thus, complex simulations of real populations can be added to the repertoire of the experimental geneticist. Second, electrophoresis has provided a means of gathering the large amount of empirical data necessary to check some of the assumptions used in mathematical models. The information and interpretation of the electrophoretic data have generated some controversy about the role of "neutral" evolutionary changes in natural populations. Last, newer techniques of molecular genetics are being used to analyze the relationships among species and the rate of evolutionary processes. We consider these studies later.



Sir Ronald A. Fisher (1890–1962). (Courtesy of The National Portrait Gallery, England.)

Sewall Wright (1889–1988). (Courtesy of Dr. Sewall Wright.)

HARDY-WEINBERG EQUILIBRIUM

Let us begin with a few definitions. For the most part, we define a *species* as a group of organisms potentially capable of interbreeding. Most species are made up of **populations,** interbreeding groups of organisms that are usually subdivided into partially isolated breeding groups called **demes.** As we will see, it is these demes, or local populations, that can evolve.

In 1908, G. H. Hardy, a British mathematician, and W. Weinberg, a German physician, independently discovered a rule that relates allelic and genotypic frequencies in a population of diploid, sexually reproducing individuals if that population has random mating, large size, no mutation or migration, and no selection. The rule has three aspects:

1. The allelic frequencies at an autosomal locus in a population will not change from one generation to the next (allelic-frequency equilibrium).
2. The genotypic frequencies of the population are determined in a predictable way by the allelic frequencies (genotypic-frequency equilibrium).
3. The equilibrium is neutral. That is, if it is perturbed, it will be reestablished within one generation of random mating at the new allelic frequencies (if all the other requirements are maintained).

Calculating Allelic Frequencies

If we consider an autosomal locus in a diploid, sexually reproducing species, allelic frequencies can be measured in either of two ways. The first is simply by counting genes:

$$\text{frequency of the } a \text{ allele, } q, = \frac{\text{number of } a \text{ alleles}}{\text{total number of alleles}}$$

The expression "frequency of" can be shortened to $f(\ )$. For example, the frequency of the a allele is written as $f(a)$. Since the homozygotes have two of a given allele and heterozygotes have only one, and since the total number of alleles is twice the number of individuals (each individual carries two alleles), we can calculate allelic frequencies in the following manner. Consider, for example, the phenotypic distribution of MN blood types (controlled by the codominant M and N alleles) among two hundred persons chosen randomly in Columbus, Ohio:

$$\begin{aligned} \text{type M } (MM \text{ genotype}) &= 114 \\ \text{type MN } (MN \text{ genotype}) &= 76 \\ \text{type N } (NN \text{ genotype}) &= \underline{10} \\ & \quad 200 \end{aligned}$$

Then,

$$p = f(M) = \frac{2(114) + 76}{2(200)} = \frac{304}{400} = 0.76$$

Similarly,

$$q = f(N) = \frac{2(10) + 76}{2(200)} = \frac{96}{400} = 0.24$$

Alternatively, because the frequencies of the two alleles, *M* and *N*, must add up to unity ($p + q = 1$, $q = 1 - p$, and $p = 1 - q$), if we know that $p = 0.76$, then $q = 1 - 0.76 = 0.24$.

Another way of calculating allelic frequencies is based on knowledge of the genotypic frequencies. In this example, the frequencies are

$$f(MM) = \frac{114}{200} = 0.57$$

$$f(MN) = \frac{76}{200} = 0.38$$

$$f(NN) = \frac{10}{200} = 0.05$$

We derive an expression for calculating *p* and *q* based on genotypic frequencies as follows:

$$p = f(M) = \frac{2 \times \text{number of } MM + \text{number of } MN}{2 \times \text{total number}}$$

$$= \frac{2 \times \text{number of } MM}{2 \times \text{total number}} + \frac{\text{number of } MN}{2 \times \text{total number}}$$

$$= f(MM) + (1/2)f(MN)$$

and,

$$q = f(N) = \frac{2 \times \text{number of } NN + \text{number of } MN}{2 \times \text{total number}}$$

$$= \frac{2 \times \text{number of } NN}{2 \times \text{total number}} + \frac{\text{number of } MN}{2 \times \text{total number}}$$

$$= f(NN) + (1/2)f(MN)$$

Thus, allelic frequencies can be calculated as the frequency of homozygotes, plus half the frequency of heterozygotes, as follows:

$$p = f(M) = f(MM) + (1/2)f(MN)$$

$$= 0.57 + (1/2)0.38 = 0.76$$

$$q = f(N) = f(NN) + (1/2)f(MN)$$

$$= 0.05 + (1/2)0.38 = 0.24$$

or

$$q = 1 - p = 1 - 0.76 = 0.24$$

Note that these two methods (counting alleles and using genotypic frequencies) are algebraically identical and thus give identical results.

Assumptions of Hardy-Weinberg Equilibrium

We will consider a population of diploid, sexually reproducing organisms with a single autosomal locus segregating two alleles (i.e., every individual is one of three genotypes—*MM, MN*, or *NN*). Later on, we generalize the discussion to include multiple alleles and multiple loci. For the moment, the focus is on a genetic system such as the *MN* locus in human beings. The following major assumptions are necessary for the Hardy-Weinberg equilibrium to hold.

Random Mating

The first assumption is **random mating,** which means that the probability that two genotypes will mate is the product of the frequencies (or probabilities) of the genotypes in the population. If the *MM* genotype makes up 90% of a population, then any individual has a 90% chance (probability = 0.9) of mating with a person with an *MM* genotype. The probability of an *MM* by *MM* mating is (0.9)(0.9), or 0.81.

Deviations from random mating come about for two reasons: choice or circumstance. If members of a population choose individuals of a particular phenotype as mates more or less often than at random, the population is engaged in **assortative mating.** If individuals with similar phenotypes are mating more often than at random, *positive assortative mating* is in force; if matings occur between individuals with dissimilar phenotypes more often than at random, *negative assortative mating,* or **disassortative mating,** is at work.

Deviations from random mating also arise when mating individuals are either more closely related genetically or more distantly related than individuals chosen at random from the population. **Inbreeding** is the mating of related individuals, and **outbreeding** is the mating of genetically unrelated individuals. Inbreeding is a consequence of pedigree relatedness (e.g., cousins) and small population size.

One of the first counterintuitive observations of population genetics is that deviations from random mating alter genotypic frequencies but not allelic frequencies. Envision a population in which every individual is the parent of two children. On the average, each individual will pass on one copy of each of his or her alleles. Assortative mating and inbreeding will change the zygotic (genotypic) combinations from one generation to the next, but will not change which alleles are passed into

the next generation. Thus genotypic, but not allelic, frequencies change under nonrandom mating.

Large Population Size

Even when an extremely large number of gametes is produced in each generation, each successive generation is the result of a sampling of a relatively small portion of the gametes of the previous generation. A sample may not be an accurate representation of a population, especially if the sample is small. Thus, the second assumption of the Hardy-Weinberg equilibrium is that the population is infinitely large. A large population produces a large sample of successful gametes. The larger the sample, the greater the probability that the allelic frequencies of the offspring will accurately represent the allelic frequencies in the parental population. When populations are small or when alleles are rare, changes in allelic frequencies take place due to chance alone. These changes are referred to as **random genetic drift,** or just *genetic drift.*

No Mutation or Migration

Allelic and genotypic frequencies may change through the loss or addition of alleles through mutation or migration (immigration or emigration) of individuals from or into a population. The third and fourth assumptions of the Hardy-Weinberg equilibrium are that neither mutation nor migration causes such allelic loss or addition in the population.

No Natural Selection

The final assumption necessary to the Hardy-Weinberg equilibrium is that no individual will have a reproductive advantage over another individual because of its genotype. In other words, no natural selection is occurring. (Artificial selection, as practiced by animal and plant breeders, will also perturb the Hardy-Weinberg equilibrium of captive populations.)

In summary, the Hardy-Weinberg equilibrium holds (is exactly true) for an infinitely large, randomly mating population in which mutation, migration, and natural selection do not occur. In view of these assumptions, it seems that such an equilibrium would never be characteristic of natural populations. However, this is not the case. Hardy-Weinberg equilibrium is approximated in natural populations for two major reasons. First, the consequences of violating some of the assumptions, such as no mutation or infinitely large population size, are small. Mutation rates, for example, are on the order of one change per locus per generation per 10^6 gametes. Thus, there is virtually no measurable effect of mutation in a single generation. In addition, populations do not have to be infinitely large to act as if they were. As we will see, a relatively small population can still closely approximate Hardy-Weinberg equilibrium. In other words, minor deviations from the other assumptions can still result in a good fit to the equilibrium; only major deviations can be detected statistically. Second, the Hardy-Weinberg equilibrium is extremely resilient to change because, regardless of the perturbation, the equilibrium is usually reestablished after only one generation of random mating. The new equilibrium will be, however, at the new allelic frequencies—the Hardy-Weinberg equilibrium does not "return" to previous allelic values.

Proof of Hardy-Weinberg Equilibrium

The three properties of the Hardy-Weinberg equilibrium are that (1) allelic frequencies do not change from generation to generation, (2) allelic frequencies determine genotypic frequencies, and (3) the equilibrium is achieved in one generation of random mating. We will concentrate for a moment on the second property. In a population of individuals segregating the *A* and *a* alleles at the *A* locus, each individual will be one of three genotypes: *AA*, *Aa*, or *aa*. If $p = f(A)$ and $q = f(a)$, then we can predict the genotypic frequencies in the next generation. If all the assumptions of the Hardy-Weinberg equilibrium are met, the three genotypes should occur in the population in the same frequencies at which gametes would be randomly drawn in pairs from a **gene pool.** A gene pool is defined as all of the alleles available among the reproductive members of a population from which gametes can be drawn. Thus,

$$f(AA) = (p \times p) = p^2$$
$$f(Aa) = (p \times q) + (q \times p) = 2pq$$
$$f(aa) = (q \times q) = q^2$$

demonstrates the second property of the Hardy-Weinberg equilibrium (fig. 19.1).

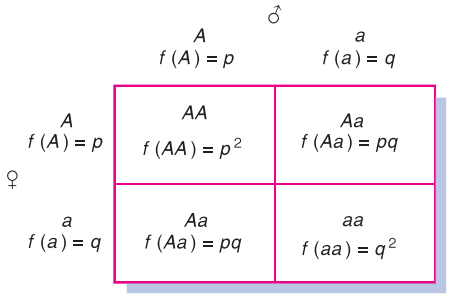


Figure 19.1 Gene pool concept of zygote formation. Males and females have the same frequencies of the two alleles: $f(A) = p$ and $f(a) = q$. After one generation of random mating, the three genotypes, *AA*, *Aa*, and *aa*, have the frequencies of p^2, $2pq$, and q^2, respectively.

Another way of demonstrating the properties of the Hardy-Weinberg equilibrium for the one-locus, two-allele case in sexually reproducing diploids is by simply observing the offspring of a randomly mating, infinitely large population. Let the initial frequencies of the three genotypes be any values that sum to one; for example, let X, Y, and Z be the proportions of the AA, Aa, and aa genotypes, respectively. The proportions of offspring after one generation of random mating are as shown in table 19.1. For example, the probability that an AA individual will mate with an AA individual is $X \times X$, or X^2. Since all the offspring of this mating are AA, they are counted only under the AA column of offspring in table 19.1. When all possible matings are counted, the offspring with each genotype are summed. The proportion of AA offspring is $X^2 + XY + (1/4)Y^2$, which factors to $(X + [1/2]Y)^2$. Recall that the frequency of an allele is the frequency of its homozygote plus half the frequency of the heterozygote. Hence, $X + (1/2)Y$ is the frequency of A, since $X = f(AA)$ and $Y = f(Aa)$. If $p = f(A)$, then $(X + [1/2]Y)^2$ is p^2. Thus, after one generation of random mating, the proportion of AA homozygotes is p^2. Similarly, the frequency of aa homozygotes after one generation of random mating is $Z^2 + YZ + (1/4)Y^2$, which factors to $(Z + [1/2]Y)^2$, or q^2. The frequency of heterozygotes when summed and factored (table 19.1) is $2(X + [1/2]Y)(Z + [1/2]Y)$, or $2pq$. Therefore, after one generation of random mating, the three genotypes (AA, Aa, and aa) occur as p^2, $2pq$, and q^2.

Looking at the first property of the Hardy-Weinberg equilibrium, that allelic frequencies do not change generation after generation, we can ask, Have the allelic frequencies changed from one generation to the next (from the parents to the offspring)? Before random mating, the frequency of the A allele is, by definition, p:

$$f(A) = p = f(AA) + (1/2)f(Aa) = X + (1/2)Y$$

After random mating, the frequency of the A homozygote is p^2, and the frequency of the heterozygote is $2pq$. Thus, the frequency of the A allele, the frequency of its homozygote plus half the frequency of the heterozygotes, is

$$\begin{aligned} f(A) &= f(AA) + (1/2)f(Aa) \\ &= p^2 + (1/2)(2pq) \\ &= p^2 + pq = p(p + q) \\ &= p \text{ (remember, } p + q = 1) \end{aligned}$$

Thus, in a randomly mating population of sexually reproducing diploid individuals, the allelic frequency, p, does not change from generation to generation. Here, by observing the offspring of a randomly mating population, we have proven all three properties of the Hardy-Weinberg equilibrium.

Generation Time

Although generation interval is commonly thought of as the average age of the parents when their offspring are born, the statistical concept of a generation is more complex. Demographers use formulas that relate generation time to the age of reproducing females, the reproductive level of each age group, and the probability of survival in each age group. Here, to avoid these complexities, we will use **discrete generations,** unless otherwise noted. That is, we will assume that all the individuals drawn in a sam-

Table 19.1 **Proportions of Offspring in a Randomly Mating Population Segregating the *A* and *a* Alleles at the *A* locus: $X = f(AA)$, $Y = f(Aa)$, and $Z = f(aa)$**

		Offspring		
Mating	**Proportion**	***AA***	***Aa***	***aa***
$AA \times AA$	X^2	X^2		
$AA \times Aa$	XY	$(1/2)XY$	$(1/2)XY$	
$AA \times aa$	XZ		XZ	
$Aa \times AA$	XY	$(1/2)XY$	$(1/2)XY$	
$Aa \times Aa$	Y^2	$(1/4)Y^2$	$(1/2)Y^2$	$(1/4)Y^2$
$Aa \times aa$	YZ		$(1/2)YZ$	$(1/2)YZ$
$aa \times AA$	XZ		XZ	
$aa \times Aa$	YZ		$(1/2)YZ$	$(1/2)YZ$
$aa \times aa$	Z^2			Z^2
Sum	$(X + Y + Z)^2$	$(X + [1/2]Y)^2$	$2(X + [1/2]Y)(Z + [1/2]Y)$	$(Z + [1/2]Y)^2$

ple, for purposes of determining allelic and genotypic frequencies, are drawn from the same generation, and that, in resampling the population, the second sample represents the offspring of the first generation. The discrete-generation model holds for organisms such as annual plants and fruit flies maintained under laboratory conditions, with no breeding among individuals of different generations. Generations that overlap, as in populations of human beings and many other organisms, usually are better described by somewhat more complex mathematical models.

Testing for Fit to Hardy-Weinberg Equilibrium

There are several ways to determine whether a given population conforms to the Hardy-Weinberg equilibrium at a particular locus. However, the question usually arises when there is just a single sample from a population, representing only one generation. Can the existence of the Hardy-Weinberg equilibrium be determined with just one sample? The answer is that we can determine whether the three genotypes (*AA*, *Aa*, and *aa*) occur with the frequencies p^2, $2pq$, and q^2. If they do, then the population is considered to be in Hardy-Weinberg proportions; if not, then the population is not considered to be in Hardy-Weinberg proportions.

MN Blood Types

To determine whether observed and expected allelic frequencies are the same, we can use the chi-square statistical test. In a chi-square test, we compare an observed number with an expected number. In this case, the observed values are the actual numbers of the three genotypes in the sample, and the expected values come from the prediction that the genotypes will occur in the p^2, $2pq$, and q^2 proportions. An analysis for the Ohio MN blood-type data is presented in table 19.2. The agreement between observed and expected numbers is very good, obvious even before the calculation of the chi-square value. Since the critical chi-square for one degree of freedom at the 0.05 level is 3.841 (see table 4.4), we find that the Ohio population does not deviate from Hardy-Weinberg proportions at the *MN* locus.

Earlier (chapter 4), we used the chi-square statistic to test how well real data fit an expected data set based on a ratio predicted before the test. For example, we tested the data against a 3:1 ratio in table 4.2. In that case, the number of degrees of freedom was simply the number of independent categories: the total number of categories minus one. Here, however, our expected ratio is derived from the data set itself. The values p^2, $2pq$, and q^2 came from p and q, which were estimated from the data. In this case, we lose one additional degree of freedom for every independent value we estimate from the data. If we calculate p from a sample, we lose one degree of freedom. However, we do not lose a degree of freedom for estimating q, since q is no longer an independent variable: $q = 1 - p$. So in the previous case, we lose two degrees of freedom—one for estimating p and one for independent categories. The general rule of thumb in using chi-square analysis to test for data fit to Hardy-Weinberg proportions is that the number of degrees of freedom must equal the number of phenotypes minus the number of alleles (in this case, $3 - 2 = 1$).

The chi-square analysis in table 19.2 may seem paradoxical. Because the observed allelic frequencies calculated from the original genotypic data are used to calculate the expected genotypic frequencies, it may appear to some individuals that the analysis must, by its very nature, show that the population is in Hardy-Weinberg proportions. To demonstrate that this is not necessarily the case, a counterexample appears in table 19.3. We use data similar to the Ohio sample, except that the original number of heterozygotes has been distributed equally among the two homozygote classes. The same allelic frequencies are maintained, yet the genotypic distribution differs. The chi-square value of 200.00 for these data demonstrates that the population represented in table 19.3 is not in Hardy-Weinberg proportions. Thus, a chi-square analysis of fit to the Hardy-Weinberg proportions by no means represents circular reasoning.

Table 19.2 **Chi-Square Test of Goodness-of-Fit to the Hardy-Weinberg Proportions of a Sample of 200 Persons for MN Blood Types for Which $p = 0.76$ and $q = 0.24$**

	MM	MN	NN	Total
Observed Numbers	114	76	10	200
Expected Proportions	p^2	$2pq$	q^2	1.0
	(0.5776)	(0.3648)	(0.0576)	1.0
Expected Numbers	115.52	72.96	11.52	200.0
$\chi^2 = (O - E)^2/E$	0.020	0.127	0.201	0.348

Table 19.3 **Chi-Square Test of Goodness-of-Fit to the Hardy-Weinberg Proportions of a Second Sample of 200 Persons for MN Blood Types for Which $p = 0.76$ and $q = 0.24$ and Heterozygotes Are Absent**

	MM	MN	NN	Total
Observed Numbers	152	0	48	200
Expected Proportions	p^2	$2pq$	q^2	1.0
	(0.5776)	(0.3648)	(0.0576)	1.0
Expected Numbers	115.52	72.96	11.52	200.0
$\chi^2 = (O - E)^2/E$	11.52	72.96	115.52	200.00

PKU

Circumstances sometimes do not allow us to test for Hardy-Weinberg proportions. In the case of a dominant trait, for example, allelic frequencies cannot be calculated from the genotypic classes because the homozygous dominant individuals cannot be distinguished from the heterozygotes. However, we can estimate allelic frequencies by assuming that the Hardy-Weinberg equilibrium exists and, thereby, assuming that the frequency of the recessive homozygote is q^2, from which q and then p can be estimated.

If, for example, Hardy-Weinberg equilibrium is assumed for a disease such as phenylketonuria (PKU), which is expressed only in the homozygous recessive state, it is possible to calculate the proportion of the population that is heterozygous (carriers of the PKU allele). But is it fair to assume Hardy-Weinberg equilibrium here? Until recent medical advances allowed intervention, there was a good deal of selection against individuals with PKU, who were usually mentally retarded. Thus the assumption of no selection, required for equilibrium, is violated. However, only one child in ten thousand live births has PKU. When a genotype is as rare as one in ten thousand, selection has a negligible effect on allelic frequencies. Therefore, because of the rarity of the trait, we can assume Hardy-Weinberg equilibrium and calculate

$$\text{frequency of recessive homozygote} = q^2 = 1/10{,}000 = 0.0001$$

so,

$$q = \sqrt{0.0001} = 0.01$$

and

$$p = 1 - q = 0.99$$

Therefore,

$$\text{frequency of normal homozygote} = p^2 = (0.99)^2 \cong 0.98 \text{ or } 98 \text{ in } 100$$

$$\text{frequency of heterozygote} = 2pq = 2(0.01)\,(0.99) \cong 0.02 \text{ or } 2 \text{ in } 100.$$

By assuming the Hardy-Weinberg equilibrium, we have discovered something not intuitively obvious: A recessive gene causing a trait as rare as one in ten thousand is carried in the heterozygous state by one individual in fifty. Obviously, the chi-square test cannot be used to verify the Hardy-Weinberg proportions since we derived the allelic frequencies by assuming Hardy-Weinberg proportions to begin with. In statistical terms, the number of phenotypes minus the number of alleles $= 2 - 2 = 0$ degrees of freedom, which precludes doing a chi-square test.

EXTENSIONS OF HARDY-WEINBERG EQUILIBRIUM

The Hardy-Weinberg equilibrium can be extended to include, among other cases, multiple alleles and multiple loci.

Multiple Alleles

Multinomial Expansion

The expected genotypic array under Hardy-Weinberg equilibrium is p^2, $2pq$, and q^2, which form the terms of the binomial expansion $(p + q)^2$. If males and females each have the same two alleles in the proportions of p and q, then genotypes will be distributed as a binomial expansion in the frequencies p^2, $2pq$, and q^2 (see fig. 19.1). To generalize to more than two alleles, one need only add terms to the binomial expansion and thus create a multinomial expansion. For example, with alleles a, b, and c with frequencies p, q, and r, the genotypic distribution should be $(p + q + r)^2$, or

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

Homozygotes will occur with frequencies p^2, q^2, and r^2, and heterozygotes will occur with frequencies $2pq$, $2pr$, and $2qr$. The *ABO* blood-type locus in human beings is an interesting example because it has multiple alleles and dominance.

ABO Blood Groups

The *ABO* locus has three alleles: I^A, I^B, and i, with the I^A and I^B alleles codominant, and both dominant to the i allele. These alleles control the production of a surface antigen on red blood cells (see fig. 2.13). Table 19.4 contains blood-type data from a sample of five hundred persons from Massachusetts. Is the population in Hardy-Weinberg proportions? The answer is not apparent from the data in table 19.4 alone, since there are two possible genotypes for both the A and the B phenotypes. No estimate of the allelic frequencies is possible without making assumptions about the number of each genotype within these two phenotypic classes. Is it possible to estimate the allelic frequencies? The answer is yes, if we assume that Hardy-Weinberg equilibrium exists.

One procedure follows. Let us assume that $p = f(I^A)$, $q = f(I^B)$, and $r = f(i)$. Blood type O has the ii genotype; if the population is in Hardy-Weinberg proportions, this genotype should occur at a frequency of r^2. Thus

$$f(ii) = 231/500 = 0.462 = r^2$$

and

$$r = f(i) = \sqrt{0.462} = 0.680$$

From table 19.4, we see that blood type A plus blood type O include only the genotypes I^AI^A, I^Ai, and ii. If the population is in Hardy-Weinberg proportions, these together should be $(p + r)^2$, in which $p^2 = f(I^AI^A)$, $2pr = f(I^Ai)$, and $r^2 = f(ii)$:

$$(p + r)^2 = (199 + 231)/500 = 0.860$$

Then, taking the square root of each side

$$p + r = \sqrt{0.860} = 0.927$$

and

$$p = 0.927 - r = 0.927 - 0.680 = 0.247$$

Table 19.4 **ABO Blood-Type Distribution in 500 Persons from Massachusetts**

Blood Type	Genotype	Number
A	I^AI^A or I^Ai	199
B	I^BI^B or I^Bi	53
AB	I^AI^B	17
O	ii	231
Total		500

The frequency of allele I^B, q, can be obtained by similar logic with blood types B and O, or simply by subtraction:

$$q = 1 - (p + r) = 1 - 0.927 = 0.073$$

Thus, the Hardy-Weinberg equilibrium can be extended to include multiple alleles and can be used to make estimates of the allelic frequencies in the ABO blood groups. With ABO, it is statistically feasible to do a chi-square test because there is one degree of freedom (number of phenotypes − number of alleles = 4 − 3 = 1). We are really testing only the AB and B categories; if we did our calculations as shown, the observed and expected values of phenotypes A and O must be equal.

Multiple Loci

The Hardy-Weinberg equilibrium can also be extended to consider several loci at the same time in the same population. This situation deserves mention because the whole genome is likely involved in evolutionary processes and we must, eventually, consider simultaneous allelic changes in all loci segregating alleles in an organism. (Even with a high-speed computer, simultaneous consideration of many loci is a bit far off in the future.) When two loci, A and B, on the same chromosome are in equilibrium with each other, the combinations of alleles on a chromosome in a gamete follow the product rule of probability. Consider the A locus with alleles A and a and the B locus with alleles B and b, respectively, with allelic frequencies p_A and q_A for A and a, respectively, and p_B and q_B for B and b, respectively. Given completely random circumstances, the chromosome with the A and B alleles should occur at the frequency $p_A\,p_B$. This is referred to as **linkage equilibrium.** When alleles of different loci are not in equilibrium (i.e., not randomly distributed in gametes), the condition is referred to as **linkage disequilibrium.** The approach to linkage equilibrium is gradual and is a function of the recombination distance between the two loci.

For example, let's start with a population out of equilibrium so that all chromosomes are AB (70%) or ab (30%). Then $p_A = 0.7$, $q_A = 0.3$, $p_B = 0.7$, and $q_B = 0.3$. We expect the Ab chromosome to occur $0.7 \times 0.3 = 0.21$, or 21% of the time. The frequency of the Ab chromosome is zero. Assume the map distance between the two loci is 0.1; in other words, 10% of chromatids in gametes are recombinant. Initially, we consider that each locus is in Hardy-Weinberg proportions, or the frequency of AB/AB individuals = 0.49 (0.7×0.7); the frequency of ab/ab individuals is 0.09 (0.3×0.3); and the frequency of AB/ab individuals is 0.42 ($2 \times 0.7 \times 0.3$).

After one generation of random mating, gametes will be as follows:

from *AB*/*AB* individuals (49%): only *AB* gametes, 49% of total

from *ab*/*ab* individuals (9%): only *ab* gametes, 9% of total

from *AB*/*ab* individuals (42%):

AB gametes, 18.9% of total (0.45 × 0.42)

ab gametes, 18.9% of total (0.45 × 0.42)

Ab gametes, 2.1% of total (0.05 × 0.42)

aB gametes, 2.1% of total (0.05 × 0.42)

(The values of 18.9% and 2.1% for the dihybrids result from the fact that since map distance is 0.1, 10% of gametes will be recombinant, split equally between the two recombinant classes—5% and 5%. Ninety percent will be parental, split equally between the two parental classes—45% and 45%. Each of these numbers must be multiplied by 0.42 because the dihybrid makes up 42% of the total number of individuals.)

Although we expect 21% of the chromosomes to be of the *Ab* type, only 2.1%, 10% of the expected, appear in the gene pool after one generation of random mating. You can see that linkage equilibrium is achieved at a rate dependent on the map distance between loci. Unlinked genes, appearing 50 map units apart, also gradually approach linkage equilibrium.

Although we will not derive these extensions here, we note two others. If the frequencies of alleles at an autosomal locus differ in the two sexes, it takes two generations of random mating to achieve equilibrium. In the first generation, the allelic frequencies in the two sexes are averaged so that each sex now has the same allelic frequencies. Genotypic frequencies then come into Hardy-Weinberg proportions in the second generation. However, if the allelic frequencies differ in the two sexes for a sex-linked locus, Hardy-Weinberg proportions are established only gradually. The reasoning is straightforward. Females, with an X chromosome from each parent, average the allelic frequencies from the previous generation. However, males, who get their X chromosomes from their mothers, have the allelic frequencies of the females in the previous generation. Hence, the allelic frequencies are not the same in the two sexes after one generation of random mating, and equilibrium is achieved slowly.

NONRANDOM MATING

The Hardy-Weinberg equilibrium is based on the assumption of random mating. Deviations from random mating come about when phenotypic resemblance or relatedness influences mate choice. When phenotypic resemblance influences mate choice, either *assortative* or *disassortative* mating occurs, depending on whether individuals choose mates on the basis of similarity or dissimilarity, respectively. For example, in human beings, assortative mating occurs for height—short men tend to marry short women, and tall men tend to marry tall women. When relatedness influences mate choice, either *inbreeding* or *outbreeding* occurs, depending on whether mates are more or less related than two randomly chosen individuals from the population. An example of inbreeding in human beings is marriage between first cousins. Both types of nonrandom mating (assortative-disassortative mating and inbreeding-outbreeding) have the same qualitative effects on the Hardy-Weinberg equilibrium: assortative mating and inbreeding increase homozygosity without changing allelic frequencies, whereas disassortative mating and outbreeding increase heterozygosity without changing allelic frequencies.

Two differences are apparent, however, between the effects of phenotypic resemblance and relatedness on mate choice. First, assortative or disassortative mating disturbs the Hardy-Weinberg equilibrium only when the phenotype and genotype are closely related. That is, if assortative mating occurs for a nongenetic trait, then the Hardy-Weinberg equilibrium will not be distorted. Inbreeding and outbreeding affect the genome directly. A second difference between the two types of mating is that the effects of inbreeding or outbreeding are felt across the whole genome, whereas the disturbances to the equilibrium caused by assortative and disassortative mating occur only for the particular trait being considered (and for closely linked loci). Given the similarities in the consequences of the two types of matings, we will concentrate our discussion on inbreeding.

Inbreeding

Inbreeding comes about in two ways: (1) the systematic choice of relatives as mates and (2) the subdivision of a population into small subunits, leaving individuals little choice but to mate with relatives. We will concentrate on inbreeding as the systematic choice of relatives as mates. The consequences of both are similar.

Common Ancestry

An inbred individual is one whose parents are related—that is, there is **common ancestry** in the family tree. The extent of inbreeding thus depends on the degree of common ancestry that the parents of an inbred individual share. When mates share ancestral genes, each may pass on copies of the same ancestral allele to their offspring. An inbred individual can then carry identical copies of a single ancestral allele. In other words, an in-

dividual of *aa* genotype is homozygous and, if it is possible that the *a* allele from each parent is a length of DNA originally copied from a common ancestor, the *aa* individual is said to be inbred.

The first observable effect of inbreeding is the expression of hidden recessives. In human beings, each individual carries, on the average, about four **lethal-equivalent alleles,** alleles that kill when paired to form a homozygous genotype (box 19.1). In many, and probably most, human societies, zygotes are generally heterozygous for these lethal alleles because of a cultural pattern of outbreeding, mating with nonrelatives. Rarely does an outbred zygote receive the same recessive lethal from each parent. Dominance acts to mask the expression of deleterious recessive alleles. But, in the process of inbreeding, when the zygote may receive copies of the same ancestral allele from each parent, there is a substantial increase in the probability that a deleterious allele will pair to form a homozygous genotype (fig. 19.2). Inbreeding can result in spontaneous abortions (miscarriages), fetal deaths, and congenital deformities. In many species, however, inbreeding—even self-fertilization—occurs normally. These species usually do not have the problem with lethal equivalents that species that normally outbreed do. Through time, species that normally inbreed have had these deleterious alleles mostly eliminated, presumably by natural selection. Inbreeding has even been used successfully for artificial selection in livestock and crop plants.

From our previous discussion, you can see that there are two types of homozygosity—**allozygosity,** in which two alleles are alike but unrelated (not copies of the same ancestral allele) and **autozygosity,** in which two alleles have **identity by descent** (i.e., are copies of the same ancestral allele). An **inbreeding coefficient, *F*,** can be defined as the probability of autozygosity, or the probability that the two alleles in an individual at a given locus are identical by descent. This coefficient can range from zero, at which point there is no inbreeding, to one, at which point it is certain an individual is autozygous.

Increased Homozygosity from Inbreeding

What are the effects of inbreeding on the Hardy-Weinberg equilibrium? Let us for a moment return to the gene pool concept to produce zygotes. Assume that an allele drawn from this gene pool is of the *A* type, drawn with a probability of p. On the second draw, the probability of autozygosity, that is, of drawing a copy of the same allele *A*, is F, the inbreeding coefficient. Thus the probability of an autozygous *AA* individual is pF. On the second draw, however, with probability $(1 - F)$, either the *A* or *a* allele can be drawn, with probabilities of $p^2(1 - F)$ and $pq(1 - F)$, respectively. Note that a second *A* allele produces a homozygote that is not inbred

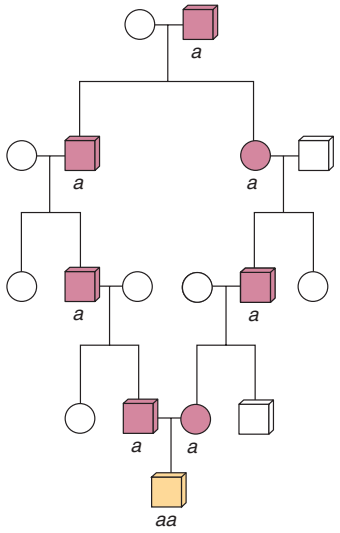


Figure 19.2 Homozygosity by descent of copies of the same ancestral allele, *a*. The individual at the bottom of the pedigree is inbred with the *aa* genotype.

(allozygous). If the first allele drawn was an *a* allele, with probability q, then the probability of drawing the same allele (copy of the same ancestral allele) is F, and thus the probability of autozygosity is qF. However, the probability of drawing an *a* or *A* allele that does not contribute to inbreeding is $(1 - F)$ and, therefore, the probability of an *aa* or *Aa* genotype is $q^2(1 - F)$ and $pq(1 - F)$, respectively. These calculations are summarized in table 19.5, a summary of the genotypic proportions in a population with inbreeding.

Several points emerge from table 19.5. First, when the inbreeding coefficient is zero (completely random mating), the table reduces to Hardy-Weinberg proportions. Second, compared with Hardy-Weinberg proportions, inbreeding increases the proportion of homozygotes in the population (identity by descent implies homozygosity). With complete inbreeding ($F = 1$), only homozygotes will occur in the population.

How does inbreeding affect allelic frequencies? Recall that an allelic frequency is calculated as the frequency of homozygotes for one allele plus half the frequency of the heterozygotes. Here we let p_{n+1} be the frequency of the *A* allele after one generation of inbreeding:

$$\begin{aligned} p_{n+1} &= p^2(1 - F) + pF + (1/2)(2pq)(1 - F) \\ &= p^2(1 - F) + pF + pq(1 - F) \\ &= p^2 + pq + F(p - p^2 - pq) \\ &= p(p + q) + pF(1 - p - q) \\ &= p(1) + pF(0) \\ &= p \end{aligned}$$

BOX 19.1

Experimental Methods

The Determination of Lethal Equivalents

The average person carries about four lethal-equivalent alleles that are hidden because they are recessive. Four lethal equivalents means four alleles that are lethal when homozygous, or eight alleles conferring a 50% chance of mortality when homozygous, or any similar combination of lethal and semilethal alleles. The exact arrangement cannot be determined with current analytical methods. We arrive at the estimate of hidden defective and lethal alleles by using inbreeding data.

J. Crow and M. Kimura, in 1970, analyzed data showing that in Swedish families in which marriages occurred between first cousins, between 16 and 28% of the offspring had genetic diseases. For unrelated parents, the comparable figure is between 4 and 6%. Therefore, it is estimated that the offspring of first cousins have an added risk of 12 to 22% of having a genetic defect. The children of first cousins have an inbreeding coefficient of one-sixteenth. Hence, a theoretical individual who is completely inbred has the risk of genetic defect increased sixteenfold over an individual whose parents are first cousins. If 100% risk is considered 1 lethal equivalent, then a completely inbred individual would carry 2 to 3.5 lethal equivalents ($16 \times 12\%$–$16 \times 22\%$). However, a completely inbred individual is, in essence, a doubled gamete. Since our interest is in the number of deleterious alleles a normal person carries, it is necessary to further multiply the risk by a factor of two to determine the number of lethal-equivalent alleles carried by a normal individual. The conclusion is that the average person carries the equivalent of four to seven alleles that would, in the homozygous state, cause a genetic defect.

A similar calculation can be made using viability data rather than genetic defects to determine the occurrence of lethal equivalents. A study from rural France, also analyzed by Crow and Kimura, showed that the mortality rate of offspring of first cousins was 25%, whereas the analogous figure for the offspring of unrelated parents was about 12%, an increased risk of 13% for the offspring of cousins. Multiplying this risk figure of 0.13 by 32 (16×2) presents a figure of four lethal equivalents per average person in the population. In 1971, L. Cavalli-Sforza and W. Bodmer, using data primarily from Japanese populations, reported an estimate of about two lethal equivalents per average person. Despite some inter-population differences in these estimates, they are about the same order of magnitude—two to seven lethal equivalents per person.

Thus, inbreeding does not change allelic frequencies. We can also see intuitively that inbreeding affects zygotic combinations (genotypes), but not allelic frequencies: Although inbreeding may determine the genotypes of offspring, inbreeding does not change the numbers of each allele that an individual transmits into the next generation.

In summary, inbreeding causes an increase in homozygosity, affects all loci in a population equally, and, in itself, has no effect on allelic frequencies, although it can expose deleterious alleles to selection. The results of inbreeding are evident in the appearance of recessive traits that are often deleterious. Inbreeding increases the rate of fetal deaths and congenital malformations in human beings and in other species that normally outbreed. In outbred agricultural crops and farm animals, decreases in size, fertility, vigor, and yield often result from inbreeding. Once deleterious traits appear due to inbreeding, natural selection can cause their removal from the population. However, in species adapted to inbreeding, including many crop plants and farm animals, inbreeding does not expose deleterious alleles because those alleles have generally been eliminated already.

Pedigree Analysis

Path Diagram Construction

The inbreeding coefficient, F, of an individual (the probability of autozygosity) can be determined by pedigree analysis. This is done by converting a pedigree to a **path diagram** by eliminating all extraneous individuals, those who cannot contribute to the inbreeding coefficient of the individual in question. A path diagram shows the direct line of descent from common ancestors. An example of the conversion of a pedigree to a path diagram is shown in figure 19.3, in which individuals C and F are omitted from the path of descent because they are not related to anyone on the other side of the family tree and, therefore, do not contribute to the "common ancestry" of individual I. The pedigree in figure 19.3 shows an offspring who is the daughter of first cousins. Since first

Table 19.5 Genotypic Proportions in a Population with Inbreeding

Genotype	Due to Random Mating ($1 - F$)		Due to Inbreeding (F)		Observed Proportions
AA	$p^2(1 - F)$	+	pF	=	$p^2 + Fpq$
Aa	$2pq(1 - F)$			=	$2pq(1 - F)$
aa	$q^2(1 - F)$	+	qF	=	$q^2 + Fpq$
Total	$(p^2 + 2pq + q^2)(1 - F)$	+	$(p + q)F$	=	
	$(1 - F)$	+	F	=	1

cousins are the offspring of siblings, they share a set of common grandparents. Thus, individual I can be autozygous for alleles from either ancestor A or B, her great-grandparents. The path diagram shows the only routes by which autozygosity can occur.

The inbreeding coefficient of the offspring of first cousins can be calculated as follows. The path diagram of figure 19.3 is shown again in figure 19.4, with lowercase letters designating gametes. Two paths of autozygosity appear in this diagram, one path for each grandparent as a common ancestor: A to D and E, then to G and H, and finally to I; or B to D and E, then to G and H, and finally to I.

In the path with A as the common ancestor, A contributes a gamete to D and a gamete to E. The probability is one-half that D and E each carry a copy of the same allele. That is, there are four possible allelic combinations for the two gametes, a_1 and a_2: *A-A*; *A-a*; *a-A*; and *a-a*. Of these combinations, the first and last (*A-A* and *a-a*) give a copy of the same allele to the two offspring, D and E, and can thus contribute to autozygosity. The probability that gametes a_1 and d carry copies of the same allele is one-half, and the probability that d and g carry copies of the same allele is also one-half. Similarly, on the other side of the pedigree, the probability is one-half that a_2 and e carry copies of the same allele and one-half that e and h carry copies of the same allele. Thus, the overall probability that the alleles that g and h carry are identical by descent (autozygous) is $(1/2)^5$. In general, it would be $(1/2)^n$ for each path, where n is the number of ancestors in the path.

You may have spotted an additional factor here. Of the possible combinations of allelic copies passed on to D and E, one-half (*A-A* and *a-a*) are autozygous combinations. However, the other half of the combinations, *A-a* and *a-A*, can lead to autozygosity if A is itself inbred. If we let F_A be the inbreeding coefficient of A (the probability that any two alleles at a locus in A are identical by descent), then F_A is the probability that the *A-a* and *a-A* combinations are also autozygous. Thus, the probability that a common ancestor, A, passes on copies of an

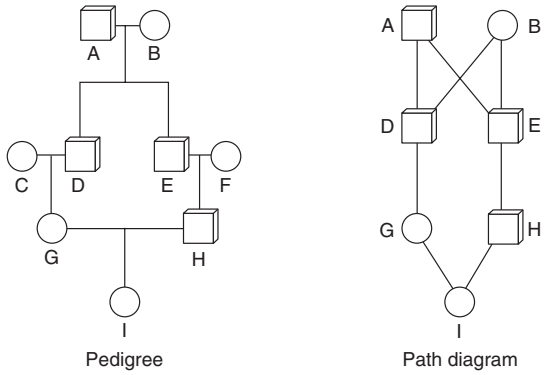


Figure 19.3 Conversion of a pedigree to a path diagram. This pedigree depicts the mating of first cousins. In the path diagram, all extraneous individuals are removed, leaving only those who could contribute to the inbreeding of individual I. Individuals in the line of descent are connected directly with *straight lines,* indicating the paths along which gametes are passed.

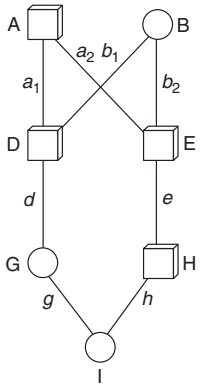


Figure 19.4 The path diagram of the mating of first cousins with gametes labeled in *lowercase letters.*

identical ancestral allele is $1/2 + (1/2)F_A$, or $(1/2)(1 + F_A)$. In other words, there is a one-half probability that the alleles transmitted from A to D and E are copies of the same allele. In the other half of the cases, these alleles can be identical if A is inbred. The probability of identity of A's two alleles is F_A. The expression for the inbreeding coefficient of I, F_I, can now be changed from $(1/2)^n$ by substituting $(1/2)(1 + F_A)$ for one of the (1/2)s to

$$F_I = (1/2)^n(1 + F_A)$$

This equation accounts only for the inbreeding of I by the path involving the common ancestor, A, and does not account for the symmetrical path with B as the common ancestor. To obtain the total probability of inbreeding, the values from each path must be added (because these are mutually exclusive events; see chapter 4). Thus the complete formula for the inbreeding coefficient of the offspring of first cousins is

$$F_I = \Sigma[(1/2)^n(1 + F_J)] \qquad \textbf{(19.1)}$$

in which F_I is the probability that the two alleles in I are identical by descent, n is the number of ancestors in a given path, F_J is the inbreeding coefficient of the common ancestor of that path, and all paths are summed.

In the example of the mating of first cousins (fig. 19.4)

$$F_I = (1/2)^5(1 + F_A) + (1/2)^5(1 + F_B)$$

If we assume that F_A and F_B are zero (which we must assume when the pedigrees of A and B are unknown), then

$$F_I = 2(1/2)^5 = (1/2)^4 = 0.0625$$

This can be interpreted to mean that about 6.25% of individual I's loci are autozygous, or that there is a 6.25% chance of autozygosity at any one of I's loci.

The inbreeding coefficient of the offspring of siblings (fig. 19.5) can also be calculated, assuming that A and B are not themselves inbred (F_A and F_B are zero), as

$$F_I = 2(1/2)^3 = 0.25$$

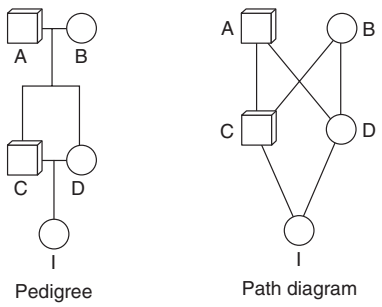


Figure 19.5 Conversion of a sib-mating pedigree to a path diagram. Individual I is inbred.

Thus, about 25% of the loci in an offspring of siblings are autozygous.

Path Diagram Rules

The following points should be kept in mind when calculating an inbreeding coefficient:

1. All possible paths must be counted. A path is possible if gametes can actually pass in that direction. Paths that violate the rules of inheritance cannot be used. For example, in figure 19.4, the following path is unacceptable: I G E A D H I.
2. In any path, an individual can be counted only once.
3. Every path must have one and only one common ancestor. The inbreeding coefficient of any other individual in the path is immaterial.

In figure 19.6, we present a complex pedigree produced from repeated sib mating, a pattern found in livestock and laboratory animals. This pedigree has two interesting points. First, common ancestors occur in several different generations. Second, some of the paths are complex. Thus, we must be sure to count all paths (paths 5 and 6 might not be immediately obvious). Although not shown in figure 19.6, one of the common ancestors, A, is also inbred ($F_A = 0.05$)—a fact that we must take into consideration in paths 3 and 5. Thus, F_I is as follows:

From path 1: $(1/2)^3$	$= 0.1250$
From path 2: $(1/2)^3$	$= 0.1250$
From path 3: $(1/2)^5(1 + 0.05)$	$= 0.0328$
From path 4: $(1/2)^5$	$= 0.0313$
From path 5: $(1/2)^5(1 + 0.05)$	$= 0.0328$
From path 6: $(1/2)^5$	$= 0.0313$
	$F_I = 0.3782$

Population Analysis

It is also possible to define the inbreeding coefficient, F, of a population as the relative reduction in heterozygosity in the population due to inbreeding. In an individual, F is the probability of autozygosity; it represents an increase in homozygosity, which is therefore a decrease in heterozygosity. In a population, it also represents the reduction in heterozygosity. From the definition, we can calculate the population F as follows:

$$F = \frac{(2pq - H)}{2pq}$$

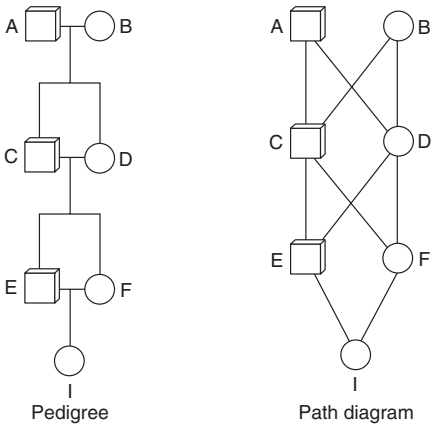
where H is the actual proportion of heterozygotes in a population, and $2pq$ is the expected proportion of het-

erozygotes based on Hardy-Weinberg proportions. This equation reduces to

$$F = 1 - \frac{H}{2pq} \qquad \textbf{(19.2)}$$

This equation shows that when $H = 2pq$, F is zero, meaning that there is no decrease in heterozygotes and therefore, apparently, no inbreeding. When there are no heterozygotes, $F = 1$. This could be the case in a completely inbred population—for example, a self-fertilizing plant species.

As an example of an intermediate case, take the sample of one hundred individuals segregating the A_1 and A_2 alleles at the A locus: A_1A_1, fifty-four; A_1A_2, thirty-two; and A_2A_2, fourteen. In this example, $p = 0.7$, $q = 0.3$, and $H = 0.32$. Since $2pq = 0.42$, $H/2pq = 0.32/0.42 = 0.76$, and $F = 1 - 0.76$, or 0.24. Thus, the inbreeding coefficient of this population is 0.24; there is a 24% reduction in heterozygotes, due presumably to inbreeding.



Paths

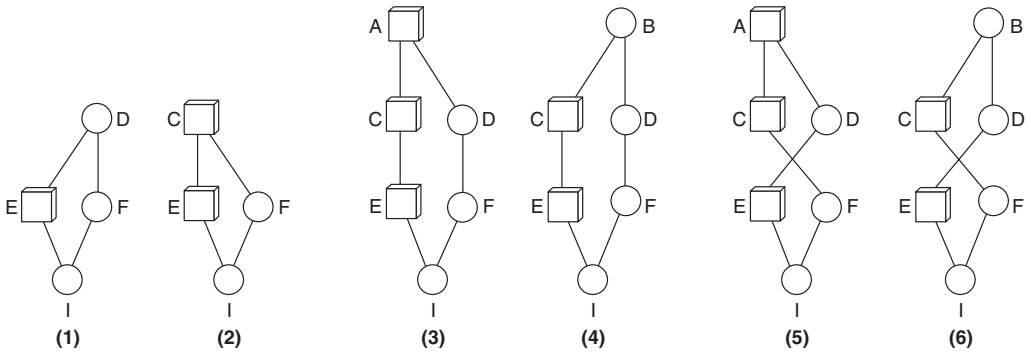


Figure 19.6 Pedigree and path diagram of two generations of sib matings. The six paths involving the potential for autozygosity are shown. $F_A = 0.05$. The paths involve common ancestors in two generations.

SUMMARY

STUDY OBJECTIVE 1: To understand the concept of population-level genetic processes 553–554

In a large, randomly mating population of sexually reproducing diploid organisms, not subject to the influences of mutation, migration, or selection, an equilibrium will be achieved for an autosomal locus with two alleles.

STUDY OBJECTIVE 2: To learn the assumptions and nature of the Hardy-Weinberg equilibrium and its extensions 554–557

The Hardy-Weinberg equilibrium predicts that (1) allelic frequencies (p, q) will not change from generation to generation; (2) genotypes will occur according to the binomial distribution $p^2 = f(AA)$, $2pq = f(Aa)$, and $q^2 = f(aa)$; and (3) if perturbed, equilibrium will reestablish itself in just one generation of random mating.

STUDY OBJECTIVE 3: To test whether a population is in Hardy-Weinberg equilibrium 557–560

To determine whether a population is in Hardy-Weinberg proportions, the observed and expected distribution of genotypes can be compared by the chi-square statistical test. In some circumstances, when it is reasonable to assume equilibrium, we can estimate allelic and genotypic frequencies even when dominance occurs. The Hardy-Weinberg equilibrium is easily extended to the prediction of the frequencies of multiple alleles, multiple loci, and different frequencies of alleles in the two sexes, for both sex-linked and autosomal loci.

STUDY OBJECTIVE 4: To analyze the process and consequences of nonrandom mating in diploid populations 560–565

Random mating is required for the Hardy-Weinberg equilibrium to hold. Deviations from random mating fall into two categories, depending on whether phenotypic resemblance or relatedness is involved in mate choice. Phenotypic resemblance is the basis for assortative and disassortative mating, in which individuals choose similar or dissimilar mates, respectively. Assortative mating causes increased homozygosity only among loci controlling the traits that influence mate choice. There are no changes in allelic frequencies. Similarly, disassortative mating causes increased heterozygosity without changing allelic frequencies.

Mating among relatives, or inbreeding, is represented by F, the inbreeding coefficient, which measures the probability of autozygosity (homozygosity by descent). It can be calculated from pedigrees by using the formula

$$F = \Sigma[(1/2)]^n (1 + F_J)]$$

where n is the number of ancestors in a given path and F_J is the inbreeding coefficient of the common ancestor of that path. Inbreeding exposes recessive deleterious traits already present in the population and causes homozygosity throughout the genome. It does not, by itself, change allelic frequencies. F can also be calculated from the reduction in heterozygosity in a population.

SOLVED PROBLEMS

PROBLEM 1: One hundred fruit flies (*Drosophila melanogaster*) from California were tested for their genotype at the alcohol dehydrogenase locus using starch-gel electrophoresis. Two alleles were present, *S* and *F*, for slow and fast migration, respectively. The following results were noted: *SS*, sixty-six; *SF*, twenty; *FF*, fourteen. What are the allelic and genotypic frequencies in this population?

Answer: Since the sample size is one hundred, the proportions of the three genotypes, *SS*, *SF*, and *FF*, are 0.66, 0.20, and 0.14, respectively. We can calculate allelic frequencies directly from these genotypes, remembering that the frequency of an allele is the frequency of its homozygote plus half the frequency of the heterozygote, or

$$p = f(S) = f(SS) + (1/2)f(SF)$$
$$= 0.66 + (1/2)(0.20) = 0.76$$

$$q = f(F) = f(FF) + (1/2)f(SF)$$
$$= 0.14 + (1/2)(0.20) = 0.24$$

Alternatively, we could get allelic frequencies by counting alleles. Thus,

$$p = \frac{2 \times \text{number of } SS + \text{number of } SF}{2 \times \text{total number}} = \frac{2(66) + 20}{2(100)} = \frac{152}{200} = 0.76$$

$$p = \frac{2 \times \text{number of } FF + \text{number of } SF}{2 \times \text{total number}} = \frac{2(14) + 20}{2(100)} = \frac{48}{200} = 0.24$$

PROBLEM 2: Is the population described in problem 1 in Hardy-Weinberg equilibrium?

Answer: We can determine whether the numbers of the three genotypes (*SS*, *SF*, and *FF*) are in Hardy-Weinberg proportions through the chi-square statistical test. The observed numbers of the three genotypes are sixty-six, twenty, and fourteen, respectively. Using allelic frequencies of $p = f(S) = 0.76$ and $q = f(F) = 0.24$, we expect p^2, $2pq$, and q^2, respectively, of the three genotypes. That is,

$$p^2 = (0.76)^2 = 0.5776, \text{ or } 57.76 \text{ in } 100$$

$$2pq = 2(0.76)(0.24) = 0.3648, \text{ or } 36.48 \text{ in } 100$$

$$q^2 = (0.24)^2 = 0.0576, \text{ or } 5.76 \text{ in } 100$$

We can now set up a chi-square table as follows:

	SS	*SF*	*FF*	Total
Observed Numbers	66	20	14	100
Expected Proportions	p^2	$2pq$	q^2	1.0
	(0.5776)	(0.3648)	(0.0576)	1.0
Expected Numbers	57.76	36.48	5.76	100
$\chi^2 = (O - E)^2/E$	1.176	7.445	11.788	20.408

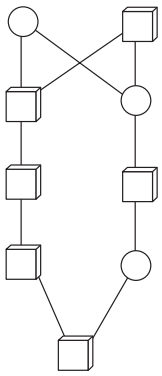
The critical chi-square value (0.05 at one degree of freedom) is 3.841, so we reject the hypothesis that this population is in Hardy-Weinberg proportions. From inspection of the table, it appears that there are too few heterozygotes and too many homozygotes, indicating that inbreeding could be the cause of the discrepancy.

PROBLEM 3: Convert the pedigree in figure 19.2 into a path diagram, and determine the inbreeding coefficient of the inbred individual, assuming that the common ancestors are not themselves inbred.

Answer: There are two paths (see the figure), each with seven ancestors. Thus, the inbreeding coefficient is

$$F = \Sigma[(1/2)^n (1 + F_j)] = 2(1/2)^7 = 0.016$$

Hence, the inbreeding coefficient is 0.016; about 1.6% of the loci of the inbred individual are autozygous.



EXERCISES AND PROBLEMS*

HARDY-WEINBERG EQUILIBRIUM

1. One hundred persons from a small town in Pennsylvania were tested for their MN blood types. Is the population they represent in Hardy-Weinberg proportions? The genotypic data are: *MM*, forty-one; *MN*, thirty-eight; and *NN*, twenty-one.
2. From the following two sets of data, calculate allelic and genotypic frequencies, and determine whether the populations are in Hardy-Weinberg proportions. Do a statistical test if one is appropriate.
   a. Allele *A* is dominant to *a*; *A*-, 91; *aa*, 9.
   b. Electrophoretic alleles *F* and *S* are codominant at the malate dehydrogenase locus in *Drosophila;* *FF*, 137; *FS*, 196; *SS*, 87.
3. The dominant ability to taste PTC comes from the allele *T*. Among a sample of 215 individuals from a population in Vancouver, 150 could detect the taste of PTC, and 65 could not. Calculate the allelic frequencies of *T* and *t*. Is the population in Hardy-Weinberg proportions?

* Answers to selected exercises and problems are on page A-21.

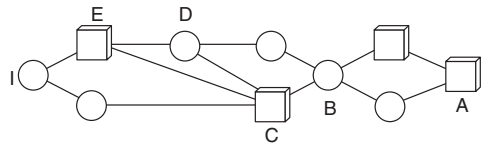
4. The frequency of children homozygous for the recessive allele for cystic fibrosis is about one in twenty-five hundred. What is the percentage of heterozygotes in the population?
5. PTC tasting is dominant in human beings.
   a. Should most human populations be heading toward a 3:1 ratio of tasters to nontasters? Explain.
   b. Confronted with a population sample of human beings of unknown origin, would you expect more or less than half the sample to be tasters?
6. Graph the relationship of the proportions of genotype (*AA, Aa, aa*) as allelic frequencies change.
7. A particular recessive disorder is present in one in ten thousand individuals. If the population is in Hardy-Weinberg equilibrium, what are the frequencies of the two alleles?
8. What allelic frequency will generate twice as many recessive homozygotes as heterozygotes?
9. Assume brown eye color is the result of a dominant allele at one locus. Attack or defend mathematically the following statement: With time, the frequency of brown-eyed individuals will increase, until about three out of four individuals are brown-eyed.
10. A particular human population has five hundred *MM* individuals, three hundred *MN*, and seven hundred *NN*. Calculate the allelic frequencies, and determine whether the population is in Hardy-Weinberg equilibrium.
11. Assume random mating occurs among the individuals of the population described in problem 10. What will be the frequency of each type of individual in the next generation?
12. On a small island, 235 mating individuals are all true-breeding for brown eyes. An epidemic eliminates all the population except ten young women, two young men, and four older (postmenopausal) women. A boatload of foreigners arrives; the foreign population consists of six heterozygous brown-eyed females, four homozygous brown-eyed males, and ten blue-eyed males. Assuming that one locus controls eye color, that mating is random with respect to eye color, and that each male and female capable of breeding does so, calculate the genotypic frequencies of their offspring.
13. In a given population, only the I^A and I^B alleles are present in the ABO system; there are no individuals with type O blood or with *i* alleles. If two hundred people have type A blood, seventy-five have type AB blood, and twenty-five have type B blood, what are the allelic frequencies in this population?

EXTENSIONS OF HARDY-WEINBERG EQUILIBRIUM

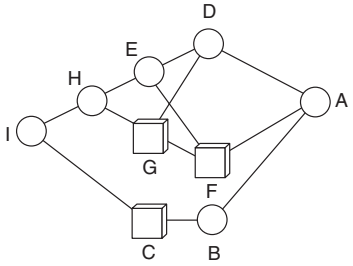
14. The following data are *ABO* phenotypes from a population sample of one hundred persons. Determine the frequencies of the three alleles: type A, seven; type B, seventy-two; type AB, twelve; type O, nine. What do you have to assume? Is the population in Hardy-Weinberg proportions?
15. How quickly and in what manner is Hardy-Weinberg equilibrium achieved under the following initial conditions (assuming a diploid, sexually reproducing population)?
   a. One locus, five alleles
   b. Two unlinked loci, two alleles each
16. A sample of fruit flies was testcrossed to determine the allelic arrangements of two linked loci in the gametes of that generation. With the following data, can you determine whether linkage equilibrium holds? Gametic arrangements are *AB*, fifty-eight; *ab*, eight; *Ab*, twelve; and *aB*, twenty-two.
17. In a large, randomly mating human population, the frequencies of the I^A, I^B, and *i* alleles are 0.7, 0.2, and 0.1, respectively. Calculate the expected frequencies for each blood type.
18. In a human population of one hundred people, seventeen have type A blood, seventeen have type B, two have type AB, and sixty-four have type O. If this population is in equilibrium, what are the allelic frequencies?

NONRANDOM MATING

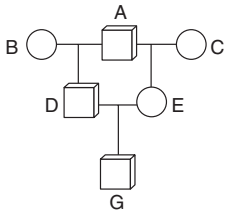
19. Under what circumstances is inbreeding deleterious?
20. What is the inbreeding coefficient of I in the following pedigree? Assume that the inbreeding coefficients of other members of the pedigree are zero unless other information tells you differently.



21. What is the inbreeding coefficient of individual I in this pedigree? $F_A = 0.01$; $F_B = 0.02$; $F_C = 0.02$.



22. The following is the pedigree of an offspring produced by the mating of half siblings. Individuals A and C have inbreeding coefficients of 0.2; all others are zero. Convert the pedigree to a path diagram and determine the inbreeding coefficient of individual G.



23. Given the population in Exercises and Problems problem 1, what is its inbreeding coefficient?

24. In a sample of one hundred people, are fourteen *MM*, thirty-two *MN*, and fifty-four *NN* individuals. Calculate the inbreeding coefficient.

25. If, in a population with two alleles at an autosomal locus, $p = 0.8$, $q = 0.2$, and the frequency of heterozygotes is 0.20, what is the inbreeding coefficient?

CRITICAL THINKING QUESTIONS

1. Prove that two generations are needed for the establishment of Hardy-Weinberg proportions when an autosomal locus with two alleles in a sexually reproducing species has frequencies of the two alleles that differ in the two sexes.

2. What might the ramifications to conservation efforts be of zoos maintaining captive breeding programs for rare and endangered species?

Suggested Readings for chapter 19 are on page B-19.