æ

# 7

# INTERACTION OF SELECTION, MUTATION, AND DRIFT

*9 March 2013*

"In recent years, there has been some tendency to revert to more or less mystical conceptions revolving about such phrases as "emergent evolution" and "creative evolution." The writer must confess to a certain sympathy with such viewpoints philosophically but feels that they can have no place in an attempt at scientific analysis of the problem." Wright (1931)

In the previous chapters, we treated the response to selection as an effectively deterministic process, making the assumption that the stochastic force of random genetic drift is negligible relative to the power of selection, and also ignoring the origin of new variation by mutation. Such an approach often works well when the focus is on short-term evolutionary issues. However, on longer time scales, selection, mutation, and drift can interact to pattern variation both within and among populations in significant and sometimes counterintuitive ways. As all populations are finite in size, and all genomes are subject to mutation, these matters must be incorporated into any general theory of evolution. Thus, although the material in this chapter is confined to one- and two-locus systems, the resultant principles provide the basic building blocks for more complex models for the evolution of quantitative traits presented in subsequent chapters.

Whereas mutation and drift respectively introduce and remove variation from populations, selection can have either effect, depending on whether it is directional or purifying in nature. Of special interest is the degree to which all three forces interact to define the distribution of allele frequencies in an equilibrium population (or more precisely, in a quasi-equilibrium population, as with drift there is always some stochastic wandering of allele frequencies around a long-term expectation). One of the key issues considered in the following pages concerns the amount of variation maintained in the face of opposing pressures. We initially address this matter by retaining the assumption of an effectively infinite population size, considering the issue of selection-mutation balance and the fitness load that recurrent mutation always imposes upon a population. We then evaluate the situation in which drift is sufficiently strong to compete with or even overpower the effects of selection. The latter issue is of special interest when we consider selection on a quantitative trait, as strong selection at the phenotypic level does not necessarily translate into strong selection on any particular underlying locus. However, we also show that even

when completely penetrant, only a small fraction of advantageous mutations are successfully fixed in a population, owing to the overwhelming influence of stochastic forces when alleles are rare.

Because the ways in which genes evolve often depend on the background context, we also use this chapter to introduce some key issues regarding the evolution of multilocus systems. First, drawing on results outlined in Chapter 3 for the effects of linkage on the effective population size for a chromosomal region, we explore how this translates into a reduction in the efficiency of selection for advantageous alleles. Second, using compensatory mutations as an entrée into the matter of epistasis, we evaluate the extent to which such pairwise changes are promoted in small vs. large populations. Third, we evaluate the situation in which two or more key mutations are required for a new adaptation, showing that some relatively simple scalings apply to the time to establishment with respect to population sizes and mutation rates.

## SELECTION AND MUTATION AT SINGLE LOCI

Many of the central questions in population and quantitative genetics concern the mechanisms responsible for the maintenance of genetic variation in natural populations. Here, we introduce a few classical models for the balance between the opposing forces of mutation and directional selection. Our preliminary focus will be on the simple case of two alleles, as this serves as the foundation for more complex models for the maintenance of quantitative variation covered in later chapters.

Consider a locus with advantageous allele **A** and deleterious allele **a**, with respective frequencies $1-p$ and $p$. Let $u$ be the mutation rate from **A** to **a**, and $v$ be the rate of back mutation to **A**, and assume random mating, constant selection, and an effectively infinite population size. From Chapter 5 the new frequency of **a** after a generation of viability selection is

$$p' = p\,\frac{W_{\rm a}}{\overline{W}} \tag{7.1}$$

where $W_{\rm a}$ is the marginal fitness of **a**, and $\overline{W}$ is the mean population fitness. Letting $p''$ be the allele frequency following mutation, we then have

$$p'' = (1-v)p' + u(1-p') = (1-u-v)p' + u \tag{7.2}$$

This follows because $1-v$ is the fraction of **a** that remains unchanged following mutation, while a fraction $u$ of all **A** alleles (with frequency $1-p'$) mutate to **a**. Thus, under the joint action of selection and mutation, the new frequency of **a** is

$$p'' = (1-u-v)p\,\frac{W_{\rm a}}{\overline{W}} + u \tag{7.3}$$

Haldane (1927) was the first to consider the stable equilibrium frequencies that are eventually reached under this model of opposing mutational and selection pressures. Letting the fitnesses of genotypes **AA**, **Aa**, and **aa** be 1, $1-hs$, and $1-s$, the

equilibrium frequencies $\widetilde{p}$ satisfying $\Delta p = p' - p = 0$ are given by the solutions of the rather complicated cubic equation

$$(1-\widetilde{p})^3 s(2h-1) + (1-\widetilde{p})^2[2-3h+uh+v(1-h)] + (1-\widetilde{p})[-s(1-h)+u(1-hs)+v(1-2s+hs)] - v(1-s) = 0 \qquad (7.4)$$

(Bürger 2000). Provided $0 < s < 1$ and $h \leq 0.5$, this expression has a single stable equilibrium, and considerable simplification is possible in a number of biologically realistic cases. For example, for the case of neutrality ($s = 0$), the equilibrium is simply defined by the opposing forces of mutation

$$\widetilde{p} = \frac{u}{u+v} \qquad (7.5)$$

The situation of most interest here concerns the polymorphism maintained by a balance between selection and mutation when allele **a** is at a selective disadvantage. To simplify the solution, it is generally assumed that back mutation to the advantageous allele is a negligible force. There are two justifications for such an assumption, one mathematical and one biological. First, unless the selection coefficient is small relative to the mutation rate, the frequency of the mutant allele will generally be low enough that back mutation will be a second-order effect. Second, although functional genes may mutate to deleterious alleles by numerous mechanisms, precise back-mutations to normal alleles will necessarily be much rarer events, i.e, we expect $v \ll u$. Letting $v = 0$, Equation 7.4 reduces to a more manageable, quadratic equation, with solution

$$\widetilde{p} = \frac{\sqrt{[hs(1+u)]^2 + 4(1-2h)us + (1+u)hs}}{2(2h-1)s} \qquad (7.6a)$$

assuming $s > u$. For the extreme (and unlikely) situation in which **a** is a completely dominant deleterious mutation ($h = 1$),

$$\widetilde{p} = \frac{u}{s} \qquad (7.6b)$$

whereas if **A** is recessive ($h = 0$),

$$\widetilde{p} = \sqrt{\frac{u}{s}} \qquad (7.6c)$$

For the general case of intermediate dominance ($0 < h \leq 0.5$),

$$\widetilde{p} = \frac{u}{hs}, \quad \text{provided} \quad h \gg \sqrt{u/s} \qquad (7.6d)$$

A number of other special cases are presented in Nagylaki (1992) and Bürger (2000).

The multiple-allele version of this model can be obtained in a straight-forward manner. Suppose there are $k$ alleles $\mathbf{A}_1, \cdots, \mathbf{A}_k$ and let $u_{ij}$ be the probability that allele $\mathbf{A}_i$ mutates to allele $\mathbf{A}_j$. Letting $u_i = \sum_{j \neq i} u_{ij}$ be the total mutation rate from allele $\mathbf{A}_i$ to any other allele, and assuming constant viability selection followed by mutation and then random mating, the allele-frequency change equations become

$$p_i'' = \frac{1}{\overline{W}} \left( (1-u_i)\, W_i\, p_i + \sum_{j \neq i} u_{ji} W_j\, p_j \right) \qquad (7.7)$$

where $W_i$ is the marginal fitness of allele $\mathbf{A}_i$. The equilibrium behavior of this system can be quite complex, and with sufficiently strong mutation, the possibility of stable cycles exists (Bürger 2000).

Clark (1998) examined a special case of the multiple-allele model in which there is one optimal allele, and all heterozygotes for single mutations have fitness $1 - hs$, while those for two different mutant alleles have fitness $1 - ks$, where $k$ is a measure of complementation between two deleterious alleles (with $k = 0$ implying that each allele compensates for the other allele's deficiencies). Under this model, multiple deleterious alleles are maintained by mutation pressure, and provided $k < 1$, the sum of their frequencies is higher than expected under the two-allele model. The latter result arises as interallelic complementation reduces the magnitude of selection operating on mutant alleles jointly present in the same genotype.

---

**Example 7.1.** How much variation can mutation maintain when a mutant allele is lethal ($s = 1$)? The equilibrium frequency of a dominant lethal allele is

$$\widetilde{p} = u$$

(Equation 7.6b), whereas for a recessive lethal

$$\widetilde{p} = \sqrt{u}$$

(Equation 7.6c). Thus, because $u \ll 1$ (Chapter 3), recessive lethals are expected to be much more common than dominant lethals, a pattern that is seen for numerous human genetic disorders (Cavalli-Sforza and Bodmer 1971). Drawing from a tradition starting with Haldane (reviewed in Nachman 2004), these expressions are often used to estimate the lethal mutation rate for monogenic human diseases under the assumption that the observed frequencies of lethal alleles are at mutation-selection equilibrium (e.g., Kondrashov 2003).

For a dominant lethal, the frequency of selected individuals in the equilibrium population is

$$\text{freq}(\mathbf{aa}) + \text{freq}(\mathbf{Aa}) = u^2 + 2u(1 - u) \simeq 2u$$

whereas for a recessive, the frequency of selected individuals is

$$\text{freq}(\mathbf{aa}) = (\sqrt{u}\,)^2 = u$$

Thus, despite the great disparity in allele frequencies for dominant and recessive lethals, because $u$ is expected to be very small, there is only a twofold difference in the expected frequencies of affected individuals.

What about the equilibrium mean fitness of the population? With a dominant lethal

$$\overline{W} = \text{freq}(\mathbf{AA}) = (1 - u)^2 \simeq 1 - 2u$$

while for a recessive lethal,

$$\overline{W} = 1 - \text{freq}(\mathbf{aa}) = 1 - u$$

**Example 7.2.** Albinism in humans is caused by a recessive allele, with an estimated frequency of albinos of around 1/20,000 (Cavalli-Sforza and Bodmer 1971). If we assume that albinos are at a slight selective disadvantage ($s = 0.1$) and at mutation-selection equilibrium, what is the estimated mutation rate to albino alleles? Assuming Hardy-Weinberg, so that $\widetilde{p}^2 = 1/20,000$, from Equation 7.6c,

$$\widetilde{p}^2 = \frac{1}{20,000} = \left(\sqrt{\frac{u}{0.1}}\right)^2$$

which implies $u = 5 \times 10^{-6}$. Conversely, if we were to assume a mutation rate of $u = 10^{-5}$, the strength of selection against albinism would be inferred from

$$\widetilde{p}^2 = \frac{1}{20,000} = \left(\sqrt{\frac{10^{-5}}{s}}\right)^2$$

implying $s = 0.2$, i.e., a 20% reduction of fitness in albinos.

## SELECTION AND DRIFT AT SINGLE LOCI

In the preceding section, we assumed a situation in which the forces of selection and mutation are powerful enough to ignore the stochastic consequences of random genetic drift, at least in the short term. This deterministic approach to population genetics yields explicit equilibrium solutions for allele frequencies within populations, usually with no oscillatory behavior. In reality, however, drift plays a significant role in all long-term population-genetic contexts. For example, even when selection against deleterious mutations is strong, the defective alleles segregating in a population today will generally be descendants of entirely different mutations than those millenia in the past. All mutations eventually experience one of two alternative fates, complete loss or fixation.

Our focus now becomes the probability of fixation of an allele by the spread of its descendants to a total frequency of 1.0. In general, drift reduces the efficiency of selection in that the sampling of gametes to form each consecutive generation results in random deviations in allele frequencies from the expectations based on selection alone. If drift is strong relative to selection, a favored allele may stochastically decrease in frequency and sometimes eventually become lost, while a disadvantageous allele may increase in frequency and sometimes become fixed. Throughout the following subsections, we ignore the effects of recurrent mutation, focusing instead on the fate of a pre-existing allele or newly arisen mutation.

Most of the theory of the interaction between selection and drift was developed for a single diallelic locus under viability selection, in which case the change in allele frequency per generation can be treated as the sum of changes resulting from selection and drift,

$$\Delta p = \Delta p_s + \Delta p_d$$

where $\Delta p_s$ is given by Equation 5.1b, and $\Delta p_d$ (the per generation change due to drift) is a random variable. Drift causes no directional tendency in the change in allele frequency, and hence $E(\Delta p_d) = 0$. Thus, the simplest measure of the strength of drift is the expected variance in allele-frequency change due to gamete sampling, which under the standard Wright-Fisher model (Chapter 2) is defined by the binomial distribution,

$$\sigma^2(\Delta p_d) = \frac{p(1-p)}{2N_e} \tag{7.8}$$

where $p$ is the allele frequency prior to sampling, $N_e$ is the variance effective population size, and the 2 accounts for diploidy (Chapter 3). If $\sigma^2(\Delta p_d)$ is small relative to $\Delta p_s$, allele-frequency changes will not be dramatically different from their expectations under selection in an infinite population, but when $\sigma^2(\Delta p_d) > \Delta p_s$, drift can substantially obscure the deterministic force of selection.

Consider the situation in which alleles have additive fitness effects, with genotypes **AA**, **Aa**, and **aa** having respective fitnesses 1, $1 + s$, and $1 + 2s$. Letting $p$ be the frequency of allele **a**, then from Equation 5.2, $\Delta p_s \simeq s\,p(1-p)$, assuming weak selection ($|s| \ll 1$). Comparing this result with Equation 7.18, it is clear that directional selection dominates drift when $2N_e|s| \gg 1$, whereas drift dominates when $2N_e|s| \ll 1$.

Because the intensity of drift scales with $1/(2N_e)$, a useful heuristic is that $2N_e s$ approximates the ratio of the power of selection to drift. This argument is not quite precise because the variance of allele-frequency change is only a rough indicator of the sampling properties of the allele-frequency distribution. However, **diffusion theory**, which gives an essentially complete description of the dynamics of a diallelic locus under drift and selection, upholds this general conclusion (Appendix 1). We will frequently encounter the composite parameter $2N_e s$ in the following paragraphs.

### Probability of Fixation Under Additive Selection

There is no possibility of a perfectly stable polymorphism when drift and selection interact. Indeed, even in the case of overdominant selection (where there is a stable equilibrium in an infinite population, Chapter 5), one allele will eventually drift to fixation unless both homozygotes are lethal. Under this view, all new mutations ultimately become either lost or fixed at the population level, and those that become fixed will themselves be subject to replacement by subsequently arising mutations. Thus, when finite populations are considered, we need to think in terms of fixation probabilities and sojourn times of mutations. Even highly favorable alleles have fixation probabilities less than 1.0 to a degree that depends on the initial frequency $p_0$, the strength of selection, and the effective population size $N_e$.

Denote by $p_f(p_0)$ the probability that an allele starting at initial frequency $p_0$ becomes fixed. As noted in Chapter 2, under neutrality, the probability of fixation depends only on an allele's initial frequency regardless of population size,

$$p_f(p_0) = p_0 \tag{7.9}$$

Depending on its magnitude and direction, selection will cause this probability to increase or decrease. When allelic effects on fitness behave additively, such that each

copy of allele **a** changes fitness by $s$ (giving fitnesses of 1, $1+s$, and $1+2s$),

$$p_f(p_0) \simeq \frac{1-e^{-4N_e s p_0}}{1-e^{-4N_e s}} \tag{7.10a}$$

$$\simeq p_0 + 2N_e s p_0(1-p_0) \qquad \text{when } 2N_e|s| \leq 1 \tag{7.10b}$$

Equation 7.10a, due to Kimura (1957) with a slightly improved version given by Cash (1977), is derived using diffusion theory in Appendix 1. The simplified version, Equation 7.10b, was developed by Robertson (1960) using the Taylor-series approximation $e^{-x} \simeq 1 - x + x^2/2$ for $|x| \ll 1$, and an alternative derivation is given below. Although these approximations apply to both beneficial ($s > 0$) and deleterious ($s < 0$) alleles, and work especially well favorable alleles (Carr and Nassar 1970), they can significantly *overestimate* the fixation probabilities of highly deleterious alleles ($N_e\, s \leq -1$), an issue examined in detail by Bürger and Ewens (1995).

It is critical to note that even when an allele is under strong selection, drift still plays a powerful role when allele frequencies are near zero. Starting with a single copy of an advantageous allele (with frequency $p_0 = 1/(2N)$, where $N$ is the absolute size of the population), Equation 7.10a implies that the probability of fixation of a new mutation is approximately $2s\,(N_e/N)$ when $4N_e s \gg 1$. As we expect $N_e$ to generally be $\ll N$ (Chapter 3), this implies that a newly arisen favorable mutation is usually lost by drift, no matter how beneficial. However, once the frequency of a strongly beneficial allele becomes sufficiently high, fixation is almost certain. For example, if $N_e s p_0 > 0.5$, the probability of fixation exceeds 0.70, while if $N_e s p_0 > 1$, the probability of fixation exceeds 0.93.

For mutations of weak effect, it is informative to consider the probability of fixation of a newly arisen mutation relative to the neutral expectation of $1/(2N)$. Returning to Equation 7.10a, and approximating the numerator as $4N_e s p_0$, with $p_0 = 1/(2N)$, the scaled probability of fixation

$$p'_f(p_0) = \frac{p_f(p_0)}{1/(2N)} \simeq \frac{4N_e s}{1-e^{-4N_e s}} \tag{7.11}$$

is found to be entirely a function of the composite parameter $S = 4N_e s$, which as noted above is a measure of the strength of selection ($2s$ in favor of homozygotes) relative to that of drift, $1/(2N_e)$ (Figure 7.1). For positive selection with $S = 0.01$, 0.1, and 1.0, respectively, $p'_f(p_0) \simeq 1.005$, 1.05, and 1.58, whereas with negative selection with the same absolute values, $p'_f(p_0) \simeq$, 0.995, 0.95, and 0.58. This shows that the fixation probability of a mutant allele is very close to the neutral expectation of $p_f(p_0) \simeq p_0$ provided $|\,S\,| \ll 1$. This domain of **effectively neutrality** is potentially significant in a number of different contexts. For example, populations of sufficiently small size are unable to purge deleterious mutations or promote beneficial mutations with $|s| < 1/(4N_e)$.

**–Insert Figure 7.1 Here–**

A number of other useful approximations for alleles with additive effects on fitness have been derived from diffusion theory. For example, Kimura (1969) found

that the average cumulative contribution of a new mutation to the population-level heterozygosity (summed over all generations until lost or fixed) is equal to

$$H_T = \left(\frac{4N_e}{N}\right)\left(\frac{S-1+e^{-S}}{S(1-e^{-S})}\right) \tag{7.12}$$

Although this measure may seem somewhat abstract, the product of $H_T$ and the number of new mutations arising in the population per generation, $2Nu$, is equal to the expected heterozygosity under selection-mutation-drift equilibrium. For neutral mutations ($S = 0$), $H_T = 2N_e/N$, implying an expected heterozyosity of $4N_eu$ (which assuming $4N_eu \ll 1$ is consistent with results in Chapter 2 obtained by a different method). For large positive $S$ (strongly beneficial mutations), $H_T$ approaches a limiting value of $4N_e/N$, implying that on a per-mutation basis, such mutations make twice the contribution to the heterozygosity as neutral mutations. Finally, for deleterious mutations with strong enough effects to be eliminated by selection, $H_T \simeq 2/(N|\,s\,|)$.

As in the case of the fixation probability, the expected heterozygosity at a locus scaled to the neutral expectation (dividing $2NuH_T$ by $4N_eu$) is a simple function of $S$ (Figure 7.1). Viewed in this way, it can be seen that although both the relative fixation rate and the contribution to heterozygosity increase with $S$, the former responds much more sharply. This is because deleterious mutations that essentially never fix in a population nevertheless make transient contributions to the heterozygosity prior to elimination by selection, whereas positively selected mutations that are driven through the population relatively rapidly contribute to heterozygosity for only a relatively short period.

A useful approximation for newly arisen mutations with additive effects is that, conditional upon fixation, the expected number of generations spent at frequency $x$ is

$$\Phi_f(x) = \frac{2N_e(1-e^{-Sx})(1-e^{-S(1-x)})}{SNx(1-x)(1-e^{-S})} \tag{7.13a}$$

(from Equation 8.66 in Kimura 1983). There are two notable points with respect to this residence-time relationship (Figure 7.2). First, provided $|S| < 1.0$, conditional upon fixation, a new mutant allele spends approximately $2N_e/N$ generations in each frequency class. Second, the occupancy features of a deleterious mutation en route to fixation are exactly the same as those for a beneficial mutation with the same absolute fitness effects, implying that both have the same mean time to fixation, even though the probability of fixation is lower in the former case. First pointed out by Maruyama and Kimura (1974), this counterintuitive behavior results from the fact that if a deleterious allele is to become fixed, it must do so as a consequence of some fortuitously rapid and extreme sampling errors.

It is also sometimes useful to know the expected residence times of mutations that eventually become lost, $\Phi_l(x)$. From Equation 8.70 in Kimura (1983), the unconditional mean residence times for mutations (regardless of being fixed or lost) is

$$\Phi(x) = \frac{2N_e(1-e^{-S(1-x)})}{Nx(1-x)(1-e^{-S})} \tag{7.13b}$$

and using the fact that

$$\Phi(x) = p_f(1/2N)\cdot\Phi_f(x) + [1-p_f(1/2N)]\cdot\Phi_l(x) \tag{7.13c}$$

$$\Phi_l(x) = \frac{N_e e^{Sx}(e^{S(1-x)} - 1)^2}{N^2 x(1-x)(e^S - 1)(e^{S[1-(1/2N)]} - 1)} \tag{7.13d}$$

Again, we see that the residence times conditional upon loss are essentially the same for positive and negative selection coefficients of the same magnitude (Figure 7.2). This is not true for the unconditional residence times, $\Phi(x)$, which are functions of $\Phi_f(x)$ and $\Phi_l(x)$ weighted by the probabilities of fixation and loss (Equation 7.13c).

For effectively neutral mutations destined to loss, $|S| < 1.0$,

$$\Phi_l(x) \simeq \frac{N_e(1-x)}{N\lambda x} \tag{7.14a}$$

where $\lambda = 1 - [1/(2N)]$, whereas the unconditional residence time is

$$\Phi(x) \simeq \frac{N_e}{Nx} \tag{7.14b}$$

i.e., the average time spent in frequency class $x$ is inversely proportional to $x$.

**–Insert Figure 7.2 Here–**

The preceding expressions are useful in a number of applications. For example, the mean numbers of generations to fixation, loss, or either can be obtained respectively by summing Equations 7.13a, 7.13d, and 7.13c over all frequency classes. Simplifications are possible in some cases. For example, as noted above, a neutral mutation destined to fixation spends an average of $2N_e/N$ generations in each frequency class, and because there are $2N-1$ classes, the time to fixation of effectively neutral alleles is essentially $4N_e$ generations, an outcome obtained in Chapter 2 by different means. The conditional time to loss of a neutral mutation is

$$t_l = \frac{2N_e \ln(2N)}{N\lambda} \tag{7.15}$$

(derived in Appendix 1). The mean number of generations until complete loss of a new mutation with deleterious heterozygous effect $s$ is

$$t_l = 2(N_e/N)[\ln(2N/S) + 0.423] \tag{7.16}$$

provided $S \gg 1$ (Kimura and Ohta 1969b; Nei 1971). More general expressions, which require some numerical integration can be found in Kimura and Ohta (1969a).

The mean total number of copies descendant from a mutation prior to loss or fixation is useful in a number of contexts, e.g., determination of the number of individuals affected by a deleterious mutation. This is defined as

$$\overline{n} = \sum_{y=1}^{2N-1} \Phi(y/2N) \cdot y \tag{7.17a}$$

with a shift of the function $\Phi$ to $\Phi_l$ or $\Phi_f$, respectively, leading to the expected numbers conditional on loss or fixation. For the case of neutral mutations,

$$\overline{n} = 4N_e\lambda \tag{7.17b}$$

$$\overline{n}_f = 4N_eN\lambda \tag{7.17c}$$

$$\overline{n}_l = 2N_e\lambda \tag{7.17d}$$

The mean frequency prior to absorption is simply $\overline{n}/(2N)$ divided by the average absorption time.

---

**Example 7.3.** Although it is generally thought that selection will increase the determinism of a system, this is not necessarily the case. Cohan (1984) showed that starting with identical allele frequencies, the probability of divergence between replicate populations can *increase* relative to the situation under pure drift if the initial frequency of the advantageous allele is sufficiently small. This point can easily be seen as follows. Supposing two replicate populations are segregating alleles **A** and **a** at a locus, with the frequency of **A** being $p = 0.25$, then under pure drift, the probability that one replicate becomes fixed for **A** and the other for **a** is $2 \cdot 0.25 \cdot (1 - 0.25) = 0.375$. Now suppose that **A** is favored by selection, with $N_e s = 0.5$. Again assuming $p_0 = 0.25$, Equation 7.10a gives the fixation probability of **A** as 0.46, implying that the probability of fixing alternative alleles is $2 \cdot 0.46 \cdot 0.54 = 0.496$. Thus, in this case, divergence is substantially *increased* by the interaction between selection and drift.

In general, the probability of fixing alternative alleles in two replicates is $2p_f(p)\,[\,1 - p_f(p)\,]$, which is maximized when $p_f(p) = 1/2$. Thus, the probability of divergence is increased by selection if $p_f(p)$ under selection is closer to 1/2 than $p_f(p) = p$ under drift, and because $p_f(p) > p$ for a selectively-favored allele, a minimum requirement for increased divergence under pan-selection is that the starting frequency of the advantageous allele be $< 1/2$. More specifically, the probability of divergence under drift plus selection exceeds that under drift when the initial frequency is smaller than $\widehat{p} = 1 - p_f(\widehat{p})$. Figure 7.3 shows that the conditions for this to occur are not very restrictive under additive selection.

This observation has a number of practical implications. For example, an elevated level of population subdivision for a quantitative trait relative to the neutral expectation is often taken to imply divergent selective regimes across subpopulations (Chapter 12). But here we see that under identical selection pressures, populations that initiate with low-frequency, advantageous alleles can exhibit levels of divergence more conventionally interpreted as being associated with diversifying selection. Whether allele frequencies, selection coefficients, and drift intensities commonly have the right mixes for uniform selection to enhance the magnitude of phenotypic divergence remains to be seen, but a wide range of conditions appear to yield divergence levels that would be difficult to discriminate from the neutral expectation (Lynch 1986).

---

**–Insert Figure 7.3 Here–**

## Probability of Fixation Under Arbitrary Selection

We now consider the more general model, allowing for dominance, with the genotypes **aa**, **Aa**, and **AA** having fitnesses 1, $1 + s(1 + h)$, and $1 + 2s$. Diffusion theory

(as developed in Appendix 1) then gives the fixation probability of allele **A** as

$$p_f(p_0 \mid s, h) \simeq \frac{\int_0^{p_0} e^{G(x)}\, dx}{\int_0^1 e^{G(x)}\, dx} \tag{7.18a}$$

where

$$G(x) = -4N_e sx[1 + h(1 - x)] \tag{7.18b}$$

For a new mutant introduced as a single copy, $p_0 = 1/(2N)$, under random mating and at least partial dominance,

$$p_f\left(\frac{1}{2N}\right) \simeq \frac{2N_e s(1 + h)}{N[1 - e^{-4N_e s(1+h)}]} \tag{7.19a}$$

This shows that the probability of fixation of a new mutation is largely determined by the heterozygous effect, as almost all copies of a mutation remain in this state until the allele frequency has achieved a moderately high level. For a complete recessive ($h = -1$), the approximation leading to Equation 7.19a breaks down, and higher-order terms in the approximation of Equation 7.18a are required. However, for strong positive selection on homozygotes of a completely recessive allele ($4N_e s \gg 1$), a close approximation is given by

$$p_f\left(\frac{1}{2N}\right) \simeq \frac{\sqrt{4N_e s/\pi}}{N} \tag{7.19b}$$

(see Example A1.7 for details).

If there is direct inbreeding due to mating of close relatives (beyond the amount of long-term inbreeding that is naturally generated by drift), Equation 7.18a still holds, but now with

$$G(x) = -4N_e sx\{2f + (1 - f)[1 + h(1 - x)]\} \tag{7.20a}$$

where $f$ is a measure of the departure of genotypes from Hardy-Weinberg expectations, defined (in Chapter 2) by the frequency of heterozygotes, $2p(1 - p)(1 - f)$ (Caballero and Hill 1992). Using Equation 7.18a, the fixation probability now becomes

$$p_f\left(\frac{1}{2N}\right) \simeq \frac{2N_e s[2f + (1 - f)(1 + h)]}{N} \tag{7.20b}$$

(Caballero and Hill 1992; Caballero 1996), which for a complete recessive ($h = -1$) reduces to

$$p_f\left(\frac{1}{2N}\right) \simeq \frac{4N_e fs}{N} \tag{7.20c}$$

Thus, with even a small amount of inbreeding, the probability of fixation of a beneficial recessive allele is considerably higher than under random mating (Equation 7.19b) due to the elevated exposure in homozygotes (Caballero et al. 1991). In contrast, inbreeding has much more moderate effects on the fixation probabilities of alleles with additive ($h = 0$) or dominant ($h = 1$) fitness effects.

By indirectly causing localized inbreeding, population subdivision can also influence the probability of fixation. Whitlock (2003) found that for a wide variety of

population structures, the global probability of fixation of a new beneficial mutation is well approximated by

$$p_f\left(\frac{1}{2N}\right) = \frac{2N_e s(1+h)(1-F_{ST})}{N} \tag{7.21}$$

where the effective and total population sizes ($N_e$ and $N$) are defined at the metapopulation level, and $F_{ST}$ is an index of population subdivision (defined as the fraction of metapopulation variation for neutral allele frequencies that is distributed among populations; see Chapter 2). Note that with complete population subdivision ($F_{ST} = 1$), fixation is impossible at the metapopulation level, as mutations are permanently confined to the demes in which they arise.

One cannot immediately infer from Equation 7.21 whether population subdivision enhances or reduces the probability of fixation because subdivision influences both $F_{ST}$ and $N_e$. Expressions for effective population sizes under a number of metapopulation structures were presented in Chapter 3, and parallel expressions for $F_{ST}$ can be found in most of the literature cited there. In the case of the ideal island model with symmetric migration between demes and equal contributions of all demes to the entire metapopulation (Chapter 3), $N_e = N/(1-F_{ST})$, and Equation 7.21 reduces to $2(1+h)s$, showing that in this particular case the probability of fixation is independent of the magnitude of population subdivision and simply equal to twice the selective advantage in heterozygotes (Maruyama 1970). Analyses of more complex population structures (Slatkin 1981; Barton 1993) are all special cases of Whitlock's (2003) expression provided the assumption of equal deme productivity is met; and the modifications necessary when this condition are violated are developed in Whitlock (2003) as well. The more complex situation in which the strength of selection varies among demes has been taken up by Whitlock and Gomulkiewicz (2005).

Otto and Whitlock (1997) provide results for fixation probabilities in populations of changing size, showing that selection is more effective in growing populations (increasing the probabilities that favorable alleles are fixed and deleterious alleles are lost) than in declining populations. This result has obvious implications for managed populations. Fortuitously, the limiting expression for the fixation probability of alleles with additive effects (given above as $2sN_e/N$) applies to populations that are changing in size, provided appropriate modifications are made in the definition of $N_e$ (Otto and Whitlock 1997). The much more complex issue of jointly varying population sizes and selection coefficients is taken up by Uecker and Hermisson (2011). A number of additional diffusion results are given for a diallelic locus in Appendix 1, but simple expressions are generally unavailable for multiple alleles.

### Fixation of Overdominant and Underdominant Alleles

A case of special interest is the effect of drift on a locus experiencing selective overdominance, where the heterozygote has higher fitness than either homozygote. Whereas in an infinite population, such balancing selection permanently maintains both alleles (Example 5.4), drift will ultimately fix one allele in a finite population provided the homozygote has nonzero fitness. Although it might seem that balancing

selection will still always magnify the longevity of a polymorphism, contrary to intuitive expectations, selection in a finite population sometimes *increases* the rate of fixation at an overdominant locus (Robertson 1962; Ewens and Thomson 1970; Chen et al. 2008).

If the equilibrium frequency expected in an infinite population is extreme (roughly $\widetilde{p} < 0.2$ or $\widetilde{p} > 0.8$), a polymorphism starting at $\widetilde{p}$ in a finite population is usually lost *more rapidly* under balancing selection than under drift alone, thereby accelerating the removal of heterozygosity. Such behavior arises because selection keeps allele frequencies fairly close to their equilibrium values. If such values are near 0.0 or 1.0, the minor allele will be impeded from occasionally drifting to more protective states of moderate frequencies, thereby increasing the likelihood of loss by drift.

Nei and Roychoudhury (1973) evaluated this issue further with newly arisen overdominant alleles with initial frequency $1/(2N)$. In this case, the mutant allele is initially confined to the heterozygous state, so its early fate is largely independent of its own homozygous effect, but highly dependent on the magnitude of its heterozygous advantage over the resident homozygote. Fixation probabilities can only be obtained by numerical analysis in this case, but the results depend only on two parameters, $N_e(s_1 + s_2)$ and the infinite-population equilibrium frequency $\widetilde{p} = s_2/(s_1 + s_2)$, where $s_1$ and $s_2$ are respectively the selection coefficients against the homozygotes associated with the mutant and resident alleles. If $\widetilde{p}$ for the allele under consideration is much less than 0.5, the fixation probability is less than the neutral expectation for the reasons noted above. However, if $\widetilde{p}$ is larger than 0.5 (the fitness of the resident homozygote is lower than that of the mutant allele), the fixation probability is always greater than the neutral expectation, even though fixation results in the loss of the optimal (heterozygous) genotype. Moreover, in this case, the fixation probability of the mutant allele is only slightly smaller than that predicted by Equation 7.10a when $s_2$ is used as a selection coefficient (Nei and Roychoudhury 1973). If $2N_e(s_1 + s_2) \ll 1$, selection is uniformly overpowered by drift, and the system behaves in an effectively neutral fashion.

The fixation times for newly arisen overdominant mutations parallel the patterns of loss of variation that Robertson (1962) first noted (Nei and Roychoudhury 1973). When the equilibrium frequency is outside of the range of (0.2, 0.8), the mean fixation time is lower than the neutral expectation of $4N_e$ generations, whereas for $0.2 < \widetilde{p} < 0.8$ the time is elevated, with more extreme behaviors seen at high $N_e(s_1 + s_2)$ (Figure 7.4). Particularly intriguing is the fact that the fixation time of an overdominant mutation is symmetrical around $\widetilde{p} = 0.5$, i.e., for a given strength of selection $N_e(s_1 + s_2)$, the time to fixation is the same at equilibrium frequencies $\widetilde{p}$ and $1 - \widetilde{p}$. Consistent with the situation for mutants with additive effects noted above, this means that when an overdominant mutant allele is associated with the least fit homozygous type, for the rare occasions in which fixation occurs, it does so just as rapidly on average as when it is associated with the most fit homozygote (and therefore fixes more frequently). Further considerations for the situation in which populations are subdivided are given in Nishino and Tajima (2004).

**–Insert Figure 7.4 Here–**

Important situations also exist in which a new mutation is underdominant with

respect to the resident allele, i.e., has reduced fitness when in the heterozygous state, but equal or higher fitness as a homozygote. In an infinite population, such an allele would always be driven from the population if its marginal fitness at low frequency is less than that of the resident allele. In a finite population, however, there is some chance that the mutant allele might drift to high frequency, transiently taking the population through a reduction in mean fitness (during the period in which heterozygotes are common), but possibly eventually becoming fixed.

Such a scenario has generated considerable interest in the area of speciation biology, as the fixation of an underdominant mutation in a subpopulation will lead to a situation in which hybrids between subpopulations have reduced fitness. In principle, such a condition can constitute the first stage in the development of reproductive isolation.

For the situation in which the two homozygotes have equal fitness and heterozygotes experience a reduction in fitness $s$, Lande (1979) found that if $sN_e/N \ll 1$ (a condition likely to be met based on empirical information on $N_e/N$; Chapters 3 and 4)

$$p_f(1/2N) \simeq \frac{\sqrt{N_e s/\pi}}{N \cdot e^{N_e s} \cdot \mathrm{erf}(\sqrt{N_e s})} \tag{7.22}$$

where the error function

$$\mathrm{erf}(x) = (2/\sqrt{\pi}) \int_0^x e^{-y^2} dy \tag{7.23}$$

is the cumulative frequency of a unit normal, which can be calculated by various numerical approximations (Abramowitz and Stegun 1972). If the efficiency of selection is sufficiently low ($N_e s \ll 2$), $p_f(1/2N) \simeq 1/(2N)$, as expected for an effectively neutral allele. However, if the efficiency of selection is high ($N_e s > 2$), so that $\mathrm{erf}(\sqrt{N_e s}) \simeq 1$,

$$p_f(1/2N) \simeq \frac{\sqrt{N_e s/\pi}}{N e^{N_e s}} \tag{7.24}$$

Of special interest in the study of speciation are chromosomal rearrangements that cause problems during meiosis in chromosomal heterozygotes, with $s$ as large as 0.5 being quite plausible (Lande 1979, 1984). With $N_e s = 2$, 5, and 10, Equation 7.24 predicts fixation rates that are 0.22, 0.017, and 0.00016 times the neutral expectation. Such results imply that if heterozygote fitness is greatly reduced, transitions to alternative allelic states (with equivalent homozygous fitness) are only possible if $N_e$ is very small. However, when such fixations do occur, they proceed much more rapidly than the neutral expectation of $4N_e$ generations (Lande 1979).

Walsh (1982) generalized the above results to the situation in which the fitness in the novel homozygote is elevated to $1 + t$, such that after passage through a fitness bottleneck, fixation of the underdominant allele leads to an increase in mean population fitness. Letting $\theta = N_e s$, and $\omega = 1 + (t/2s)$,

$$p_f(1/2N) = \frac{\mathrm{erf}\{[(1/2N) - (0.5/\omega)]\sqrt{4\theta\omega}\} + \mathrm{erf}\{\sqrt{\theta/\omega}\}}{\mathrm{erf}\{[1 - (0.5/\omega)]\sqrt{4\theta\omega}\} + \mathrm{erf}\{\sqrt{\theta/\omega}\}} \tag{7.25}$$

For $t < 2s$, the fixation probability is close to that predicted by Equation 7.22, whereas for very large $t$, $p_f(1/2N)$ can moderately exceed the neutral expectation

provided $N_e s$ is not so strong that the allele is incapable of drifting to a high enough frequency to be favored by selection (Figure 7.5).

The latter case is of special interest, as one can identify a critical effective population size ($N_e^*$) above which the efficiency of selection is so strong that there is essentially no possibility of the population passing through the fitness bottleneck imposed by heterozygotes. With heterozygotes having a fitness reduction of $s$, homozygotes an advantage of $t$, and $p$ being the frequency of the mutant allele, the mean population fitness is $\overline{W} = 1 - 2p(1-p)s + p^2 t$, which reaches a minimum at $\widehat{p} = s/(t+2s) = 0.5\omega$, with $p < \widehat{p}$ implying net selection against and $p > \widehat{p}$ net selection in favor of the mutant allele. Thus, the key issue is whether the mutant allele can drift from initial frequency $1/(2N)$ to $\widehat{p}$, at which point selection can pull it to fixation. When $p$ is small, the frequency of mutant homozygotes is negligible, and the new allele effectively behaves like a deleterious mutation being removed from the population at rate $s$, and it can be shown that there is essentially no chance of the allele drifting to $\widehat{p}$ if

$$N_e^* > \frac{t+2s}{s^2} \tag{7.26}$$

(Lynch 2012a). For example, with a mutant allele with disadvantage $s = 0.01$ in the heterozygous state but advantage $t = 0.01$ in the homozygous state, an effective population size above 300 imposes a very strong barrier to establishment. Lande (1979, 1985) shows that such selective valleys are much more likely to be vaulted in subdivided populations, where local extinction and recolonization permit individual demes to make transitions to an alternative genotypic state and then export such a fixed change to a newly opened habitat.

**–Insert Figure 7.5 Here–**

### Expected Allele Frequency in a Particular Generation

A number of applications, including attempts to predict the response to selection, arise where it is useful to know the expected allele frequency at time $t$, $E(p_t)$. While exact results can be obtained from probability transition matrices (Carr and Nassar 1970; Hill 1969a) and good approximations can be derived from diffusion theory (Appendix 1; Maruyama 1977; Ewens 2004) and other approaches (Curnow and Baker 1968, 1969; Pike 1969), these methods tend to be numerically intensive. Fortunately, simple approximations have been developed for weak selection.

In a finite population, drift can reduce the selection response by progressively diminishing the expected heterozygosity each generation. Consider a locus with additive selection, with genotypes **aa**, **Aa**, and **AA** having fitnesses 1, $1+s$, and $1+2s$. If we assume weak selection, such that changes in allele frequencies associated with selection are relatively minor, compared to those induced by drift, from Equation 5.1b, the expected per-generation frequency change for an allele in the $j$th generation of additive selection can be described as

$$E(\Delta p_j) \simeq sE[\,p_j(1-p_j)\,] \simeq sp_0(1-p_0)\left(1 - \frac{1}{2N_e}\right)^j \tag{7.27}$$

where $p_0$ is the initial allele frequency. The last approximation follows directly from the expression for the expected heterozygosity for a neutral locus in a finite population after $j$ generations with a starting allele frequency of $p_0$, Equation 2.5. Summing over generations, the expected frequency after $t$ generations of selection and drift is

$$E(p_t) = p_0 + \sum_{j=0}^{t} E(\Delta p_j) \simeq p_0 + sp_0(1-p_0) \sum_{j=0}^{t} \left(1 - \frac{1}{2N_e}\right)^j$$

$$\simeq p_0 + 2N_e s\, p_0(1-p_0) \left(1 - e^{-t/2N_e}\right) \tag{7.28a}$$

where the last step follows from the useful approximation

$$\sum_{j=0}^{t} \left(1 - \frac{1}{2N_e}\right)^j \simeq 2N_e \left(1 - e^{-t/2N_e}\right) \tag{7.28b}$$

More generally, if the genotypes **aa**, **Aa**, and **AA** have fitnesses 1, $1 + s(1 + h)$, and $1 + 2s$, then for small $N_e|s|$ and $N_e|sh|$, the expected frequency of **A** is

$$E(\, p_t) \simeq p_0 + 2N_e s p_0(1-p_0) \left[ \left(1 - e^{-t/2N_e}\right) + \frac{h(1-2p_0)}{3} \left(1 - e^{-3t/2N_e}\right) \right] \tag{7.29}$$

These approximations provide a remarkably simple route to obtaining fixation probabilities under weak selection ($N_e s \ll 1$). Because an allele is ultimately either fixed ($p_\infty = 1$) or lost ($p_\infty = 0$), the asymptotic mean frequency as $t \to \infty$ is equal to the fixation probability,

$$E(\, p_\infty) = 1 \cdot p_f(p_0) + 0 \cdot [1 - p_f(p_0)\,] = p_f(p_0)$$

Thus, taking the limit of Equation 7.29 as $t \to \infty$ yields a useful expression for the probability of fixation under weak selection and arbitrary dominance,

$$f(p_0) \simeq p_0 + 2N_e s p_0(1-p_0) \left(1 + \frac{h(1-2p_0)}{3}\right) \tag{7.30}$$

For additive fitness effects ($h = 0$), this expression is identical to Equation 7.10b. Hill (1969a,b) found this approximation to be reasonable provided $N_e|s| < 1$. The more general versions (Equations 7.29 and 7.30) were produced by Silvela (1980).

## JOINT INTERACTION OF SELECTION, DRIFT, AND MUTATION

We now turn to the situation in which selection, drift, and mutation operate simultaneously. Under these conditions, alleles are not simply permanently lost or fixed. Rather, the allele frequencies in a population of constant size eventually reach a stochastic equilibrium (or **stationary distribution**), $\phi(x)$, where $x$ denotes the allele frequency. Recall from Chapter 2 that we can interpret such an equilibrium in two different ways. First, given a conceptually large number of replicate populations, $\phi(x)$ closely approximates the frequency histogram of the numbers of populations

with specific allele frequencies at the locus. Conversely, if we were to follow a single population temporally and construct a histogram of the historical record of allele frequencies at the locus over a very large number of time points, we would again recover $\phi(x)$.

Diffusion theory provides a general solution to this problem (Appendix 1). For the simple biallelic case in which mutations from allele **A** to **a** occur at rate $u$, and $v$ is the reciprocal rate, Wright (1949) found that the equilibrium distribution for the advantageous **A** allele is given by

$$\phi(x) = C\overline{W}^{2N_e} x^{4N_e v-1} (1-x)^{4N_e u-1} \quad \text{for } 0 < x < 1 \tag{7.31a}$$

where $C$ is a normalization constant such that Equation 7.31a integrates to one and hence is a proper probability density (Example A1.3 provides a derivation of this expression). Here, $\overline{W}$ is the mean population fitness, which is itself a function of $x$ and the selection coefficients associated with different gametic states. Note that when both mutation rates are substantially $< 1/(4N_e)$, conditions that may frequently be met for single nucleotide sites (Chapter 4),

$$\phi(x) \simeq \frac{C\overline{W}^{2N_e}}{x(1-x)} \tag{7.31b}$$

showing that with weak mutation pressure, the expected allele frequencies *conditional upon the population being polymorphic* are independent of both the mutation rate and the mutation bias. This result, which represents still another counterintuitive consequence of the influence of drift on gene frequencies, can be understood in the following way.

Suppose that allele **A** has a selective advantage $s$ over allele **a**, and let the rate of mutation from allele $i$ to $j$ be $u_{ij}$. At stationary state, the ratio of times that a population is completely fixed for optimal and suboptimal alleles is

$$\frac{\widetilde{P}_A}{\widetilde{P}_a} = \left(\frac{v}{u}\right) e^S \tag{7.32}$$

where $S = 4N_e s$ (Wright 1931; Li 1987; Bulmer 1991; McVean and Charlesworth 1999). Note that $(v/u)$ and $e^S$ are, respectively, the mutation and selection biases in favor of allele **A**, with the latter being equivalent to the ratio of fixation probabilities of beneficial and detrimental alleles with the same absolute $s$ (obtainable from Equation 7.10a).

Equation 7.32 illustrates two key points. First, although the distribution of allele frequencies conditional on polymorphism can be independent of mutational properties, the frequency of alternative fixed classes is not. Second, the ratio at which the two monomorphic classes produce polymorphisms $(u/v)$ is perfectly compensated by the differential densities of the two classes, and provided the population is sufficiently small that each new mutation is either lost or fixed before another one is produced at the locus, this effect is not influenced by secondary mutations. Equation 7.31b breaks down, however, when population sizes are large enough that the waiting times for new mutations are smaller than the sojourn times of mutant alleles.

Because Equation 7.31a treats allele frequencies as continuously distributed variables, they may behave aberrantly at the absorbing boundaries of frequencies $x = 0$ and 1. However, a rough approximation for the absolute frequencies of the fixed classes can be obtained by noting that

$$R_p = 2N[\widetilde{P}_A u \bar{t}_a + \widetilde{P}_a v \bar{t}_A] \tag{7.33a}$$

is the rate of production of polymorphisms from the fixed classes, with $\bar{t}_a$ $\bar{t}_A$ being, respectively, the mean sojourn times of mutations to alleles **a** and **A**. Using the relationship in Equation 7.32 and the fact that $\widetilde{P}_a + \widetilde{P}_A + \widetilde{P}_p = 1$, where $P_p$ is the probability that the population is polymorphic, the probabilities that the population is fixed for either allele or polymorphic for both can be solved starting with

$$\widetilde{P}_p \simeq 1 - e^{R_p} \tag{7.33b}$$

By multiplying the values of Equation 7.31a by $P_p$ over the range of $x = 1/(2N)$ to $1 - [1/(2N)]$, we then obtain the spectrum of alternative population states.

Figure 7.6 provides some examples of the form of the stationary distribution for biallelic loci experiencing bidirectional mutation. For neutral mutations, the distribution is highly u- or j-shaped (depending on the magnitude of mutation bias) at low population mutation rates ($4Nu$ and $4Nv \ll 1$), as the population is almost always in a nearly fixed state. The distribution becomes flat with values of $4Nu$ and $4Nv$ near 1.0, and then more peaked as $4Nu$ and $4Nv$ become progressively larger (with the mean centered on the infinite-population expectation given by Equation 7.5). For populations that are sufficiently small as to seldom harbor polymorphisms, Equation 7.5 also represents the probability of the alternative fixed states. Selection skews the distribution towards the more favorable allele, but even with $S$ as large as 10, a moderate frequency of the deleterious allele can be expected (even though fixation of the latter would essentially never occur).

**–Insert Figure 7.6 Here–**

Equation 7.31a is useful in a number of applications. Consider, for example, the case of a deleterious recessive allele maintained by mutation (with $u$ being the mutation rate to deleterious alleles, and $s$ being the selective disadvantage of mutant homozygotes). Letting $x$ be the frequency of the deleterious allele, the mean population fitness is $\overline{W} = 1 - sx^2$, using the approximation $(1-y)^{2N_e} \simeq e^{-2N_e y}$ for small $y$, so that $\overline{W}^{2N_e} \simeq e^{-2N_e s x^2}$, and ignoring back mutation to the advantageous allele, the equilibrium distribution is

$$\phi(x) = Ce^{-2N_e s x^2} x^{4N_e u - 1} (1-x)^{-1} \quad \text{for } 0 < x < 1 \tag{7.33}$$

a result originally due to Wright (1938).

Nei (1969) provides a broad overview of the allele-frequency spectrum for lethal mutations, including those that are entirely recessive or overdominant. As neither of these conditions are commonly observed (LW Chapter 10), we note only some of the results for partially recessive lethals. In this case, the average expected frequency

at selection-mutation balance is given by Equation 7.6d, essentially independent of population size, and provided $2N_ehs \gg 1$ (i.e., the power of selection against heterozygotes exceeds the power of drift), the variance in allele-frequency is approximately

$$\sigma^2(p) = \widetilde{p}/(4N_ehs) \tag{7.34}$$

Nei (1971) and Li and Nei (1972) give expressions for the expected numbers of individuals affected by a newly arisen deleterious mutation prior to its elimination by selection.

An area of special interest is the behavior of the four possible nucleotides at a particular site. Denoting the four frequencies as $x_i$ (where $i = 1, \ldots, 4$) and their selection coefficients as $s_i$ (here assumed to be weak and additive), under the assumption that all nucleotides mutate to each other type at the same rate $u$, Equation 7.31a generalizes to

$$\phi(x_1, x_2, x_3, x_4) = C\overline{W}^{\,2N_e} (x_1\, x_2\, x_3\, x_4)^{4N_eu-1} \tag{7.35}$$

where $\overline{W} = 1 + 2\sum_{i=1}^{4} x_i s_i$ is the mean population fitness. Not surprisingly, the solution to this trivariate expression ($x_4$ being defined as $1 - x_1 - x_2 - x_3$) is quite cumbersome (Li 1987; Zeng et al. 1989; Bulmer 1991; McVean and Charlesworth 1999).

Consider, however, the situation in which there is one optimal nucleotide, the frequency of which is denoted by $x$, with the three others having an equal selective disadvantage $s$ in the heterozygous state. Scaling the fitness of the less-fit alleles to be 1, the mean population fitness is then $\overline{W} = 1+2xs$, which is closely approximated by $e^{2xs}$ under the assumption of small $s$. Letting the mutation rate of all nucleotides to the optimal state be $v$ and the total mutation rate of the optimal nucleotide to the other states be $u$, it follows from Equation 7.32 that the expected frequency of the optimal nucleotide is

$$\widetilde{P}_{opt} \simeq \frac{(v/u)e^S}{1 + (v/u)e^S} \tag{7.36}$$

(Li 1987; Bulmer 1991; McVean and Charlesworth 1999). Strictly speaking, this expression applies to the weak-mutation limit (where $N(u + v) \ll 1$ ensures that polymorphisms are rare), so that $\widetilde{P}_{opt}$ denotes the frequency of time the site is fixed for the optimal nucleotide. Equation 7.36 makes a simple, intuitive statement – the frequency of the optimal nucleotide at a site is a function of a single composite quantity, $(v/u)e^S$, which as noted above denotes the net pressure towards the optimal state. As $N_e \to 0$, the expected frequency of the optimal allele approaches the expectation under pure mutation pressure, $v/(u + v)$. For populations that are sufficiently large to maintain substantial heterozygosity, Equation 7.36 is no longer a strict definition of the probability of sampling an optimal allele, as prior to fixation the descendants of a new mutation will themselves have time to acquire secondary mutations. In this case, $P_{opt}$ is more appropriately viewed as the probability that the most recent common ancestor of the alleles currently segregating in a population is an allele of the optimal type.

Sella and Hirsh (2005) and Lynch (2012b) expanded the model leading to Equation 7.36 to allow for multiple alleles with different fitness states. Both models assumed a stepwise-mutation model, with allele $i$ mutating to $i-1$ with rate $u$ and to

$i+1$ with rate $v$, and again are strictly valid as indicators of average allele frequency only in the weak-mutation limit where the population is expected to be typically nearly monomorphic for a single allele at most points in time. Sella and Hirsh assigned fitness $W_i = 1 + s_i$ to allele $i$, and assumed symmetric mutation ($u = v$). Letting $S_i = 4N_e s_i$ (assuming diploidy), the equilibrium probability that $i$ is the fixed (or nearly so) allele is completely independent of the mutation rate,

$$\widetilde{p}_i = \frac{e^{S_i}}{T}, \quad \text{where} \quad T = \sum_{i=1}^{n} e^{S_i} \tag{7.37}$$

and $n$ is the number of alleles. Whereas the Sella-Hirsh model makes no assumptions about fitness ordering between alleles, Lynch's model assumes an ordered fitness increase in a series of alleles, such that $W_i = 1 - e^{-ki}$, with the constant $k$ setting the granularity of fitness change between adjacent alleles, a fitness of 1.0 being approached asymptotically as $i \to \infty$. In this case, the stationary distribution is

$$\widetilde{p}_i = \frac{(v/u)^i e^{-S_i}}{T}, \quad \text{where} \quad T = \sum_{i=1}^{\infty} (v/u)^i e^{-S_i} \tag{7.38}$$

and $S_i = 4N_e e^{-ki}$.

Formulae such as these, which can readily be modified to alternative fitness schemes. Among other things, they are useful for determining the extent to which drift limits the level of adaptation attainable by a population. For example, assuming higher mutation rates to unfavorable states ($u > v$), the advancement toward ever-higher (and fitter) allelic states stalls around a critical value in the allelic series, above which $s_i \simeq e^{-ki}$ is sufficiently small that drift (combined with mutation pressure) overwhelms selection, thereby preventing further adaptive progress (Lynch 2012b). Although alleles in a fitness state above this critical point might arise by mutation, because they are effectively neutral, they are subject to regressive evolution. On the other hand, alleles with sufficiently large disadvantages are incapable of proceeding to fixation, and are purged by selection. Thus, as further discussed in the following section, under virtually all models of adaptation, a **drift barrier** ultimately prevents a population from achieving a perfect state of adaptation, even in a constant environment.

## HALDANE'S PRINCIPLE AND THE MUTATION LOAD

Having established the expected allele frequencies at a locus jointly influenced by mutation, selection, and drift, we now consider in more detail the price that all organisms pay for the privilege of evolving. Because most mutations are deleterious, and many unconditionally so, for every beneficial allele created by mutation, many more detrimental mutations will be introduced to a population. In populations of sufficiently large size, the majority of such mutations will be kept at low frequency and eventually purged, but the relentless flux of new mutations will nevertheless result in an equilibrium load on the mean fitness in the population (Muller 1950; Crow 1993). Remarkably, under reasonably general conditions, this load is often essentially independent of the effects of individual mutations.

In an elegant display of population-genetic reasoning, Haldane (1937) proposed that the reduction in fitness resulting from recurrent deleterious mutations is a function of the deleterious mutation rate alone, an observation that has come to be known as **Haldane's principle**. Consider a deleterious recessive allele **a** with selective disadvantage $s$ in homozygotes. Recalling Equation 5.6d, the mean population fitness when this locus is in selection-mutation balance is

$$\overline{W} = 1 - s \cdot \text{freq}(\mathbf{aa}) = 1 - s\left(\sqrt{\frac{u}{s}}\right)^2 = 1 - u \tag{7.39a}$$

Because the expected frequency of recessive homozygotes is inversely proportional to the selective disadvantage, the reduction in mean fitness (the **mutation load**) is independent of the strength of selection and simply equal to the deleterious mutation rate per allele.

For a deleterious dominant allele with equilibrium frequency $u/s$,

$$\begin{aligned} \overline{W} &= 1 - s\,[\,\text{freq}(\mathbf{aa}) + \text{freq}(\mathbf{Aa})\,] \\ &= 1 - s \cdot \left[\left(\frac{u}{s}\right)^2 + 2\left(\frac{u}{s}\right)\left(1 - \frac{u}{s}\right)\right] \\ &= 1 - 2u + \frac{u^2}{s} \end{aligned} \tag{7.39b}$$

Assuming $s \gg u$, the term $u^2/s$ is negligible, and the mean fitness is again essentially independent of the strength of selection and simply a function of the mutation rate (in this case, the per-locus rate $2u$).

Finally, consider an allele with partial dominance, with heterozygote fitness $1 - hs$. Recalling from Equation 5.6d that the equilibrium allele frequency is $\widetilde{p} = u/(hs)$, the mean population fitness is

$$\begin{aligned} \overline{W} &= 1 - 2hs\,\widetilde{p}(1 - \widetilde{p}) - s\widetilde{p}^2 \\ &\simeq 1 - 2hs\,\widetilde{p} = 1 - 2hs\left(\frac{u}{hs}\right) = 1 - 2u \end{aligned} \tag{7.39c}$$

so that the expected mean fitness is independent of both $h$ and $s$. Bürger (2000) explores these expressions in considerable detail, confirming that the error in ignoring secondary terms in the preceding expressions is of order $u^2/s$ or smaller. With multiple deleterious alleles per locus, these same expressions apply if $u$ is interpreted as the total mutation rate of the most beneficial allele to all classes of deficient alleles at a locus (Crow and Kimura 1964; Clark 1998).

One potential caveat to these results is that the derivation assumes a situation in which there are negligible epistatic effects on fitness. Kimura and Maruyama (1966) examined this issue by considering a quadratic fitness function of the form $w_i = 1 - h_1 i - h_2 i^2$, where $i$ is the number of mutations carried by the individual. With $h_2 = 0$, the model of additive effects assumed above is closely approximated, and Haldane's principle continues to hold, with mean fitness being approximately equal to $e^{-U}$, where $U$ is the deleterious mutation rate per diploid genome. However, at the opposite extreme with $h_1 = 0$, fitness declines with the square of the number of mutations, and mean fitness is elevated to $\sim e^{-U/2}$ regardless of the magnitude of $h_2$. A more general expression that allows for nonzero values of both $h_1$ and $h_2$,

provided by Kimura and Maruyama (1966), demonstrates that this type of **synergistic epistasis** always reduces the mutational load on a sexual population. In contrast, with **diminishing-returns epistasis**, where the decline in fitness with increasing numbers of deleterious mutations becomes progressively shallower, the mutation load is elevated beyond the Haldane expectation.

Fitness functions involving epistasis have played a significant role in our attempt to understand the evolution of sexual reproduction, primarily because the behavior just noted does not extend to asexual genomes, as first shown by Kimura and Maruyama (1966) in a remarkably simple way. Consider an asexual population of mixed clones, with $p_0$ being the frequency of the clone with the minimum number of mutations in one generation and $p_0'$ being its frequency in the next generation. Then, accounting for selection and mutation,

$$p_0' = \frac{p_0 W_0 e^{-U}}{\overline{W}} \tag{7.40}$$

where $\overline{W}$ is the mean population fitness, $W_0 = 1$ is the fitness of the optimal genotype, and $e^{-U}$ is the fraction of the members of this class that do not acquire mutations. Note that no assumptions have been made here with respect to the mode of gene action or on the form of the fitness distribution, and yet at equilibrium ($p_0' = p_0$) we obtain the very general result that mean fitness $\overline{W} = e^{-U}$. Thus, if synergistic epistasis among deleterious mutations is important, a matter on which there is little empirical consensus (Rice et al. 2002; Barton and Otto 2005; Kouyos et al. 2007; Keightley and Halligan 2009), a sexual population will have a long-term advantage in terms of mean fitness. Substantial additional work exists on this subject (e.g., Kondrashov 1984, 1988; Charlesworth 1990; Agrawal and Chasnov 2001; Otto 2003; Haag and Roze 2007).

An additional issue with respect to Haldane's principle is that $N_e$ must be several fold greater than $1/(hs)$ for Haldane's principle to be closely approximated. If this is not the case, deleterious alleles will be capable of drifting to frequencies higher than expected under selection-mutation balance alone. Although this observation led Kimura et al. (1963) to conclude that the mutational load due to segregating mutations will monotonically increase with decreasing $N_e$, their study invoked a relatively high level of back mutation in order to maintain a quasi-equilibrium allele frequency. If instead, one treats back mutation as negligible force (for reasons stated above), it can be shown that the load associated with segregating mutations is nonmonotonic with respect to $N_e$. The segregational load reaches a maximum (in excess of the Haldane expectation) at the point where $1/(2N_e) \simeq hs$, as it is at this point that mutations have a maximum deleterious effect that is still consistent with being highly vulnerable to random genetic drift (Lynch et al. 1995a,b). As $N_e$ declines below this point, the segregational load approaches zero simply because drift is so strong that few segregating polymorphisms of any kind are maintained, and at this point permanent damage simply accrues via the fixation of deleterious alleles, i.e., there is a **fixation load** in addition to any segregational load. Indeed, once a population enters this small-population-size domain, the mutation load may no longer even be maintained at a quasi-equilibrium state as a continual flux of new rounds of weakly deleterious mutations leads to further fixations. If unopposed for a sufficiently long time, such a condition can eventually reduce mean population fitness

to the point at which the average individual is incapable of replacing itself, leading to population extinction via a **mutational meltdown** (Lynch et al. 1995a,b).

Even populations large enough to avoid extinction by a mutational meltdown must experience some fixation load, as there must often be mutationally derived alleles with small enough deleterious effects to be immune to selection. The issue has been explored by a number of investigators using a variety of models for mutational passage between allelic classes (Hartl and Taubes 1998; Poon and Otto 2000; Sella and Hirsh 2005; Lynch 2012b). Although the exact results vary somewhat among studies, in every case the load resulting from fixation of suboptimal alleles is inversely proportional to the effective population size, often with an upper bound on the order of $1/(4N_e)$.

One way to arrive at this result is to recall the two-allele model given above as Equation 7.36. Noting that the load for a fixed deleterious mutation with heterozygous effect $s$ is $2s$ times the expected fraction of time that the deleterious allele is fixed, we then have

$$\begin{aligned} L &= \frac{2su/v}{e^S + (u/v)} \\ &\simeq \frac{2su/v}{1 + 4N_e s + (u/v)} \end{aligned} \tag{7.41a}$$

with the approximation arising when $S = 4N_e s < 1.0$, which must be the case for there to be a significant chance of fixation of a deleterious allele. Under the latter conditions, with symmetrical mutation rates ($u = v$),

$$L = \frac{1}{2N_e + (1/s)} < \frac{1}{4N_e} \tag{7.41b}$$

Mutational bias in the direction of deleterious alleles ($u/v > 1$) will elevate this load, but the point remains the same. Finite population size imposes an ultimate barrier to adaptational refinements that can be maintained in a population. Although this load may appear to be small, as noted in Chapter 4, in all known cases, $u < 1/(2N_e)$, suggesting that the drift load per locus is likely to be typically greater than Haldane's segregational load. In addition, the previous derivations apply to single loci, whereas the cumulative load over all $n$ loci contributing to a trait will be roughly $n$ times the single-locus load. Thus, drift appears to generally impose a nontrivial barrier to adaptive perfection.

There has been considerable debate about the meaning and consequences of the genetic load (Wallace 1991; Crow 1993; Kondrashov and Crow 1993; Reed and Aquadro 2006). As deleterious mutations are removed via reduced survival or reproduction, they must have some demographic consequences. Taken literally though, if the deleterious mutation-free genotype is viewed as the standard ($W_0 = 1$), an equilibrium load $L$ would imply approximately $e^{-L}$ viability (not including mortality unassociated with genetic variation) if its entire influence was born by survivorship. This would then require an inflation of family sizes by a factor $e^L$ relative to the minimum value of two necessary to maintain population-size stability. Under this view, the load concept is paradoxical in that a low-fecundity organism such as a vertebrate would never be able to bear the demographic costs should the genome-wide

deleterious mutation rate exceed $\sim 1.0$, which is likely the case in animal species (Chapter 4). Lesecque et al. (2012) show, however, that the magnitude of selective death is greatly diminished if the fitness of individuals is scaled relative to the actual mean fitness in the population rather than to the idealized $W_0 = 1$. Such a situation would be expected if selection operates mainly through competition of the actual members of the population, rather than by comparison to a nonexistent genotype.

## FIXATION ISSUES INVOLVING TWO LOCI

Populations and species diverge from each other through successive fixations of new mutations, which can be effectively neutral, advantageous, or even slightly deleterious. The relative contributions from these classes is of considerable interest, especially the question of what fraction of substitutions is advantageous and hence adaptive (Kimura 1983; Gillespie 1994). Our goal here is to broaden our outline of fixation theory by considering the influence of the genetic background on expected substitution rates.

There are a number of contexts in which fixation probabilities of alleles are influenced by factors operating at other loci. For example, as discussed in Chapter 3, selection operating on any locus, either positive or negative, results in a reduction in the effective population size in the local chromosomal region, thereby reducing the efficiency of selection operating on all loci linked to the target of selection. Such effects will reduce the fixation probabilities for beneficial alleles, while enhancing the likelihood of fixation of deleterious alleles. In addition, for mutations with contextual (epistatic) effects, fixation probabilities depend critically on the genetic background, and hence on the frequencies of alternative alleles at interacting loci. All of these factors depend very much on the effective population size, which defines the baseline level of variation expected in a population.

### The Hill-Robertson Effect

We first consider the matter of selective interference created by linked variation involving beneficial alleles. Suppose that the gamete with the highest fitness, **AB**, is initially absent and can only be generated by recombination in **Ab**/**aB** double heterozygotes. Letting $x_2$ and $x_3$ denote the frequencies of the **Ab** and **aB** gametes, and $c$ be the recombination frequency between the two loci, then the probability of **AB** being generated in the population is related to the product of the expected frequency of **Ab**/**aB** heterozygotes and the probability that a random gamete from such individuals is **AB**, $(2x_2x_3)\,(c/2)$. Because $x_2x_3 \leq 1/4$ and a population with stable size must produce $2N$ successful gametes, the upper bound to the expected number of **AB** gametes generated in any generation is then $(2N)(c/4)$. Thus, if $Nc < 2$, fewer than one **AB** gametes will be produced each generation by recombination, so unless there is a strong advantage to **AB**, one of the intermediate gamete types will most likely become fixed before **AB** can reach an appreciable enough frequency to be deterministically promoted by selection. Such fixation of one of the intermediate types will then leave new mutation as the only mechanism for the generation of **AB**.

For this special case where the optimal gamete is initially absent, Latter (1966b) developed approximate expressions for the mean time to the first appearance of the **AB** gamete by recombination and for its subsequent fixation probability.

Although there is no general expression for the probability of fixation when alleles at two or more loci are competing for fixation, a number of important results were developed by Hill and Robertson (1966). Most notably, they obtained a weak-selection approximation for the probability of fixation for the following case. Let two diallelic loci (with designated alleles **A**/**a** and **B**/**b**) have recombination frequency $c$, $p_0$ be the initial frequency of **A**, and $D_0$ be the initial gametic-phase disequilibrium (as defined in Chapter 2). Assuming completely additive selection (no dominance or epistasis), with each copy of **A** adding $s_1$ and each copy of **B** adding $s_2$ to total fitness, the probability that **A** becomes fixed is

$$p_f(p_0) \simeq p_0 + 2N_e s_1 p_0(1-p_0) + \frac{2N_e s_2}{2N_e c + 1} D_0 \tag{7.42}$$

provided that $2N_e|s_1|$ and $2N_e|s_2| < 1$. Comparing this two-locus approximation to the single-locus result (Equation 7.10b) shows that the probability of fixation can be increased or decreased depending on the sign of the initial gametic-phase disequilibrium, $D_0$.

Computer simulations show that when selection is strong ($N_e|s_1|$ and/or $N_e|s_2| \gg 1$), linkage (i.e., $c < 0.5$) generally *decreases* the probability of fixation of an advantageous allele relative to the single-locus result (Hill and Robertson 1966). If **A** and **B** are favored alleles, linkage has little effect on the probability of fixation of the **ab** gamete, but the probabilities of fixation of the **Ab** and **aB** gametes increase at the expense of the optimal **AB** gamete (Latter 1965; Hill and Robertson 1966). This decrease is maximized when $N_e c$ is small and both loci have the same effect (e.g., $s_1 = s_2$), as then there is no selective distinction between the two intermediate gametes, rendering them neutral with respect to each other. This is a significant point, as most theoretical investigations on the effects of linkage on the selection response have assumed loci with equal effects (e.g., Fraser 1957; Latter 1965, 1966a,b; Gill 1965a,b,c; Qureshi 1968; Qureshi and Kempthorne 1968; Qureshi et al. 1968), thereby inflating the perceived importance of linkage.

This general phenomenon of selective interference between linked loci was subsequently nicknamed the **Hill-Robertson effect** by Felsenstein (1974). As discussed in Chapter 3, the primary implication of the Hill-Robertson effect is that selection renders the behavior of linked loci closer to that expected under neutrality by reducing the effective population size for the chromosomal region (Birky and Walsh 1988; Charlesworth 1994; Peck 1994). This effect applies to the efficiency of selection on all non-neutral alleles, both advantageous and deleterious. For example, sometimes a moderately beneficial mutation will arise in tight linkage to a highly detrimental allele at another locus, resulting in the former's rapid elimination from the population if the net fitness of the chromosomal region is still lower than that of the population mean. In addition, the average substitution rate at a locus generating deleterious alleles is *increased* if that locus is linked to another locus generating either deleterious or beneficial alleles (Birky and Walsh 1988). In other words, the net effect of linkage is to reduce the overall efficiency of selection for fitness-enhancing mutations, magnifying the accumulation of mildly deleterious mutations at the expense of fixing more advantageous alleles.

This realization that the broad spectrum of Hill-Robertson effects is equivalent to a reduction in $N_e$ greatly facilitates the estimation of fixation probabilities of new mutations subject to background selection and occasional selective sweeps. Indeed, in most contexts that have been examined so far, the standard fixation expressions given above still apply provided the appropriate modifications are made to the definition of $N_e$ (Stephan et al. 1999), as has also been found for subdivided and growing/declining populations. These redefinitions, which have already been outlined at the end of Chapter 3, again point to the great technical utility of the concept of effective population size.

### Mutations with Contextual Effects

To this point, we have generally been assuming that the magnitude of selection operating directly on an allele is independent of the genetic background (other than effects associated with linkage disequilibrium) on which it resides. However, there are numerous situations in which this will not be the case. Most notable is the broad category of **compensatory mutation**, wherein specific single mutations at either of two loci cause a reduction in fitness, while their joint appearance restores fitness or even elevates it beyond the ancestral state. Such epistatic interactions play a prominent role in Wright's (1931, 1932) **shifting balance theory** for adaptive evolution, under which an adaptive valley is traversed in a local subpopulation, with the locally fixed advantageous genotype then being exported to surrounding demes by migration. At the intramolecular level, compensatory mutations appear to be important in a variety of changes in protein sequences and in the composition of nucleotides in the stems of RNA molecules (Stephan and Kirby 1993; Kondrashov et al. 2002; Kulanthinal et al. 2004; Azevedo et al. 2006; Breen et al. 2012).

Ascertaining the conditions under which evolution by compensatory mutation is most likely to occur is challenging because unlike the situation in which a single mutation fixes at a rate depending only on its own initial frequency, the success of a mutation involved in an interlocus interaction depends on the frequency of alleles at the interacting locus, on the fitnesses associated with the nine possible two-locus genotypes, and on the recombination rate between the two loci. Consequently, no general theory for the long-term evolution of interacting loci has yet been developed, although considerable progress has been made in a number of special cases.

As the matter of fixation probability becomes less clear in the case of adaptations involving more than one mutation, in this final section, we will slightly shift our focal point to the rate and mean time to establishment of an adaptation. The latter is defined to be the expected arrival time of the final multi-site adaptation destined to be fixed in the population, starting from a state in which all participating mutations are absent. This excludes the additional time required for fixation, which can generally be obtained from the expressions given above and will often be considerably smaller than the first arrival time. When considering the response to a long-term regular regime of selection, the steady-state rate of evolution is expected to be close to the rate of establishment, as the extra time to fixation simply stretches out each individual event leaving the intervals between them the same. Assuming a constant influx of adaptive mutations, the steady-state rate of adaptation is then

simply the inverse of the time to establishment.

As a benchmark for the following theoretical results, we start with the rate of establishment of a single-site adaptation, with mutations having additive fitness effects. Given a per-site mutation rate of $u$, $2Nu$ new mutations are expected to arise each generation, each at frequency $1/(2N)$. As noted above, if the population size is sufficiently large that $4N_eu \gg 1$, with fixation probability $p_f(1/2N) \simeq 2s(N_e/N)$, the rate of establishment is

$$r_e = 4N_eus \tag{7.43}$$

which is directly proportional to the effective population size, the mutation rate to adaptive changes, and the selective advantage. This approach, of course, assumes that the response to selection is limited by the appearance of new adaptive alleles, and in subsequent chapters we will consider in detail the situation in which part or all of the selection response is a consequences of preexisting variation. It also ignores the point made in the previous section that if $2Nu > 1$ (more than one favorable mutation arises per generation), the simultaneous presence of multiple segregating mutations will reduce the effectiveness of selection, lowering the expected substitution rate (Chapters 8, 10).

As the simplest possible model for the rate of adaptation by new mutations, Equation 7.43 also relies on the rather naive assumption that fixations have no bearing on subsequent events. This assumption can be violated for at least two reasons. First, the fixation of a mutation can alter the selection coefficients of future mutations by, for example, moving the mean phenotype closer to the optimal state and consequently reducing the magnitude of selection for further change. This point is implicit in the drift barrier to adaptation noted above, and relates to the idea of Hartl et al. (1995) that the ultimate consequence of the relentless improvement of traits by natural selection is the evolution of effective neutrality among the remaining pool of segregating alleles. Second, when mutations have epistatic effects on fitness, i.e., depend on the genetic background, the possibility exists of neutral or even deleterious mutations becoming beneficial in certain contexts. We refer to multi-site traits exhibiting the latter types of genetic behavior as **complex adaptations**, as the scenario for their evolution is much less obvious than that under conditions of additive fitness effects.

How do such compensatory changes and other more complex adaptations become established? One possibility is simply that double mutations, while extremely rare, will still arise, with one eventually being carried to fixation by selection. If, however, the mutation rate at a nucleotide site is $10^{-9}$ (Chapter 4), a population size in excess of $10^{18}$ is required to routinely see such double mutations, making this route unlikely for most populations. Conversely, in very small populations, the path towards adaptation must involve successive fixations via drift, which is also likely to be a very long process. In contrast, moderately large populations offer a dual problem in that fixation of key intermediate mutations can be problematic if neutral (owing to the very long time to drift to fixation) and highly unlikely if deleterious. However, starting with Gillespie (1984), it became clear that another pathway, often referred to as **stochastic tunneling** (Komarova et al. 2003; Iwasa et al. 2004), offers a route for the establishment of complex adaptations in large populations even when the intermediate states are deleterious. Under this scenario, secondary mutations arise within the pool of segregating first-step mutations, resulting in fixation of the

double mutant without either single mutation becoming common and hence without a bottleneck in mean population fitness.

The power of stochastic tunneling is that it allows selection to explore (and exploit) the fitness surface more than is possible by single-step mutations, and there is a growing, technical body of work on the subject (Carter and Wagner 2002; Komarova et al. 2003; Iwasa et al. 2004; Weinreich and Chao 2005; Weissman et al. 2009, 2010; Gokhale et al. 2009; Lynch and Abegg 2010; Lynch 2010). Drawing from this literature, our goal here is to provide approximate answers to three basic questions regarding complex adaptations. First, what is the critical population size below which sequential fixation dominates tunneling as a mechanism for adaptation? Second, what is the expected rate (time) to establishment of such double mutations? Third, how does recombination influence these processes?

To put the first question in context, we note that there must be a critical population size $N^*$ below which adaptations are essentially only acquired via sequential fixations, owing to the extreme rarity of occasions in which multiple mutations are simultaneously segregating at key sites. Below this threshold value, selection is restricted to exploring the fitness landscape by single mutational steps from the currently fixed genotype. While a single chance fixation can place a population one step closer to a distant adaptive peak, it can also move it even further way. Conversely, for population sizes exceeding $N^*$, stochastic tunneling allows selection to explore the consequences of genotypes two (and in large enough populations, even more) mutational steps away from the currently most common genotypes. This simple argument suggests that adaptation in small populations will typically occur by simple single-step hill climbing, occasionally supplemented by fortuitous drift across a sufficiently shallow adaptive valley (with a reduction in fitness incurred during such a phase). In contrast, large populations should experience episodes in which adaptive events involve the simultaneous fixation of two (or more) mutations, without any intervening period of fitness loss.

A simple argument on the critical population size for the situation in which first-step mutations are neutral (a fitness plateau) follows from Walsh (1995) and Lynch and Abegg (2010). Consider a complex adaptation requiring two mutations, with the two sites in complete linkage, and suppose that an **A** mutation destined to fixation has arisen. How likely is it that a **B** mutation will arise within a member of this lineage on its way to fixation? Assuming neutrality of the first-step mutation, on average, the second mutation has a window of $4N_e$ generations in which it can arise on an **A** background, and during this period the average frequency of **A** is 0.5. Thus, given an **A** mutation destined to fixation, the expected number of alleles acquiring the second-site mutation is $4N_e \cdot (2Nu) \cdot (1/2) = 4N_eNu$. Hence, when $N_e \simeq N$, there is essentially no chance of a two-mutation haplotype even arising during the fixation of a one-step mutation if the population size is smaller than $1/\sqrt{4u}$. Obviously, if the first-step lineage is destined to become lost, even fewer copies of the double mutation are produced. Now suppose that the double mutation has selective advantage $s$, so that the fixation probability of the **AB** haplotype is $\simeq 2s$. Again assuming $N_e \simeq N$, the adaptation will almost certainly arise by stochastic tunneling rather than by sequential fixation if the population size exceeds

$$N^* \simeq \frac{1}{2\sqrt{2us}} \tag{7.44}$$

Note that this is not a terribly stringent condition, as with $u = 10^{-9}$ and $s = 0.01$, $N^* \simeq 112,000$. The critical population size is larger by a factor of $1/\sqrt{x}$ if $N_e = xN$ (recall from Chapter 3 that $x$ is usually $\ll 1$). When the intermediate step is strongly deleterious (with effect $s_d$), then provided $4N_e s_d \gg 1$, first-step mutations are almost never fixed, with tunneling dominating over sequential fixation.

We now turn to the matter of rates of establishment, focusing again on the situation in which two loci are fixed for alleles **A** and **B** respectively, and inquiring as to the time to reach an alternative state of fixation at both loci, with respective alleles **a** and **b**. We will assume equivalent mutation rates ($u$) from **A** to **a** and **B** and **b**. The simplest selection scenario in this case, first explored by Kimura (1985), assumes that gametes **Ab** and **aB** have equivalent fitness $1 - s$ and gametes **AB** and **ab** have equivalent fitnesses of 1.0. Thus, although transitions between pure population states of **AB** and **ab** may occur, nothing is gained in terms of fitness. Within the sequential fixation domain, such that mutations are limiting and the efficiency of selection is weak ($4N_e u$ and $4N_e s \ll 1$), the degree of linkage can be ignored (as only one locus is polymorphic at a time), and the mean time to establish the novel **ab** type (or vice versa) is the sum of the waiting times for the two mutational steps,

$$\bar{t}_e = \frac{1}{2Nu}\left(\frac{1}{2p_{fd}} + \frac{1}{p_{fb}}\right) \tag{7.45}$$

where $p_{fd}$ and $p_{fb}$ are, respectively, the probabilities of fixation of deleterious (first-step) and beneficial (second-step) alleles (obtained by applying selection coefficients $-s$ and $s$ to Equation 7.10a). Transitions to state **Ab** or **aB** occur at rate $(4Nu)(p_{fd})$ (from **ab** or **AB**) the product of the population mutation rate and twice the rate of first-step fixation (because there are two ways to produce first-step mutations), and then conditional on the first change, the second occurs at rate $2Nup_{fb}$. Because the probability of fixation of a deleterious allele is $e^{-4N_e s}$ that of a beneficial allele (above), the establishment time in this case is expected to be primarily determined by the time required for fixation of first-step alleles, so that

$$r_e \simeq 4Nup_{fd} \tag{7.46}$$

If, on the other hand, selection against the intermediate haplotypes is much stronger than drift so that fixation of the intermediate state is unlikely (the stochastic-tunneling domain), the most likely scenario for a transition to the **ab** type is a population initially residing in a state of selection-mutation balance at both loci. Assuming complete linkage, and a selection coefficient $s$ associated with the **a** and **b** alleles when not combined, the **Ab** and **aB** gametes, each with initial frequency $u/s$ (from Equation 7.6d), would then serve as staging grounds for mutations to the **ab** type. Mutant **ab** gametes arise at rate $u$ from each of the $2Nu/s$ intermediate types, and fix in an essentially neutral fashion with probability $1/(2N)$ (as most resident gametes are of type **AB**, with equivalent fitness). Thus, the rate of establishment of the **ab** type is

$$\bar{r}_e \simeq (2u/s)(u) = \frac{2u^2}{s} \tag{7.47}$$

(Gillespie 1984; Stephan 1996), which is essentially independent of population size.

When mutations are reversible, the question also arises as to the long-term stationary distribution of alternative states. Adhering to the reasoning that **Ab** and

**aB** gametes will generally be maintained at low levels by selection-mutation balance, and assuming equal back and forward mutation rates, Higgs (1998) elegantly showed that the stationary distribution for the frequency ($x_0$) of the **AB** gamete is

$$\phi(x_0) = \frac{1}{(1-z)^{2\alpha-1}} \frac{\Gamma(2\alpha)}{\Gamma(\alpha)^2} \left[ x_0(1-z-x_0) \right]^{\alpha-1} \tag{7.48}$$

where $\alpha = 8Nu^2/s$ is the population rate of mutational production of **ab** gametes, $z = 2u/s$ is the summed frequency of the **Ab** and **aB** gametes, and $\Gamma$ denotes the gamma function (Equation 2.25b). The frequency of the **ab** gamete is simply $1-x_0-z$. With $\alpha < 1$, the distribution of $x_0$ is highly U-shaped, with the probabilities of the population being fixed for alternative **AB** and **ab** states being nearly equal. A more general analysis, which allows for weaker efficiency of selection ($4N_e s < 1$), and differential selection and mutation operating on the intermediate states is presented by Innan and Stephan (2001).

Now suppose that the secondary mutation has advantage $s_b$, and denote the disadvantage of first-step mutations as $s_d$. The general Equation 7.45 still applies in the sequential-fixation domain, and we again expect the rate of establishment to be approximated closely by Equation 7.46 owing to the long waiting time for the fixation of a first-step mutation. For the stochastic-tunneling domain, however, Equation 7.47 must be modified to account for the fact that the fixation probability of the double mutant is $\sim 2s_b(N_e/N)$,

$$\overline{r}_e \simeq (2u/s_b)(2Nu)[2s_b(N_e/N)] = \frac{8N_e u^2 s_b}{s_d} \tag{7.49}$$

The key observations here are that the rate of establishment now depends on the effective population size, while also scaling linearly with the square of the mutation rate and the ratio of selection coefficients associated with first- and second-step mutations. The rate of establishment in the reverse direction is obtained by substituting $s_b + s_d$ for $s_d$ and $s_d$ for $s_b$.

Finally, we consider the special situation in which first-step mutations are effectively neutral. Again, Equation 7.45 provides an accurate description for the sequential-fixation domain, and with substitution of the appropriate fixation probabilities reduces to

$$\bar{t}_e = \frac{1}{2Nu} \left( \frac{1}{2(1/2N)} + \frac{1}{p_{fb}} \right) \simeq \frac{1}{2u} \left( 1 + \frac{1}{2N_e s} \right) \tag{7.50}$$

with the last approximation obtained by using $p_{fb} \simeq 2s(N_e/N)$. Thus, provided $2N_e s \gg 1$, when the intermediate mutation is effectively neutral, the expected time to establishment is $\simeq 2u$ and only weakly dependent on the size of a population.

To obtain the expected rate of tunneling for the case of neutral intermediates, we require the probability that tunneling occurs within the descendant lineage of a first-step mutation before it becomes lost from the population. By various methods, this probability has been found to be approximately $\sqrt{2us}$ in large populations (Komarova et al. 2003, Iwasa et al. 2004; Weissman et al. 2009, 2010; Lynch and Abegg 2010). With $4Nu$ first-step mutations arising per generation, the rate of establishment via tunneling is then

$$r_e \simeq 4Nu\sqrt{2usN_e/N} = 4u\sqrt{2usN_eN} \tag{7.51}$$

If the mutation rates at the two steps are different, $u$ inside and outside of the square-root expression should be treated as mutation rates at the first- and second steps. The key observation here is that when the intermediate stages are neutral, tunneling occurs at a higher rate in larger populations, contrary to the situation with deleterious intermediates. Moreover, although two mutations are required for the final adaptation, the rate of establishment depends on the 3/2s power of the mutation rate, unlike the square scaling with deleterious intermediates.

The above analyses assume an evolutionary path to a final adaptation through just a single intermediate step, and actual fitness surfaces are likely to be more complex, with a variety of potential pathways through any number of mutations. The rates of establishment of complex adaptations under these alternative scenarios has been examined by Gokhale et al. (2009), Weissman et al. (2009), and Lynch and Abegg (2010). Simple analytical expressions have been found in only a few cases, two of which we now summarize. As complex adaptations involving more than two mutations are unlikely to evolve by sequential fixation, owing to the long time necessary for cumulative fixations, we restrict our attention to the stochastic tunneling domain, focusing on the issue of how $r_e$ scales with the underlying features of population size, mutation rate, and selection intensity.

For the case of neutral intermediates with increasing numbers ($d$) of mutations required for the final adaptation (and the order of events assumed to be irrelevant), the rate of establishment can be viewed as a series of nested tunneling events. For example, for the case of $d = 3$ (two neutral mutations required before the final adaptation is assembled with a third mutation), Equation 7.51 expands to

$$r_e = 6Nu\sqrt{2u\sqrt{2usN_e/N}} \tag{7.52a}$$

Note that the first term is now $6Nu$ because first-step mutations can arise at three sites. The nest step then initiates at either of the two remaining two sites, with the final stage initiated at the one remaining site involves tunneling within the sublineage containing the first two mutations. For arbitrary $d$, this expression generalizes to

$$r_e = d\phi u(2Nu)^{1-0.5^{d-1}} S^{0.5^{d-1}} \tag{7.52b}$$

where $S = 4N_e u$, and

$$\phi = \prod_{i=1}^{d-1} (d-i)^{0.5^i} \tag{7.53c}$$

This result shows that, with neutral intermediates, the rate of establishment by tunneling scales with no more than the square of the mutation rate and with no less than linearly with the absolute population size, these extremes being approached at high $d$. Thus, the rate of establishment of complex adaptations can be much more rapid than expected under the naive assumption that independently arising mutations would lead to a scaling with the $d$th power of the mutation rate.

For the case of deleterious intermediates, suppose that all haplotypes involving one to $d-1$ mutations are equally deleterious (with fitness $1-s_d$), with the final mutation conferring an advantage $s_b$. First step mutations then arise at rate $2Ndu$, but owing to selection have an expected survivorship time of $1/s_d$ generations, during which period $d-2$ additional intermediate step mutations must be acquired, followed

by the appearance of a final-step mutation destined to fixation. This leads to a rate of establishment via tunneling of

$$r_e \simeq 4N_e d!(u/s_d)^d s_d s_b \tag{7.54}$$

which reduces to Equation 7.49 when $d = 2$. Here we see that $r_e$ now scales with the $d$th power of the mutation rate owing to the limited opportunities for mutation during the short sojourn times of deleterious mutations, whereas there is a linear scaling with the effective population size. One cautionary note with respect to all of the above-mentioned scaling features is that mutation rates appear to generally evolve to be inversely related to the effective size of a population, which will tend to reduce the dependence of rates of establishment on $u$ and measures of population size, as these two factors typically enter as products of each other (Lynch 2010).

Finally, we note that all of the above analyses assume an absence of recombination. This is a matter of significance, as it is often surmised that recombination facilitates the evolution of complex adaptations. In the sequential-fixation regime, recombination can be ignored simply because multiple polymorphic sites are never present simultaneously. However, in the stochastic tunneling domain, opportunities will exist for both the creation and breakdown of optimal haplotypes. For the case of deleterious intermediates but selectively equivalent end states (above), Higgs (1998) provides more general expressions, allowing for arbitrary levels of recombination. Strong linkage substantially accelerates the rate of peak shifts with this fitness landscape because the frequencies of the low-fitness intermediates remain nearly unchanged during transitions to alternative high-fitness states, ensuring that the population does not pass through a phase of reduced mean fitness (Kimura 1985; Michalakis and Slatkin 1996; Stephan 1996; Innan and Stephan 2001). In contrast, recombination between the high-fitness **AB** and **ab** gametes during a peak shift produces low-fitness intermediates, imposing a bottleneck on mean population fitness, thereby inhibiting the movement from one state to the other.

Lynch (2010) and Weissman et al. (2010) examined this problem with a broader class of models, reaching the conclusion that recombination is most likely to have either a minor or an inhibitory effect on the establishment of a complex adaptation. Consider, for example, the case of a two-site adaptation, starting with a population fixed for the suboptimal **ab** haplotype. The overall influence of recombination on the rate of establishment of the **AB** haplotype is a function of two opposing effects – the rate of origin of **AB** gametes by recombination within doubly heterozygous (**aB** / **Ab**) parents is proportional to the rate of recombination between the sites ($c$), whereas the net selective advantage of the resultant **AB** haplotypes is discounted from $s$ to $s-c$ by subsequent recombinational breakdown (as in the early stages, **ab** haplotypes still predominate, and are the primary partners in recombination events with **AB**). Thus, because the product $c(s-c)$ is maximized at $c = s/2$, two-site adaptations are expected to emerge most rapidly in chromosomal settings where the recombination rate is half the selective advantage of the final adaptation.

For the case of neutral intermediates, details in Lynch (2010) suggest that even at the optimal recombination rate, the rate of establishment is generally enhanced by much less than an order of magnitude relative to the situation with complete linkage, whereas $c > (s/2)$ is not greatly inhibitory. In contrast, when first-step mutations are deleterious, even though the promotional effect of recombination at the

optimal recombination rate (again $\sim s_b/2$) is negligible unless $s_b \gg s_d$, if the rate of recombination exceeds the selective advantage of the **AB** haplotype, recombination presents an extremely strong barrier to establishment of the **AB** haplotype (Lynch 2010). The latter result arises because almost all recombinational events involving a newly arisen **AB** haplotype will involve an **ab** participant, generating the maladaptive **Ab** and **aB** products.

Taken together, these results suggest that only a narrow range of recombination rates (in the neighborhood of $s_a/2$) can enhance the rate of establishment of a complex adaptation from *de novo* mutations. Moreover, because the role that recombination plays in the origin of specific adaptations depends on both the selective advantage of the final product and the physical distance between the genomic sites of the underlying sites, the issue cannot be reduced to a simple generalization. With a highly context-dependent optimal recombination rate (per nucleotide site), it becomes unclear whether selection is likely to have any general influence on the promotion of recombination-rate modifiers (Chapter 4).

These kinds of observations, in which a two-locus system stochastically shifts from one semi-stable state to another through evolutionary time, appear to be closely related (albeit not transparently) to the features of a number of models of complex traits. For example, diallelic models of quantitative traits under stabilizing selection often exhibit multiple equilibria for allele frequencies (including alternative monomorphic and polymorphic states), depending on the effects of alleles and the ways of assembling a multilocus phenotype that most closely resembles the optimum (Bulmer 1972; Barton 1986, 1989; Bürger 1989; Gavrilets and Hastings 1994). One can easily imagine that finite populations would wander from one local equilibrium to another through time depending on the history of mutation and drift, although no formal theory on the rate of such internal shifts has been worked out.

**Literature Cited**

Abramowitz, M., and I. A. Stegun. 1972. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables.* Dover Publications, New York, NY. [7]

Agrawal, A. F., and J. R. Chasnov. 2001. Recessive mutations and the maintenance of sex in structured populations. *Genetics* 158: 913–917. [7]

Azevedo, L., G. Suriano, B. van Asch, R. M. Harding, and A. Amorim. 2006. Epistatic interactions: how strong in disease and evolution? *Trends Genet.* 22: 581-585. [7]

Barton, N. H. 1986. The maintenance of polygenic variation through a balance between mutation and stabilizing selection. *Genet. Res.* 47: 209–216. [7]

Barton, N. H. 1989. The divergence of a polygenic system subject to stabilizing selection, mutation and drift. *Genet. Res.* 47: 209–216. [7]

Barton, N. H. 1993. The probability of fixation of a favored allele in a subdivided population. *Genet. Res.* 62: 149–157. [7]

Barton, N. H., and S. P. Otto. 2005. Evolution of recombination due to random drift. *Genetics* 169: 2353–2370. [7]

Birky, C. W., and J. B. Walsh. 1988. Effects of linkage on rates of molecular evolution. *Proc. Natl. Acad. Sci. USA* 85: 6414–6418. [7]

Breen, M. S., C. Kemena, P. K. Vlasov, C. Notredame, and F. A. Kondrashov. 2012. Epistasis as the primary factor in molecular evolution. *Nature* 490: 535-538. [7]

Bulmer, M. G. 1972. The genetic variability of polygenic characters under optimizing selection, mutation and drift. *Genet. Res.* 19: 17–25. [7]

Bulmer, M. 1991. The selection-mutation-drift theory of synonymous codon usage. *Genetics* 129: 897–907. [7]

Bürger, R. 1989. Linkage and the maintenance of heritable variation by mutation-selection balance. *Genetics* 121: 175-184. [7]

Bürger, R. 2000. *The mathematical theory of selection, recombination, and mutation.* John Wiley & Sons, Ltd., New York, NY. [7]

Bürger, R., and W. J. Ewens. 1995. Fixation probabilities of additive alleles in diploid populations. *J. Math. Biol.* 33: 557–575. [7]

Caballero, A. 1996. A note on the change in gene frequency of a selected allele in partial full-sib mating populations. *Genetics* 142: 649–650. [7]

Caballero, A., and W. G. Hill. 1992. Effects of partial inbreeding on fixation rates and variation of mutant genes. *Genetics* 131: 493–507. [7]

Caballero, A., P. D. Keightley, and W. G. Hill. 1991. Strategies for increasing fixation probabilities of recessive mutations. *Genet. Res.* 58: 129–138. [7]

Carr, R. N., and R. F. Nassar. 1970. Effects of selection and drift on the dynamics of finite populations. *Biometrics* 26: 41–49. [7]

Carter, A. J., and G. P. Wagner. 2002. Evolution of functionally conserved enhancers can be accelerated in large populations: a population-genetic model. *Proc. Biol. Sci.* 269: 953–960. [7]

Cash, W. S. 1977. An improved solution for the ultimate probability of fixation of a favorable allele. I. Ultimate probability of fixation of a favorable allele. *Biometrics* 33: 528–532. [7]

Cavalli-Sforza, L. L., and W. F. Bodmer. 1971. *The genetics of human populations*. W. H. Freeman and Co., San Francisco, CA. [7]

Charlesworth, B. 1990. Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genet. Res.* 55: 199–221. [7]

Charlesworth, B. 1994. The effect of background selection against deleterious alleles on weakly selected, linked variants. *Genet. Res.* 63: 213–228. [7]

Chen, C. T., Q. S. Chi, and S. A. Sawyer. 2008. Effects of dominance on the probability of fixation of a mutant allele. *J. Math. Biol.* 56: 413–434. [7]

Cohan, F. M. 1984. Can uniform selection retard random genetic divergence between isolated populations? *Evolution* 38: 495–504. [7]

Clark, A. G. 1998. Mutation-selection balance with multiple alleles. *Genetica* 102/103: 41–47. [7]

Crow, J. F. 1993. Mutation, mean fitness, and genetic load. *Oxford Surveys Evol. Biol.* 9: 3–42. [7]

Crow, J. F., and M. Kimura. 1964. The theory of genetic loads. *Proc. XIth Internat. Congr. Genetics* 2: 495–505. [7]

Curnow, R. N., and L. H. Baker. 1968. The effect of repeated cycles of selection and regeneration in populations of finite size. *Genet. Res.* 11: 105–112. [7]

Curnow, R. N., and L. H. Baker. 1969. A correction to our earlier paper on "The effect of repeated cycles of selection and regeneration in populations of finite size". *Genet. Res.* 13: 105–106. [7]

Ewens, W. J. 2004. *Mathematical population genetics.* 2nd Edition. Springer-Verlag, New York, NY. [7]

Ewens, W. J., and G. Thomson. 1970. Heterozygote selective advantage. *Ann. Hum. Genetics* 33: 365–376. [7]

Felsenstein, J. 1974. The evolutionary advantage of recombination. *Genetics* 78: 737–756. [7]

Fraser, A. S. 1957. Simulation of genetic systems by automatic digital computers. II. Effects of linkage on rates of advance under selection. *Aust. J. Biol. Sci.* 10: 492–499. [7]

Gavrilets, S., and A. Hastings. 1994. Dynamics of genetic variability in two-locus models of stabilizing selection. *Genetics* 138: 519–532. [7]

Gill, J. L. 1965a. Effects of finite size on selection advance in simulated genetic populations. *Aust. J. Biol. Sci.* 18: 599–617. [7]

Gill, J. L. 1965b. A Monte Carlo evaluation of predicted selection response. *Aust. J. Biol. Sci.* 18: 999–1007. [7]

Gill, J. L. 1965c. Selection and linkage in simulated genetic populations. *Aust. J. Biol. Sci.* 18: 1171–1187. [7]

Gillespie, J. H. 1984. Molecular evolution over the mutational landscape. *Evolution* 38: 1116–1129. [7]

Gillespie, J. H. 1994. *The causes of molecular evolution.* Oxford Univ. Press, Oxford, UK. [7]

Gokhale, C. S., Y. Iwasa, M. A. Nowak, and A. Traulsen. 2009. The pace of evolution across fitness valleys. *J. Theor. Biol.* 259: 613–620. [7]

Haag, C. R., and D. Roze. 2007. Genetic load in sexual and asexual diploids: segregation, dominance and genetic drift. *Genetics* 176: 1663–1678. [7]

Haldane, J. B. S. 1927. A mathematical theory of natural and artificial selection. Part V. *Proc. Cambridge Phil. Soc.* 23: 838–844. [7]

Haldane, J. B. S. 1937. The effect of variation on fitness. *Amer. Natur.* 71: 337–349. [7]

Hartl, D. L., D. E. Dykhuizen, and A. M. Dean. 1995. Limits of adaptation: the evolution of selective neutrality. *Genetics* 111: 655-674. [7]

Hartl, D. L., and C. H. Taubes. 1998. Towards a theory of evolutionary adaptation. *Genetica* 102/103: 525–533. [7]

Higgs, P. G. 1998. Compensatory neutral mutations and the evolution of RNA. *Genetica* 102/103: 91–101. [7]

Hill, W. G. 1969a. On the theory of artificial selection in finite populations. *Genet. Res.* 13: 143–163. [7]

Hill, W. G. 1969b. The rate of selection advance with nonadditive loci. *Genet. Res.* 13: 165–173. [7]

Hill, W. G., and A. Robertson. 1966. The effects of linkage on limits to artificial selection. *Genet. Res.* 8: 269–294. [7]

Innan, H., and W. Stephan. 2001. Selection intensity against deleterious mutations in RNA secondary structures and rate of compensatory nucleotide substitutions. *Genetics* 159: 389–399. [7]

Iwasa, Y. F. Michor, and M. A. Nowak. 2004. Stochastic tunnels in evolutionary dynamics. *Genetics* 166: 1571–1579. [7]

Keightley, P. D., and D. L. Halligan. 2009. Analysis and implications of mutational variation. *Genetica* 136: 359–369. [7]

Kimura, M. 1957. Some problems of stochastic processes in genetics. *Ann. Math. Stat.* 28: 882–901. [7]

Kimura, M. 1969. The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics* 61: 893–903. [7]

Kimura, M. 1983. *The neutral theory of molecular evolution.* Cambridge Univ. Press, Cambridge, UK. [7]

Kimura, M. 1985. The role of compensatory neutral mutations in molecular evolution. *J. Genet.* 64: 7–19. [7]

Kimura, M., and T. Maruyama. 1966. The mutational load with epistatic gene interactions in fitness. *Genetics* 54: 1337–1351. [7]

Kimura, M., T. Maruyama, and J. F. Crow. 1963. The mutation load in small populations. *Genetics* 48: 1303–1312. [7]

Kimura, M., and T. Ohta. 1969a. The average number of generations until fixation of a mutant gene in a finite population. *Genetics* 61: 763–771. [7]

Kimura, M., and T. Ohta. 1969b. The average number of generations until extinction of an individual mutant gene in a finite population. *Genetics* 63: 701–709. [7]

Komarova, N. L., A. Sengupta, and M. A. Nowak. 2003. Mutation-selection networks of cancer initiation: tumor suppresser genes and chromosomal instability. *J. Theor. Biol.* 223: 433–450. [7]

Kondrashov, A. S. 1984. Deleterious mutations as an evolutionary factor. 1. The advantage of recombination. *Genet. Res.* 44: 199–217. [7]

Kondrashov, A. S. 1988. Deleterious mutations and the evolution of sexual reproduction. *Nature* 336: 435–440. [7]

Kondrashov, A. S. 2003. Direct estimates of human per nucleotide mutation rates at 20 loci causing Mendelian diseases. *Hum. Mutat.* 21: 12–27. [7]

Kondrashov, A. S., and J. F. Crow. 1993. A molecular approach to estimating the human deleterious mutation rate. *Hum. Mutat.* 2: 229–234. [7]

Kondrashov, A. S., S. Sunyaev, and F. A. Kondrashov. 2002. Dobzhansky-Muller incompatibilities in protein evolution. *Proc. Natl. Acad. Sci. USA* 99: 14878–14883. [7]

Kouyos, R. D., O. K. Silander, and S. Bonhoeffer. 2007. Epistasis between deleterious mutations and the evolution of recombination. *Trends Ecol. Evol.* 22: 308–315. [7]

Kulathinal, R. J., B. R. Bettencourt, and D. L. Hartl. 2004. Compensated deleterious mutations in insect genomes. *Science* 306: 1553–1554. [7]

Lande, R. 1979. Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement. *Evolution* 33: 234–251. [7]

Lande, R. 1984. The expected fixation rate of chromosomal inversions. *Evolution* 38: 743–752. [7]

Lande, R. 1985. The fixation of chromosomal rearrangements in a subdivided population with local extinction and colonization. *Heredity* 54: 323–332. [7]

Latter, B. D. H. 1965. The response to artificial selection due to autosomal genes of large effect. II. The effects of linkage on the limits to selection in finite populations. *Aust. J. Biol. Sci.* 18: 1009–1023. [7]

Latter, B. D. H. 1966a. The response to artificial selection due to autosomal genes of large effect. III. The effects of linkage on the rate of advance and approach to fixation in finite populations. *Aust. J. Biol. Sci.* 19: 131–146. [7]

Latter, B. D. H. 1966b. The interaction between effective population size and linkage intensity under artificial selection. *Genet. Res.* 7: 313–323. [7]

Lesecque, Y., P. D. Keightley, and A. Eyre-Walker. 2012. A resolution of the mutation load paradox in humans. *Genetics* 191: 1321–1330. [7]

Li, W. H. 1987. Models of nearly neutral mutations with particular implications for nonrandom usage of synonymous codons. *J. Mol. Evol.* 24: 337–345. [7]

Li, W. H., and M. Nei. 1972. Total number of individuals affected by a single deleterious mutation in a finite population. *Amer. J. Hum. Genet.* 24: 667–679. [7]

Lynch, M. 1986. Random drift, uniform selection, and the degree of population differentiation. *Evolution* 40: 640–643. [7]

Lynch, N. 2010. Scaling expectations for the time to establishment of complex adaptations. *Proc. Natl. Acad. Sci. USA* 38: 16577–16582. [7]

Lynch, M., and A. Abegg. 2010. The rate of establishment of complex adaptations. *Mol. Biol. Evol.* 27: 1404–1414. [7]

Lynch, M. 2012a. The evolution of multimeric protein assemblages. *Mol. Biol. Evol.* 29: 1353-1366. [7]

Lynch, M. 2012b. The evolutionary layering of cellular functions and the limits to molecular perfection. *Proc. Natl. Acad. Sci. USA* 109: 18851-18856. . [7]

Lynch, M., J. Conery, and R. Bürger. 1995a. Mutational meltdowns in sexual populations. *Evolution* 49: 1067–1080. [7]

Lynch, M., J. Conery, and R. Bürger. 1995b. Mutation accumulation and the extinction of small populations. *Amer. Natur.* 146: 489–518. [7]

Maruyama, T. 1970. Fixation probabilities of genes in a subdivided population. *Genet. Res.* 15: 221–225. [7]

Maruyama, T. 1977. *Stochastic problems in population genetics*. Springer-Verlag, Berlin. [7]

Maruyama, T., and M. Kimura. 1974. A note on the speed of gene frequency changes in reverse direction in a finite population. *Evolution* 28: 161–163. [7]

McVean, G. A., and B. Charlesworth. 1999. A population genetic model for the evolution of synonymous codon usage: patterns and predictions. *Genet. Res.* 74: 145–158. [7]

Michalakis, Y., and M. Slatkin. 1996. Interaction of selection and recombination in the fixation of negative-epistatic genes. *Genet. Res.* 67: 257–269. [7]

Muller, H. J. 1950. Our load of mutations. *Amer. J. Hum. Genet.* 2: 111–176. [7]

Nachman, M. W. 2004. Haldane and the first estimates of the human mutation rate. *Indian Acad. Sci.* 83: 231-233. [7]

Nagylaki, T. 1992. *Introduction to theoretical population genetics.* Springer-Verlag, New York, NY. [7]

Nei, M. 1969. Heterozygous effects and frequency changes of lethal genes in populations. *Genetics* 63: 669–680. [7]

Nei, M. 1971. Extinction time of deleterious mutant genes in large populations. *Theor. Popul. Biol.* 2: 419–425. [7]

Nei, M., and A. K. Roychoudhury. 1973. Probability of fixation and mean fixation time of an overdominant mutation. *Genetics* 74: 371–380. [7]

Nishino, J., and F. Tajima. 2004. Effect of dominance on heterozygosity and the fixation probability in a subdivided population. *Genes Genet. Syst.* 79: 41–48. [7]

Otto, S. P. 2003. The advantages of segregation and the evolution of sex. *Genetics* 164: 1099–1118. [7]

Otto, S. P., and M. C. Whitlock. 1997. The probability of fixation in populations of changing size. *Genetics* 146: 723–733. [7]

Peck, J. 1994. A ruby in the rubbish: beneficial mutations, deleterious mutations, and the evolution of sex. *Genetics* 137: 597–606. [7]

Pike, D. J. 1969. A comparison of two methods for predicting changes in the distribution of gene frequency when selection is applied repeatedly to a finite population. *Genet. Res.* 13: 117–126. [7]

Poon, A., and S. P. Otto. 2000. Compensating for our load of mutations: freezing the meltdown of small populations. Evolution 54: 1467–1479. [7]

Qureshi, A. W. 1968. The role of finite population size and linkage in response to continued trucation selection. II. Dominance and overdominance. *Theor. Appl. Genet.* 38: 264–270. [7]

Qureshi, A. W., and O. Kempthorne. 1968. On the fixation of genes of large effects due to continued truncation selection in small populations of polygenic systems with linkage. *Theor. Appl. Genet.* 38: 229–255. [7]

Qureshi, A. W., O. Kempthorne, and L. N. Hazel. 1968. The role of finite population size and linkage in response to continued truncation selection. I. Additive gene action. *Theor. Appl. Genet.* 38: 256–263. [7]

Reed, F. A., and C. F. Aquadro. 2006. Mutation, selection and the future of human evolution. *Trends Genet.* 22: 479–84. [7]

Rice, W. R. 2002. Experimental tests of the adaptive significance of sexual recombination. *Nature Rev. Genet.* 3: 241–251. [7]

Robertson, A. 1960. A theory of limits in artificial selection. *Proc. Royal Soc. Lond. B* 153: 234–249. [7]

Robertson, A. 1962. Selection for heterozygotes in small populations. *Genetics* 47: 1291–1300. [7]

Sella, G., and A. E. Hirsh. 2005. The application of statistical physics to evolutionary biology. *Proc. Natl. Acad. Sci. USA* 102: 9541–9546. [7]

Silvela, L. 1980. Genetic changes with generations of artificial selection. *Genetics* 95: 769–782. [7]

Slatkin, M. 1981. Fixation probabilities and fixation times in a subdivided population. *Evolution* 35: 477–488. [7]

Stephan, W. 1996. The rate of compensatory evolution. *Genetics* 144: 419–426. [7]

Stephan, W., B. Charlesworth, and G. McVean. 1999. The effect of background selection at a single locus on weakly selected, partially linked variants. *Genet. Res.* 73: 133–146. [7]

Stephan, W., and D. A. Kirby. 1993. RNA folding in *Drosophila* shows a distance effect for compensatory fitness interactions. *Genetics* 135: 97–103. [7]

Uecker, H., and J. Hermisson. 2011. On the fixation process of a beneficial mutation in a variable environment. *Genetics* 188: 915–930. [7]

Wallace, B. 1991. *Fifty years of genetic load: an odyssey.* Cornell Univ. Press, Ithaca, NY. [7]

Walsh, J. B. 1982. Rate of accumulation of reproductive isolation by chromosome rearrangements. *Amer. Natur.* 120: 510–532. [7]

Walsh, J. B. 1995. How often do duplicated genes evolve new functions? *Genetics* 139: 421–428. [7]

Weinreich, D. W., and L. Chao. 2005. Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. *Evolution* 59: 1175–1182. [7]

Weissman, D. B., M. M. Desai, D. S. Fisher, and M. W. Feldman. 2009. The rate at which asexual populations cross fitness valleys. *Theor. Pop. Biol.* 75: 286–300. [7]

Weissman, D. B., D. S. Fisher, and M. W. Feldman. 2010. The rate of fitness-valley crossing in sexual populations. *Genetics* 186: 1389–1410. [7]

Whitlock, M. C. 2003. Fixation probability and time in subdivided populations. *Genetics* 164: 767–779. [7]

Whitlock, M. C., and R. Gomulkiewicz. 2005. Probability of fixation in a heterogeneous environment. *Genetics* 171: 1407–1417. [7]

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159. [7]

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc. 6th Internat. Cong. Genetics* 1: 356–366. [7]

Wright, S. 1938. The distribution of gene frequencies under irreversible mutation. *Proc. Natl. Acad. Sci. USA* 24: 253–259. [7]

Wright, S. 1949. Adaptation and selection. *In* G. L. Jepson, G. G. Simpson, and E. Mayr (ed.), *Genetics, paleontology and evolution,* pp. 365–389. Princeton Univ. Press, Princeton, NJ. [7]

Zeng, Z. B., H. Tachida, and C. C. Cockerham. 1989. Effects of mutation on selection limits in finite populations with multiple alleles. *Genetics* 122: 977–984. [7]

**Figure 7.1.** Probability of fixation (solid line) and lifetime contribution to heterozygosity (dashed line) of a new mutant allele with additive effects on fitness as a function of $4N_e s$ (using Equations 7.11 and 7.12), both relative to the neutral expectation.

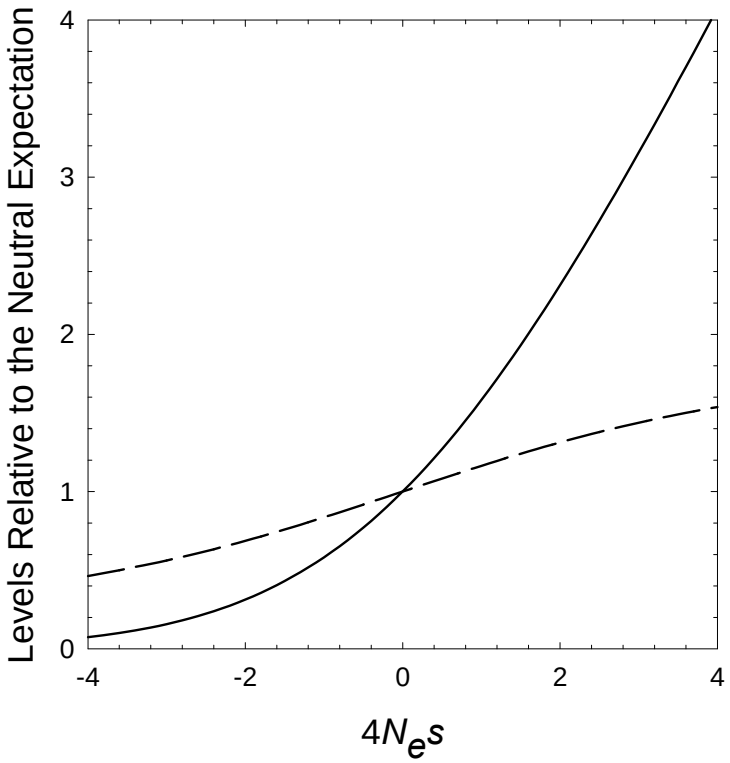

**Figure 7.2**. Average number of generations that a new mutation spends within different frequency classes conditional on going to fixation (above) or conditonal on being lost (below), given as a function of the scaled selection parameter $S = 4N_e s$ (inset values), obtained using Equations 7.13a,d, with $N = N_e = 1000$. Note that in each case, the results are identical for beneficial and deleterious mutations with the same absolute values of $s$. With $N_e \neq N$, the results must be multiplied by $N_e/N$.

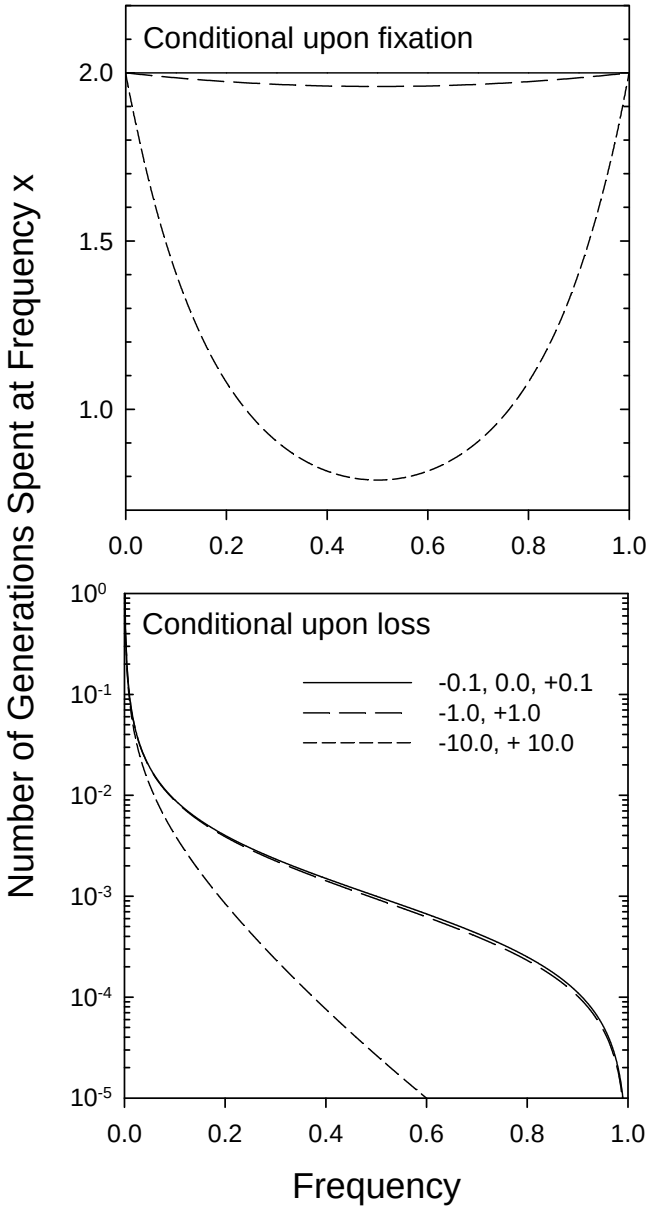

**Figure 7.3**. The influence of drift on the probability of fixation of alternative alleles in a pair of populations starting from an identical state. A diallelic locus under additive selection is considered. The shaded area is the region of $p_0$ (the initial frequency of **A**) and $4N_e s$ space where the probability that isolated populations are fixed for alternative alleles under selection and drift is higher than under drift alone. In this region, parallel selection increases the amount of evolutionary indeterminism relative to drift alone.

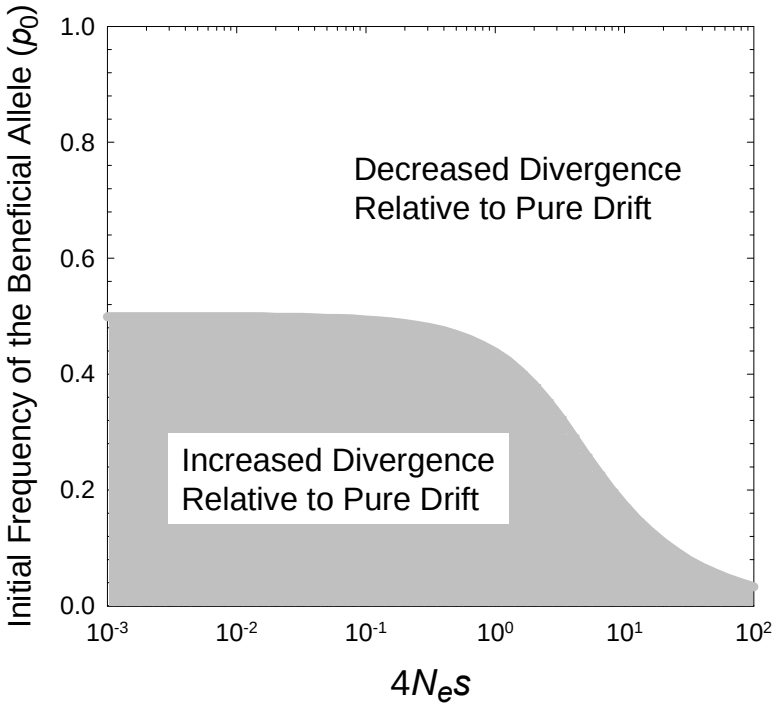

**Figure 7.4**. Ratios for the fixation probabilities and expected times to fixation for a newly arisen overdominant mutation relative to the expectation for a neutral mutation, as a function of the equilibrium frequency expected in a population of infinite size $\widetilde{p} = s_2/(s_1 + s_2)$, where the fitnesses are $1 - s_1$, 1, and $1 - s_2$ (the former being for the mutant homozygote). Each curve gives results for a different value of $N_e(s_1 + s_2)$, a measure of the ratio of the overall power of selection to drift, where $N_e$ is the effective population size. For any value of $N_e(s_1 + s_2)$, the probability of fixation increases with the magnitude of selection against the alternative homozygote, as this defines the selective advantage of the novel allele in the heterozygous state. From Nei and Roychoudhury (1973).

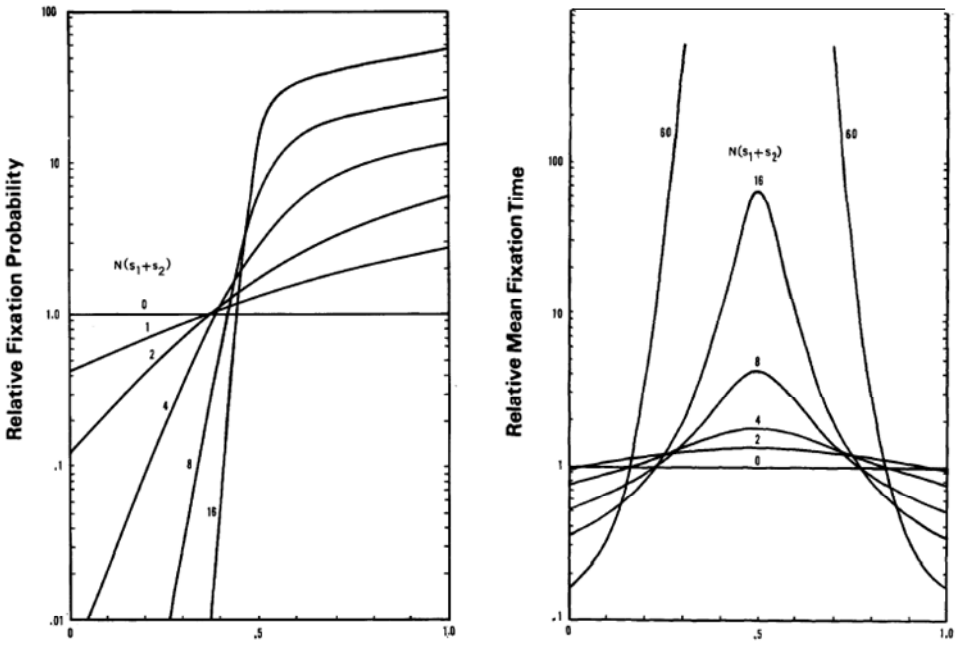

**Figure 7.5**. The probability of fixation of a newly arisen underdominant mutation, relative to the neutral expectation of $1/(2N)$, with selective disadvantage $s$ in the heterozygous state and advantage $t$ in the homozygous state. From Walsh (1982).

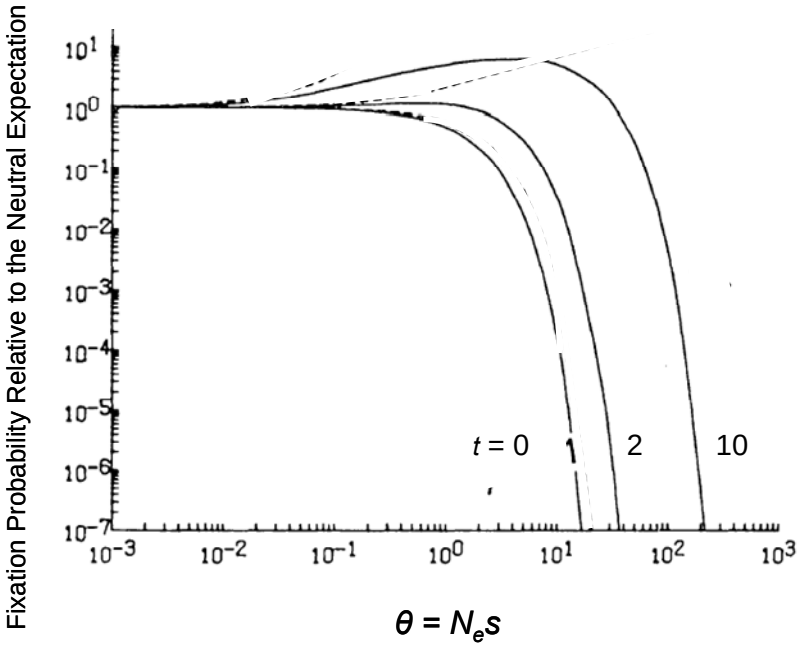

**Figure 7.6**. Stationary distributions of allele frequencies under the joint forces of mutation, selection, and random genetic drift. An absolute population size of $N = 2000$ is assumed with $N_e = N$.