

2

NEUTRAL EVOLUTION IN ONE- AND TWO-LOCUS SYSTEMS

Draft Version 10 November 2008

Although the majority of research in evolutionary biology is focused on issues related to natural selection, the rigor of all such analyses depends critically on our understanding of expected patterns of evolution in the absence of selection. The simple reasoning here is that if we are to have much confidence in any adaptive argument, it ought to be possible to firmly reject a simpler neutral hypothesis. Thus, prior to exploring various aspects of adaptive evolution, we first embark on a broad overview of neutral models of evolution. The theory underlying such models brings us into immediate contact with the issue of **genetic drift**, the random fluctuation in allele frequencies that necessarily results from the sampling of a finite number of gametes each generation. The magnitude of such deviations increases with decreasing population size, and combined with the input of new alleles by mutation, ensures that populations will evolve even in the absence of selection (Kimura 1983).

Consider, for example, a single heterozygous **Bb** parent that produces two progeny. There is a 50% probability that one offspring will inherit the **A** and the other the **a** allele, in which case no net change in allele frequency has been transmitted from parent to offspring. However, there is also a 50% probability that both offspring will inherit the same allele. In extremely large populations, these random changes resulting from gamete sampling tend to average out, leaving the allele frequency in the offspring population very close to that in the parental generation. However, on long enough time scales, the cumulative effects of even small single-generation changes in allele frequencies can become quite pronounced. As will be seen below, if the time scale of interest (t , in generations) is much less than the average number of reproductive adults in the population (N), then random fluctuations in allele frequencies can usually be ignored. This justifies the assumption of an effectively infinite population size as a good first approximation in many applications of population genetics. However, for situations in which t is on the order of N or larger, evolution can no longer be viewed as a deterministic process. Rather, any

observed evolutionary change must be viewed as one realization of many possible outcomes.

In the following pages, it will be shown that although finite population size induces stochastic evolutionary change, random genetic drift has several predictable effects. First, even if mating is completely random, there will still be some long-term trend toward matings among relatives. Because all members of a population must ultimately descend from a narrow ancestral base, the smaller the population size, the greater this tendency will be. Thus, in a tiny **dioecious** (separate-sex) population with a stable adult number of two, all matings must be between full sibs, even though the reproductive pair itself may be a random draw from a larger progeny pool in the preceding generation. It follows that the genetic consequences of finite population size must be similar to those of inbreeding – the average homozygosity at a locus is expected to increase, although as noted below, this can be offset in part by replenishment from mutationally derived variation. Second, gamete sampling causes allele frequencies to gradually drift toward zero or one, with the probability of ultimate fixation of any particular allele being equal to the initial frequency in the absence of selection. Third, subdivision of a population into isolated demes results in allele-frequency divergence among demes. The greater the degree of isolation of the subgroups, the more pronounced this differentiation will be.

This chapter and the following two are concerned with establishing a more formal theory for these ideas. We first consider matters of one- and two-locus evolution in the context of a population with an idealized mating system and no influence from selective forces. As will be seen in subsequent chapters, such models can often be of great utility even for situations in which selection is operating, provided the forces of selection are weaker than those associated with random genetic drift. In addition, such theory provides the underlying logic for the development of molecular marker-based methods for estimating the power of mutation, recombination, and random genetic drift, the subject of Chapter 4. Chapter 3 provides a critical link between Chapters 2 and 4 by demonstrating how the basic results derived under the assumption of an ideal random mating population can be extended to a wide variety of alternative reproductive systems and population structures. In subsequent chapters, these one- and two-locus results will be used to develop neutral models for the evolution of quantitative traits.

THE WRIGHT-FISHER MODEL

Because the number of possible types of population structure is literally infinite (involving, for example, local inbreeding, geographic subdivision, age-specific mortality and fecundity, and temporal variation in deme size), it is impossible for us to consider the dynamics of neutral alleles in a fully general sense. Instead, we will focus initially on single finite populations within which mating is random. Even this simple structure admits to many possible variants, depending for example, on whether there are separate sexes, whether there is variation in family size, and whether generations are overlapping. All of these matters will be taken up in Chapter 3, where it will become clear that most of the results in the current chapter still hold with

an appropriate redefinition of the concept of population size.

Most population geneticists adhere to the **Wright-Fisher model**, the roots of which trace to Fisher (1922) and Wright (1931). Here we assume a diploid population with a fixed number (N) of **monoecious** (hermaphroditic) adults, complete random mating (including the possibility of self-fertilization), and discrete generations. The gamete pool produced by the adults is assumed to be effectively infinite, such that the $2N$ gametes that actually contribute to the next generation can be effectively viewed as being sampled with replacement.

Consider a locus with two alleles, **B** and **b**, neither having a selective advantage with respect to the other. If there are i copies of allele **B** in generation t , the probability that the number in generation $t + 1$ is equal to j follows the binomial sampling distribution,

$$P_{ij} = \binom{2N}{j} (i/2N)^j [1 - (i/2N)]^{2N-j} \quad (2.1)$$

This expression holds for all possible values of $i, j = 0, 1, \dots, 2N$. Letting \mathbf{P} be the $(2N + 1) \times (2N + 1)$ matrix of these coefficients, the entire probability distribution of the frequency of allele **B** can be expressed succinctly by the equation,

$$\mathbf{X}(t + 1) = \mathbf{P}\mathbf{X}(t) \quad (2.2a)$$

where the elements of the column vector $\mathbf{X}(t)$ are the probabilities that the allele is present in $i = 0, 1, \dots, 2N$ copies in generation t . If the transition matrix \mathbf{P} remains constant from generation to generation, as it does under the assumptions given above, Equation 2.2a generalizes to

$$\mathbf{X}(t) = \mathbf{P}^t \mathbf{X}(0) \quad (2.2b)$$

When considering a single population starting with allele frequency $i/2N$, all of the entries in the initial vector $\mathbf{X}(0)$ are equal to zero, except the $(i + 1)$ th element, which is equal to one. Equation 2.2b then yields the historical development of the probability distribution of the allele frequency over time. That is, the elements of $\mathbf{X}(t)$ denote the frequencies of hypothetical replicate populations at time t that are expected to have allele frequency $i/2N$, $i = 0, 1, \dots, 2N$. The first ($i = 0$) and final ($i = 2N$) elements of $\mathbf{X}(t)$ are of special interest, as they are **absorbing states** – once an allele becomes lost (frequency = 0.0) or fixed (frequency = 1.0) in a population, it remains at that state indefinitely (barring reintroduction via mutation or migration). Thus, as Equation 2.2b is iterated, all of the interior elements of $\mathbf{X}(t)$ eventually converge on zero, and the sum $X_0(t) + X_{2N}(t)$ converges to one. The ultimate probability of fixation of allele **B** is given by $X_{2N}(\infty)$, whereas the ultimate probability of loss of allele **B** (or equivalently, the ultimate probability of fixation of allele **b**) is given by $X_0(\infty)$.

From the elements of $\mathbf{X}(t)$, it is straightforward to compute the expected allele frequency, the variance of allele frequencies among replicate populations, probabilities of fixation by generation t , etc. However, although this approach is exact, because it becomes unwieldy when N is large, many useful approximations have been developed. Some of these approaches will be covered below, with a most powerful alternative approach (the diffusion approximation) being covered extensively in Appendix 1.

The basic **transition-matrix approach** outlined above forms the foundation of many models of finite populations with discrete generations, the primary modification being in the form of the elements of the matrix (Ewens 2004). Nevertheless, it should be noted that the Wright-Fisher model is just one of many possible conceptual frameworks for approximating a randomly mating population. For example, Moran (1962) developed a treatment whereby at each point in time a single random individual is chosen to reproduce and then a single random individual is chosen to die. Because allele frequencies can change by only a single step during each time interval under the **Moran model**, it is more analytically tractable than the Wright-Fisher model, although it is also restricted to haploid populations. Finally, we reemphasize that although the Wright-Fisher model makes some rather narrow assumptions about the reproductive features of the population, extensions to alternative systems of mating can be obtained by utilizing the concept of effective population size (Chapter 3).

Example 1. Consider an initially heterozygous individual **Bb** in a self-fertilizing line maintained by single-progeny descent. In this case, $N = 1$, and the only three possible subsequent allele-frequency states in the population are 0, 1, or 2 **B** alleles. Denoting the initial state of the population by $\mathbf{X}^T(0) = [0, 1, 0]$, the probability that the population is in states 0, 1, or 2 at some future generation t is given by Equation 2.2b with

$$\mathbf{P} = \begin{pmatrix} 1 & 0.25 & 0 \\ 0 & 0.50 & 0 \\ 0 & 0.25 & 1 \end{pmatrix}$$

Although the numerical values for the elements of \mathbf{P} can be obtained directly from Equation 2.1, for this simple example they can also be arrived at intuitively. For example, the elements in the first column of \mathbf{P} denote the probabilities that the population will be in states $j = 0, 1, 2$ in generation $t + 1$ given that it is in state 0 in generation t . The only nonzero element in this column is $P_{00} = 1$ because the 0th state is absorbing, i.e., once the population enters this state, it remains there indefinitely. The following table illustrates the progression of the elements of $\mathbf{X}(t)$ over time.

t	$X_0(t)$	$X_1(t)$	$X_2(t)$
0	0.00000	1.00000	0.00000
1	0.25000	0.50000	0.25000
2	0.37500	0.25000	0.37500
3	0.43750	0.12500	0.43750
4	0.46875	0.06250	0.46875
5	0.48438	0.03125	0.48438
...
10	0.49951	0.00098	0.49951
...
15	0.49998	0.00003	0.49998
...
∞	0.50000	0.00000	0.50000

The first and last elements of $\mathbf{X}^T(t)$ respectively denote the probabilities that the line has become fixed for the **B** or the **b** allele by time t . Thus, with equal probability, the line eventually becomes completely monomorphic for either the **B** or the **b** allele. When there are no directional forces operating on a locus, the two probabilities of fixation are equal to the initial frequencies of the respective alleles.

LOSS OF HETEROZYGOSITY BY RANDOM GENETIC DRIFT

The stochastic effects of random genetic drift can be succinctly defined in terms of the sampling variance of an allele frequency. Consider a large pool of gametes, a fraction p of which carry the **B** allele, and let $2N$ gametes be randomly drawn to produce a new generation of N individuals. Defining the expected frequencies of genotypes **BB**, **Bb**, and **bb** in the progeny generation by the Hardy-Weinberg proportions p^2 , $2p(1-p)$, and $(1-p)^2$, the expected number of **B** alleles contained in a random offspring is simply $[2 \cdot p^2] + [1 \cdot 2p(1-p)] + [0 \cdot (1-p)^2] = 2p$. However, the expected square of the number of **B** alleles carried is $[2^2 \cdot p^2] + [2^2 \cdot p(1-p)] + [0^2 \cdot (1-p)^2] = 2p(1+p)$. Thus, the variance of the number of **B** alleles carried by individuals is $2p(1+p) - (2p)^2 = 2p(1-p)$, while the variance of the total number of **B** alleles carried by the N offspring is N times this, $2Np(1-p)$. Because the frequency of allele **B** is the number of copies divided by $2N$, the sampling variance of the frequency is $2Np(1-p)/(2N)^2 = p(1-p)/(2N)$, is directly proportional to the heterozygosity and inversely proportional to the population size. The expression $p(1-p)/(2N)$ defines the dispersion in allele frequency resulting from a single generation of gamete sampling, conditional on allele frequency p in the parental population.

In the absence of any counteracting evolutionary forces, the dispersive effects of genetic drift will continue each generation, leading to a progressive erosion of population-level heterozygosity, until all loci have eventually become fixed for just a single allele. To evaluate the long-term impact of finite population size on the expected heterozygosity of a locus, we make use of the properties of the inbreeding coefficient, f , which denotes the probability that two alleles at a locus in an individual are identical by descent (LW Chapter 7). There is always a small chance that uniting gametes will derive from related individuals, even in a randomly mating population. For example, in a monoecious population containing only two individuals, there are only four genes residing at each locus, so the probability that one gamete will randomly unite with another containing a direct descendant of the same gene is $1/4$. With four individuals, there are eight genes, and this probability becomes $1/8$. Thus, under the idealized Wright-Fisher model, the probability that two direct copies of any parental gene will randomly unite in an offspring is $1/(2N)$. Barring a rare mutation, all such offspring are homozygotes.

Although the quantity $1/(2N)$ may be thought of as the new inbreeding that is incurred each generation, this does not fully describe the build-up of homozygosity in a population. For even if uniting gametes do not carry genes that are direct

copies of a parental gene, they may still be identical by descent through inbreeding in a previous generation. Under random mating, the probability of the latter event is simply the inbreeding coefficient of the parental generation. Thus, because the probability of drawing genes that are not direct copies of the same parental gene is $[1 - (1/2N)]$, the expected inbreeding coefficient in generation t is

$$f_t = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right) f_{t-1} \quad (2.3)$$

Subtracting both sides from one yields the recursion formula

$$(1 - f_t) = \left(1 - \frac{1}{2N}\right) (1 - f_{t-1}) \quad (2.4a)$$

which generalizes to

$$(1 - f_t) = \left(1 - \frac{1}{2N}\right)^t (1 - f_0) \quad (2.4b)$$

and finally to

$$(1 - f_t) = \left(1 - \frac{1}{2N}\right)^t \quad (2.4c)$$

if we assume a noninbred base population ($f_0 = 0$). Again, we see the central role that population size plays in the dynamics of genetic variation. As $t \rightarrow \infty$, the fraction of the population that is not inbred, $1 - f_t$, approaches zero at a rate that is inversely proportional to N . This rate of decay of heterozygosity of $1/(2N)$ was first obtained by Wright (1931). It may be a source of encouragement to the non-mathematically inclined that Fisher (1922), using a rather different approach, obtained the wrong answer.

To see the connection between the inbreeding coefficient and the expected heterozygosity in a population, consider a diallelic locus for which the heterozygosity in the base population is $2p(1-p)$. In the descendent population with inbreeding coefficient f , individuals can only be heterozygotes if they carry alleles that are not identical by descent, the probability of which is $(1-f)$. If two alleles are not identical by descent, they must have been acquired independently, so the probability that a genotype containing a pair of such alleles is a heterozygote is $2p(1-p)$. Thus, the expected heterozygosity of a population with inbreeding coefficient f and initial allele frequency p is $2p(1-p)(1-f)$, showing that relative to the base population, the fractional reduction in heterozygosity is equal to f . Because this argument applies regardless of the initial heterozygosity (and regardless of the number of segregating alleles), Equation 2.4c may be rewritten to describe the expected population heterozygosity at time t ,

$$H_t = H_0 \left(1 - \frac{1}{2N}\right)^t \quad (2.5)$$

The time course for the loss of heterozygosity can be clarified by using an exponential approximation to Equation 2.5. Because $(1-x)^t \simeq e^{-xt}$ for $|x| \ll 1$, for N greater than 10 or so

$$H_t = H_0 e^{-t/(2N)} \quad (2.6a)$$

To find the time to reach a certain reduction in heterozygosity, note that

$$t = -2N \ln(H_t/H_0) \quad (2.6b)$$

which shows that the heterozygosity is reduced to half of its initial value in $\sim 1.4N$ generations and to 5% of H_0 in $\sim 6N$ generations. Thus, a population twice the size of another requires twice the number of generations to reach the same expected state.

With a temporally varying population size, Equation 2.6a becomes

$$H_t = H_0 \prod_{i=1}^t \left(1 - \frac{1}{2N_i}\right) \simeq H_0 \exp \left[- \sum_{i=1}^t 1/(2N_i) \right] \quad (2.7)$$

where the \prod sign denotes a product of terms. This expression illustrates an important point. Because each of the generation-specific terms, $[1 - (1/2N_i)]$, is necessarily less than one, an expansion of population size can only reduce the rate of erosion of heterozygosity; it cannot eliminate it.

One significant limitation of the preceding expressions is that they only provide information on the behavior of the average heterozygosity in a population. In reality, fluctuations in allele frequencies resulting from random genetic drift also ensure that variation in heterozygosity will arise among loci that start in the same state. In a finite population of size N , the heterozygosity of a diallelic locus can take on $N + 1$ discrete values, $0, 2(1/2N)[1 - (1/2N)], \dots, 2(N/2N)[1 - (N/2N)]$. Using the transition-matrix approach (Equations 2.2a,b), one can obtain the exact probability distribution of heterozygosity for a locus starting with allele frequency $i/2N$, using the fact that $X_j(t) + X_{2N-j}(t)$ is the probability that the population has heterozygosity $2(j/2N)[1 - (j/2N)]$. However, as noted above, this approach becomes computationally difficult as N becomes large. An alternative approach utilizes a remarkable achievement by Kimura (1955), who used diffusion theory (Appendix 1) to obtain an analytical expression for the probability density of allele frequency at time t , given the starting value p_0 ,

$$\begin{aligned} \varphi(p_t|p_0) &= p_0(1-p_0) \sum_{i=1}^{\infty} i(2i+1)(i+1) \cdot \\ & F(1-i, i+2, 2, p_0) \cdot F(1-i, i+2, 2, p_t) \cdot e^{-i(i+1)t/(4N)} \end{aligned} \quad (2.8)$$

where $F(1-i, i+2, 2, p_0)$ and $F(1-i, i+2, 2, p_t)$ are specific variants of the hypergeometric function (see Abramowitz and Stegun 1972; Equation 15.1.1). Using this expression, $[\varphi(p_t|p_0) + \varphi(1-p_t|p_0)]$ is the probability of heterozygosity $2p_t(1-p_t)$ at time t . We will make more use of Equation 2.8 in the next sections, illustrating in particular its implications for the dispersion of allele frequencies among isolated populations.

PROBABILITIES AND TIMES TO FIXATION OR LOSS

Equation 2.8 denotes the probability density of allele frequency p given that the population is still polymorphic. Thus,

$$\Omega(p_0, t) = \int_0^1 \varphi(p_t|p_0) dp_t \quad (2.9a)$$

is the probability that both alleles are still present in generation t . The probability that an allele with initial frequency p_0 has been fixed by generation t is

$$p_f(p_0, t) = p_0 + p_0(1 - p_0) \sum_{i=1}^{\infty} (2i + 1)(-1)^i \cdot F(1 - i, i + 2, 2, p_0) \cdot e^{-i(i+1)t/4N} \quad (2.9b)$$

whereas the probability of loss of the allele, $p_l(p_0, t)$, is given by Equation 2.9b with $(1 - p_0)$ exchanged for p_0 (Kimura 1955). Summing up,

$$\Omega(p_0, t) + p_f(p_0, t) + p_l(p_0, t) = 1 \quad (2.10)$$

As can be seen from the negative exponential terms in the previous expressions, as $t \rightarrow \infty$, $\Omega(p_0, t) \rightarrow 0$, $p_f(p_0, t) \rightarrow p_0$, and $p_l(p_0, t) \rightarrow (1 - p_0)$. Thus, under neutrality, the probability that the population ultimately becomes fixed for an allele is simply equal to the initial allele frequency. It follows that averaging over a very large number of replicate populations, the allele frequency is expected to remain constant at p_0 , with p_0 replicates ultimately having allele frequency 1 and $1 - p_0$ having allele frequency 0.

An issue of special interest is the mean time until an allele is absorbed into either state $p = 0$ or $p = 1$. Using diffusion theory, Kimura and Ohta (1969) obtained expressions for both quantities, and Kimura (1970) presented a description of the entire probability distributions for absorption times. The following example uses a somewhat simpler approach to arrive at results identical to those of Kimura and Ohta (1969), and provides yet another illustration of how the effects of random genetic drift scale with population size.

Example 2. Ewens (2004) used the following line of reasoning to derive the expected time to absorption of a neutral allele under the Wright-Fisher model. Letting δp denote the change in allele frequency in one unit of time, the mean time to absorption for an allele with frequency p may be rewritten as

$$\bar{t}_a(p) = E[\bar{t}_a(p + \delta p)] + 1$$

In words, this expression states that, starting at frequency p , the mean absorption time is the mean absorption time one time unit later when the allele frequency is $p + \delta p$, plus one. Approximating the expectation with the first three terms of a Taylor expansion (see LW Equation A1.2),

$$\bar{t}_a(p) \simeq \bar{t}_a(p) + E(\delta p) \frac{\partial \bar{t}_a(p)}{\partial p} + E[(\delta p)^2] \frac{\partial^2 \bar{t}_a(p)}{2\partial p^2} + 1$$

Under neutrality, the expected change in allele frequency is $E(\delta p) = 0$, and as derived above, the expected variance in allele-frequency change is $E[(\delta p)^2] = p(1 - p)/(2N)$. The preceding expression then reduces to

$$\frac{\partial^2 \bar{t}_a(p)}{\partial p^2} \simeq -\frac{4N}{p(1 - p)}$$

Performing a double integration with respect to p leads to the solution

$$\bar{t}_a(p_0) \simeq -4N[p_0 \ln(p_0) + (1 - p_0) \ln(1 - p_0)] \quad (2.11a)$$

which is the mean time until an allele with initial frequency p_0 is either lost or fixed in a population.

A similar approach can be used to estimate the mean time to fixation for those alleles that do indeed become fixed, $\bar{t}_f(p_0)$. The essential modification here is that in estimating $\bar{t}_f(p_0)$, $E(\delta p)$ is no longer equal to zero, because in order for an allele to become fixed, at least one copy must be produced each generation. That is, in the case of conditional fixation, of the $2N$ genes drawn each generation, one is definitely a **B** allele, whereas the remaining $2N - 1$ genes can be viewed as random, leading to $E(\delta p) = \{(1/2N) + [1 - (1/2N)]p\} - p = (1 - p)/(2N)$. Similarly, because the states of only $2N - 1$ genes are random, $E[(\delta p)^2] = (2N - 1)p(1 - p)/(2N)^2$. Unless N is very small, the approximation $E[(\delta p)^2] = p(1 - p)/(2N)$ still holds quite well, and following the procedures utilized above, we then have

$$\left(\frac{2}{p}\right) \frac{\partial \bar{t}_f(p)}{\partial p} + \frac{\partial^2 \bar{t}_f(p)}{\partial p^2} \simeq -\frac{4N}{p(1 - p)}$$

The solution of this first-order linear differential equation requires several steps, which we omit. The final result is

$$\bar{t}_f(p_0) \simeq -\frac{4N(1 - p_0) \ln(1 - p_0)}{p_0} \quad (2.11b)$$

The mean time to loss of an allele conditional upon loss is identical to the previous expression, but with $(1 - p_0)$ interchanged with p_0 ,

$$\bar{t}_l(p_0) \simeq -\frac{4Np_0 \ln(p_0)}{1 - p_0} \quad (2.11c)$$

Finally, because the probability of ultimate fixation of an allele is equal to its initial frequency (p_0) and the probability of ultimate loss is $(1 - p_0)$, it follows that

$$\bar{t}_a(p_0) = p_0 \bar{t}_f(p_0) = (1 - p_0) \bar{t}_l(p_0)$$

Example 8 in Appendix 1 shows how diffusion theory can be used to obtain these same results.

ALLELE-FREQUENCY DIVERGENCE AMONG POPULATIONS

A natural consequence of allele-frequency drift within populations is the divergence of isolated replicate populations. Suppose a monoecious base population with allele frequency p_0 is suddenly split into several completely isolated subpopulations, each of size N , with random mating within each subpopulation and an absence of selection, migration, and mutation. The variance in allele frequency among subpopulations in generation t is

$$\sigma_p^2(t) = E(p_t^2) - E^2(p_t)$$

Adding and subtracting $E(p_t)$,

$$\begin{aligned}\sigma_p^2(t) &= [E(p_t) - E^2(p_t)] + [E(p_t^2) - E(p_t)] \\ &= E(p_t)[1 - E(p_t)] - E[p_t(1 - p_t)]\end{aligned}$$

Because there are no systematic forces causing the allele frequency to increase or decrease, $E(p_t) = p_0$, and the first quantity on the right is $p_0(1 - p_0)$. The quantity $E[p_t(1 - p_t)]$ is half the expected heterozygosity in a population in generation t . Thus, substituting Equation 2.5,

$$\sigma_p^2(t) = p_0(1 - p_0) \left[1 - \left(1 - \frac{1}{2N} \right)^t \right] \quad (2.12a)$$

which is well approximated by

$$\sigma_p^2(t) \simeq p_0(1 - p_0)(1 - e^{-t/2N}) \quad (2.12b)$$

for $N > 10$. This shows that the among-population variance asymptotically approaches $p_0(1 - p_0)$.

Equations 2.12a,b do not fully describe the pattern of development of differences among lines. For example, they yield little insight into the actual form of the distribution of population allele frequencies, and they do not specify the probability of fixation of alleles. However, all of this information is contained in the formulations presented above on the probability distribution of allele frequencies within populations. For example, the transition-matrix approach (Equations 2.2a,b) and the diffusion approximation (Equation 2.8) yield the expected temporal dynamics of the distribution of allele frequencies in different replicate populations, all starting from an identical frequency, p_0 (Figure 2.1). With increasing time, the total area under the curves declines and the distributions become flatter, because the proportions of populations that have experienced gene fixation or loss (i.e., have arrived at absorbing states) are not included. Moreover, the expected distribution is a function of t/N generations, as can also be seen from the exponential terms in Equation 2.8. As should be clear by now, this scaling of the temporal dynamics of random genetic drift to the reciprocal of population size is a natural consequence of the fact that the variance of allele-frequency change is inversely proportional to N .

-Insert Figure 2.1 Here-

Example 3. Because all populations are finite in size, the theory of random genetic drift is of central significance to all areas of population genetics. It may therefore come as a surprise that highly replicated experiments, examining the chance dynamics of within- and among-population allele-frequency change, are extremely rare. However, the results of one massive experiment nicely affirm the theoretical expectations outlined above. Starting with two homozygous lines of *Drosophila melanogaster*, one of which

was fixed for allele bw^{75} and the other for allele bw at the brown locus, Buri (1956) established 212 F_1 hybrid populations. For the following 19 generations, he randomly mated 8 males and 8 females within each population and monitored the changes in allele frequencies across the entire set of lines. This could be done visually because the genotype at the brown locus determines eye color: $bw^{75}bw^{75}$ = bright red-orange, $bw^{75}bw$ = deep red-brown, and $bwbw$ = white. (Two separate experiments, one with 107 and the other with 105 populations, were performed, but the results are so similar that they have been pooled in the following analysis).

Buri's observations can be used to verify that the bw^{75} and bw alleles are effectively neutral with respect to each other as follows (Figure 2.2, top). In the absence of selection, the expected frequency of the bw^{75} allele averaged over all populations should equal its initial frequency, 0.50, in all generations. Nevertheless, just as the frequency within any population is expected to deviate from 0.5 because of drift, so will the mean allele frequency in the total aggregate of populations. The sampling variance of the overall mean frequency is equal to the sum of the expected within- and among-population allele-frequency variances divided by the number of populations, 212. The latter quantity has already been defined in Equation 2.12, while the former is the expected binomial sampling variance divided by the sample size ($2N$), or $p_0(1 - p_0)[1 - (1/2N)]^t/(2N)$. The figure shows that although the frequency of the bw^{75} allele averaged over all populations increased to 0.525, it generally remained within two standard errors of the expectation under pure drift. The overall pattern of change in mean allele frequency is therefore compatible with the expectations for a neutral locus subject to random genetic drift.

The dynamics of the among-population divergence (Figure 2.3) are qualitatively very similar to the expected pattern illustrated in the left panel of Figure 2.1. As the population allele frequencies diverge, the initial bell-shaped distribution becomes flatter, and then begins to acquire a U-shape as populations that are fixed for the bw^{75} or bw alleles accumulate. Had the experiment been extended further in time, the distribution would have eventually consisted of only two classes, those fixed for bw^{75} and those fixed for bw , with roughly equal frequencies.

Despite the qualitative agreement with theoretical expectations, the rate of divergence illustrated in Figure 2.3 is somewhat greater than that expected for randomly mating populations of 16 individuals. However, this does not necessarily invalidate the theory outlined above, as it is possible that not all 16 potential parents reproduced each generation, and/or that the distribution of family sizes deviated from randomness. From the standpoint of genetic drift, either condition would cause the populations to behave genetically as though they were smaller than the actual size.

With the massive amount of data in Buri's experiment, it is possible to obtain an empirical estimate of this **effective population size** in the following way. Not including fixed classes, there are 31 possible allele frequencies in Buri's populations ($1/32$ to $31/32$), and each of these 31 classes was observed at various times in one or more of the 212 populations. Focusing on any one allele-frequency class, the sampling variance conditional on the initial allele frequency can then be calculated from the allele frequencies observed in the subsequent generation. Recall that the sampling variance of allele frequency from one generation to the next is $p(1-p)/(2N)$. The 31 points shown in Figure 2.4 provide an empirical description of this function, and an excellent fit is obtained if it is assumed that the average effective population size was $N \simeq 10.2$ rather than the idealized 16. In other words, the sampling variance of allele frequencies is in very close accord with that expected for an average ideal population of 10.2 randomly

mating individuals. Once this change in scale from $N = 16$ to $N = 10.2$ is taken into account, the erosion of average heterozygosity within populations and the build-up of among-population variance of allele frequencies are quite consistent with the theory outlined above (Figure 2.2, middle and bottom).

-Insert Figures 2.2, 2.3 and 2.4 Here-

HIGHER-ORDER ALLELE-FREQUENCY MOMENTS

In the previous sections, we evaluated the expected values of various population features under neutrality. However, in applying such expressions to empirical studies, it is important to keep in mind that the random sampling of allele frequencies across generations will cause the exact behavior of any particular population or group of populations to deviate from the expected pattern. Thus, there is a practical need for expressions for the variance of various population parameters resulting from genetic sampling, which in turn requires an understanding of the behavior of higher-order allele-frequency moments. For a population obeying the features of the idealized Wright-Fisher model, this can be accomplished by noting that the expected value of an allele-frequency moment in generation $t + 1$ conditional on allele frequency p_t in the previous generation is

$$E(p_{t+1}^k | p_t) = E[(p_t + \delta p)^k] \quad (2.13)$$

where δp denotes the change in allele frequency in the previous generation resulting from gamete sampling. For binomial sampling, expressions are available for all expected values of powers of δp , so Equation 2.13 can be solved recursively starting with the lower-order moments. For example, in the absence of any directional forces of selection, $E(\delta p) = 0$ and the expected frequency of an allele remains perpetually at its initial level (p_0),

$$E(p_t) = p_0 \quad (2.14a)$$

The second moment is obtained by noting that

$$E(p_{t+1}^2 | p_t) = E(p_t^2 + 2p_t\delta p + \delta p^2)$$

Because $E(\delta p) = 0$ and $E(\delta p^2) = p_t(1 - p_t)/(2N)$ under binomial sampling,

$$E(p_{t+1}^2 | p_t) = E\left(p_t^2 + \frac{p_t - p_t^2}{2N}\right)$$

Letting $\lambda_1 = 1 - (1/2N)$, this expression can be rearranged to give the recursion equation

$$E(p_{t+1}^2 | p_t) - p_0 = [E(p_t^2) - p_0]\lambda_1$$

the general solution of which is

$$E(p_t^2) = p_0 - [p_0(1 - p_0)]\lambda_1^t \quad (2.14b)$$

Using expectations for higher-order δp^k , expressions for additional moments can be acquired, two of which prove to be particularly useful (Crow and Kimura 1970),

$$E(p_t^3) = p_0 - \frac{3}{2}p_0(1 - p_0)\lambda_1^t - \frac{1}{2}p_0(1 - p_0)(2p_0 - 1)(\lambda_1\lambda_2)^t \quad (2.14c)$$

$$E(p_t^4) = p_0 - \frac{18N - 11}{10N - 6}p_0(1 - p_0)\lambda_1^t - p_0(1 - p_0)(2p_0 - 1)(\lambda_1\lambda_2)^t \\ + p_0(1 - p_0) \left(p_0(1 - p_0) - \frac{2N - 1}{10 - 6} \right) (\lambda_1\lambda_2\lambda_3)^t \quad (2.14d)$$

where $\lambda_2 = 1 - (2/2N)$, and $\lambda_3 = 1 - (3/2N)$. Modifications for these expressions for populations with separate sexes and 1:1 sex ratios are given by Lynch and Hill (1986).

Example 4. The preceding expressions can be used to derive the evolutionary variance of heterozygosity at a locus under the assumption of Hardy-Weinberg equilibrium. This quantity is most readily attained for the situation in which there are only two alleles segregating at the locus. Letting $H_t = 2p_t(1 - p_t)$ denote the heterozygosity at generation t , the expected variance of heterozygosity is

$$\sigma^2(H_t) = E\{[2p_t(1 - p_t)]^2\} - \{E[2p_t(1 - p_t)]\}^2 \\ = E(4p_t^2) - E(8p_t^3) + E(4p_t^4) - [E(2p_t - p_t^2)]^2$$

The solution is obtained by substituting Equations 2.14a-d for the expectations of allele-frequency moments, with further simplification made possible by using the approximations $\lambda_1 \simeq e^{-t/2N}$, $\lambda_1\lambda_2 \simeq e^{-3t/2N}$, and $\lambda_1\lambda_2\lambda_3 \simeq e^{-6t/2N}$,

$$\sigma^2(H_t) \simeq H_0 \left[\frac{2}{5}e^{-t/2N} + \left(H_0 - \frac{2}{5} \right) e^{-3t/2N} - H_0 e^{-t/2N} \right] \quad (2.15)$$

This quantity can be viewed as either the variance in heterozygosity that develops at a particular neutral locus among replicate populations starting from the same initial allele frequencies or as the variance in heterozygosity among a pool of loci within the same population with identical initial allele frequencies. As in the case of the expected heterozygosity, the temporal dynamics of the evolutionary variance of heterozygosity scale inversely with the size of the population. Moreover, the variation in heterozygosity resulting from genetic drift alone can be quite high, with the standard deviation always exceeding the expected heterozygosity beyond $t = 2N$ generations (Figure 2.5). With more than two alleles per locus, the preceding expressions would need to be modified to account for the negative evolutionary sampling covariance between different alleles at the locus.

-Insert Figure 2.5 Here-

LINKAGE DISEQUILIBRIUM

In the study of multilocus traits, we are naturally interested in combinations of alleles both within and between loci. If the alleles at two loci are independently distributed, then the expected frequency of a particular gamete type can be predicted from the products of the allele frequencies at the two loci. For example, with two alternative alleles (**A** and **a**) at one locus having frequencies p and $1 - p$, and those (**B** and **b**) at another locus having frequencies q and $1 - q$, the expected frequencies of gamete types **AB**, **Ab**, **aB**, and **ab** are pq , $p(1 - q)$, $(1 - p)q$, and $(1 - p)(1 - q)$ respectively. A natural measure of the deviation of the frequency of a gametic type from such expectations can be represented by the **coefficient of linkage disequilibrium**

$$D_{AB} = p(AB) - pq \quad (2.16)$$

where $p(AB)$ denotes the observed frequency of the **AB**th gamete. This definition has the useful feature of being equivalent to the covariance of the distribution of alleles **A** and **B** in the same gametes.

In the absence of selection, there will be no tendency for the alleles at different loci to be associated positively versus negatively, although historical forces such as migration may cause some such correlations. Letting D_0 denote an initial level of disequilibrium, c denote the frequency of recombination between loci, and $\lambda_1 = 1 - (1/2N)$, the expected disequilibrium resulting from the joint forces of recombination and gamete sampling is

$$\begin{aligned} E(D_t) &= [(1 - c)\lambda_1]^t D_0 \\ &\simeq D_0 e^{-(2Nc+1)t/(2N)} \end{aligned} \quad (2.17)$$

(Hill and Robertson 1966), showing that disequilibrium declines toward zero in the absence of any replenishing forces. In contrast, the variance of D can be quite substantial even when its expected value is zero. The problem can be evaluated by use of the following set of recursion equations for fourth-order moments of allele frequencies,

$$\begin{pmatrix} E[p(1-p)q(1-q)] \\ E[D(1-2p)(1-2q)] \\ E(D^2) \end{pmatrix}_{t+1} = \lambda_1 \cdot \begin{pmatrix} \lambda_1 & \lambda_1(1-c)/(2N) & 2(1-c)^2/(4N^2) \\ 0 & \lambda_2^2(1-c) & 4\lambda_2(1-c)^2/(2N) \\ 1/(2N) & \lambda_1(1-c)/(2N) & [\lambda_2^2 + (1/4N^2)](1-c)^2 \end{pmatrix} \cdot \begin{pmatrix} E[p(1-p)q(1-q)] \\ E[D(1-2p)(1-2q)] \\ E(D^2) \end{pmatrix}_t \quad (2.18)$$

(Hill and Robertson 1968). Using terms from above, the evolutionary variance of D associated with drift among replicated populations or among loci starting from the same allele frequencies is

$$\sigma^2(D) = E(D^2) - E^2(D) \quad (2.19)$$

Provided $D_0 = 0$, then $E(D_t) = 0$, and $\sigma^2(D_t) = E(D_t^2)$. Ohta and Kimura (1969; their Equations 20-25) obtained a closed-form solution to this expression, and also introduced the use of the dimensionless squared correlation coefficient

$$r_D^2 = \frac{E(D^2)}{E[p(1-p)q(1-q)]} \quad (2.20)$$

This standardized measure of linkage disequilibrium is equivalent to the square of the within-gamete correlation of allele frequencies at the two loci.

MUTATION-DRIFT EQUILIBRIUM

In the preceding pages, we were largely concerned with the dynamics of genetic variance owing to the effects of random genetic drift alone. Under this model, finite population size eventually results in the complete loss of genetic variation (and covariation) within populations, at which point all loci are fixed for ancestral alleles with probabilities equal to their initial frequencies. In reality, mutation will generally reintroduce variation at a low rate, which not only offsets some of the loss resulting from drift, but also ensures that neutral loci will continue to diverge among isolated populations. If the time scale of the problem under consideration is short ($t \ll 2N$) and the initial level of within-population variation is high (relative to the mutational rate of production of new heterozygosity per generation), the contribution from mutation will be negligible, and the preceding expressions will be quite adequate. However, for longer-term evolutionary issues, such as the maintenance of variation in natural populations and interspecific divergence, mutation cannot be ignored.

The incorporation of mutation into a neutral model of evolution is relatively straight-forward. Suppose there are k possible alleles at a locus, each with a mutation rate of u per generation. The dynamics of heterozygosity can then be obtained by recalling from above that the expected frequency of heterozygotes in generation $t + 1$ in the absence of mutation is $\lambda_1 H_t$, whereas the expected frequency of homozygotes is $1 - \lambda_1 H_t$. Following mutation, the heterozygous state will be retained if neither allele mutated, the probability of which is $(1 - 2u)$ ignoring the very small probability of double mutations to the same state, or if one of the alleles mutated to a different state than the other, the probability of which is $[2u(k - 2)/(k - 1)]$ assuming that all allelic types are equally mutationally exchangeable. On the other hand, homozygotes will be mutationally converted to heterozygotes at rate $2u$. Thus, the expected dynamics of heterozygosity can be expressed as

$$H_{t+1} = H_t \lambda_1 \left(1 - 2u + \frac{2u(k - 2)}{k - 1} \right) + 2u(1 - \lambda_1 H_t) \quad (2.21)$$

The equilibrium heterozygosity under drift-mutation balance, obtained by setting $H_{t+1} = H_t$, is

$$E(H) = \frac{4Nu}{1 + [4Nuk/(k - 1)]} \quad (2.22a)$$

a result first given by Kimura (1968). If a large number of alternative possible alleles ($k \gg 1$) is assumed, as is reasonable when the unit of analysis is an entire gene, Equation 2.22a reduces to

$$E(H) \simeq \frac{4Nu}{1 + 4Nu} \quad (2.22b)$$

which is equivalent to the result for the infinite-alleles model of Kimura and Crow (1964). On the other hand, if the unit of analysis is a nucleotide site, then $k = 4$, and

$$E(H) = \frac{12Nu}{3 + 16Nu} \quad (2.22c)$$

where u is now the mutation rate per nucleotide site. Equation 2.22a needs to be modified if not all alleles mutate at the same rate or are equally mutationally accessible (Kimura 1983; Nei and Kumar 2000), but provided the number of mutations entering the population per generation ($2Nu$) is $\ll 1$, then $\hat{H} \simeq 4Nu$ regardless of the model assumed. As $2Nu \rightarrow \infty$, the infinite-alleles model implies $E(H) \rightarrow 1.0$, whereas the $k = 4$ model implies $E(H) \rightarrow 0.75$. The latter result is a simple consequence of the predicted presence of four segregating alleles with equal frequencies (0.25) when the power of drift is overwhelmed by mutation. A very general treatment is given by Cockerham (1984), who considered the transient approach to equilibrium and allowed for unequal mutation rates.

As the above drift-mutation models play a central role in the neutral theory of molecular evolution (Kimura 1983), substantial attention has been given to additional informative details that are obscured by the summary statistic of average heterozygosity. For example, because of the stochastic nature of both mutation and drift, the allele frequencies at any neutral locus are expected to wander stochastically over time, with some loci being transiently fixed for one particular allele, and others experiencing the full spectrum of allele frequencies. Kimura (1968) obtained an expression analogous to Equation 2.8, which yields the complete probability distribution of the frequency of an allele under the symmetric mutation model described above, starting from an arbitrary allele frequency. Although this expression is quite complicated, a highly useful result is that regardless of the starting point, an equilibrium distribution of allele frequencies is eventually attained

$$\phi(p) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} (1 - p)^{\alpha-1} p^{\beta-1} \quad (2.23)$$

where $\alpha = 4Nu$, and $\beta = 4Nu/(k - 1)$. Equation 2.23 may be viewed as either the expected distribution of allele frequencies over all neutral loci within a single population in mutation-drift equilibrium or as the distribution of allele frequencies at a particular locus among replicate populations with identical N and u . Nei and Li (1976) present the theory necessary for predicting the approach of a nonequilibrium population to the equilibrium state.

The expected value of any function of population allele frequencies (e.g., homozygosity) can be obtained by simply integrating the function over the density distribution $\phi(p)$. However, because Equation 2.23 defines a beta distribution, many of its properties are already well known. For example, the expected allele frequency is

$$E(p) = \frac{\beta}{\alpha + \beta} = \frac{1}{k} \quad (2.24a)$$

and the variance of allele frequencies among replicates is

$$\sigma^2(p) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} = \frac{k - 1}{k^2[2Nuk(k - 1) + 1]} \quad (2.24b)$$

Expressions for the variance of heterozygosity for a population en route to equilibrium have been derived by Li and Nei (1975) and Lessard (1981), and at equilibrium

$$\sigma^2(H) = \frac{2\theta[1 + (\theta/l)]}{[1 + \theta + (\theta/l)]^2[2 + \theta + (\theta/l)][3 + \theta + (\theta/l)]} \quad (2.25)$$

where $l = k - 1$ (Stewart 1976).

Ohta and Kimura (1971) and Hill (1975) have obtained expressions for the expected values of the two-locus moments, analogous to those described above as Equation 2.18, for the infinite-allele model under stochastic drift-mutation equilibrium. Letting $\rho = 4Nc$, $\theta = 4Nu$, Hill's expressions reduce to

$$\begin{aligned} E[p(1-p)q(1-q)] &= M(22 + 13\rho + 32\theta + \rho^2 + 6\rho\theta + 8\theta^2) \\ E[D(1-2p)(1-2q)] &= 8M \\ E[D^2] &= M(10 + \rho + 4\theta) \end{aligned} \quad (2.26a)$$

where

$$M = \theta^2 / [(\theta + 1)(18 + 13\rho + 54\theta^2 + \rho^2 + 19\rho\theta + 40\theta^2 + 6\rho\theta^2 + 8\theta)], \quad (2.26b)$$

and the standardized linkage disequilibrium is given by

$$r_D^2 = \frac{10 + \rho + 4\theta}{22 + 13\rho + 32\theta + \rho^2 + 6\rho\theta + 8\theta^2} \quad (2.27a)$$

As will be seen in Chapter 4, θ is generally substantially smaller than one (in multicellular species, it is rarely larger than 0.10), whereas for nucleotide sites separated by hundreds of base pairs or more, ρ is generally greater than one. If these conditions are met,

$$r_D^2 \simeq \frac{10 + \rho}{22 + 13\rho + \rho^2} \quad (2.27b)$$

which asymptotically approaches $1/\rho$ for large ρ . An even more daunting problem is the procurement of expressions for the evolutionary variances of these statistics. As the measures in Equation 2.26 involve fourth-order moments of allele frequencies, within and between loci, their variances require expressions for moments up to the eighth order. Hill and Weir (1988) have tackled this problem.

THE GENEALOGICAL STRUCTURE OF A POPULATION

The preceding analyses suggest a few measurable summary statistics, such as average levels of heterozygosity and linkage disequilibrium, are defined by the processes of drift, mutation, and recombination in predictable ways. However, provided we retain our focus on neutral regions of the genome, it is possible to go quite a bit

further, even to the extent of predicting the expected features of the genealogical relationships between different sequences sampled within populations. The basic issues, laid out in detail by Kingman (1982a,b), relate directly to the sampling properties of descendent alleles.

Because all of the genes currently residing within a population are direct products of past processes of gametic sampling and mutation, they are all ultimately related in a genealogical sense. Thus, if one were to sample two alleles in a current population and then follow them back in time through ancestors, a point would eventually be reached at which both copies can be traced to a single copy in an ancestral individual, at which point the two alleles are said to have **coalesced**. A key principle is that the form of the expected gene genealogy for neutral genes, in particular the expected times to coalescence, is completely independent of the mutational process.

Here we consider a sample of n alleles drawn from a current population, assumed to obey all the properties of the idealized Wright-Fisher model. Focusing initially on just a single pair of randomly sampled alleles, we first consider the probability that both members of the pair are direct copies of a single allele in the preceding generation. As there are $2N$ gene copies in the population each generation, this probability is simply $1/(2N)$, whereas $\lambda_1 = [1 - (1/2N)]$ is the probability that coalescence occurs at some earlier generation. Conditional on coalescence not having occurred in generation one, the probability of coalescence two generations in the past is again equal to $1/(2N)$, yielding the unconditional probability of coalescence in generation two of $\lambda_1(1/2N)$. This simple rule can be generalized to give the probability of coalescence exactly t generations in the past,

$$P_c(t) = \lambda_1^{t-1}(1/2N) \quad (2.28)$$

which defines a geometric distribution, with the sum of $P_c(t)$ over the interval $t = 1$ to ∞ being equal to one. One simple related point is that the probability of a common ancestor between two sampled alleles in the last t generations is simply $1 - \lambda_1^t \simeq 1 - e^{-t/2N}$.

The average time to coalescence of two randomly sampled genes is simply

$$\bar{t}_c(2) = (1/2N) \sum_{t=0}^{\infty} (t+1) \lambda_1^t = 2N \quad (2.29)$$

Thus, the expected number of generations required for any two random alleles to trace back to an ancestral copy is simply equal to twice the population size (more precisely, twice the effective population size as defined in Chapter 3).

The logic used to derive this result can be readily extended to the genealogical structure of the entire sample of n gene copies. There are $n_p = n(n-1)/2$ possible pairs of n copies, each of which will or will not coalesce in the preceding generation with respective probabilities $1/(2N)$ and $[1 - (1/2N)]$. If the sample size is much smaller than the population size, the probability of coalescence for any pair in the sample in the preceding generation is simply the product $n_p/(2N)$. Thus, the probability distribution for the coalescence of one pair within a set of n sequences is

$$P_c(n_p, t) = [1 - (n_p/2N)]^{t-1} [n_p/(2N)] \quad (2.30)$$

The mean time to coalescence of the first pair is then $2N/n_p$ generations (as opposed to $2N$ generations with a single pair). Because at this point two copies have coalesced into one, the sample size has been reduced by one, and the mean time to coalescence of the next pair is found by resetting n_p to $(n-2)(n-1)/2$. This procedure can be followed recursively down to the final pair ($n_p = 1$), which again has an expected coalescence time of $2N$ generations (Figure 2.6).

The total expected genealogical depth of a sample is obtained by summing the expectations of each coalescence event, yielding

$$\bar{t}_c(n) = 4N \left(1 - \frac{1}{n} \right) \quad (2.31)$$

Thus, under neutrality, the expected time to the most recent common ancestor of all alleles residing at a locus is $4N$ generations. This is equivalent to the mean time to fixation of a neutral mutation, as can be verified by substituting $p_0 = 1/(2N)$ into Equation 2.11b. As an approximately exponentially distributed variable, the evolutionary variance of each individual coalescence time is equal to the square of its expected value, and the variance of the time to fixation is equal to the sum of the variances of the individual events.

-Insert Figure 2.6 Here-

Note that all of the preceding results in this section were derived without regard to any underlying genetic features of the sampled alleles. However, having determined the expected genealogical features of neutral gene sequences, it is straightforward to incorporate genetic issues, as mutations will arise randomly along the branches of the genealogy in numbers proportional to time. For example, given the average $2N$ generations separating randomly sampled gene copies, the average number of mutations separating such genes is $2 \cdot 2N \cdot u = 4Nu$, the two arising because each copy is $2N$ generations descendant from the common ancestor. Note that unlike the heterozygosity, which has a maximum value of 1.0, this measure of mutational divergence is unbounded, as it allows for the possibility of multiple mutations per copy.

Literature Cited

- Abramowitz, M., and I. A. Stegun. 1972. *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*. Dover Publications, Inc., New York, NY. [2]
- Buri, P. 1956. Gene frequency in small populations of mutant *Drosophila*. *Evolution* 10: 367–402. [2]
- Cockerham, C. C. 1984. Drift and mutation with a finite number of allelic states. *Proc. Natl. Acad. Sci. USA* 81: 530–534. [2]
- Crow, J. F., and M. Kimura. 1970. *An Introduction to Population Genetics Theory*. Harper and Row, Publ., New York, NY. [2]
- Ewens, W. J. 2004. *Mathematical Population Genetics*. 2nd Edition. Springer-Verlag, Berlin, Germany. [2]
- Fisher, R. A. 1922. On the dominance ratio. *Proc. Royal Soc. Edinburgh* 42: 321–341. [2]
- Hill, W. G. 1975. Linkage disequilibrium among multiple neutral alleles produced by mutation in finite population. *Theor. Pop. Biol.* 8: 117–126. [2]
- Hill, W. G., and A. Robertson. 1966. The effect of linkage on limits to artificial selection. *Genet. Res.* 8: 269–294. [2]
- Hill, W. G., and A. Robertson. 1968. Linkage disequilibrium in finite populations. *Theor. Appl. Genetics* 38: 226–231. [2]
- Hill, W. G., and B. S. Weir. 1988. Variances and covariances of squared linkage disequilibria in finite populations. *Theor. Pop. Biol.* 33: 54–78. [2]
- Kimura, M. 1955. Solution of a process of random genetic drift with a continuous model. *Proc. Natl. Acad. Sci. USA* 41: 144–150. [2]
- Kimura, M. 1968. Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles. *Genet. Res.* 11: 247–269. [2]
- Kimura, M. 1970. The length of time required for a selectively neutral mutant to reach fixation through random frequency drift in a finite population. *Genet. Res.* 15: 131–133. [2]
- Kimura, M. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge Univ. Press, Cambridge, UK. [2]
- Kimura, M. and J. F. Crow. 1964. The number of alleles that can be maintained in a finite population. *Genetics* 49: 725–738. [2]
- Kimura, M., and T. Ohta. 1969. The average number of generations until fixation of a mutant gene in a finite population. *Genetics* 61: 763–771. [2]
- Kingman, J. F. C. 1982a. The coalescent. *Stoch. Proc. Appl.* 13: 235–248. [2]
- Kingman, J. F. C. 1982b. On the genealogy of large populations. *J. Appl. Prob.* 19: 27–43. [2]
- Lessard, S. 1981. Is the between-population variance negligible in the total variance of heterozygosity? Case of a finite number of loci subject to the infinite-allele model in finite monoecious populations. *Theor. Pop. Biol.* 20: 394–410. [2]

- Li, W.-H., and M. Nei. 1975. Drift variances of heterozygosity and genetic distance in transient states. *Genet. Res.* 25: 229–248. [2]
- Lynch, M., and W. G. Hill. 1986. Phenotypic evolution by neutral mutation. *Evolution* 40: 915–935. [2]
- Moran, P. A. P. 1962. *The Statistical Processes of Evolutionary Theory*. Clarendon Press, Oxford, UK. [2]
- Nei, M., and S. Kumar. 2000. *Molecular Evolution and Phylogenetics*. Oxford Univ. Press, Oxford, UK. [2]
- Nei, M., and W.-H. Li. 1976. The transient distribution of allele frequencies under mutation pressure. *Genet. Res.* 28: 205–214. [2]
- Ohta, T., and M. Kimura. 1969. Linkage disequilibrium due to random genetic drift. *Genet. Res.* 13: 47–55. [2]
- Ohta, T., and M. Kimura. 1971. Linkage disequilibrium between two segregating nucleotide sites under the steady flux of mutations in a finite population. *Genetics* 68: 571–580. [2]
- Stewart, F. M. 1976. Variability in the amount of heterozygosity maintained by neutral mutations. *Theor. Pop. Biol.* 9: 188–201. [2]
- Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159. [2]

Figure 2.1. Expected probability distributions for the frequencies of neutral alleles in replicate, randomly mating populations of size N after t generations of divergence. The initial allele frequency in the base population is 0.5 on the left and 0.1 on the right. The abscissa is the population allele frequency, while the ordinate is proportional to the probability of occurrence of that frequency. Note that the time scale is in units of N generations, where N is the population size, so that $t = N$ generations implies 100 generations for a population of size 100 and 10,000 generations for a population of size 10,000. (From Kimura 1955).

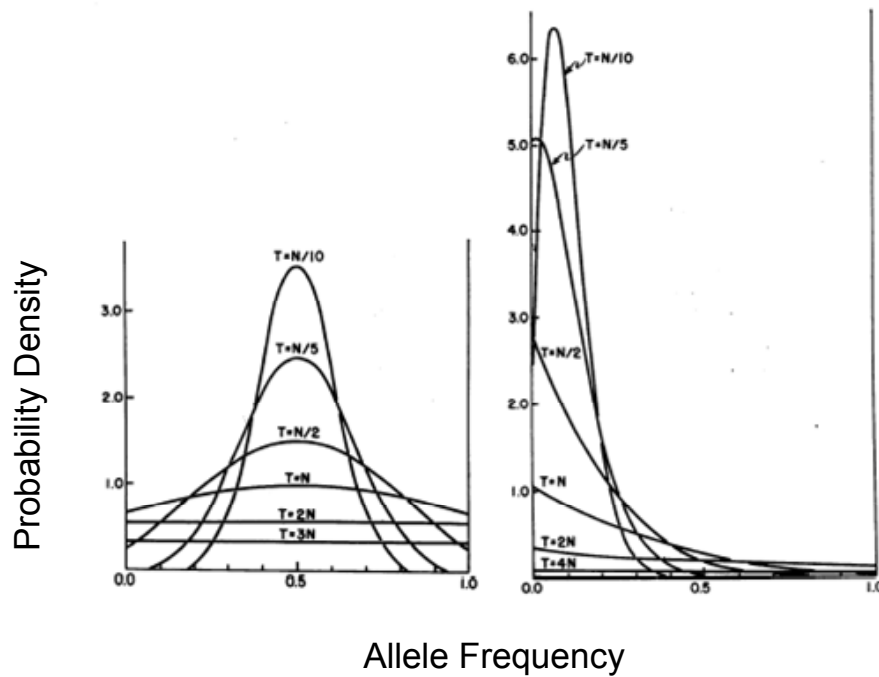


Figure 2.2. Patterns of change in the frequencies of the bw^{75} allele in 212 isolated populations of *Drosophila melanogaster*, each consisting of 8 breeding males and 8 breeding females. Top: The average allele frequency over the entire pool of populations. The dotted and solid lines respectively denote deviations of two standard errors from the expected value of $p_0 = 0$ under the assumption of effective population sizes of 16 and 10.2 individuals. Middle: Mean observed heterozygosity compared to the expectations assuming an effective population size of 10.2. The expected heterozygosity is 0.5 in generations 1 and 2 because the base population (generation 0) consisted entirely of heterozygotes, and with separate sexes, an additional generation is required for the unification of alleles that are identical by descent. Bottom: Among-line variance of allele frequencies compared with the expectations assuming an effective population size of 10.2. (From Buri 1956).

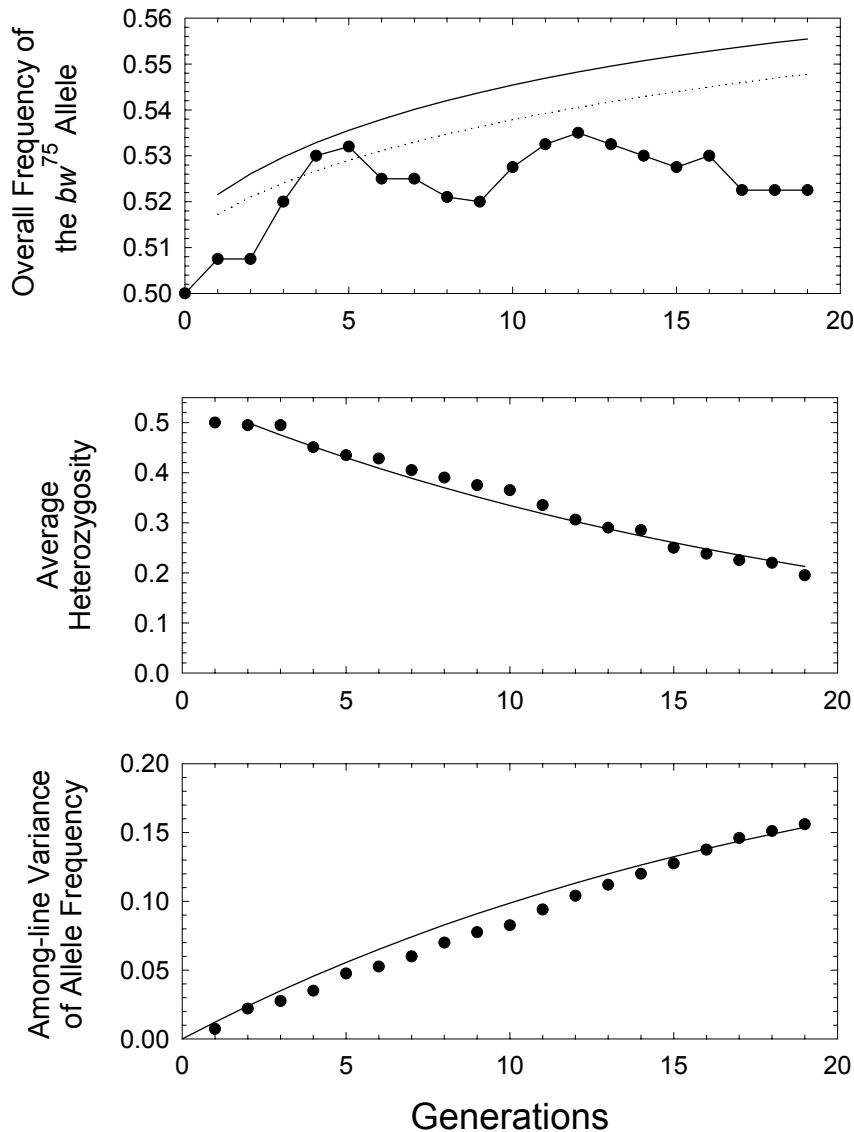


Figure 2.3. Distribution of the number of bw^{75} alleles in 212 populations of *D. melanogaster* each initiated with a frequency of 0.5. (From Buri 1956).

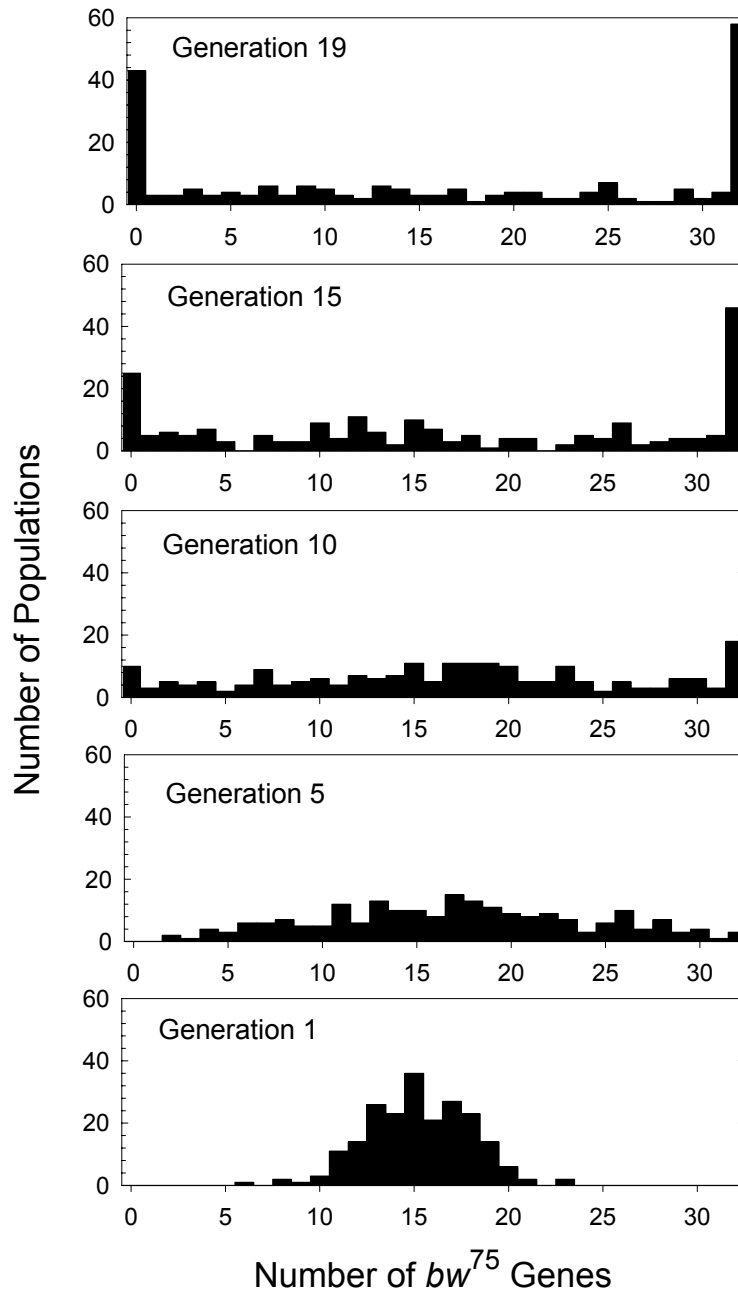


Figure 2.4. Observed sampling variances of allele frequencies for situations in which the donor population contained 1 to 31 bw^{75} genes. The dashed line is the expected pattern, $p(1-p)/2N$, if the actual populations of 8 males and 8 females were randomly mating with equal chances of contributing offspring. The solid line describes the pattern for an average effective population size of 10.2. (From Buri 1956).

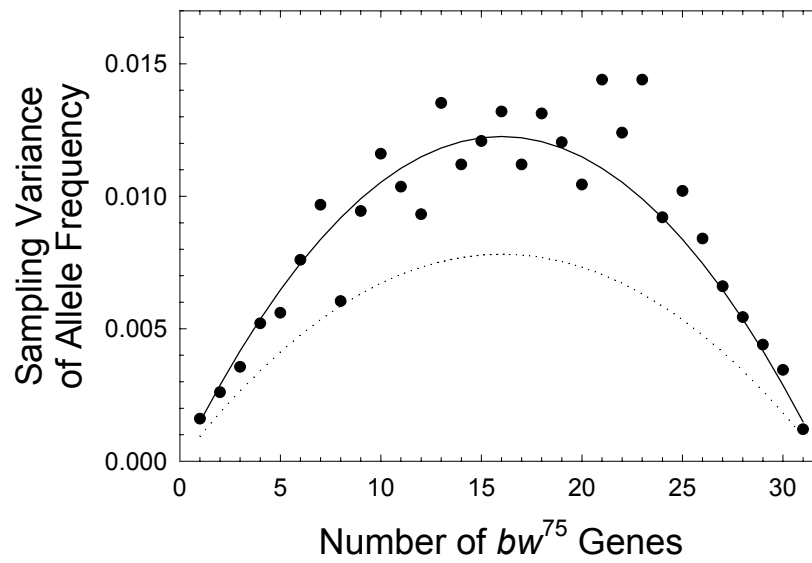


Figure 2.5. Average heterozygosity and its standard deviation among replicate populations as a function of time (scaled in units of $2N$ generations). The upper curves assume an initial heterozygosity of $H_0 = 0.5$ and the lower curves of $H_0 = 0.2$. A diallelic locus is assumed, and the origin of new variation by mutation is ignored. Experimental error resulting from sampling of a finite number of individuals is ignored as well, i.e., we consider only the variance of true population-level heterozygosities resulting from gamete sampling.

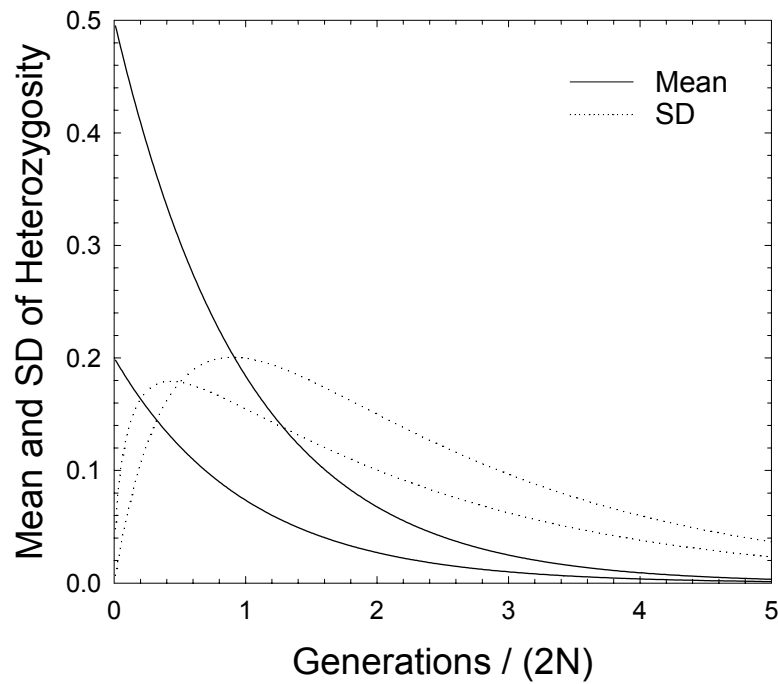


Figure 2.6. Expected coalescence times for a sample of $n = 5$ neutral genes taken from an idealized Wright-Fisher population of size N . The numbers of gene pairs in each consecutive step of the coalescent process are denoted by n_p , and the expected times to coalescence at each step are equal to $2N/n_p$ generations. The particular lineages that join during each step are arbitrary.

