Genetic drift is of considerable importance in most artificial selection experiments, which tend to have very small effective population sizes. We have previously considered the short-term implications of drift in generating variance around the expected response (Chapter 6), which occurs even in the absence of major changes in allele frequencies. This chapter examines the implications for long-term response, considering how drift interacts with selection and how new mutations contribute to the selection response. Throughout, we generally restrict attention to directional selection, considering stabilizing selection in Chapter 13. Given that both initial variation and new mutation can contribute to a selection response, we will use the term long-term response to refer to that component of the response expected from the initial variation in the population at the start of selection. Eventually, all the initial variation will be exhausted and the response from this component will have reached a selection limit. The actual response, however, can continue past this limit due to the input from new mutation, eventually reaching (under constant selection) an asymptotic response, wherein the response reaches a steady-state value due to the contribution to the genetic variance from new mutation is balanced by its removal from drift and selection.

This chapter covers a wide variety of topics, starting with a review of allele frequency changes under the joint actions of drift and selection. Next, we examine the subtle (but very important) effect of selection in decreasing the effective population size over that expected from the number of reproducing individuals. This reduction occurs because with selection on a heritable character, a few families (and their relatives) end up making most of the contribution to the total response. Thus, individuals that survive selection to reproduce are closer relatives than expected by chance, increasing the rate of inbreeding. With the background provided by these two areas in hand, we then turn to considerations of the ultimate contribution from the initial genetic variation in a population and expected asymptotic response due to the contribution of new mutations.

Historically, the theory of the expected long-term response ignored the effects of mutation by considering only the expected response that can be generated given the initial genetic variation present, and we consider long-term response (in the absence of mutation) first. Our treatment follows this historical development by first considering Robertson’s theory for the expected response using only the initial variation and then review how well Robertson’s predictions are supported by experimental results. One important corollary of Robertson’s theory is that there is an optimal selection intensity, as with a fixed number of measured individuals as we increase the selection intensity (and hence the short-term response), we do so at the expense of reducing the effective population size, which reduces the long-term response. We finish our discussion of Robertson’s theory by considering extensions to allow for linkage and various aspects of population structure (such as family selection, drift in the base population, and selection in a subdivided population). We conclude this chapter by considering the effects of mutation contribution.
CHAPTER 26
THE POPULATION GENETICS OF SELECTION AND DRIFT

Drift and Selection at a Single Locus

Most of the theory of the interaction between selection and drift was developed for a single diallelic locus under viability selection. We start by reviewing these results, which can be applied directly to a QTL under selection by using the fitnesses given by Equation 11.4. Throughout we ignore the effects of mutation. Under this model the change per generation in allele frequency can be thought of as the sum of the change under selection plus the change under drift,

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

where $\Delta p_s$ is given by Equation 9.1b. Since allele frequencies change by random sampling, $\Delta p_d$ (the per generation change due to drift) is a random variable. Under drift there is no tendency for a net change in allele frequency, hence $E(\Delta p_d) = 0$. The simplest measure of the strength of drift is the expected variance in gene frequency change due to drift, $\sigma^2(\Delta p_d)$.

If $p$ is the frequency of allele A in gametes, then under the standard Wright-Fisher model of genetic drift, the sample variance follows from the binomial distribution,

$$\sigma^2(\Delta p_d) = p(1-p)/2N_e$$

(26.1)

where $N_e$ is the variance effective population size (Chapter 2). If $\sigma^2(\Delta p_d)$ is small relative to $\Delta p_s$, allele frequency changes are not dramatically different from changes under selection in an infinite population, while drift dominates when $\sigma^2(\Delta p_d)$ is large relative to $\Delta p_s$. Consider additive fitnesses, where the genotypes aa : Aa : AA have fitnesses 1 : 1 + s : 1 + 2s, and $p = \text{freq}(A)$. If selection is weak ($|s| << 1$), then $\Delta p_s \simeq s p(1-p)$, which implies $|\Delta p_s| >> \sigma^2(\Delta p_d)$, and hence selection dominates drift, when $2N_e |s| >> 1$. Conversely, drift is expected to dominate selection when $2N_e |s| << 1$.

This argument is not quite precise because the variance is only a rough indicator of the sampling properties of a distribution. Diffusion theory, an approach used to approximate the behavior of certain well-behaved random processes, gives an essentially complete solution to the dynamics of a diallelic locus under drift and selection. Appendix 1 provides an introduction to diffusion theory, which will also be quite useful when we analyze phenotypic evolution models (Chapters 21 and 22).

A quantity of particular interest is $u$, the probability that an allele is fixed. In an infinite population, $u = 1$ for an allele favored by selection (provided it is not overdominant). In a finite population, $u < 1$ and depends on (among other things) the initial frequency $p_0$ and the effective population size $N_e$. Denote by $u(p_0)$ the probability that an allele starting at initial frequency $p_0$ becomes fixed. The probability of fixation of a neutral alleles depends only on its initial frequency,

$$u(p_0) = p_0$$

(26.2)

and hence is independent of population size. This is not the case for an allele under selection. For additive selection,

$$u(p_0) \simeq \frac{1 - e^{-4N_e sp_0}}{1 - e^{-4N_es}}$$

(26.3a)

$$\simeq p_0 + 2N_e sp_0(1-p_0) \quad \text{when } 2N_e |s| \leq 1$$

(26.3b)

Equation 26.3a, due to Kimura (1957), is derived using diffusion theory in Appendix 1 and values are plotted in Figure 26.1. Equation 26.3b, due to Robertson (1960), uses the
approximation $e^{-x} \simeq 1 - x$ for $|x| \ll 1$ to further simplify Kimura’s result. Finally, note that from Equation 26.3b that

$$u(p_0) \simeq p_0 \quad \text{if} \quad 2N_e|s| << 1 \quad (26.3c)$$

Hence (recalling Equation 26.2), the allele essentially behaves as if were neutral over all allele frequencies. Alleles with $s \neq 0$, but $2N_e|s| << 1$, are called effectively neutral to reflect this fact.

Even when an allele is strongly selected, drift is important when its frequency is near zero or one. Equation 26.3b implies that the probability of fixation starting with a single copy ($p_0 = 1/[2N]$) of an advantageous allele is approximately $2s(N_e/N)$ when $4N_e s >> 1$. Hence a favored allele introduced as a single copy is usually lost by drift. However, once the frequency of a strongly favored allele becomes sufficiently large, fixation is almost certain. Substituting into Equation 26.3a, it is found that if $p \geq 1/(2N_e s)$, the probability of fixation exceeds 0.70, while if $p \geq 1/(N_e s)$, the probability of fixation exceeds 0.93 (Robertson 1960). If $A$ initially increases by drift, it reaches a threshold frequency above which deterministic selection takes over, rapidly increasing its frequency towards one. Near frequency one, drift again takes over, fixing the allele much more rapidly than expected under deterministic selection.

![Figure 26.1](image.png)

**Figure 26.1.** Probability of fixation of an additive allele ($h = 0$) as a function of $N_e s$ and initial allele frequency $p_0$ (using Equation 26.3a).

More generally, if the genotypes $aa : Aa : AA$ have fitnesses $1 : 1 + s(1 + h) : 1 + 2s$, the fixation probability of allele $A$ is approximately

$$u(p_0 | s, h) \simeq \frac{\int_{0}^{p_0} e^{G(x)} \, dx}{\int_{0}^{1} e^{G(x)} \, dx} \quad (26.4a)$$

where

$$G(x) = -4N_e s x [1 + h(1 - x)] \quad (26.4b)$$

If there is inbreeding due to mating of close relatives (in addition to the amount of inbreeding generated by drift), Caballero and Hill (1992) show that Equation 26.4a still holds, but now with

$$G(x) = -4N_e s x [2f + (1 - f)(1 + h(1 - x))] \quad (26.4c)$$
Here $f$ is a measure of the amount of departure of genotypes from Hardy-Weinberg expectations given the current allele frequency, determined from the expected frequency of heterozygotes, $2p(1 - p)(1 - f)$. Kimura and Crow (1963) show that finite population size introduces a departure of genotype frequencies from the expected Hardy-Weinberg values with $f = -1/(2N - 1)$, which (for large $N$) is generally taken as being zero.

As mentioned above, the most frequently required fixation probability is that for a new mutant introduced as a single copy, $p_0 = 1/(2N)$. For a random-mating population,

$$u \left( \frac{1}{2N} \right) \approx \frac{2N_e s (1 + h)}{N} \quad (26.5a)$$

For a complete recessive ($h = -1$), the approximation leading to Equation 26.5a breaks down and higher order terms in the approximation of Equation 26.4a must be considered. This leads to

$$u \left( \frac{1}{2N} \right) \approx \sqrt{\frac{N_e s}{\pi}} \quad (26.5b)$$

Caballero et al (1991), Equations 26.5a and b assumed random mating, while Caballero and Hill (1992), and Caballero (1996) show that with partial inbreeding, Equation 26.4c implies that the fixation probability now becomes

$$u \left( \frac{1}{2N} \right) \approx \frac{2N_e s [2f + (1 - f)(1 - h)]}{N} \quad (26.5c)$$

In particular, with partial inbreeding the fixation probability for a complete recessive becomes

$$u \left( \frac{1}{2N} \right) \approx \frac{4N_e f s}{N} \quad (26.5d)$$

Note that even with a small amount of inbreeding, this probability is considerably higher than that under random mating (Equation 26.5b). Using this result, Caballero et al (1991) show that mating sibs following selection gives a much higher probability of fixing beneficial recessive alleles than random mating individuals following selection. This mating scheme has little effect on the fixation probabilities of beneficial additive alleles and hence can be used without reducing the fixation probabilities for other types of alleles.

Otto and Whitlock (1997) review results for fixation probabilities in populations of changing size, finding 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. Finally, a number of additional diffusion results for a diallelic locus are given in Appendix 1. Unfortunately, diffusion results for multiple alleles and/or multiple loci are generally unavailable.

**Fixation Probabilities for Alleles at a QTL**

We can translate the above results into fixation probabilities at QTLs by using the results from Chapters 9 and 11 to obtain the corresponding fitnesses associated with a particular QTL of interest. When the locus has only a small effect on the character, the fitnesses given by Equation 11.4 for a QTL under directional selection can be used in conjunction with Equation 26.4a to obtain fixation probabilities. If the allele displays no dominance in the character ($k = 0$), then (assuming no linkage effects and/or epistasis) our previous results imply that the probability of fixation exceeds 0.7 when

$$N_e s p = N_e |\bar{\tau}| p \left| \frac{|d|}{\sigma_z} \right| \geq 1/2 \quad (26.6a)$$
and exceeds 0.93 when this quantity exceeds 1. We can rearrange Equation 26.6a to show that the fixation probability exceed 0.7 when the allele frequency is sufficiently large,

\[ P > \frac{\sigma_z}{|a|} \frac{1}{2N_e \tau} \]  

(26.6b)

Hence, if the product of initial allele frequency and the standardized allelic effect \( p|a|/\sigma_z \) is sufficiently small, drift dominates even if selection on the character is strong. With low values of \( N_e \tau \), only alleles of large effect and/or high frequency are likely to be fixed by selection. As \( N_e \tau \) increases, favored alleles with smaller effects and/or lower frequencies are more likely to become fixed. Much of the theory developed in this chapter assumes the infinitesimal model, in which case \( a/\sigma_z \) approaches zero (in the infinitesimal limit) and each individual allele essentially behaves as if it is neutral, which greatly simplifies the dynamics of the genetic variances under selection. Finally, the careful reader might recall from Chapter 4 that with truncation selection in a finite population, sampling causes fluctuations in \( \tau \). This additional complication need not overly concern us, as Hill (1969a, 1985) and Kojima (1961) show that the error introduced by assuming a constant \( \tau \) is generally small.

### Expected Allele Frequency in a Particular Generation

For predicting the response to selection, we are especially interested in \( E(p_t) \), the expected allele frequency at time \( t \). While exact results can be obtained from probability transition matrices (Hill 1969a), and good approximations from diffusion theory (Maruyama 1977, Ewens 1979) 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. Assume the infinitesimal model: selection-induced changes in allele frequencies are negligible. In a finite population, drift reduces the selection response by reducing the expected heterozygosity each generation. Thus, the expected allele frequency change for an allele under additive selection when drift dominates is

\[ E(\Delta p_t) \approx sE[p_t(1 - p_t)] \approx sp_0(1 - p_0) \left( 1 - \frac{1}{2N_e} \right)^t \]  

(26.7)

where \( p_0 \) is the initial allele frequency. The last approximation follows by noting that \( E[p_t(1 - p_t)] = H_t \) (the expected heterozygosity in generation \( t \)) and applying Equation 3.6. Note that, strictly speaking, in the infinitesimal model \( s \) approaches zero so that selection-induced changes in \( p \) are also be close to zero. Drift, on the other hand, changes \( H_t \) each generation. Thus, this analysis assumes that the changes in \( H_t \) due to drift are much more significant than changes in \( p \) due to selection. Summing over generations,

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

\[ \approx p_0 + 2N_e \left( 1 - e^{-t/2N_e} \right) s p_0(1 - p_0) \]  

(26.8a)

where the last step follows using the approximation

\[ \sum_{j=0}^{t} \left( 1 - \frac{1}{2N_e} \right)^j \approx 2N_e \left( 1 - e^{-t/2N_e} \right) \]  

(26.8b)

More generally, if the genotypes \( \text{aa: Aa: AA} \) have fitnesses \( 1:1+s(1+h):1+2s \), then for \( N_e|s| \) and \( N_e|h| \) small,

\[ E(p_t) \approx p_0 + 2N_e sp_0(1 - p_0) \left[ 1 - e^{-t/2N_e} \right] + \frac{h(1 - 2p_0)}{3} \left( 1 - e^{-3t/2N_e} \right) \]  

(26.9)
Note that fixation probabilities can be directly obtained from these approximations. An allele is either ultimately fixed \( (p_\infty = 1) \) or is lost \( (p_\infty = 0) \), giving

\[
E(p_\infty) = 1 \cdot u(p_0) + 0 \cdot [1 - u(p_0)] = u(p_0)
\]

Taking the limit (as \( t \to \infty \)) of Equation 26.9 gives a general expression for the probability of fixation under weak selection,

\[
u(p_0) \simeq p_0 + 2N_e s p_0 (1 - p_0) \left( 1 + \frac{h(1 - 2p_0)}{3} \right) \tag{26.10}\]

This can also be obtained by using a Taylor series expansion of Equation 26.4a. Equations 26.8a and 26.9 (the later for the special case of complete dominance) are due to Robertson (1960). Hill (1969a,b) found that Robertson's approximations are reasonable when \( N_e |s| \) is small. The general versions (Equations 26.9 and 26.10) are due to Silvela (1980).

A case of some interest is the effect of drift on a selectively overdominant locus. Recall that overdominance can be generated when directional artificial selection is balanced by natural selection (see Chapter 11). In an infinite population, selection retains both alleles, while drift ultimately fixes one of the alleles in a finite population. How much does selection retard the expected time to fixation relative to a population under only drift? Surprisingly, Robertson (1962) found that in some cases selection decreases the expected time to fixation of an allele at an overdominant locus. If the expected equilibrium frequency is extreme, e.g. \( \hat{p} < 0.2 \) or \( \hat{p} > 0.8 \), an allele starting at \( \hat{p} \) is often fixed faster under selection plus drift than under drift only (Figure 26.2). Selection for overdominants in this case speeds, rather than retards, the removal of heterozygosity. While this appears counterintuitive, it makes sense if we consider that the consequence of selection is to keep the allele frequency fairly close to its equilibrium value. With an extreme value (say near 0) of \( \hat{p} \), this means that instead of the allele drifting to moderate frequencies, it stays at low frequencies, decreasing the time required to be lost by drift. This point is discussed in greater detail in Hill and Robertson (1968).

---

**Figure 26.2.** Ratio of the expected time to fixation at an overdominant locus relative to the time to fixation under drift only. The fitnesses are \( 1 - s_1 : 1 : 1 - s_2 \), and the population starts at the equilibrium frequency \( \hat{p} = s_2 / (s_1 + s_2) \). The retardation factor is defined as the ratio of the expected time to fixation with drift and selection over the expected time with drift alone.

Note that this can be less than one (the time to fixation is faster with selection and drift than...
with drift alone) with extreme values of $\hat{p}$ (less than 0.2 or greater than 0.8). After Robertson (1962).

The Cohan Effect: Increased Divergence Under Uniform Selection

Selection is generally considered to increase the determinism of a system. Suppose two replicate populations are segregating alleles $A$ and $a$ at a particular locus with $p = \text{freq}(A) = 0.25$. Under pure drift, Equation 26.2 implies that the probability replicate one is fixed for $A$ and replicate two for $a$ is $0.25 \times (1 - 0.25) = 0.1875$, with the same probability for the alleles being fixed in the opposite pattern. Hence the probability that two replicate lines are fixed for alternate alleles is 0.375. Now suppose that $A$ is favored by selection. Cohan (1984b) showed that when selection is weak to moderate the probability of divergence between replicate populations often increases relative to pure drift. For example, if $N_e s = 1/2$, Equation 26.3a gives the fixation probability of $A$ as 0.46, implying the probability of fixing alternate alleles is $2 \times 0.46 \times 0.54 = 0.496$. Divergence in this case is increased by the interaction between selection and drift. We refer to this as the Cohan effect.

In general, the probability of fixing alternate alleles in two replicates is $2u(p) \left[1 - u(p)\right]$. Under pure drift, $u(p) = p$, giving $2p(1 - p)$, which is maximized when $p = 1/2$. Thus the probability of divergence is increased by selection if $u(p)$ is closer to $1/2$ than $p$, which can be formally stated as

$$|u(p) - 1/2| < |p - 1/2|$$

Since $u(p) > p$ for a selectively-favored allele, $p < 1/2$ is a required condition for increased divergence. The probability of divergence under drift plus selection equals that under pure drift only when the initial frequency satisfies $\bar{p} = 1 - u(\bar{p})$. If $p$ is below $\bar{p}$, selection interacting with drift increases the amount of divergence relative to pure drift. Figure 26.3 shows that under additive selection the conditions for this to occur are not very restrictive. See Lynch (1986) for further discussion of how often the Cohan effect might occur.

![Figure 26.3](image)

**Figure 26.3.** The Cohan effect. 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 lines are fixed for alternative alleles is higher under selection and drift than under strict drift. In this region, selection increases the amount of indeterminism relative to drift alone.

Results for Two Loci: the Hill-Robertson Effect

Linkage introduces additional complications in computing fixation probabilities. Suppose
the AB gamete has the highest fitness, but is initially rare. If a population is formed by sampling a finite number of individuals, there is a positive probability that the sample lacks AB gametes. If such gametes are absent, they can only be generated by recombination in Ab/ab heterozygotes. Let \( x_2 \) and \( x_3 \) denote the gamete frequencies of Ab and aB. If the recombination frequency between loci is \( c \), then the expected probability of AB being generated is \( (2x_2x_3) (c/2) \), the frequency of the Ab/aB heterozygote times the probability that a randomly chosen gamete from that heterozygote is AB. Since \( x_2x_3 \leq 1/4 \), the expected number of AB gametes generated in any generation is bounded by \( (2N)(c/4) \). It follows that if \( Ne < 2 \), then less than one AB gamete is expected each generation, implying that this gamete must have a significant fitness advantage over other gametes to overpower drift. Thus, if \( Ne \) is small (or \( x_2 \) and/or \( x_3 \ll 1 \)), one gamete type can become fixed before any AB gametes are generated by recombination, even if the AB gamete has a major effect on fitness. Latter (1966b) develops approximate expressions for the probability that a population in which a favored gamete is initially absent becomes fixed for that gamete and for the mean waiting time to the first appearance of such a gamete.

While there is no general expression for the probability of fixation when two or more loci effect fitness, a number of results were developed in the important (and often neglected) paper of Hill and Robertson (1966). They were able to obtain a weak-selection approximation for the probability of fixation for the following case. Let two diallelic loci (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. Assuming completely additive selection (no dominance or epistasis), where each copy of A adds \( s_1 \) and each copy of B adds \( s_2 \) to total fitness, the probability that A is fixed is approximately

\[
u(p_0) \approx p_0 + 2Ne s_1 p_0 (1-p_0) + \frac{2Ne s_2}{2Ne c + 1} D_0\]  

(26.11)

provided that \( 2Ne s_1, 2Ne s_2 < 1 \). Comparing the two-locus approximation to the single-locus approximation (Equation 26.3b) so that the probability of fixation can be increased or decreased (relative to the single-locus result) depending on the sign of the initial gametic-phase disequilibrium, \( D_0 \).

When selection is strong (\( Ne s_1 \) and/or \( Ne s_2 \ll 1 \)), things get rather interesting. By simulation, Hill and Robertson found that the effect of linkage is to decrease the probability of fixation of an advantageous allele relative to the single-locus result. In particular, if A and B are favored alleles, linkage (i.e. \( c < 1/2 \)) has little effect on the probability of fixation of the ab gamete. However, the probability of fixation of AB and aB gametes increases (relative to the product of single locus fixation) at the expense of decreasing the probability of fixation of the AB gamete (Latter 1965, Hill and Robertson 1966). This decrease is maximized when \( Ne c \) is small and both loci have the same effect (e.g., \( s_1 = s_2 \)). This is an important point, as most simulation studies of the effects of linkage on selection response in a finite population size assume loci of equal effect (e.g., Fraser 1957; Latter 1965, 1966a,b; Gill 1965a,b,c; Qureshi and Kempthorne 1968; Qureshi and Qureshi 1968), which has the effect of inflating the importance of linkage.

Hill and Robertson interpret their result in terms of effective population size: selection on a locus reduces the effective population size experienced by linked loci. Felsenstein (1974) refers to this reduction in \( Ne \) due to selection on other loci as the Hill-Robertson effect. One implication of the Hill-Robertson effect is that the consequence of selection is to make linked loci more neutral due to the reduction in \( Ne \) (Birky and Walsh 1988). For example, consider a locus at which advantageous alleles are being generated continually by mutation. The consequence of linking this locus to another locus under selection (either positive or negative) is to reduce the rate of substitution of advantageous alleles. Somewhat counterintuitively, the average substitution rate of a locus generating deleterious alleles is increased by linking that locus to
another locus generating deleterious alleles (Birky and Walsh 1988). However, this makes sense in the context of the Hill-Robertson effect: the linked locus under selection reduces the effective population size of the locus being followed, reducing $4N_e|s|$ and hence making the allele behave in a more neutral fashion. Note that loci under artificial (or natural) selection for a particular character are only a subset of loci under selection, as natural selection will be acting on a wide suite of loci, for example by removing deleterious mutations. The net effect of such selection can be nontrivial in its effect on $N_e$.

We consider the cumulative reduction in $N_e$ for loci under constant directional selection in the next section. There is, however, a potential further reduction in $N_e$ due to natural selection (such as the removal of new deleterious mutations). If $\sigma^2(W)$ denotes the variance in fitness of diploid individuals at a particular site, Barton (1995) showed (for loose linkage) that the effective population size experienced at a site recombination frequency $c$ away is

$$
\frac{N_e}{N} \simeq \left(1 + \frac{\sigma^2(W)}{2c^2}\right)^{-1}
$$

(26.12)

With very tight linkage, Barton showed that the fixation probability is not simply predictable from the reduction in $N_e$, being strongly influenced by the initial disequilibrium.

Following up on this observation, Otto and Barton (1997) showed that alleles at modifier loci which increase the recombination rate also increase the probability of fixation of favored alleles at a linked locus under selection. This can result in these recombination modifiers hitchhiking along to fixation with the favored mutations. Such modifiers are favored because under low recombination, the selection coefficient on a particular mutation affecting a character under selection also depends on the selection coefficients at linked loci. Thus, the fate of a particular mutant is highly dependent upon the background in which it arose. As recombination increases, the fate of a mutation becomes increasingly uncoupled from the fate of the background in which it arose. Otto and Barton suggest this provides an explanation for the evolution of increased recombination frequencies (and hence the evolution of sex). Otto and Barton’s theory makes the prediction that recombination rates may increase in selected populations relative to unselected controls. As the following example illustrates, this has indeed been seen in some experiments.

---

**Example 26.1.** Korol and Iliadi (1994) subjected a *Drosophila melanogaster* population to divergent selection for positive and negative geotaxis (positive geotaxis flies have an increased tendency to fly up in a vertical maze, negative geotaxis flies display a tendency to fly down in a maze). Recombination frequencies were scored on the unselected controls and the positively (geo$^+$) and negatively (geo$^-$) selected lines, with chromosomes II, III, and X scored after 36, 40, and 44 generation of selection (respectively). Over the scored regions, the map distance in the geo$^+$ line increased by a total of 78 cM relative to the unselected controls, while the geo$^-$ line increased by 66 cM. Presumably these increases in recombination frequencies result from the increased probability of fixation of favorable mutations that are linked to modifiers increasing recombination frequencies. Korol and Iliadi review other experiments showing increased recombination frequencies as a correlated response to artificial selection on other characters.

---

**THE EFFECT OF SELECTION ON EFFECTIVE POPULATION SIZE**

Selection has an obvious effect on effective population size, in that if a fraction $p$ of $T$ scored
individuals are allowed to reproduce, the number of parents becomes $N = pT$. For a fixed number of scored individuals, increasing the intensity of selection (i.e., decreasing $p$) decreases $N$ and hence $N_e$. Thus, (all else being equal) the stronger selection, the lower the effective population size. Selection also has a more subtle (and cumulative) effect in that it reduces the effective population size below that of an unselected control population with the same number of parents (so that $N_e < Tp$). This phenomenon was initially mentioned by Morley (1954), who noted in sheep flocks exposed to selection that “the genetically superior individuals will tend to be most inbred”. To see how selection reduces $N_e$, recall from Chapter 2 that one of the assumptions of an ideal population (where the actual size $N$ equals the effective size $N_e$) is that all parents have an equal chance of contributing offspring. With a character under selection this is no longer true, as superior families contribute more offspring to the next generation than inferior families, inflating the offspring variance, which reduces the $N_e$. This follows since (Equation 2.14) $N_e = (N - 1/2)/(\sigma_k^2/4 + 1/2)$, where $\sigma_k^2$ is the variance in offspring number. If the number of offspring follows a Poisson distribution, then $\sigma_k^2 = 2$ and $N_e = N - 1/2 \approx N$. However, if some parents contribute a disproportionate number of offspring, $\sigma_k^2 > 2$ and $N_e < N$. The more disproportionate the contribution from some families, the larger the variance and hence the smaller $N_e$. Thus, a single generation of selection reduces $N_e$ by inflating $\sigma_k^2$ over that for a population not under selection. A second factor, and the major complication in computing $N_e$ for a population under selection, is that continued selection has a cumulative effect in reducing the variance beyond the single-generation effect. This occurs because for a heritable character under selection, parents pass on some of their ability to have an increased contribution to their offspring which inflates $\sigma_k^2$, further reducing in $N_e$. This reduction becomes more pronounced as either heritability and/or selection intensity increases.

The Expected Reduction in $N_e$ from Directional Selection

While the effective population size due to artificial selection can easily be retrospectively computed from either pedigree information or from the sampling variance in marker allele frequencies (see Chapter 2), predicting the reduction of $N_e$ in advance is considerably more difficult. The exact value of $N_e/N$ depends on a variety of assumptions about both the family and population structure and on the underlying genetical model (the infinitesimal is typically assumed). Theoretical investigations of the effects of selection on reducing $N_e$ were initiated by Robertson (1961), who gave simple approximations for both the single generation change in $N_e$ and the asymptotic change following many generations of selection. Two different approaches have been used to examine the reduction in $N_e$ — computing the expected variance in gene frequency for an unselected locus in a population under selection (Robertson 1961, Nei and Murata 1966, Caballero 1994, Santiago and Caballero 1995) and computing the rate of inbreeding from the number of ancestors (Burrows 1984a,b; Wolliams 1989; Verrier et al. 1990; Wray and Thompson 1990; Wray et al. 1990, 1994; Wolliams et al. 1994). The former corresponds to the variance effective population size, the latter to the inbreeding effective population size. Strictly speaking, diffusion theory results require usage of the variance effective size (as diffusion approximations use the sample variance in allele frequency). However, as discussed in Chapter 2, inbreeding and variance size are usually equivalent unless the population size is changing over time. While these treatments considered the effective population size on a neutral locus unlinked to loci influencing the traits(s) under selection, the results should be very similar for selected loci under the infinitesimal model, as in this case drift (rather than selection) provides the major impact on allele frequency.

Our treatment follows Santiago and Caballero (1995), who give a general expression for $N_e$ under selection allowing for nonrandom mating. Assuming random mating first, their
expression for \( N_e \) after \( \tau \) generations of selection is

\[
\frac{N_{e,\tau}}{N} = \left[ 1 - \frac{\alpha_1}{2} + \left( \frac{\sigma_k^2}{4} + Q^2 C^2 \right) (1 + \alpha_1) \right]^{-1} \tag{26.13a}
\]

where \( \sigma_k^2 \) is the sampling variance in offspring number in the absence of selection, \( \alpha_1 = -1/(N - 1) \) is a measure of the departure from Hardy-Weinberg due to finite population size, \( Q \), accounts for the cumulative effects of \( \tau \) generations of selection, and \( C^2 \) is the variance in selective advantage among families. For a single generation of selection, \( Q_1 = 1 \), implying that \( \sigma_k^2/4 + C^2 \) is the effect of selection in the current generation, while \( (Q_1^2 - 1)C^2 \) is the cumulative effect of selection in previous generations. As is show below, both \( Q \) and \( C^2 \) are functions of the selection intensity, heritability, and intraclass correlation of sibs. Assuming that \( N \) is large (so that \( \alpha_1 \approx 0 \), and a Poisson distribution of offspring in the absence of selection (\( \sigma_k^2 = 2 \)), Equation 26.13a reduces to

\[
\frac{N_{e,\tau}}{N} = \frac{1}{1 + Q^2 C^2} \tag{26.13b}
\]

Robertson (1961) and Milkman (1978) show that \( C^2 \approx \tau^2 t \) where \( t = \text{Cov}(FS)/\sigma_z^2 \) is the intraclass correlation of full sibs (LW Equation 17.3).

Obtaining the value of \( t \) requires a little care, as selection changes the genetics variances, and hence \( t \). Under the infinitesimal model (Chapters 5, 10) all the change in genetic variances are due to selection-induced linkage disequilibrium. In this case, the additive genetic variance \( \sigma_A^2 = \sigma_a^2 + d \), where \( \sigma_a^2 \) (the additive genic variance) is the additive genetic variance in the absence of disequilibrium (the additive genetic variance in the base population) and \( d \) is the disequilibrium. Recall (Chapter 5) that under the infinitesimal model, the within- and between-family contributions to the additive genetic variance differ, as the within-family contribution (\( \sigma_a^2/2 \)) is not influenced by disequilibrium, while the between-family variance (\( \sigma_A^2/2 + d \)) is. Assuming (for now) the absence of dominance and shared sib environmental effects, the intraclass correlation becomes

\[
t = \frac{\sigma^2_a/2 + d}{\sigma^2_A + \sigma^2_E} = \frac{\sigma^2_a/2 + d}{\sigma^2_a + d + \sigma^2_E} = \frac{h_0^2/2 + d/\sigma^2_z(0)}{h_0^2 + d/\sigma^2_z(0) + (1 - h_0^2)} = \frac{h_0^2/2 + d/\sigma^2_z(0)}{1 - d/\sigma^2_z(0)} \tag{26.14a}
\]

where \( h_0^2 \) and \( \sigma^2_z(0) \) are the heritability and phenotypic variance in the unselected base population. The last equality in the first line of Equation 26.14a follows by dividing through by \( \sigma^2_z(0) \) (the phenotypic variance in the base population) and noting that \( \sigma_E^2/\sigma_z^2(0) = (1 - h_0^2) \), where \( \sigma_E^2 \) here includes all sources of variation besides the additive genetic variance. Thus, the larger \( \tau^2 \) and/or \( h^2 \), the larger \( C^2 \) and the small \( N_e \) becomes. When dominance and/or shared environmental effects are present, \( T \) is also a function of \( \sigma_D^2 \) and \( \sigma_{E(z)}^2 \) (see old Equation 5.17 from chpt 7).

We now have all the results needed to compute the single-generation reduction in \( N_e \). Defining \( \kappa = \sigma_z^2/\sigma_k^2 \) as the reduction in the phenotypic variance in the selected parents, Equation 5.12 gives \( d_1/\sigma^2_z(0) = \kappa h^4/2 \). Substituting into Equation 26.14a gives

\[
t_1 = \frac{h^2(1 - \kappa h^2)/2}{1 - \kappa h^4/2} \tag{26.14b}
\]
Recalling that $Q_1 = 1$ gives the reduction in $N_e$ from a single generation of selection as approximately

$$N_{e,1} \simeq N \left(1 + \frac{\tau^2 (h^2/2)(1 - \kappa h^2)}{1 - \kappa h^4}\right)^{-1}$$

(26.15)

This result was first obtained by Robertson (1961), who did not include the $(1 - \kappa h^4/2)$ term, and corrected by Wray and Thompson (1990).

With selection on a heritable character, gene frequency changes between generations are correlated. As a consequence, the simple sampling variance correction $C^2 = \tau^2 t$ is not sufficient to correct for the effects of selection as it ignores the cumulative effects ($Q_\tau$) of these correlations over $\tau$ prior generations of selection. Santiago and Caballero (1995) show that the cumulative effect $Q_\tau$ is given by

$$Q_\tau = 1 + \frac{G}{2} (1 + r) + \cdots + \left[\frac{G}{2} (1 + r)\right]^\tau = \sum_{i=0}^{\tau} \left[\frac{G}{2} (1 + r)\right]^i$$

(26.16a)

$G$ is the fraction of genetic variance remaining after selection ($G = 1 - \kappa h^2$ under the infinitesimal model) and $r$ is the correlation between the selective values of mates ($r = -1/[N - 1]$ under random matting). In the limit ($\tau \to \infty$), this sum converges to

$$Q = \frac{2}{2 - G(1 + r)}$$

(26.16b)

Robertson assumed that the limiting value of $Q$ is 2, but Equation 26.16b shows that this is an overestimate, and hence $N_e$ is underestimated. $G$, and hence $Q$, decreases as $h^2$ increases, so that the underestimation of $N_e$ using Robertson’s result is most serious when heritability (or selection intensity) is large.

**Example 26.2.** Consider directional truncation selection on a normally-distributed character by selecting the uppermost 20 percent of the population. From Example 2 of Chapter 5, we have that this value of $p$ gives a selection intensity of $\tau = 1.4$ and a reduction in variance of $\kappa = 0.21$. Assuming initial (before selection) values of $h^2 = 0.5$ and $\sigma_z^2 = 100$, Example 2 gives (under the infinitesimal model) equilibrium values of $\widehat{h}^2 = 0.43$ and $\widehat{d} = -17.59$. Hence,

$$G = 1 - \kappa \widehat{h}^2 = 1 - 0.21 \cdot 0.43 = 0.906$$

and

$$\hat{\lambda} = \frac{h^2/2 + \hat{d}/\sigma_z^2}{1 - \hat{d}/\sigma_z^2(0)} = \frac{0.5/2 - 12.59/100}{1 + 12.59/100} = 0.11$$

Hence, $\hat{C}^2 = \hat{\tau}^2 \cdot \hat{\lambda} = 1.4^2 \cdot 0.11 = 0.217$ while (assuming $r \simeq 0$)

$$Q = \frac{2}{2 - G} = \frac{2}{2 - 0.906} = 1.8$$

Giving the equilibrium reduction in $N_e$ as

$$N_e = \frac{N}{1 + Q^2 \hat{C}^2} = 0.59 N$$

Similar calculations using other $p$ values gives
The general prediction that effective population size decreases in selected populations has been examined in a number of *Drosophila* experiments, where inbreeding is estimated directly from parental pedigrees. This prediction has generally been confirmed, with a reasonable fit to Robertson’s theory (McBride and Robertson 1963; Jones 1969a,b; López-Fanjul 1989). It is generally seen that $N_e$ is lowest in lines showing the greatest response to selection, as is expected from Equation 26.15 since these lines have the highest realized heritabilities. Gallego and López-Fanjul (1983) tested a second prediction using selection on sternopleural bristles: since the reduction in $N_e$ occurs because of between-family selection (inflating the between-family variance $\sigma^2_k$), no reduction in $N_e$ is expected under within-full-sib family selection. In accordance with theory, no reduction was observed.

Reproductive fitness often declines during long-term selection experiments. This can result in a further increases the variance in fitness among individuals, which in turn further increases the variance in offspring number contributed by each parent. This increased variance can significantly decrease the effective population size below that predicted by Equation 26.15, which assumes only the variance effects associated with artificial selection. Yoo (1980c) found that differences in fertility were more important in reducing effective population size than the effects of artificial selection during a long-term selection experiment for increased abdominal bristle number in *Drosophila*.

### Example 26.3

Cohan and Hoffmann (1986) examined the divergence between replicate lines of *Drosophila melanogaster* selected for increased resistance to ethanol. The selected lines had a higher between-line variance for characters associated with increased resistance than did the unselected control replicates. This could be explained by reduction in effective population size due to selection and/or by the Cohan effect. The reduction in effective population size, by increasing drift, is expected to increase the between-line variance in any character, selected or unselected. Conversely, the Cohan effect predicts that only characters under selection, or characters controlled by loci tightly linked to these selected characters, should show increased divergence. To distinguish between these, Cohan and Hoffmann examined three unselected characters, and found no differences between the selected and control lines, suggesting the main cause of increased divergence was due to the Cohan effect.

### Molecular Variation is Reduced in Regions of Low Recombination

Early studies comparing levels of DNA sequence variation from different regions of the *Drosophila melanogaster* genomes showed a very interesting pattern. Loci on the fourth chromosome Berry et. al. 1991) and near chromosome telomeres (Begun and Aquadro 1991) showed greatly reduced levels of DNA polymorphism relative to loci in other genomic locations. What these two regions have in common is greatly reduced recombination, suggesting that regions of low recombination will also show reduced levels of genetic variation. This expected correlation was confirmed in *D. melanogaster* (Begun and Aquadro 1992, Aquardo...
et al. 1994, Moriyama and Powell 1996), *D. simulans* (Moriyama and Powell 1996), *D. ananas-sae* (Stephan et al. 1998), mice (Nachman 1997), humans (Nachman et al. 1998) and in the Sea Beet, *Beta vulgaris* (Kraft et al. 1998). Figure 26.4 displays some of this data.

**Figure 26.4.** The correlation between the level of DNA polymorphism at a locus and the recombination frequency in the region in which that locus resides. Left: Data for *Drosophila melanogaster* (after Aquadro et al. 1994). Right: Data for four X-linked locus in the mouse, *Mus domesticus* (After Nachman 1997).

There is broad agreement that this correlation results from the reduction in *N_e* due to selection on linked loci. Most of the observed DNA sequence variation is thought to be neutral or nearly so (Kimura 1983), and the expected amount of polymorphism at a neutral locus is an increasing function of *N_e* (see Equation 2.20b). Loci residing in regions of low recombination are influenced by more linked loci and hence having reduced *N_e*. While selection at linked loci is widely believed to account for the polymorphism-recombination correlation, there is currently an ongoing debate on the exact nature of this selection. It could be due to selective sweeps (Berry et al. 1991; Begun and Aquadro 1991, 1992), the recent action of directional selection driving a favored allele to fixation, which drags along with a linked region of DNA that hitchhikes to fixation along with the favored allele. Alternatively, the reduction in *N_e* could be due to background selection (Charlesworth et al. 1993), the removal of deleterious mutations by natural selection. If selective sweeps are the primary mechanism generating this correlation, one would expect to see a skewed distribution of rare polymorphisms relative to that expected under the strict neutral model, reflecting the fact that the sweep drives a single chromosome type through the population. Background selection on the other hand simply eliminates single chromosomes carrying deleterious new mutations and does not significantly skew the distribution of rare polymorphisms (Braverman et al. 1995). Using this observation to discriminate between models, the preliminary evidence seems to be in favor of background selection (Braverman et al. 1995, Hudson and Kaplan 1995, Charlesworth et al. 1995), although the debate is far from settled.

**DRIFT AND LONG-TERM SELECTION RESPONSE**

Recall that we make the distinction between long-term and asymptotic response. The former is the total response attributable to the existing variation at the time of selection, while the later is the expected eventual rate of response due to the input of new mutation. When the effective population size is small, essentially all of the observed response is due to the initial
variation, with the population reaching an apparent selection limit until the appearance of
new mutations allows further response. In larger populations, however, these two compo-
nents of response (that from initial variation and that from new mutations) become more
difficult to separate, and no selection limit may be observed when in fact all of the initial
variation has become exhausted. Much of the theory of long-term response has ignored
mutation and we examine this first. Even thought this theory is unrealistic over very long
time scales, it often provides a good description of how the population exhausts its initial
variation.

Basic Theory
We expect response to selection in very small populations to be significantly influenced by
drift, showing less total response than expected in larger populations starting with the same
initial genetic variance. A fairly extensive theory examining the effects of drift on long-term
response (the utilization of the existing genetic variation) has been developed, starting with
the extremely influential paper of Robertson (1960). Most of this theory is based on summing
over single-locus results, and we this approach unless stated otherwise (this assumes that
epistasis and linkage effects can be ignored). As before, we first consider a single diallelic
locus (which we index as locus $i$) where the genotypes $aa : Aa : AA$ have genotypic values
(for the character under selection) of $0 : a(1 + k) : 2a$. Let $p_t$ denote the frequency of $A$
at this locus at generation $t$, $\Delta_i(t)$ be the contribution to total response from this locus in generation
$t$, and $u_i(p_0)$ the probability that $A$ is ultimately fixed at this locus given it starts at frequency
$p_0$. The total response is given by summing over all loci,

$$R(t) = \sum_i \Delta_i(t)$$

Under drift, both $p_t$ and $\Delta_i(t)$ are random variables and (assuming genotypes are in Hardy-Weinberg
proportions) are related by

$$\Delta_i(t) = m_i(p_t) - m_i(p_0)$$

(26.17a)

where $m_i(p)$ is given by Equation 11.1a. The expected contribution (at generation $t$) from
this locus is

$$E[\Delta_i(t)] = 2a \left[ E(p_t) - p_0 + k \left( p_t(1 - p_t) - p_0(1 - p_0) \right) \right]$$

(26.17b)

Since $A$ is ultimately either fixed ($p_\infty = 1$) or lost ($p_\infty = 0$), $E(p_t)$ converges to

$$1 \cdot u_i(p_0) + 0 \cdot [1 - u_i(p_0)] = u_i(p_0)$$

while $E[p_t(1 - p_t)]$ converges to zero. Thus, the limiting expected contribution from locus
$i$ is

$$E[\Delta_i(\infty)] = 2a \left[ u_i(p) - p_0 - k \left( p_0(1 - p_0) \right) \right]$$

(26.18a)

Two cases of special interest are when $A$ is additive ($k = 0$), in which case

$$E[\Delta_i(\infty)] = 2a \left[ u_i(p) - p_0 \right]$$

(26.18b)

and when $A$ is recessive ($k = -1$),

$$E[\Delta_i(\infty)] = 2a \left[ u_i(p) - p_0^2 \right]$$

(26.18c)
The variance (and indeed all higher moments) of the total response at the selection limit are easily computed, as \( \Delta_i(\infty) \) takes on only two values:

\[
\Delta_i(\infty) = \begin{cases} 
2a - m_i(p_0) & \text{with probability } u_i(p_0) \\
0 - m_i(p_0) & \text{with probability } 1 - u_i(p_0)
\end{cases}
\]

(26.19)

In particular,

\[
\sigma^2[\Delta_i(\infty)] = E[\Delta_i^2(\infty)] - \left(E[\Delta_i(\infty)]\right)^2 = 4a^2u_i(p_0)[1 - u_i(p_0)]
\]

(26.20a)

With weak selection, \( u_i(p) \approx p \) (i.e., the allelic dynamics are governed by drift), implying

\[
\sigma^2[R(\infty)] \approx 4 \sum a^2 p_0(1 - p_0)
\]

(26.20b)

If all loci are additive, this is simply \( 2\sigma^2_A(0) \), the expected between-line divergence under pure drift (Chapter 3). Under sufficiently strong selection, almost all favorable alleles are fixed and the variance is close to zero as \( u_i(p) \approx 1 \). When selection is moderate to weak, some loci can show a Cohan effect, provided \( u_i(p)[1 - u_i(p)] > p(1 - p) \). If such loci are sufficiently frequent, selection increases the between-line variance relative to drift. This requires both weak selection and that most favored alleles are rare. The variance in response at the selection limit is considered in more detail by Hill and Rasbash (1986a) and Zeng and Cockerham (1990).

Robertson’s Theory of Selection Limits

Equations 26.15-26.20 are fairly general, assuming only Hardy-Weinberg, no epistasis, and that single-locus results can be added across loci. To proceed further, we need explicit expressions for \( u_i(p) \) (to describe the limit) and \( E(p_t) \) and \( E[p_t(1 - p_t)] \) (to describe the...
dynamics). The most complete description, due to Robertson (1960), is for additive alleles, where \( \Delta_i(t) = 2a[E(p_i) - p_0] \). Equation 26.8a gives an approximate expression for the expected allele frequency \( E(p_i) \) under the assumption that the allele has a small effect and phenotypes are normally distributed (i.e., the infinitesimal model). For notational ease, we will drop the expectation notation, but the reader should keep in mind that we are examining the expected response. Recalling from Equation 11.4 that \( s = a\sigma_i/\sigma_z = aS/\sigma_z^2 \), substitution into 26.8a gives the expected response from locus \( i \) after \( t \) generations of selection as

\[
\Delta_i(t) = 2a[E(p_i) - p_0] \approx 2N_e \left( 1 - e^{-t/2N_e} \right) \left( \frac{S\sigma_i}{\sigma_z^2} \right) 2a p_0(1 - p_0) \tag{26.21a}
\]

This can by simplified further by noting that \( 2a^2 p_0(1 - p_0) \) is the initial additive variance contributed by the locus. Since we assumed no gametic-phase disequilibrium, summing over all loci gives

\[
R(t) \approx 2N_e \left( 1 - e^{-t/2N_e} \right) \frac{S\sigma_i^2(0)}{\sigma_z^2} \tag{26.21b}
\]

Note that \( S\sigma_i^2(0)/\sigma_z^2 = S h^2(0) \) is the expected response in the first generation, provided that the conditions for the breeders’ equation hold. Thus,

\[
R(t) \approx 2N_e \left( 1 - e^{-t/2N_e} \right) R(1) \tag{26.22a}
\]

giving an expected limiting total response of

\[
R(\infty) \approx 2N_e R(1) \tag{26.22b}
\]

The careful reader will note that we assumed the phenotypic variance remains relatively constant over time, as would occur if \( h^2 \) is small. Provided this assumption holds, the total expected response is simply \( 2N_e \) times the initial response, as first suggested by Dempster (1955) and formally derived by Robertson (1960).

Alternatively, Equation 26.22a can be derived as follows. Assuming the main force for allele frequency change is drift, Equation 3.2 implies \( \sigma_i^2(t) \approx \sigma_i^2(0)[1 - 1/(2N_e)]^t \). Writing the response in generation \( t \) as \( h^2(t)S = \sigma_i^2(t) S/\sigma_z^2 \), summing over generations and applying Equation 26.8b recovers Equation 26.22a.

Equation 26.22b is an upper limit for total response, which may seem somewhat counterintuitive since it was derived by assuming weak selection. The key is that (everything else being equal) the initial response \( R(1) \) when selection dominates is much larger than when drift dominates, so that \( 2N_e \) times the initial response overestimates the total response when selection dominates.

To see this point, consider the maximal contribution \( \Delta_i^{max} = 2a(1 - p_0) \) from a locus (which occurs when the favored allele is fixed) relative to the predicted contribution \( \Delta_i(\infty) \). From Equation 26.21a, it follows that \( \Delta_i(\infty) = 2N_e 2a^2 p_0(1 - p_0) S/\sigma_z^2 \), giving the ratio of maximal to expected contribution as

\[
\frac{\Delta_i^{max}}{\Delta_i(\infty)} = \frac{1}{2N_e} \frac{2a(1 - p_0)}{2a^2 p_0(1 - p_0) \tau/\sigma_z} = \frac{\sigma_z}{2N_e \tau a p_0} \tag{26.23a}
\]

Thus, when

\[
2N_e \tau a p_0 > \sigma_z \tag{26.23b}
\]

it follows that \( \Delta_i^{max} < \Delta_i(\infty) \) and hence \( 2N_e R(1) \) overestimates the ultimate limit. Further note that when the inequality given by Equation 26.23b holds, further increases in effective
population size do not have a significant effect on increasing the selection limit as $\mu_i(p) \simeq 1$ and hence the contribution from the $i$th locus is $\Delta_i^{max}$. Alternatively, when the inequality given by Equation 26.23b fails, $\Delta_i^{max} > \Delta_i(\infty)$. However, in this case drift is expected to dominate (see Equation 26.6a), so that we do not expect the maximal possible response from each locus, as many favored loci will be lost, rather than fixed.

Another quantity of interest is the expected half-life of response, $t_{0.5}$, the time required to obtain half the final response. Recalling Equation 26.21a and solving $1 - e^{-t_{0.5}/2N_e} = 1/2$ gives the expected half-life as

$$t_{0.5} = N_e \ln 2 \simeq 1.4 N_e \tag{26.24}$$

Again, this is an upper limit with the half-life decreasing as the product $N_e \tau$ increases. An observed half-life considerably below that predicted by Equation 26.24 suggests that a large portion of the response is due to fixation of favorable alleles by selection, as selection (when it dominates) changes allele frequencies much faster than drift.

Equations 26.21-24 rely on a number of assumptions besides additivity: no opposing natural selection, no linkage effects, two alleles per locus, and weak selection. Several authors have examined how well these results hold up when these last two assumptions are relaxed. Hill and Rasbash (1986a) found for diallelic loci that the distribution of allelic effects is relatively unimportant, but differences in allele frequencies can be critical. In particular, increasing effective population size has much more of an effect on the selection limit when favored alleles are rare. This is expected, as common alleles are often over the threshold frequency where their dynamics are largely determined by drift (see Equation 26.6a). Increasing population size lowers this threshold, eventually capturing even rare alleles. Latter and Novitski (1969) and Zeng and Cockerham (1990) examined the effects of multiple alleles, finding that the expected limit and half-life results given by Equations 26.22b and 26.24 are reasonable when selection is weak. As $N_e \tau$ increases, $R(\infty) / R(1)$ becomes highly dependent on the number and frequencies of alleles at each locus (Chapter 11). In general, it increases as the number of alleles increases, decreases as $N_e \tau$ increases, all the while remaining bounded by $2N_e$. Likewise, $t_{0.5}$ decreases as $N_e \tau$ increases, but is rather insensitive to the number of alleles.

With dominance, analytic results for the limit and half-life ($R(\infty)$ and $t_{0.5}$) are more complicated that those for no dominance. Strictly recessive alleles have received the most study. In this case, the selection limit can considerably exceed $2N_e$ times the initial response when the character is controlled by a large number of rare recessives (Robertson 1960). Additive genetic variance increases, often considerably, as these recessives increase in frequency, so this result should not be surprising (Chapter 3). With weak selection, the half-life varies from approximately $N_e$ when $p \simeq 1$ to approximately $2N_e$ when $p \simeq 0$ (Robertson 1960). Again, as $N_e \tau$ increases, half-life decreases. Even with strictly additive loci, an temporary increase in the genetic variance (even in the face of genetic drift) can occur if there are a number of rare, but favored, alleles. As these increase in frequency, the additive variance also increases. If genetic drift strictly governs the dynamics of the variance, then these rare alleles have a small chance of increasing and do not significantly (on average) inflate the variance. However, if selection is of even modest importance on the dynamics at any particular locus, the single-generation response becomes a very poor predictor the long-term response.

James (1962), Verghese (1974), and Nicholas and Robertson (1980) extended Robertson’s theory for various models of natural selection opposing artificial selection. Not surprisingly, the selection limit is reduced by the presence of opposing natural selection. None of these models retains genetic variability, as drift eventually fixes all loci, even those overdominant in fitness. Finally, as we will see shortly, a number of studies have examined the effects of linkage.
TESTS OF ROBERTSON’S THEORY OF SELECTION LIMITS

Robertson’s theory applies to the expected response from the existing variation in the base population at the start of selection. Eventually, mutational input becomes important and will ultimately dominate the long-term response, a point we will develop in detail shortly. In the very small population sizes common in many selection experiments, the distinction between exhaustion of the initial variation and the additional response due to new mutation can be fairly clear, as the latter take many more generations to become apparent than it takes to remove the existing variation. For larger population sizes, the two sources of response become more blurred. Hence, most tests of Robertson’s theory use very small populations.

Observed limits and half-lives are usually considerably below the values predicted from Robertson’s theory (reviewed in Roberts 1966, Kress 1975, Eisen 1980, Falconer and Mackay 1996). Table 26.1 gives various results from experiments with mice. These discrepancies between observation and theory are not unexpected. Robertson’s theory assumes that the limit is reached as genetic variance is exhausted by fixation at all loci. As noted in Chapter 11, selection limits can occur in spite of significant additive genetic variance, often because natural and artificial selection are in conflict. Further, as we have stressed, the selection limit of \( 2N_e R(1) \) and half-life of \( 1.4N_e \) are expected upper limits and require that drift largely dominates. An additional complication is that the effective population size is generally overestimated by simply taking the number of parents as \( N_e \). For example, variation in male mating success in Drosophila can decrease the effective population size to less than half of the number of male parents (Crow and Morton 1955). Further, most experiments have not corrected for the expected reduction in \( N_e \) from the cumulative effects of selection (Equation 26.13).

Table 26.1. Observed and predicted selection limits (scaled in terms of base-population phenotypic standard deviations) and half-lives (scaled in terms of \( N_e \)) for a variety of characters in laboratory populations of mice. From Falconer (1977), Eisen (1975), and Hanrahan et al. (1973).

<table>
<thead>
<tr>
<th>Character Selected</th>
<th>Direction of Selection</th>
<th>Total Response</th>
<th>Observed</th>
<th>Predicted</th>
<th>Ratio</th>
<th>Half-life</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strain N</td>
<td>Up</td>
<td>3.4</td>
<td>7.2</td>
<td>0.47</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Down</td>
<td>5.6</td>
<td>15.9</td>
<td>0.35</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>Strain Q</td>
<td>Up</td>
<td>3.9</td>
<td>15.8</td>
<td>0.27</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Down</td>
<td>3.6</td>
<td>9.6</td>
<td>0.38</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td>Up</td>
<td>2.0</td>
<td>7.4</td>
<td>0.27</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Down</td>
<td>4.5</td>
<td>13.7</td>
<td>0.33</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>Litter Size</td>
<td>Up</td>
<td>1.2</td>
<td>2.3</td>
<td>0.52</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Down</td>
<td>0.5</td>
<td>7.7</td>
<td>0.06</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>Postweaning weight gain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line M4</td>
<td>Up</td>
<td>1.5</td>
<td>5.4</td>
<td>0.27</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>Line M8</td>
<td>Up</td>
<td>2.0</td>
<td>10.0</td>
<td>0.20</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>Line M16</td>
<td>Up</td>
<td>4.3</td>
<td>45.0</td>
<td>0.10</td>
<td>0.3</td>
<td></td>
</tr>
</tbody>
</table>

A more direct test of Robertson’s theory is that the selection limit should increase, and half-life should decrease, as \( N_e \) increases. In general, both these predictions hold. For
example, the estimated effective population sizes of lines M4, M8, and M16 in Table 26.1 are 7.7, 18.6, and 40.9, while each line experiences essentially the same value of $\bar{I}$. (Eisen 1975). For this data set, half-life decreases as $N_e \bar{I}$ increases as predicted by theory. In a more extensive experiment, Jones et al. (1968) examined the effects of changing $N_e$ and/or $\bar{I}$ on otherwise replicate lines of *Drosophila melanogaster*. All their populations were still responding at the end of the experiment (50 generations), so the limit and half-lives could not be estimated. Nevertheless, the data (Table 26.2) are consistent with Robertson’s qualitative predictions, as long-term response increases with $N_e \bar{I}$ (Figure 26.5).

Table 26.2. The cumulative response after 50 generations of selection for increased abdominal bristle number in *Drosophila melanogaster* as a function of effective population size and selection intensity. $N_e$ is estimated as half the number of parents. None of the lines showed an apparent plateau, but the experiment was stopped after 50 generations. After Jones et al. (1968).

<table>
<thead>
<tr>
<th>$N_e$</th>
<th>$\bar{I}$</th>
<th>$R(50)$</th>
<th>$N_e$</th>
<th>$\bar{I}$</th>
<th>$R(50)$</th>
<th>$N_e$</th>
<th>$\bar{I}$</th>
<th>$R(50)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1.6</td>
<td>16.3</td>
<td>20</td>
<td>1.7</td>
<td>20.3</td>
<td>40</td>
<td>1.7</td>
<td>31.7</td>
</tr>
<tr>
<td>10</td>
<td>1.3</td>
<td>11.2</td>
<td>20</td>
<td>1.4</td>
<td>14.7</td>
<td>40</td>
<td>1.4</td>
<td>18.8</td>
</tr>
<tr>
<td>10</td>
<td>0.9</td>
<td>8.1</td>
<td>20</td>
<td>1.0</td>
<td>12.2</td>
<td>40</td>
<td>1.0</td>
<td>16.4</td>
</tr>
</tbody>
</table>

Figure 26.5. Cumulative response at generation 50 as a function of $N_e \bar{I}$ for selection on increased abdominal bristle number is *Drosophila melanogaster*. After Jones et al. (1968).

Robertson’s theory further predicts that when effective population is sufficiently large, further increases in $N_e$ should not change the limit (provided mutational input can be ignored), as all favorable alleles initially present become fixed. This has yet to be observed, which is perhaps not surprising given that most experiments have $N_e$ below 50. By designing ingenious devices to facilitate mass selection in *Drosophila melanogaster*, Weber and colleagues (Weber 1990, 1996; Weber and Diggins 1990) have been able to examine the consequences of larger population sizes. Selection experiments on wing-tip height (Weber 1990) and ethanol tolerance (Weber and Diggins 1990) had effective population sizes on the order of $N_e \simeq 200–400$. Both characters showed an increased response with increasing $N_e$. The data for wing-tip height are given in Figure 26.6. This figure also summarizes the results of nine other experiments from previous studies, showing the ratio of response after 50 generations
to the initial response. In general, this ratio increases with $N_e$ as predicted. The implication is that there is additional “useable” genetic variation present in the base population that can be exploited by increasing the strength of selection ($N_e \tau$). In very small populations, only major alleles are influenced by selection (see Equation 26.6). That response continues to increase with $N_e$, suggests that there is a large pool of loci of smaller effects. As $N_e \tau$ increases, favorable alleles at these loci are more likely to become fixed, increasing response. Larger populations also provide a greater chance for recombination to remove deleterious linked combinations, which might be fixed in smaller populations, further increasing the potential for response. One complication is that as population size increases, the contribution from mutational input becomes increasingly important over the time scales it takes to remove the initial variation. We will address this point later in the chapter. A second complication is that when the character value is influenced by inbreeding depression (as would occur if directional dominance is present), over the same number of generations, the effects on inbreeding depression will be more dramatic in smaller populations. One test for whether inbreeding depression is reducing response is to cross divergently selected lines and look for significant changes in the mean in the resulting $F_1$ population (e.g., Eisen 1975, Kownacki 1979).

Weber’s Selection Experiment on Drosophila Flight Speed

Perhaps the largest long-term artificial selection experiment (in terms of effective population size) is the heroic effort of Weber (1996), who scored a total of over 9,000,000 Drosophila for flight speed in two replicate lines subjected to 100 generations of selection. The resulting $N_e$ was in the 500-1000 range, with a percent selected of $p = 0.045$ (for a selection intensity of $\tau = 2.11$). The average speed before selection was around 2 cm/second, while the mean speed at generation 100 was 170 cm/sec. As shown in Figure 26.7A, response continued in both lines for 100 generations, but was diminishing with time, as indicated by a significant quadratic component in the response curve.

Unlike most other selection experiments, there was little slippage upon relaxation of
selection (see Figure 26.7a) and there was only a minimal loss in fitness relative to the control populations (a fitness decrease of six and seven percent at generations 50 and 85, respectively). Weber attributes this to the larger effective population size which both reduces the level of inbreeding and allows for more efficient selection on modifiers. The later potentially allows for reducing any deleterious pleiotropic effects that accompany major allele improving flight speed, as the weak second-order selection effects on modifiers are much easier to select for in larger populations. Larger population sizes also allow recombination to be more efficient, reducing the effects deleterious alleles linked to alleles improving flight speed.

Finally, Weber gained some insight into the genetic nature of the response by examining the selection response in hybrid lines formed by crossing each replicate selection line at generation 75 (lines AA1 and AA2) back to control lines (CN1 and CN2). As Figure 26.7B shows, both the $F_1$ and $F_2$ were close to the control line values, indicating very strong dominance for reduced flight speed. Evidence for epistasis was more equivocal. From the theory of line cross analysis (LW Chapter 9), an estimate of composite epistatic effects is provided by the linear contrast of means of the parental and first two filial populations, $4\bar{x}_{F_2} - 2\bar{x}_{F_1} - \bar{x}_{P_2} - \bar{x}_{P_1}$, but the resulting value was not significantly different from zero ($-38.5 \pm 37.5$). Selection on both resulting $F_2$ lines required only six generations to recover essentially all of the response seen in the 75-generation lines.

**Figure 26.7.** Results from Weber’s selection experiments for increased flight speed in *Drosophila*. Left: Results of 100 generations of selection in two replicate lines (open circles and squares). The solid triangles denote the result of two sublines subjected to relaxed selection, one for 30 generations starting at generation 65 and the other for 10 generations starting at generation 85. Right: Response to selection in hybrid sublines formed by crossing generation 75 selection lines back to controls. Selection started on the $F_3$ lines, with only six generations of selection (the $F_2$ lines) required to recover essentially all of the initial response. After Weber (1996).

**EFFECTS OF LINKAGE ON THE SELECTION LIMIT**

When QTLs are linked, we expect some reduction in the selection limit as selection at linked loci reduces the fixation probabilities of beneficial alleles (Hill and Robertson 1966, Birky and Walsh 1988, Barton 1995). Simulation studies (Fraser 1957; Latter 1965, 1966a,b; Gill 1965a,b,c; Qureshi and Kempthorne 1968; Qureshi et al. 1968, Qureshi 1968) show that linkage has only a small effect unless loci are very close (less than 5 map units). As mentioned previously, most of these studies inflate the importance of linkage by assuming that all loci have equal effects.

An approximate analytic treatment of linkage was offered by Robertson (1970a, 1977)
(and later by Hospital and Chevalet 1993, 199x), who relied on certain normality assumptions. In the absence of recombination, selection acts on an entire chromosome, and Robertson framed his results in terms of the response contributed by a single chromosome. Robertson considered three different limits, corresponding to different amounts of recombination: $L_f$, the limit (from this chromosome) with free recombination between all loci; $L_0$, the limit under complete linkage; and $L_\ell$, the limit when the map length of the chromosome is $\ell$ — if $n$ loci occur on this chromosome, the recombination rate between adjacent pairs is approximately $\ell/n$.

As before, the completely additive model is assumed, and loci are assumed to start in gametic-phase equilibrium. Let $\sigma^2_{A*}$ be the initial additive genetic variance contributed by the particular chromosome being considered and define $h^2_{A*} = \sigma^2_{A*}/\sigma^2_A$ as the initial fraction of phenotypic variance attributable to this chromosome. The expected contribution from this chromosome following a single generation of selection is $S\sigma^2_{A*}/\sigma^2_A = \hat{h}_{A*}\sigma^2_A$. When $2N_e\hat{h}_{A*}$ is small, the expected limit for a chromosome with freely recombining loci is $2N_e\hat{h}_{A*}$ times the initial response, giving $L_f \approx 2N_e\hat{h}_{A*}$. Assuming weak selection, Robertson (1970a) found that the ratio of the free-recombination limit to the complete-linkage limit is approximately

$$\frac{L_f}{L_0} \approx 1 + \frac{2}{3}(N_e\hat{h}_{A*})^2 \quad \text{when } 2N_e\hat{h}_{A*} < 1 \quad (26.25)$$

Hence, for weak selection, complete linkage has only a trivial effect when the chromosome contains a large number of QTLs.

When selection is strong ($N_e\hat{h}_{A*} \gg 1$), the results are more complicated. Robertson assumes that there are $n$ underlying loci, each with frequency $q$ of the favored allele, which increases the character by $a$ (the difference between the homozygotes). Under these assumptions, the additive variance contributed by this chromosome is $\sigma^2_{A*} = n(a^2/2)q(1-q)$. If selection is sufficiently strong, under free recombination all favored alleles are fixed, and the total response becomes $L_f = na(1-q)$. Noting that $a = \sigma_{A*}/n\sqrt{2/[nq(1-q)]}$, this can also be restated as

$$L_f = na(1-q) = \sigma_{A*} \sqrt{\frac{2n(1-q)}{q}} \quad (26.26)$$

On the other hand, with complete linkage the limit approaches twice the value of the best of the initial $2N$ chromosomes sampled (as this chromosome is ultimate fixed). The expected value for the best chromosome is given by the expected value largest order statistic (see LW Example 6 from Chapter 9). For a unit normal, this is expressed in terms of standard deviations (here $\sigma_{A*}^2$) above the mean, so that if $x_{2N}$ is the standardized largest order statistic in a sample of $2N$ chromosomes, the limit is given by

$$L_0 = (x_{2N} \sqrt{2}) \sigma_{A*} \quad (26.27a)$$

Robertson (1970a) showed for $10 < N < 40$ that $x_{2N} \sqrt{2} \approx 3$, so that that $L_0/\sigma_{A*}^2 \approx 3$. Hence, for these values of $N$,

$$\frac{L_f}{L_0} \approx 1 + \frac{2}{3}\sqrt{\frac{2n(1-q)}{q}} \quad (26.27a)$$

The factor of 3 increases to 3.8 when $N = 80$ and to 4.6 when $N = 500$. For larger values of $N$, using the asymptotic approximation for the largest order statistic given by Kendall and Stuart (1977), the factor of 3 is replaced by

$$x_{2N} \sqrt{2} \approx \frac{0.577}{\sqrt{\ln(2N)}} + 2\sqrt{\ln(2N)} \quad (26.27b)$$
Note that the increase in the selection limit is only weakly dependent on \( N \), as the largest order statistic scales as \( \sqrt{2 \ln(2N)} \).

As the number of loci \( n \) increases, Robertson suggested that the limit under free recombination approaches a limit independent of \( n \) and \( q \), namely the infinitesimal limit \( L_f = 2N \tau h^* \sigma_A \), so that with strong selection and a large number of loci

\[
L_f / L_0 \approx \left( \frac{2}{x_{2n}} \right) N \tau h^* \quad \text{when} \quad N \tau h^* > 5 \tag{26.27c}
\]

Robertson also observed that for large values of \( N \tau h^* (> 5) \), the half-life with no recombination is approximately \( t_{1/2} \approx 2/(\tau h^*) \) generations, and that differences in response (relative to free recombination) only become apparent after this number of generations have passed.

Allowing for some recombination, Robertson (1970a) found that the limit for a chromosome of length \( \ell \) is

\[
L_f / L_0 \approx 1 + N_e \ell / 3 \quad \text{when} \quad N_e \ell << 1 \tag{26.28a}
\]

To a poorer approximation, over the entire range of \( N_e \ell \),

\[
L_f / L_0 \approx 1 + \frac{K N_e \ell / 3}{N_e \ell / 3 + K} \tag{26.28b}
\]

where \( K = L_f / L_0 \). Thus \( L_f / L_0 \) approaches \( L_f / L_0 \) as \( N_e \ell \) increases. Provided \( L_f \gg L_0 \), \( L_f \) is halfway between \( L_f \) and \( L_0 \) when \( N_e \ell = 3L_f / L_0 \). Assuming moderate to large values of \( N_e \), this result (together with Equation 26.26a) implies that if \( \ell > 2 \tau h^* \), response will be at least half that expected for free recombination. The above expressions are approximate and assume that each locus has equal effect. Any variation between loci in allelic effects reduces the effect of linkage (Hill and Robertson 1966, Robertson 1970a).

Experimental results generally confirm that suppression of recombination has only a modest effect on the selection limit, as the following example illustrates.

**Example 26.4.** By using the inversions Curly and Moiré, McPhee and Robertson (1970) were able to select for sternopleural bristles in *Drosophila* under conditions of suppressed recombination on chromosomes II and III. From previous work, \( h^2 = 0.4 \), with these chromosomes accounting for 1/3 and 1/2 (respectively) of the genetic variation in bristle number (the X chromosome accounts for the remaining 1/6). In lines suppressed for recombination at both chromosomes, the limit (on a transformed scale) was 0.166 ± 0.014 in up-selected lines and -0.134 ± 0.009 in down-selected lines, reductions of 28 ± 8% and 22 ± 7% relative to the extremes obtained when normal recombination was allowed. For these studies, \( N_e \approx 10 \) and \( \tau \approx 1 \), while \( h_{II}^* = \sqrt{0.4/3} \approx 0.37 \) and \( h_{III}^* = \sqrt{0.4/2} \approx 0.45 \). Thus, selection is strong on both chromosomes as \( N_e \tau h_{II}^* \approx 3.7 \) and \( N_e \tau h_{III}^* \approx 4.5 \). Under these conditions, Robertson’s theory predicts that the limiting contribution from each (recombination-suppressed) chromosome will be approximately 3\( \sigma_A \). Given \( \sigma_A = 0.059 \), \( \sigma_A \) and the expected contributions from chromosomes II and III are 3 · 0.059 · 0.37 ≈ 0.065 and 3 · 0.059 · 0.45 ≈ 0.080, for a total absolute contribution of 0.145, consistent with the observed limits. Robertson’s theory further predicts that the half-life in recombinationally suppressed lines is roughly \( 2/(\tau h^*) \) generations, or 2/0.37 ≈ 5.4 and 2/0.45 ≈ 4.4 for chromosomes II and III, respectively, consistent with the observed half-life of 5 generations.

Two other *Drosophila* experiments examined the consequences of suppressed recombination on selection response. Both Markow (1975) and Thompson (1977) used stocks with
inversions while selecting for increased / decreased phototactic behavior. While Markow observed that suppression reduced the limit, Thompson observed no differences. Markow did not use replicate lines, so the statistical significance of her results are in doubt. However, she observed that the most recombinationally-suppressed lines had the most reduced response, consistent with theory. In Thompson’s experiments, \( N_e \approx 50, \bar{r} \approx 1, \) and \( h^* \approx 0.1 \) (for both autosomes), giving an expected half-life of \( 2/(\bar{r}h^*) = 20 \) generations (López-Fanjul 1989). Given that Thompson’s experiments were stopped at generation 21, it is not surprising that he found no difference as the effects of linkage on total response are not readily apparent until after the half-life.

Robertson’s result largely focused on the ultimate selection limit, while Hospital and Chevalet (1993, 199x) considered the dynamics of approach to this limit. In particular, Hospital and Chevalet (199x) explicitly considered the effects of gametic-phase disequilibrium. Initially, selection generates negative gametic-phase disequilibrium, reducing the initial additive variance and decreasing response. The tighter the linkage, the more pronounced this effect (Chapters 5, 10). Surprisingly, Hospital and Chevalet showed that linkage can often result in an increase in the additive variation in later generations of selection. This seemingly counterintuitive result arises because (under strong selection and tight linkage), selection increases the frequency of those gametes carrying the most favored alleles. Any linked alleles decreasing the trait are also dragged along. This increases the probability of fixation of some unfavorable alleles and hence reduces the ultimate selection limit. On the other hand, rare recombination events among such gametes can result in the creation of new, more favored gametes. As these sweep through the population, an increase in the additive variance results. Thus the negative gametic-phase disequilibrium that suppresses the early response stores up some genetic variation that can become released in later generations. This effect is most pronounced in larger populations, as in small population gametes will become fixed before any such recombination events occur.

**OPTIMAL SELECTION INTENSITY FOR MAXIMIZING LONG-TERM RESPONSE**

When a fixed number \( M \) of individuals are scored, there is a tradeoff between the intensity of selection (\( \bar{r} \)) and the amount of drift (\( N_e \)). If \( N \) individuals are allowed to reproduce (giving \( p = N/M \) as the fraction saved), decreasing \( N \) (and hence \( p \)) increases \( \bar{r} \) but decreases \( N_e \). Recalling Equation 26.22b, Robertson’s selection limit can be expressed as

\[
N_e \bar{r} \left( \frac{2\sigma_A^2(0)}{\sigma_z^2} \right)
\]

(26.29)

showing that the ultimate response (from the initial variation) depends on the product of \( N_e \) and \( \bar{r} \). Thus while decreasing \( p \) results in a larger short-term response due to increased \( \bar{r} \), it also can result in a decreased long-term response by decreasing \( N_e \), as \( N_e \bar{r} \) decreases for sufficiently large or small values of \( p \). Table 26.3 illustrates this tradeoff. For example, while the single-generation response using \( p = 50\% \) is less than half that for \( p = 10\% \), it gives a selection limit over twice as large.

| Table 26.3 | Differences in short-term versus long-term response as a function of the number of adults saved \( N \) when \( M = 50 \). Initially \( h^2 = 0.5 \) and \( \sigma^2 = 100 \). The infinitesimal model is assumed and we further assume \( N_e = N \). The selection intensity \( \bar{r} \) is obtained using Equation 4.16a (corrected for finite population size). From Equation 4.3, \( R(1) = 5\bar{r} \), while from Equation 26.22b, \( R(\infty) = 2N R(1) \). |
Robertson (1960), supporting an earlier conjecture of Dempster (1955), found (for additive loci and a normal phenotypic distribution) that the intensity of selection giving the largest total response is \( p = 50\% \), as \( N_e \tau \) is maximized for fixed \( M \) when half the population is saved. This can be seen directly for truncation selection on a normally-distributed character. Recall from Equation 4.16a that \( \tau = \varphi(x_{[1-p]})/p \) (ignoring the correction for finite population size) where \( x_p \) satisfies \( \Pr(U < x_p) = p \) for a unit normal \( U \) and \( \varphi(x) \) is the unit normal density function. Since the number saved \( N = Mp \), we have (following Hospital and Chevalet 1993),

\[
R(t) \simeq Mp \left(1 - e^{-t/2N_e}\right) \frac{\varphi(x_{[1-p]})\sigma^2_A(0)}{p\sigma_z} \\
= \varphi(x_{[1-p]}) \left[ \frac{M\sigma^2_A(0)}{\sigma_z} \left(1 - e^{-t/2N_e}\right) \right]
\]  

(26.30)

Since the term in brackets is independent of \( p \), response (as a function of \( p \)) is maximized at the maximum value of \( \varphi(x_{[1-p]}) \). The maximal value of the unit normal density function occurs at \( x = 0 \), or a \( p \) value of 0.5.

As Figure 26.8 illustrates, the selection limit as a function of \( p \) becomes extremely flat-topped as \( M \) increases, so even fairly large deviations from \( p = 50\% \) give essentially the same limit. Cockerham and Burrows (1980), relaxing the assumption of normality, found that the optimal proportion for truncation selection is still near 50\%, unless the phenotypic distribution is extremely skewed. Hill and Robertson (1966), Robertson (1970a), and Hospital and Chevalet (1993) found that the optimal proportion increases to above \( p = 50\% \) when linkage is important.

Robertson’s prediction of the optimal selection intensity for long-term response is supported experimentally. For example, Madalena and Robertson (1975) selected for decreased sternopleural bristle number in *Drosophila*. When the best 5 of 25 were chosen, the limit was 17.98 bristles, less extreme than the limit of 17.08 when the best 10 of 25 were chosen. Similar results were seen for increased abdominal bristle number in *Drosophila* (Jones et al. 1968), increased egg-laying in *Tribolium castaneum* (Ruano et al. 1975), and increased post-weaning weight in mice (Hanrahan et al. 1973).

While \( N \) places an upper limit on \( N_e \), it often severely overestimates it, as \( N_e/N \) decreases as selection intensity increases (Equation 26.13). Hence, increasing selection intensity increases drift by both reducing \( N \) and by further reducing the ratio of \( N_e/N \). Table 26.4 illustrates this effect using the same parameters as Table 26.3. Without incorporating this further reduction in \( N_e \), the ratio of expected limits when \( p = 50\% \) versus \( p = 10\% \) is \( 200/90 = 2.2 \). When this reduction in \( N_e \) due to selection is accounted for, this increases to \( 161/41 = 3.9 \).

<table>
<thead>
<tr>
<th>( N )</th>
<th>( p )</th>
<th>( \tau )</th>
<th>( R(1) )</th>
<th>( R(\infty) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.5</td>
<td>0.8</td>
<td>4.0</td>
<td>200</td>
</tr>
<tr>
<td>10</td>
<td>0.2</td>
<td>1.4</td>
<td>7.0</td>
<td>140</td>
</tr>
<tr>
<td>5</td>
<td>0.1</td>
<td>1.8</td>
<td>9.0</td>
<td>90</td>
</tr>
</tbody>
</table>
LONG-TERM RESPONSE IN FINITE POPULATIONS

Figure 26.8. **Left:** The selection limit as a function of the proportion selected (percent of the population allowed to reproduce) for 10 and 50 individuals scored. From Robertson (1960). **Right:** The optimal proportion $p_{opt}$ of individuals selected each generation to maximize the selection advance over $t$ generations, as a function of $t/M$.

Table 26.4. As selection intensity increases, $N_e$ is increasingly less than the actual number of parents, further increasing drift. The reduction in effective population size due to selection is computed using Equation 26.13b. Parameters and assumptions are as in Table 26.3 (e.g., $M = 50, h^2 = 0.5$).

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\tau$</th>
<th>$N_e$</th>
<th>$N_e/N$</th>
<th>$2N_e R(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.8</td>
<td>20.2</td>
<td>0.81</td>
<td>161</td>
</tr>
<tr>
<td>10</td>
<td>1.4</td>
<td>5.8</td>
<td>0.58</td>
<td>81</td>
</tr>
<tr>
<td>5</td>
<td>1.8</td>
<td>2.3</td>
<td>0.47</td>
<td>41</td>
</tr>
</tbody>
</table>

More generally, Robertson (1970b) obtained the optimal selection intensity when the goal is to maximize the total response (from the initial variation in the base population) to generation $t$. Robertson’s derivation follows using Equation 26.22a. As Figure 26.8 shows, the optimal proportion is function of $t/M$. Robertson assumed the infinitesimal model and equal contributions from each sex. Jodar and Lopez-Fanjul (1977) extend these results to unequal sex ratios, finding that the maximal response occurs when the number of individuals scored and the proportions selected are the same in each sex. This follows since effective population size is reduced as the sex ratio deviates from one-to-one (Equation 2.20), increasing the effects of drift. Hospital and Chevalet (1993) examined the effects of linkage, finding that the amount by which the optimal $p$ exceeds the predicted value increases with population size. In small populations, the value predicted from drift (for any particular $t/M$ value) is close to the optimal value, while in larger populations Robertson’s value seriously underestimates the optimal $p$ value.

Ruano et al. (1975) and Frankham (1977) tested Robertson’s predictions for the optimal response at a particular generation with selection experiments for egg-laying in Tribolium and abdominal bristle number in Drosophila, respectively. The theory holds up well for $t/M \leq 0.2$, but both authors found discrepancies between the observed and predicted rank order of lines subjected to different selection intensities for $t/M$ values above this. One explanation of these discrepancies could be the presence of major alleles, resulting in additive variance declining more rapidly than expected under the infinitesimal model. This results
in the optimal proportions being larger than those predicted from Figure 26.8. Frankham (1977) also suggests that not correcting for the additional decrease in $N_e$ with increased selection intensity (e.g., Table 26.4) results in incorrect values of $N_e$ and hence incorrect optimal proportions. García-Dorado and López-Fanjul (1985) examined the consequences of unequal sex ratios using sternopleural bristle number in *Drosophila*. Equal sex ratios gave the highest response, and good agreement with the optimal values predicted by Jódar and López-Fanjul was seen with unequal sex ratios.

**EFFECTS OF POPULATION STRUCTURE ON LONG-TERM RESPONSE**

Our development of Robertson’s theory of selection limits has made a two assumptions regarding population structure: selection occurs in a large panmictic population and the initial base population is infinite in size. This section relaxes these assumptions. We first examine the consequences of founder effects in the initial base population and of passing the population through bottlenecks during selection. We conclude by examining the expected limits when the population is subdivided and when selection is entirely within families.

**Founder Effects and Population Bottlenecks**

So far, we have been considering only the effects of drift due to selecting $N$ adults each generation from an initial base population assumed to be infinite. Drift can occur prior to selection if the base population is formed from only a few individuals. By altering the initial additive variance, this initial sampling can alter response. To distinguish between these different sources of drift, let $N_0$ denote the number of founders and $N_e$ the effective population size during selection. Thus, if $\sigma_A^2(0)$ is the additive variance of the population from which the founders are drawn, then the expected initial additive variance (assuming no nonadditive effects) in the founder population is $\left[1 - 1/(2N_0)\right] \sigma_A^2(0)$.

Founder effects can have a significant effect on response. Robertson (1966), reporting on the unpublished thesis of Da Silva (1961), found that lines formed from a single parental pair had response decreased by roughly 30% relative to a non-bottlenecked line from the base population (Figure 26.9). Lines formed from taking single parental pairs for three consecutive generations showed only a modest further reduction in response, suggesting that most of the founder effect occurred in the first generation. Robertson’s interpretation was that response in this population was due largely to alleles at intermediate frequency, as alleles at low frequency are expected to be lost during the initial sampling. Segregating alleles present after this initial bottleneck of two individuals have intermediate frequencies ($1/4$, $1/2$, or $3/4$) and are thus somewhat resilient to further sampling events.
Figure 26.9. Effects of population bottlenecks on selection response. Top: Selection for sternopleural bristle number in *Drosophila melanogaster*, with the most extreme 10 pairs out of 25 scored pairs selected. The outer curves are the response using the base population. The middle curves are the response for sublines formed from a single parental pair, followed by six generations of random mating to build up population size prior to selection. The inner curves correspond to sublines formed by using a single parental pair for three consecutive generations prior to selection. From Robertson (1966). Bottom: Selection for abdominal bristle number in *Drosophila melanogaster*. The responses denoted by 40(10%), 40(20%), and 40(40%) correspond to populations where the uppermost 40 pairs of adults are selected each generation, with different selection intensities. For example, 200 pairs are scored and the uppermost 40 chosen in the 40(20%) population. Responses denoted by X(10%), X(20%), and X(40%) refer to lines split from the corresponding 40-pair lines after 16 generations of selection and selected thereafter at the same intensity with 10-pairs of parents per generation. From Jones et al. (1968).

Using the above reasoning, Robertson (1960) predicted that the effect of restricting population size after several generations of selection is expected to be small, as favored alleles are expected to be at intermediate to high frequencies. Jones et al. (1968), however, found that even after many generations of selection such bottlenecks can have a large effect. Sublines formed by taking ten pairs of adults from a parental line selected for 16 generations showed reduced response relative to their parent lines (Figure 26.9). One explanation is that there were still desirable major alleles at low frequencies following 16 generations of selection. These alleles can be lost when the population passes through a bottleneck, reducing response. One source for these rare major alleles could be new mutations. Alternate explanations are considered by Frankham (1983b).

James (1970) examined the expected reduction in response due to founder effects. As before, results are developed for a single additive locus, and extended by assuming gametic-phase equilibrium and no epistasis. Since the initial additive variance in the founder population is \(1 - 1/(2N_0)\) \(\sigma_A^2(0)\), the expected response for the first generation of selection from a bottlenecked population is \(1 - 1/(2N_0)\) times the expected response for an initially infinite population. The long-term effects of an initial bottleneck are more unpredictable, depending on initial allele frequencies and the relative strength of selection. When selection is weak at all loci (the infinitesimal model), taking the initial additive variance as \(1 - 1/(2N_0)\) \(\sigma_A^2(0)\) the arguments leading to Equation 26.22a give the expected response for a founder population at generation \(t\) as

\[
R_{N_0}(t) = R(t) \left(1 - \frac{1}{2N_0}\right)
\]

(26.31a)

where \(R(t)\) is the response expected when the initial base population is infinite (Equation...
26.22a). More generally, if two replicate populations of the same size are created using different numbers of founders \((N_1, N_2)\) from a common large base population, the ratio of expected response at any generation is given by

\[
\frac{R_{N_1}}{R_{N_2}} = \frac{1 - 1/(2N_1)}{1 - 1/(2N_2)}
\]  

(26.31b)

Thus, if selection at all loci is weak and all genetic variance is additive, the effect of a bottleneck depends only on \(N_0\).

Founder effects are most serious when rare favorable alleles of large effect are present, but predicting the magnitude of the effect in any given population is difficult. When selection on a locus is strong \((2N_e s >> 1)\), the probability that a selected line formed from a bottlenecked base population will eventually be fixed for the favored allele converges to

\[
u_{N_0}(p_0) = 1 - (1 - p_0)^{2N_0}
\]  

(26.32a)

This follows since if selection is sufficiently strong, the favored allele will become fixed if it is found in the initial sample, which occurs with probability \(1 - (1 - p_0)^{2N_0}\). The ratio of the expected limiting contribution from this locus to the expected contribution when the founding population is infinite is

\[
\frac{u_{N_0}(p_0) - p_0}{u(p_0) - p_0} \approx \frac{1 - (1 - p_0)^{2N_0} - p_0}{1 - p_0} = 1 - (1 - p_0)^{2N_0-1}
\]  

(26.32b)

Since the initial frequencies of major alleles are unknown, the long-term effects of a bottleneck, even when all genetic variance is additive, is unpredictable.

Frankham (1980) examined founder effects in \textit{Drosophila} populations selected for increased abdominal bristle number. As shown in Figure 17.10, the limit to bottlenecked populations formed from two founders was between 0.69 and 0.72 of that for non-bottlenecked populations, quite close to the value of \([1 - 1/(2N_0)] = 0.75\) predicted for additive loci under weak selection (Equation 26.31b). Frankham reports similar unpublished thesis results of Da Silva (1961) and Hammond (1973). However, while D. Robertson (1969, reported in James 1970) observed a decrease in response with decreasing number of founders when the number of selected parents \((N_e)\) was 10, there was no obvious effect when \(N_e = 40\) (which is not unexpected since \(1 - 1/80\) is negligible). We have been unable to find any reports of response increasing significantly when the population is passed through a bottleneck, as can occur if significant nonadditive variance is present (Chapter 3). Clearly, there is a need for further experiments.
An especially interesting experiment on founder effects is the work of Skibinski and Shereif (1989) who examined sternopleural bristle number in *Drosophila melanogaster*. Three initial lines were created from a large base population by taking parents from different parts of the distribution of bristle number to generate a high line, a low line, and a line from the central part of the distribution. The central line had the largest total response to divergent selection. Skibinski and Shereif suggest that these results are consistent with a few major alleles underlying the trait with the central line having higher heterozygosity at these loci (hence more useable genetic variance) than the extreme lines. One complication with this interpretation is that the central line had a larger initial size than either extreme line.

**Population Subdivision**

Thus far, we have been considering the long-term response under mass selection in a single panmictic population. How robust are our results if we subdivide the total population? When only additive variance is present, Robertson (1960) showed that population structure has little effect on the selection limit. In particular, the expected limit for a population formed by crossing *N* (replicate) plateaued lines of size *m* is the same as a single line with size *Nm*. Maruyama (1970) generalized this result by showing (for additive loci and ignoring linkage effects) that any subdivision of the population gives the same limit, independent of when and how lines are crossed, provided there is no selection between lines. Madalena and Hill (1972) further showed that linkage has only a minor effect on this conclusion. They also found (again assuming only additive variance) that while between-line selection (i.e., culling some of the lines) may increase short-term response, removing lines decreases the total genetic variance of the entire population, decreasing the limit. This reduction is most severe with free recombination, and is negligible with tight linkage.

When significant nonadditive variance is present, population subdivision may increase the selection limit. For example, when favorable rare recessives are present, subdividing the population and subsequently crossing these lines when they plateau and then reselecting gives a higher expected limit than using a single panmictic line of the same total size (Madalena and Hill 1972, Slatkin 1981). The increased inbreeding in the sublines increases the frequency of homozygotes, facilitating selection for favorable recessives.

Similarly, *Wright’s shifting balance theory* (reviewed in Wright 1951, 1978), assumes that local inbreeding due to population subdivision facilitates the accumulation of rare favorable epistatic combinations of loci. Crossing such fixed (or nearly fixed) lines increases the selection limit relative to a single panmictic population, much akin to what happens with rare recessives. Indeed, Enfield and Anklesaria (1986) found in simulation studies that when additive-by-additive epistatic variance is present, certain population subdivisions give a greater short-term and long-term response than a single panmictic population. We examine the shifting balance theory in more detail in Chapter 22.

Selection experiments with population subdivision (reviewed in Rathie and Nicholas 1980, López-Fanjul 1989) generally give results similar to those expected under the strictly additive model: subdivision usually has no effect on the selection limit. Two experiments reported exceptions to this. Madalena and Robertson (1975) selected for decreased sternopleural bristle number in *Drosophila melanogaster* under two different population structures: a single-cycle structure where sublines were crossed once, and a repeat-cycle structure where sublines were crossed multiple times. The limit under the single-cycle structure was essen-
ially the same as a panmictic population, regardless of whether or not between-line selection was practiced. The limit under the repeat-cycle structure was slightly more extreme than the panmictic population. These results are complicated by the presence of major alleles lethal as homozygotes, but nevertheless suggest the presence of some favorable recessives initially at low frequency. The second exception was the experiment of Katz and Young (1975), who selected for increased body weight in *Drosophila*. Populations that were subdivided with a small amount of migration between them gave a slightly larger response than the panmictic population.

There have been a number of contrasting views on the optimal population structure for evolution. Wright (1931, 1951, 1977, 1978) suggests that evolution is most rapid when the population is subdivided, while Fisher (1958) viewed a single large panmictic population as the optimal structure. When mostly additive gene action is present, both the Wright and Fisher structures are expected to give comparable rates of evolution, although the Fisher structure may have a slight advantage when the effects of linkage are considered (in larger populations, the probability that a deleterious allele linked to a favorable allele will hitch-hike to fixation is decreased, increasing the potential response). With non-additive gene action, the optimal structure depends on the exact nature of gene action. With recessives, population subdivision increases the response. With epistasis, the Wright structure offers an advantage if epistatic combinations are such that their formation requires intermediate genotypes that are deleterious. Conversely, in other situations the Fisher structure may offer an advantage in that it allows more gene combinations to be tested.

One must keep in mind that the optimal population structure for maximizing response under one type of gene action may not be optimal for other types. In particular, many types of population structure that increase the probability of fixation of recessive and/or epistatic genes may retard the fixation of advantageous additive genes. Some population structures that increase the fixation probabilities for recessives and/or epistatic genes can reduce the fixation probabilities for additive genes. Likewise, even structures that do not decrease the fixation probability may increase the fixation time, which in turn reduces the rate of response.

Caballero et al. (1991) examined the types of mating schemes (following selection) that increase the fixation probability of recessive alleles while not significantly reducing the fixation probabilities or increasing the fixation times for additive genes. They found that mating full sibs wherever possible following selection increases the fixation probabilities for recessives (relative to random mating following selection) without any significant effect on additive alleles. The tradeoff here is a slight reduction in $N_e$ (due to the increased inbreeding by full-sib mating following selection) versus the increased selection on recessives by inbreeding (compare Equations 26.5b and 26.5d). Recall from Equation 26.5d that the measure $f$ of departures from Hardy-Weinberg frequencies enters into the selection coefficients. Caballero et al. show

$$f = \frac{N_{FS} - 1}{4N_{TM} - 3N_{FS} + 3} + f_r \quad (26.33a)$$

where $N_{FS}$ is the number of full-sib matings, $N_{TM}$ the total number of matings, and $f_r$ the departure from Hardy-Weinberg genotype frequencies under random mating in a finite population, which is given by

$$f_r = -\left(\frac{1}{8N_f} + \frac{1}{8N_m}\right) \quad (26.33b)$$

where $N_m$ and $N_f$ are the number of reproducing males and females. Note that the negative sign implies that under random mating, there is a slight expected excessive of heterozygotes relative to the frequency expected from the allele frequencies alone.

**Within-Family Selection**
The reader will recall that the variance in number of offspring contributed by each selected parent is an important determinant in effective population size — the larger this variance, the smaller $N_e$. Exploiting this, Toro and Nieto (1984) note that deliberately assigning selected parents different probabilities of contributing offspring (according to a specific formula) gives populations with the same selection intensity but different effective population sizes. Suppose 20 individuals are measured ($M = 20$), and we wish the expected selection intensity to be 1.2. This occurs if the best 5 individuals are chosen (using Equation 4.16a and correcting for finite population size) and each parent has equal probability of contributing offspring. The same selection intensity can be obtained by choosing the best 10 individuals and assigning these individuals unequal probabilities for contributing offspring (see Toro and Nieto for details). This latter scheme (while holding both selection intensity and the number of measured individuals constant) increases effective population size from 5 to 5.9, which in turn increases the long-term response.

The most extreme example of this occurs when selection is entirely within families: the best male and female are chosen from each full-sib family and mated at random between families. This doubles the effective population size relative to selecting the same number of individuals independent of family structure (recall from Equation 2.16 that if all parents contribute the same number of offspring, there is no variance in offspring number and $N_e = 2N$). We remind the reader at this point of the important, but subtle, distinction between parents having an equal probability of contributing offspring versus parents contributing exactly the same number of offspring. In the former case, some parents will contribute no offspring and others more than one, generating a non-zero variance. In the latter case, all parents make an identical contribution and there is no variance in offspring number.

Thus, using within-family selection results in a population with twice the effective size as one undergoing mass selection with the same number of individuals selected. However, as Robertson (1960) noted, the usable additive genetic variance within full-sib families is only half that available under mass selection (Equation 7.7b). These exactly cancel the advantage of a larger $N_e$, suggesting that both methods give the same limit. Dempflé (1975) pointed out that this conclusion relies critically on $h^2$ being low. Applying Equations Equation 7.7b and Equation 7.9, the response to a generation of within-family selection is (for full-sibs)

$$R_{w_{FS}}(1) = \tau h^2_{w_{FS}} \sigma_{w_{FS}}$$

where

$$h^2_{w_{FS}} = \frac{\sigma^2_A}{2 \sigma^2_{w_{FS}}} \quad \text{and} \quad \sigma^2_{w_{FS}} = \frac{\sigma^2_A}{2} + \sigma^2_{Es}$$

If the additive variance is much larger than the within-family environmental variance ($\sigma^2_{Es}$), then $h^2_{w_{FS}} \approx 1$ and $\sigma^2_{w_{FS}} \approx \sigma^2_A/2$, giving $R_{w_{FS}}(1) \approx \tau \sigma_A/\sqrt{2}$. If the total environmental values is also much smaller than the additive variance, the expected response to individual selection becomes $R(1) \approx \tau \sigma_A$. Thus, when additive variance dominates, the ratio of expected limits is

$$\frac{4NR_{w_{FS}}(1)}{2NR(1)} \approx \sqrt{2}$$

and within-family selection increases the limit.

Three other factors can favor within-family selection.

1. Significant between-family environmental variance. If most of the environmental variance is due to between-family, rather than within-family, effects (i.e., if $\sigma^2_{Es} > \sigma^2_{Es}$), within-family selection gives a larger single-generation response than individual selection (see Chapter 7). Coupling this with the decreased loss of variation due to a larger effective population size, within-family selection is superior when the between-family component of environmental variance is sufficiently large.
2. Retardation of the cumulative reduction in \( N_e \) from selection. Recall that individual
selection reduces \( N_e \) below the actual number of parents by inflating the between-
family variance in offspring number when \( h^2 \) and/or \( r \) are large. This variance is zero
under within-family selection, giving within-family selection an effective population
size greater than twice that for individual selection.

3. Gametic-phase disequilibrium. The presence of gametic-phase disequilibrium also in-
creases the effectiveness of within-family selection relative to individual selection.
Under the assumptions of the infinitesimal model, the negative gametic-phase dis-
equilibrium generated by directional selection reduces the between-family compo-

tent of additive variance, while the within-family component remains unchanged
(Chapters 5 and 7). Hence, the usable additive variance in the mass-selection lines
is decreased, while the usable additive variance in the within-family lines is un-
changed. This effect is negligible unless selection is strong and heritability is high.

Young and Skavaril (1976) used computer simulations to examine the consequences of
major alleles and linkage on within-family selection. They found that individual selection
was superior to within-family selection in small populations, especially when major alleles
are rare and/or when \( h^2 \) is small.

On the experimental side, von Butler et al. (1984) compared individual and within-
family selection on 8-week body weight in mice. In one set of replicates, within-family
selection initially showed a reduced response, but after 18 generations had essentially the
same response as mass-selected lines. In alternative set of replicates (using a different base
population), mass selection did better than within-family selection, but both populations
were still responding after the experiment was stopped (after 18 generations). Since within-
family selection is expected to show a longer period of response due to a large effective
population size, the results for the second set of replicates are inconclusive.

ASYMPTOTIC RESPONSE DUE TO MUTATIONAL INPUT

As reviewed in Chapter 11 (and by Frankham 1980, 1983a and Weber and Diggins 1990), there
is strong evidence that new mutants contribute to selection response even during the short
time scales of many “long-term” laboratory experiments. The selection limit resulting from
drift and selection removing all initial genetic variation is thus an artifact of time scale as it
ignores this mutational contribution. Even if an observed limit is due to a balance between
natural and artificial selection, new mutations with less deleterious pleiotropic effects on
fitness can arise, resulting in further response.

If a rare recessive is initially present at low frequency, the appearance of homozygotes
involving this allele may be taken as new mutations. If a recessive is present as a single copy,
then the expected time until the first appearance of a homozygote is approximately \( 2N^{1/3} \)
generations, with the appear time following a nearly geometric distribution (Robertson 1978;
Karlin and Tavaré 1980, 1981a, 1981b; Santago 1989). Since for most selection experiments,
\( N \leq 500 \), any recessives initially present will be expresses as homozygotes by generation 15.

Our discussions of the nature of long-term response with mutational input largely follow
Hill’s pioneering treatment (1982b,c). We start by assuming complete additivity. Recall from
Chapter 3 (and LW Chapter 12) that one measure of mutational input is \( \sigma^2_m \), the amount
of new additive variance produced by mutation each generation. Consider the \( i \)th locus,
where each allele mutates to a new allele with a per-generation rate of \( \mu_i \). The incremental
mutation model is assumed: when an allele \( A \) mutates to a new allele \( A' \), the genotypic
values of \( AA' \) and \( A'A' \) are \( g_{AA} + a \) and \( g_{AA} + 2a \), where \( g_{AA} \) is the genotypic value of \( AA \).
This model assumes that the genotypic value of the new mutant depends on the state of its parental allele. However, the distribution of increments \((a)\) added to the parental allele are assumed to be independent of the value of the parental allele. For \(n\) loci we have

\[
\sigma_m^2 = 2 \sum_{i=1}^{n} \mu_i \sigma_i^2(a)
\]

We first consider the infinitesimal model before examining a more general model and the consequences of dominance.

**Results for the Infinitesimal Model**

We assume complete additivity and ignore any effects of gametic-phase disequilibrium. From Equation 3.19b, the expected additive genetic variance at generation \(t\) is given by

\[
\sigma_A^2(t) \simeq 2N_e \sigma_m^2 + \left[ \sigma_A^2(0) - 2N_e \sigma_m^2 \right] \exp(-t/2N_e)
\]

(26.34)

Setting \(\sigma_A^2(0) = 0\) gives the additive variance contributed entirely from mutation as

\[
\sigma_{A,m}^2(t) \simeq 2N_e \sigma_m^2 \left[ 1 - \exp(-t/2N_e) \right]
\]

(26.35a)

Hence, the rate of response at generation \(t\) from mutational input is

\[
r_m(t) = \tau \frac{\sigma_{A,m}^2(t)}{\sigma_z} \simeq 2N_e \tau \frac{\sigma_m^2}{\sigma_z} \left[ 1 - \exp(-t/2N_e) \right]
\]

(26.35b)

where we have made the usual assumption that the phenotypic variance \(\sigma_z^2\) does not significantly change over time (more generally, \(\sigma_z^2\) can be replaced by \(\sigma_z^2(t) = \sigma_A^2(t) + \sigma_E^2\)). For \(t >> 2N_e\), the per-generation response approaches an asymptotic limit of

\[
r_m(\infty) = 2N_e \tau \frac{\sigma_m^2}{\sigma_z}
\]

(26.36)

Assuming \(\sigma_A^2(0) = 0\), half this rate occurs when \(t \simeq 1.4N_e\) (Hill 1982b,c). There are several ways to intuit the value of the asymptotic limit. From Robertson’s theory, we expect the final response to be \(2N_e\) times the initial response \(R(0)\), which for new mutants arising in any particular generation is \(R(0) = \tau \sigma_m^2/\sigma_z\). Alternatively, note that the equilibrium additive variance is \(2N_e \sigma_m^2\), which (upon recalling Equation 4.3) gives an Equation 26.36.

Summing over generations (and using the approximation given by Equation 26.8b) gives the cumulative response due to new mutation as

\[
R_m(t) = \sum_{\tau=1}^{t} r_m(\tau) \simeq 2N_e \tau \frac{\sigma_m^2}{\sigma_z} \left( t - 2N_e \left[ 1 - \exp(-t/2N_e) \right] \right)
\]

(26.37a)

as found by Hill (1982c, 1990) and Weber and Diggins (1990). For genes of sufficiently large effects \(|a| >> \sigma_z/Nt\) which can be considered to be fixed essentially instantaneously, the response becomes

\[
R_m(t) = 2tN_e \tau \frac{\sigma_m^2}{\sigma_z}
\]

(26.37b)

as suggested by Hill (1982c). Note from Equation 26.36 that this implies the asymptotic rate of response applies from generation one.
Combining the mutational response with the response due to genetic variation originally in the base population (Equation 26.22a) gives an expected cumulative response of

\[ R(t) = 2Ne \left( t \frac{\sigma_m^2}{\sigma_z} \left[ 1 - \exp(-t/2Ne) \right] \left[ \sigma_A^2(0) - 2Ne \sigma_m^2 \right] \right) \]  

The \( t \sigma_m^2 \) term, which represents the asymptotic response, will eventually dominate (i.e., for large \( t \)). The other term in the parentheses represents the transient effect of the initial additive variance, and is zero if the population starts at the mutation-drift equilibrium (i.e., \( \sigma_A^2(0) = 2Ne \sigma_m^2 \)).

Of some interest is the expected number of generations until response from mutational input exceeds that contributed by the initial variation. Let \( t^* \) be the generation when the per-generation response from both sources is equal. Here the initial additive variance remaining at generation \( t^* \) equals the new additive variance generated by generation \( t^* \),

\[ \sigma_A^2(0) \exp(-t^*/2Ne) = 2Ne \sigma_m^2 \left[ 1 - \exp(-t^*/2Ne) \right] \]

This equation has the solution

\[ t^* = 2Ne \ln(1 + \Psi) \]  

where \( \Psi = \sigma_A^2(0)/(2Ne \sigma_m^2) \). Denoting the initial heritability by \( h^2 \), a little rearrangement give

\[ \Psi = \frac{h^2}{(1 - h^2) \left( 2Ne / \sigma_m^2 / \sigma_E^2 \right)} \]

The average value of \( \sigma_m^2 / \sigma_E^2 \) is approximately 0.005, see LW Table 17.1. Using this value, it is seen that \( t^* \) is only rather weakly dependent on \( Ne \) (see Figure 26.11). If \( \Psi << 1 \), so that the expected additive variance at the mutation-drift equilibrium exceeds the initial additive variance \( (\sigma_A^2(0) << 2Ne \sigma_m^2) \), then using the approximation \( \ln(1 + x) \simeq x \) for small \( |x| \), we have

\[ t^* \simeq 2Ne \Psi = \frac{h^2}{(1 - h^2) (\sigma_m^2 / \sigma_E^2)} \]

Again using \( \sigma_m^2 / \sigma_E^2 = 0.005 \) gives \( t^* \simeq 200h^2/(1 - h^2) \). This translates into 11, 22, and 67 generations until the rate of response from mutational input exceeds the rate of response due to initial variation for \( h^2 \) values of 0.05, 0.10, and 0.25 (respectively).

It is important to stress that our expressions for half-life of response assume that drift dominates and tend to overestimate the half-life when selection is moderate to strong. Likewise, we expect that the infinitesimal model underestimates the changes in allele frequencies of new mutants under moderate to strong selection. Thus, our expression for \( t^* \) is very likely an overestimate and we should regard Equation 26.38 as an upper bound.
Figure 26.11. The expected generation at which response due to mutational input equals the response due to initial variation in the base population. We used $\sigma_m^2/\sigma_E^2 = 0.005$, which is the average value of the experiments in LW Table 26.1. The four curves correspond to initial heritabilities of 0.05, 0.10, 0.25 and 0.50.

Example 26.5. Yoo (1980a) observed a steady and reasonably constant increase in *Drosophila* abdominal bristle number over 80 generations of selection (Figure 11.5D). In particular, he observed an increase of about 0.3 bristles per generation during generations 50 to 80. Assuming the infinitesimal model, how much of this response is due to mutational input? Yoo’s base population had $\sigma_E^2 \approx 4$, $\sigma_z^2 \approx 5$, $h^2 \approx 0.2$, $t \approx 1.4$, and 50 pairs of parents were chosen each generation. Taking $\sigma_m^2/\sigma_E^2 \approx 0.001$ (the average for abdominal bristles in LW Table 26.1) gives $\sigma_m^2 = 0.004$. Assuming $N_e \approx 60$, the limiting mutational variance is

$$2N_e \sigma_m^2 = 2 \cdot 60 \cdot 0.004 = 0.48$$

giving an expected asymptotic rate of response of

$$r = \frac{\bar{A}^2}{\sigma_z} = \frac{\sigma_m^2}{\sqrt{\sigma_A^2 + \sigma_E^2}} = 1.4 \cdot \frac{0.48}{\sqrt{4 + 0.48}} \approx 0.32$$

This is very close to the observed rate of 0.3 bristles per generation (between gens. 50 and 80). However, from Equation 26.35b the expected single-generation response from new mutational input at generation 60 is only

$$1 - e^{-t/(2N_e)} = 1 - e^{-60/120} \approx 0.40$$

of this, giving 0.13 as the expected response due to new mutants. Assuming the phenotypic variance remains relatively constant with $\sigma_z^2 \approx 5$, the expected contribution at generation 60 from initial variation is

$$\frac{\bar{A}_0(t)}{\sigma_z} = \frac{h^2(0) \cdot \sigma_A^2 \cdot e^{-t/(2N_e)}}{\sigma_z} = 1.4 \cdot \frac{0.2 \cdot 5 \cdot e^{-60/120}}{\sqrt{5}} \approx 0.38$$

Adding the two sources of response together gives an expected total rate of response of $0.38 + 0.13 = 0.51$ bristles/generation. While this is larger than the observed rate, opposing natural selection slowed down response in Yoo’s lines, as evidenced by the rather sharp decay in response upon relaxation of selection as well as the presence of segregating lethals within responding lines (Yoo 1980b). A Of the expected response, $0.38/0.51 = 75\%$ is due to the initial variation, while 25% is due to new mutation.

A complication with applying the above results is that the presence of major alleles both decreases the time to lose initial variation and increases the expected response from new mutants, resulting in a larger role for mutational input than predicted from the infinitesimal model. For example, applying the approximation for mutations of large effect (Equation 26.37b), the per-generation response from mutation is 0.32. Assuming the initial variation decays according to the infinitesimal model gives a total response of $0.38 + 0.32 = 0.70$, so that mutation now accounts for $0.32/0.70 = 0.46\%$ of the total response. However, when major genes are present, the initial variation declines even faster than predicted by Equation 26.22a, giving an even higher percentage of response from new mutations.
Expected Asymptotic Response Under More General Conditions

The infinitesimal model assumes allele frequency changes are due entirely to drift. Clearly, selection can also change allele frequencies and in this case other methods of analysis are required. One approach (Hill 1982b,c) is to consider the expected contribution resulting from the eventual fixation of some of the new mutants that arise each generation. Provided mutation and selection remain constant over time, at equilibrium the rate of response equals this expected per generation contribution. Assuming \( M \) adults are measured, the frequency of a new mutant allele \( A' \) is \( \frac{1}{2M} \). To allow for dominance, assume the genotypic values of \( AA' \) and \( A'A' \) are incremented by \( a(1 + k) \) and \( 2a \) relative to the genotypic value of \( AA \). As before we assume that the joint distribution of \( a \) and \( k \) values is independent of the genotypic value of the parental allele. Let \( f(a, k) \) denote this joint probability density function and let \( \lambda = \sum \mu_i \) be the total gametic mutation rate. The expected contribution from a new mutant appearing as a single copy is thus \( 2a \cdot u \left( \frac{1}{2M}, a, k \right) \), the change in genotypic value if the new allele is fixed times its probability of fixation (the latter can be obtained by Equation 26.4, using the fitnesses given by Equation 11.4). Since \( 2M \lambda \) new mutants appear each generation, the asymptotic rate of response is

\[
\begin{align*}
\lambda_m(\infty) &= 2M \lambda E \left[ 2a \cdot u \left( \frac{1}{2M}, a, k \right) \right] \\
&= 2M \lambda \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 2a \cdot u \left( \frac{1}{2M}, a, k \right) f(a, k) da dk
\end{align*}
\]

(26.39)

Thus, the expected asymptotic rate depends critically on the exact shape of the distribution of mutational effects. Fortunately, some fairly general results emerge by using simple approximations for the probability of fixation (similar to Equations 26.5a and 26.5b, see Hill 1982b,c for details).

Consider first the case where all new mutants are additive (\( k = 0 \)). Hill (1982b) found that, provided major alleles are not common among new mutants,

\[
\lambda_m(\infty) \simeq 2N_e \lambda \frac{E^+(a^2)}{\sigma_z} = \frac{4N_e \tau \sigma_m^2}{\sigma_z} \frac{E^+(a^2)}{E(a^2)}
\]

(26.40)

where

\[
E^+(a^2) = \int_0^{\infty} a^2 f(a) da
\]

If \( f(a) \) is symmetric about zero, then \( E^+(a^2) = E(a^2)/2 \), and the asymptotic response reduces to Equation 26.36. When major alleles are common among new mutants, correction terms involving \( E^+(a^3) \) appear; see Hill (1982b) for details. With divergent selection, effects due to asymmetry in \( f(a) \) cancel and the asymptotic rate of divergence between high and low lines is just twice the rate (for single-direction selection) predicted from the infinitesimal model,

\[
4N_e \tau \frac{\sigma_m^2}{\sigma_z}
\]

independent of the shape of \( f(a) \). The effect of linkage on asymptotic response was examined by Keightley and Hill (1983, 1987), who found it to generally be small, with the relative effects of linkage increasing as \( \sigma_m^2 \) and/or \( N_e \) increase.

Hill and Keightley (1988) incorporate natural selection, assuming new mutants also influence fitness under natural selection. If both character and fitness effects are small, the distribution of \( a \) is symmetric, and natural selection changes are also symmetric in \( a \) (e.g.,
the change in fitness is a function only of |a|, there is no change in the asymptotic rate of response. When these assumptions are violated, the asymptotic rate can be reduced.

To allow for dominance, we continue to assume the incremental mutation model. From LW Equation 4.12a, the additive variance contributed by a rare allele is

\[
2p(1-p)a^2[1 + k(1 - 2p)]^2 \simeq 2pa^2(1 + k)^2,
\]
giving the contribution from a single new mutation, with \( p = (2M)^{-1} \), as approximately

\[
a^2(1 + k)^2/M
\]

Since the expected number of new mutants per locus in any given generation is \( 2M\mu \), the expected additive variance contributed each generation by new mutants a locus is

\[
2M\mu E[a^2(1 + k)^2/M] = 2\mu E[a^2(1 + k)^2]
\]

where the expectation is taken over the joint distribution of \( a \) and \( k \) values in new mutants. Summing over all loci, the expected new additive variance contributed each generation is

\[
\sigma_m^2 = 2\sum_{i=1}^{n} \mu_i E[a^2(1 + k)^2] = 2\lambda E[a^2(1 + k)^2] \tag{26.41a}
\]

Hill (1982b). The last equality assumes that the distribution of mutational values is the same at each locus. When all mutants are additive (\( k = 0 \)) and \( E(a) = 0 \), this reduces to our previous definition of \( \sigma_m^2 \). More generally, with complete additivity

\[
\sigma_m^2 = 2\lambda E(a^2) \tag{26.41b}
\]

while with complete dominance

\[
\sigma_m^2 = 2\lambda E([2a]^2) = 8\lambda E(a^2) \tag{26.41c}
\]

Thus, all else being equal, the mutational variance with complete dominance is four times larger than that for complete additivity. Finally, it is important to note that with dominance Equation 26.35a (predicting the additive genetic variance from new mutation) no longer holds, even if the other assumptions of the infinitesimal model still do, as the additive variance can actually increase over some interval of time (see Chapter 3).

For the case of complete dominance (\( k = 1 \)), Hill (1982b) found that the asymptotic rate of response is approximately

\[
r_m(\infty) \simeq 16N_e \tau \lambda E^+ (a^2) / \sigma_z \tag{26.42a}
\]

With a symmetric distribution of mutational effects this reduces to

\[
r_m(\infty) \simeq N_e \tau \frac{\sigma_m^2}{\sigma_z} \tag{26.42b}
\]

where \( \sigma_m^2 \) is given by Equation 26.41c. Thus, for the same \( \sigma_m^2 \), the response when all mutants are completely dominant is only half the expected response when alleles are additive (compare Equations 26.36 and 26.42b). However, for fixed \( \lambda \) and \( E(a^2) \), \( \sigma_m^2 \) is larger with complete dominance (compare Equations 26.41b and 26.41c) and the response under dominance is twice as large as that expected for complete additivity.

If alleles are completely recessive, allelic effects are small and the distribution of mutational effects is symmetric, the asymptotic response is approximately

\[
r_m(\infty) \simeq 2N_e \tau \lambda E(a^2) / \sigma_z \tag{26.43a}
\]
Hill (1982b). For recessives with large effects (cf. Equation 26.5b),

\[ r_m(\infty) \approx 2 \lambda E^z \left( a^{3/2} \right) \sqrt{\frac{2N_e}{\pi \sigma_z}} \]  

(26.43b)

Thus, the limiting response when all new mutants are recessive is not predictable from \( \sigma_m^2 \), even if mutational effects are symmetrically distributed. With recessive major alleles, response scales as \( \sqrt{N_e \tau} \), and hence increases much more slowly than with complete dominance and/or additivity.

When loci are linked, the asymptotic response is reduced, but the effect is small unless linkage is tight, as might occur with a few small chromosomes (Keightley and Hill 1983). As mentioned previously, reduction in response also occurs if loci influencing the trait are also linked to loci under natural selection.

An alternative mutational model, where only \( n \) alleles are present (as opposed to our assumption of an infinite number of possible alleles), has been examined by Zeng et al. (1989). As expected, such a model results in an ultimate selection limit as mutation cannot continue indefinitely to generate better alleles. Yet another mutational model was considered in simulation studies by Li and Enfield (1992), wherein the new mutational value assigned a new value from a distribution of know gene effects. Thus, the new value is given by \( a \), rather than \( G_{AA} + a \), so that the new genotypic value is completely independent of the original value. Under this model, Li and Enfield found (assuming a finite number of loci) that genetic variance increases to a maximum value, and then declines somewhat afterwards.

**Optimizing Asymptotic Selection Response**

Since the asymptotic response is a function of \( N_e \tau \) (see Equations 26.34-26.37, 26.40, 26.42-26.43), response is maximized by selection strategies that maximize this product. Thus, as was the case for maximizing long-term response (the total response using only the initial variation), there is a tradeoff in that the optimal short-term response (maximizing \( \tau \) ) is at conflict with the optimal asymptotic response (as increasing \( \tau \) decreases \( N_e \)). Thus if our choice is simply the fraction of individuals to save, our discussions above on the optimal selection intensity for long-term response also applies to considerations of the asymptotic response.

The selection intensity is not the only choice a breeder or experimentalist has in terms of possible selection schemes. We have generally been assuming individual (or mass) selection, which is based solely on an individual’s phenotype. There are, however, numerous other selection schemes, such as those incorporating information on the phenotypes of relatives as well (e.g., family and index/BLUP selection). Schemes incorporating such information can improve the accuracy of the estimate of an individual’s breeding value, and hence improve the accuracy of short-term response. This can be seen by recalling (Equation 4.4d) that the single-generation response \( R \) for any particular selection scheme is given by \( R/ (\tau \sigma_A A) = \rho(x, A) \), where selection occurs on the measure \( x \) and \( \rho(x, A) \) is the accuracy of the method (the correlation between \( x \) and an individual’s breeding value \( A \)). Holding \( \tau \) constant, the single-generation response increases with the accuracy \( \rho(x, A) \) of the selection method. Thus while different schemes can improve the short-term response over mass selection, what is their effect on asymptotic response? The answer is that, once again, schemes improving the short-term response often do so at the expense of the asymptotic response.

Optimal asymptotic response occurs by maximizing the fixation probabilities of favorable QTLs, which amounts to maximizing \( N_e s \), where \( s \) is the selection coefficient on the QTL. For an additive trait, Hill (1985) and Caballero et al. (1996) generalize Equation 11.4 to show that

\[ s = \left( \frac{\tau}{\sigma_z} \right) \frac{\rho(x, A)}{h} \]  

(26.44)
LONG-TERM RESPONSE IN FINITE POPULATIONS

331

(Under individual selection, \( \rho(x, A) = h \), recovering Equation 11.4.) Fixation probabilities under different selection schemes with the same selection intensities are thus functions of the product \( N_e \rho(x, A) \). The tradeoff is that increasing \( \rho(x, A) \) typically decreases \( N_e \) by increasing the between-family variance. Hence, as was the case for the optimal selection intensity, the optimal selection scheme for short-term response may differ from the optimal long-term one.

The accuracy \( \rho \) depends on the genetic variance and hence changes over time as these variances change. As we have see these changes in variances can be extremely difficult to predict (Chapters 9 -11). Once again, the analysis is greatly simplified by assuming the infinitesimal model. Under this model, the additive genetic variance eventually converges to a value of \( \sigma^2_{A,\infty} = 2N_e \sigma^2_m \). Here the effect of different selection schemes on the equilibrium additive variance (and \( \rho \)) is entirely determined by the effective population size that the scheme generates. In comparing two different selection schemes (i and j) with the same selection intensity, Wei et al. (1996) show that the ratio of asymptotic responses becomes

\[
\frac{\bar{R}_i}{\bar{R}_j} = \frac{\bar{\rho}(i) \tilde{\sigma}_A(i)}{\bar{\rho}(j) \tilde{\sigma}_A(j)} = \frac{\bar{\rho}(i)}{\bar{\rho}(j)} \sqrt{\frac{N_e(i)}{N_e(j)}}
\]

(26.45)

The careful reader will note that the effect of \( N_e \) is two-fold — a direct effect (the square root of their ratio) and an indirect effect through the ratio of \( \rho \) (which is a function of \( \tilde{\sigma}_A \), and hence of \( N_e \)).

**Example 26.6.** Consider the asymptotic response to mass \((m)\) versus within-family \((w)\) selection. Under within-family selection, \( N_{e(w)} \simeq 2N_e \), as the between-family variance is zero (see Equation 2.xx). In contrast, \( N_{e(m)} < N_e \), with the difference between \( N_{e(m)} \) and \( N_e \) increasing with the selection intensity and heritability (Equation 26.15). Thus,

\[
\sqrt{\frac{N_{e(w)}}{N_{e(m)}}} \geq \sqrt{2}
\]

Wei et al. (1996) found that the asymptotic accuracies are

\[
\bar{\rho}_{(m)} = \frac{2N_{e(m)} \sigma^2_m}{\sqrt{2N_{e(m)} \sigma^2_m [2N_{e(m)} \sigma^2_m + \sigma^2_e]}}
\]

\[
\bar{\rho}_{(w)} = \frac{N_{e(w)} \sigma^2_m}{\sqrt{2N_{e(w)} \sigma^2_m [2N_{e(w)} \sigma^2_m + \sigma^2_e]}}
\]

Using the above inequality and these two identities, it can be shown that

\[
\frac{\rho_{(w,\infty)}}{\rho_{(m,\infty)}} \geq \frac{1}{\sqrt{2}}
\]

Thus,

\[
\frac{\bar{R}_w}{\bar{R}_m} = \left[ \sqrt{\frac{N_{e(w)}}{N_{e(m)}}} \right] \left[ \frac{\bar{\rho}(w)}{\bar{\rho}(m)} \right] \geq \sqrt{2} \frac{1}{\sqrt{2}} = 1
\]

and hence \( \bar{R}_w \geq \bar{R}_m \). Thus for the same selection intensity, the long-term response is greater under within-family selection than under mass selection, even though mass selection is initially twice as accurate as within-family selection (and hence the initial response is twice as large).
The effects of different selection schemes on the effective population size can be seen by considering the general weighted index of within- and between-family information,

\[ I = (z - \bar{z}_f) + \lambda(\bar{z}_f - \bar{z}) = (\text{within-family}) + \lambda(\text{between-family}) \quad (26.46) \]

where \( z \) is an individual’s value, \( \bar{z}_f \) is the mean of its family, and \( \bar{z} \) is the grand mean. A number of selection schemes can be represented (either exactly or to a good approximation) by this index. For example, \( \lambda = 1 \) corresponds to individual selection \( (I = z) \), while \( \lambda = 0 \) corresponds to strict within-family selection \( (I = z_w) \). By choosing the appropriate \( \lambda \), the accuracy of selection using this index is greater than the accuracy of individual selection \( (\rho(I,A) > \rho(z,A), \text{see Chapter 20}) \), and hence selection using the optimal index gives a greater short-term response than mass selection. To a first approximation, BLUP selection corresponds to this optimal index (see Chapter xx).

Since the reduction in effective population size occurs by inflating the between-family variance, the greater the \( \lambda \) value in the index given by Equation 26.46, the greater the reduction in \( N_e \). Larger values of \( \lambda \) place more weight on family information, resulting in more individuals from the best families being coselected. Somewhat ironically, the reduction in \( N_e \) is greatest when heritability is small, as in these cases the index places the most weight on the between-family component. Yet it is exactly these cases where index/BLUP selection has the greatest short-term advantage over individual selection. While index/BLUP selection gives the largest single-generation response, when care is taken to equalize the amount of inbreeding across methods, individual selection can produce a larger single-generation response than index selection or BLUP (Quinton et al. 1992, Andersson et al. 1998).

Can one balance this tradeoff between increased accuracy for short-term response versus inflation of the between-family variance versus the reduction in the long-term response via reduction in \( N_e \)? Several authors have proposed schemes for reducing the between-family variance following selection. Toro and colleagues (Toro and Nieto 1984, Toro et al. 1988, Toro and Pérez-Enciso 1990) suggested that selected individuals be mated in ways that minimize the coancestry between them. A slightly different strategy, compensatory mating, was suggested by Grundy et al. (1994). Here individuals from families with many representatives following selection are mated to individuals with few family members. This has the effect of reducing the cumulative effect \( (Q \text{ in Equation 26.13}) \) of selection towards decreasing the effective population size by reducing the variance in family contribution. Grundy et al. also suggested a more subtle approach. They note that by using biased selection parameters in the index (for example, using upwardly biased estimates of \( h^2 \) when computing the optimal \( \lambda \)), the slight reduction in the accuracy of the adjusted index from its optimal value is more than offset by a much larger decrease in the reduction in \( N_e \). They suggest that this approach, combined with compensatory mating provides a simple way for ameliorating the reduction in \( N_e \).

This tradeoff between optimal short-term versus optimal asymptotic response certainly has economical consequences for breeders. While breeders are ultimately better off (in terms of total response) using selection schemes that are initially less accurate, competing breeders using the initially more accurate schemes will achieve large short-term response. Breeders must thus decide between staying in business over the short haul versus a larger payoff (in terms of a greater response) over the long run. As we will discuss in Chapter 20, these same considerations arise when considering how much weight to place on marker-assisted selection.
Literature Cited


Lynch, M. 1986. [26]


LONG-TERM RESPONSE IN FINITE POPULATIONS


