Evolution, 43(8), 1989, pp. 1767–1780

THE MAINTENANCE OF POLYGENIC VARIATION IN FINITE POPULATIONS

DAVID HOULE[1]
Department of Ecology and Evolution, State University of New York, Stony Brook, NY 11794

Abstract.—Models of the maintenance of genetic variance in a polygenic trait have usually assumed that population size is infinite and that selection is weak. Consequently, they will overestimate the amount of variation maintained in finite populations. I derive approximations for the equilibrium genetic variance, $\hat{V}_G$ in finite populations under weak stabilizing selection for triallelic loci and for an infinite "rare alleles" model. These are compared to results for neutral characters, to the "Gaussian allelic" model, and to Wright's approximation for a biallelic locus under arbitrary selection pressures. For a variety of parameter values, the three-allele, Gaussian, and Wrightian approximations all converge on the neutral model when population size is small. As expected, far less equilibrium genetic variance can be maintained if effective population size, N, is on the order of a few hundred than if N is infinite. All of the models predict that comparisons among populations with N less than about 10^4 should show substantial differences in $\hat{V}_G$. While it is easier to maintain absolute $\hat{V}_G$ when alleles interact to yield dominance or overdominance for fitness, less additivity also makes $\hat{V}_G$ more susceptible to differences in N. I argue that experimental data do not seem to reflect the predicted degree of relationship between N and $\hat{V}_G$. This calls into question the ability of mutation-selection balance or simple balancing selection to explain observed V_G. The dependence of $\hat{V}_G$ on N could be used to test the adequacy of mutation-selection balance models.

Received May 25, 1988. Accepted May 22, 1989

The effects of both mutation and selection on phenotypes are ubiquitous and relatively easily observed, making mutation-selection balance an appealing, parsimonious explanation for the persistence of additive genetic variance in traits undergoing selection. In the last 20 years a spate of models of traits undergoing stabilizing selection have generally supported this possibility (for a review, see Turelli [1984]). These have usually made the assumption that population size is infinite. Lande (1980*a*) asserted that his infinite-population approximation would remain adequate for most natural populations.

All of these mutation-selection balance models make the reasonable prediction that stronger selection will result in less genetic variance at equilibrium. However, in a finite population, the weaker selection is, the greater effect drift will have on gene frequencies and genetic variance. This suggests that parameters that lead to high genetic variance in an infinite population will lead to a strong dependence of genetic variance on effective population size. In this paper, I describe the magnitude of this dependence for several simple models of the maintenance of genetic variance.

Three basic kinds of approaches have been used for modeling the maintenance of genetic variance by mutation-selection balance when there is stabilizing selection on the modeled character. The first two of these model loci that have an infinite number of alleles with a continuous range of effects. The "Gaussian allelic" approximation to this model assumes that the distribution of genotypic values at each locus is normal (Kimura, 1965; Latter, 1970; Lande, 1976; Fleming, 1979; Barton and Turelli, 1987) and therefore necessitates the assumption that mutation is a stronger force than selection (Turelli, 1984). The "rare alleles" approximation (Turelli, 1984; Barton and Turelli, 1987) assumes that selection is stronger than mutation, keeping the frequencies of mutant alleles low.

The third approach is to consider only a finite number of alleles at each locus. Models have been constructed using two (Latter, 1960; Bulmer, 1972, 1980; Barton, 1986), three (Turelli, 1984), and five (Slatkin, 1987) alleles. These finite-allele models may either be viewed as more tractable approximations to the continuum-of-alleles model, or

[1] Present address: Department of Statistics, North Carolina State University, Box 8203, Raleigh, NC 27695-8203.

as models of loci for which variation within a range of allelic effects is only weakly selected, while variation outside this range is very deleterious. Slatkin (1987) showed that the finite-allele models may adequately reflect genetic variance in an infinite-alleles model when mutation and selection are nearly equal in strength, thereby plugging the gap between the parameters appropriate to the Gaussian and rare-alleles approximations. All three approaches predict that mutation-selection balance at plausible numbers of loci could explain observed levels of genetic variance.

Several authors have introduced finite population size into these models. Nei and Imaizumi (1966) showed that genetic variance is very sensitive to population size for a variety of selection schemes using Wright's (1937) equation for the probability-density distribution of allele frequencies at a biallelic locus. Bulmer (1972), also starting from Wright's equation, again noted the power of drift to decrease the equilibrium genetic variance for one combination of parameters. Latter (1970) derived an expression for the equilibrium genetic variance relative to the average effect of a mutation for his discrete-time Gaussian model but did not discuss its implications. In spite of this work, discussion of the effect of finite population size has rarely entered the debate about the efficacy of mutation-selection balance for maintaining polygenic variation. In part, this is because none of these authors emphasized the relevance of this to natural populations. Other relevant work includes the extension of the slightly deleterious near-neutral model to loci contributing variance to a character under stabilizing selection (Kimura, 1981; Foley, 1987), showing that substitution rate and, therefore, heterozygosity will be a function of population size.

Alternative models to those based on mutation-selection balance have been little considered. Several authors have modeled genetic variance maintained by overdominance (Robertson, 1956; Bulmer, 1973; Gillespie, 1984). Robertson modeled overdominant loci with additive effects on a neutral phenotypic character, while Bulmer and Gillespie combined overdominance with stabilizing selection on the phenotypic character. Thus, these models involve pleiotropy, as genotypes must differ in fitness independent of their effects on the phenotype of interest. These models can easily maintain genetic variance with a small number of loci in an infinite population. Bulmer (1973), again using Wright's (1937) allelic probability-density distribution, derived an approximation for genetic variance in a finite population under such a selection scheme and stated that a "considerable amount" of genetic variance could be maintained in "moderate-sized" populations, although he gave no actual values. Several authors have considered the important special case of neutral phenotypic characters (Clayton and Robertson, 1955; Lande, 1979; Chakraborty and Nei, 1982; Lynch and Hill, 1986). Lynch and Hill performed the most thorough analysis, showing that the genetic variance of neutral characters will be extremely dependent on population size under a wide variety of linkage and dominance relationships and mating systems.

The relative lack of attention to finite population size in mutation-selection balance models has recently been corrected. Keightley and Hill (1988), Bürger et al. (1989), and the present paper all address finite population size, using different approaches but with comparable results. In this paper, I derive the expected genetic variance under mutation, stabilizing selection, and drift for a three-allele model (Turelli, 1984). I then compare this model to Latter's (1970) "Gaussian" model, to a simple neutral model, and to Wright's (1937) model of a biallelic locus with arbitrary dominance for fitness. This allows consideration of alleles that are recessive or overdominant for fitness. All of these models predict that, under standard assumptions about the average effect of a mutation and the strength of selection, populations with effective sizes less than about 10^4 that differ in effective population size will also differ markedly in their equilibrium genetic variances.

Models

All of the models discussed here share the assumptions of a diploid, random-mating population. I assume that departures from gametic-phase equilibrium can be ignored. This is likely to be a valid assumption either when population size is very small (Keightley and Hill, 1988), when alleles are effectively neutral (Lynch and Hill, 1986), or

when population size is effectively infinite (Lande, 1980*b*; Turelli, 1984), as long as loci are not strongly clumped in their distribution. In other situations, gametic disequilibrium will tend to lead to slight overestimates of the equilibrium genetic variance (Keightley and Hill, 1988). I return to this issue in the Discussion. Alleles are assumed to interact additively in determining the phenotype, and nonallelic interactions (epistasis) are assumed to be absent. I assume that there are n equivalent loci which influence the phenotype of a character Z. The genotypic value of an individual, G, is the sum of the genotypic values at each locus, g. The phenotype of an individual is determined by its genotypic value, plus a normally distributed random environmental deviation E, which has mean 0 and variance 1. I assume that the mean phenotype in the population is 0 ($\bar{Z} = 0$). Therefore the genetic variance at equilibrium, $\hat{V}_G$, is

$$\hat{V}_G = n\hat{V}_g = \widehat{ng^2} \tag{1}$$

where $\hat{V}_G$, $\hat{V}_g$, and $\widehat{g^2}$ are the expected values of V_G (the genetic variance), V_g (the genetic variance at a single locus), and g^2. Mutation occurs at the rate μ, and the average effect of a mutation on the phenotype is c. The increase in genetic variance by mutation per zygote per generation is then

$$V_M = 2n\mu c^2. \tag{2}$$

To model stabilizing selection, I assume that the fitness function is nor-optimal, or Gaussian, with the optimum at $Z = 0$, so

$$W(Z) = \exp\left[\frac{-Z^2}{2\omega^2}\right], \tag{3}$$

where ω^2 determines the strength of stabilizing selection. The larger the value of ω^2, the weaker the selection is. The mean fitness of individuals with genotypic value G is

$$W(G) = \exp\left[\frac{-G^2}{2V_s}\right], \tag{4}$$

where V_s is the sum of ω^2 and the environmental variance, which is assumed to equal 1. A convenient summary of the strength of selection is the average amount of selection against a homozygote for an allele that is an average mutational step away from the optimum, $s = (2c)^2/2V_s$.

Three-Allele Model.—Turelli (1984) developed a simple three-allele model of genetic variance for characters under stabilizing selection. The parameters for each locus are given in Table 1. Mutation among the three alleles depends on the rate μ and obeys the following scheme

$$A_{-1} \underset{\frac{\mu}{2}}{\overset{\mu}{\rightleftharpoons}} A_0 \underset{\frac{\mu}{2}}{\overset{\frac{\mu}{2}}{\rightleftharpoons}} A_{+1}$$

Additive genetic variance is proportional to heterozygosity (H) at loci with additive allelic interactions ($V_G = c^2H$), and therefore, the change in genetic variance due to drift alone will be

$$\Delta V_G = \frac{-V_G}{2N} = -nc^2\frac{H}{2N} \tag{5}$$

where N is the effective population size (Clayton and Robertson, 1955). Due to the symmetry of the three-allele model, I can assume $\hat{p}_{+1} = \hat{p}_{-1}$, and

$$\hat{H} \approx 4\hat{p}_{+1}. \tag{6}$$

This assumption necessarily constrains the mean effect of each locus to be 0. In a finite population, this expectation will remain true; however, it will almost never be realized.

The change in p_{+1} due to mutation and selection is

$$\Delta p_{+1} \approx \frac{p_{+1}\left(1 - p_{+1}\frac{s}{2} - \frac{s}{4}\right)}{1 - p_{+1}s} + \frac{\mu}{2} - 2\mu p_{+1} - p_{+1} \tag{7}$$

assuming that selection is weak (i.e., $s \ll 1$). After substituting $H/4$ for p_{+1}, and solving for ΔH, the effect of drift can be added. This gives

$$\Delta H \approx \frac{H\left(1 - H\frac{s}{8} - \frac{s}{4}\right)}{1 - H\frac{s}{4}} - H + 2\mu - 2\mu H - \frac{H}{2N}. \tag{8}$$

At equilibrium, this yields a quadratic in H. The appropriate solution to this, when multiplied by nc^2 gives the approximation to the genetic variance that appears as Equation (9) at the bottom of the page, where $\theta = 4N\mu$. Results obtained using Equation (9) are nearly identical to simulation results reported by Bürger et al. (1989). If $\hat{p}_{+1}$ is much less than one, which is equivalent to assuming that μ is much less than $s/4$, Equation (9) may be approximated as

$$\hat{V}_G \approx \frac{V_M}{\frac{s}{4} + \frac{1}{2N}}. \quad (10)$$

Equation (10) has recently been derived independently by Keightley and Hill (1988) and suggested by Bürger et al. (1989). When N is infinite and μ is much less than $s/4$, both Equations (9) and (10) reduce to $\hat{V}_G \approx 4n\mu V_s$, which is a standard approximation valid both for finite-allele models (Latter, 1960; Bulmer, 1972; Turelli, 1984) and for infinite-allele models (Turelli, 1984). Note that while the average effect of a mutation does not affect $\hat{V}_G$ in an infinite population, it does in a finite one.

Equation (10) may also be derived for an infinite-allele model, using Barton and Turelli's (1987) rare-alleles assumptions, which necessitate the relation $\mu \ll s/4$: using Barton and Turelli's equation 5.3,

$$\frac{dV_G}{dt} \approx V_G\left(\frac{-c^2}{2V_s}\right) + V_G\left(-\frac{1}{2N}\right) + V_M \quad (11)$$

which at equilibrium gives Equation (10).

Gaussian Allelic Approximation.—An alternative to the rare-alleles approximation to an infinite-alleles model may be obtained by assuming that the distribution of allelic effects is always Gaussian at each locus (Kimura, 1965). Latter (1970) derived a recursion equation for the genetic variance with finite population size under a discrete-time version of this Gaussian allelic model. For the n-locus case, Latter's recursion has the equilibrium solution

$$\hat{V}_G \approx \sqrt{D^2 + \frac{nN}{N+n} 2V_M V_s} - D \quad (12)$$

where

$$D = \frac{nV_s - 2nNV_M}{2(N+n)}. \quad (13)$$

TABLE 1. Parameters for each locus in the three-allele model of Turelli (1984).

Allele	Frequency	Phenotypic effect
A_{-1}	p_{-1}	$-c$
A_0	p_0	0
A_{+1}	p_{+1}	c

Wright's Two-Allele Model.—If alleles have effects on many characters, then their effects on fitness may be essentially independent of their effects on any one phenotype. Thus, while an allele may act additively on phenotype, its dominance for fitness may be arbitrary. In order to incorporate arbitrary dominance, I consider the standard two-allele model in Table 2. The average effect of a mutation, c, may be either negative or positive for a locus. Note that I assume that the effect of alleles on the phenotype is additive. Under nor-optimal selection, the dominance, h, of a mutant allele with respect to fitness is 0.25. While this is plausible if stabilizing selection on phenotypes accounts for the majority of selection on polygenes, other cases, such as complete dominance ($h = 0$), or overdominance ($h > 0$) are also possible.

In a finite population of size N, with reversible mutation between the two alleles at rate μ, Wright (1937, 1969) showed that the probability distribution of q is approximately

$$\phi(q) = K\bar{W}^{2N}q^{\theta-1}(1-q)^{\theta-1} \quad (14)$$

where K is an integration constant which

$$\hat{V}_G \approx nc^2\left(\frac{2 + Ns + 2\theta - \sqrt{(2 + Ns + 2\theta)^2 - 4s\theta(N+1)}}{s(N+1)}\right) \quad (9)$$

TABLE 2. Standard two-allele model with arbitrary dominance for fitness.

Genotype	Frequency	Genotypic value	Relative fitness
A_0A_0	$(1 - q)^2$	0	1
A_0A_1	$2q(1 - q)$	c	$1 - hs$
A_1A_1	q^2	$2c$	$1 - s$

need not be evaluated and $\theta = 4N\mu$. I calculated the expected genetic variances from this distribution using equations 3 and 4 of Nei and Imaizumi (1966). In order to do this, the first two moments of Equation (14) had to be calculated by numerical integration, for which I used method II of Kimura et al. (1963) when $\theta < 1$, and direct integration when $\theta > 1$.

Neutral Phenotypic Characters.—For a character whose variance is solely determined by neutral or effectively neutral alleles, a drift-mutation balance determines the equilibrium variance. Therefore, for unlinked loci capable of mutating to a continuum of alleles and with additive allelic interactions,

$$\hat{V}_G = 2NV_M \qquad (15)$$

(Clayton and Robertson, 1955; Lynch and Hill, 1986; Chakraborty and Nei, 1982; Lande, 1979). Lynch and Hill (1986) showed that this will remain approximately true under a wide variety of linkage relationships and kinds of allelic interactions. As Ns becomes less than one, all of the approximations will approach Equation (15).

Results

Three-Allele Model vs. Wright's Model.—A check on Equation (9), the equilibrium genetic variance under the three-allele model, is provided by the Wright model for a biallelic locus when the dominance (h) is 0.25. When μ is much less than $s/4$, nonoptimal alleles will be rare in the three-allele model, and the genotype $A_{+1}A_{-1}$ may be ingored, reducing the three-allele model to a two-allele model. Values of $\hat{V}_G$ for the three-allele and Wrightian models should be identical for parameter combinations in which this condition is met, and the three-allele model should exceed the Wrightian models when $\mu \approx s/4$. Figure 1 is a graph of the ratio between the Wright approximation and the three-allele model for $\mu = 10^{-5}$. The two models agree well when drift is very powerful ($Ns \ll 1$), as in the front corner of the surface, and when selection is powerful ($Ns \gg 1$), as in the back corner. As expected, when $\mu \approx s/4$, at the right corner, the approximation of the Wright model predicts more equilibrium genetic variance than the three-allele model. The trough seen running across the surface, however, is not expected. This trough indicates that the Wright model predicts that heterozygosity will be higher than that in the three-allele model as one makes the transition from mutation-drift balance ($Ns \ll 1$) to mutation-selection balance ($Ns \gg 1$). Thus, the low point of the trough falls where $Ns = 1$. However, this discrepancy does not exceed 25% of the prediction based on the Wright model. This amount of error is relatively insignificant for the conclusions that may be drawn from such crude models.

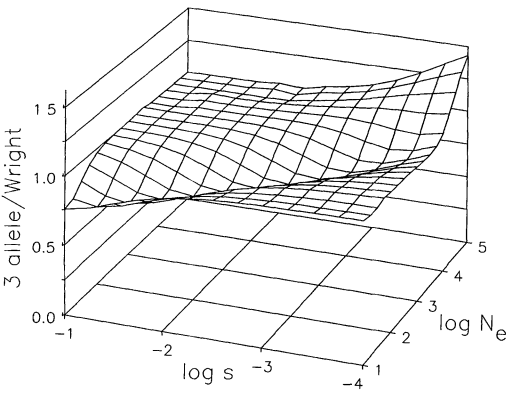

FIG. 1. Comparison of $\hat{V}_G$ in the three-allele model and the biallelic Wright model when $h = 0.25$ and $\mu = 10^{-5}$. The surface is the ratio (three-allele prediction)/(Wright prediction).

Absolute Genetic Variance.—In this section, I will focus on parameter combinations that yield a new mutational variance (V_M) of $0.001V_E$. Experiments suggest that this is a reasonable median value for weakly selected characters (Lynch, 1988).

Figure 2 gives contour plots of heritability (h^2) over a range of effective population sizes and numbers of loci for the three-allele, Gaussian allelic and neutral models. With the assumption that the environmental variance is 1.0, $h^2 = V_G/(1 + V_G)$. The average effect of a mutation was chosen to yield $V_M = 0.001$ with $n = 100$. Both the

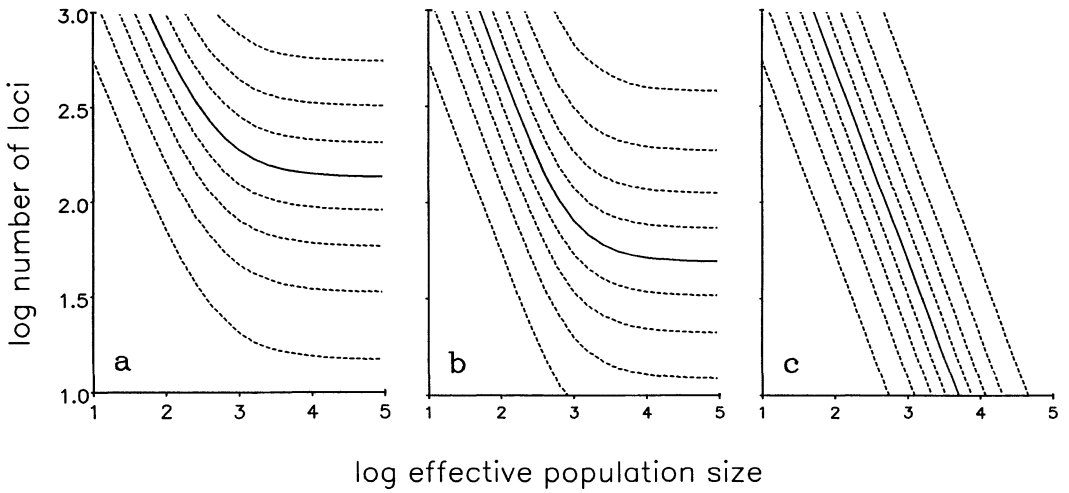

FIG. 2. Contour plots of heritability (h^2) in relation to effective population size and the number of loci: a) The three-allele model; b) the Gaussian model; and c) the neutral model. The solid lines indicate a heritability of 0.5, and the contour interval is 0.1. Low heritabilities fall in the lower left-hand corner of each panel. In each case, the parameters $\mu = 10^{-4}$, $c^2 = 0.05$, and $V_s = 20$ were used.

three-allele and Gaussian models yield substantial heritabilities if population size is greater than 10,000. At $N = 100$, both models have converged on the neutral model. In such a population, a heritability of 0.50 could only be maintained with nearly 1,000 loci potentially contributing variance. For the parameters of Figure 2, $s = 0.005$, quite a bit larger than the mutation rate, so the three-allele model is probably a more accurate predictor of $\hat{V}_G$. However, the Gaussian model converges more quickly on the neutral expectation, so the difference between the two models becomes less with decreasing population size.

In Figure 3, I graph the heritability produced by the three-allele, neutral, and Gaussian models under a wider variety of mutation rates and average effects. For each combination of mutation rate and average effect, I chose n to preserve the constraint $V_M = 0.001$. Panels at the upper right are well within the assumptions appropriate to the rare-alleles approximation. Only in the lower left panel is $\mu = s$. In each panel, the Gaussian approximation gives a higher heritability than the three-allele model, but in the eight panels where $\mu < s/4$, the three-allele model is probably a more accurate approximation of a continuum-of-alleles model.

The effect of drift changes predictably from relatively weak in the upper left panels of Figure 3, where mutation and selection are both powerful, to relatively strong in the lower right panels, where mutation and selection are weak. This is reflected in the larger effective population size at which genetic variance begins to fall off due to drift. This pattern of results suggests that, if mutation-selection balance does maintain a large proportion of the genetic variance in natural populations, one should expect the amount of genetic variance to vary with population size.

Nei and Imaizumi (1966) presented response curves for genetic variance with respect to population size for the Wrightian model; consequently, these are not presented here. The effects of varying dominance for fitness, h, for a given value of s are well known: the smaller h is, the more heterozygosity, genetic variance, and heritability is maintained at equilibrium. Since h for the nor-optimal selection models is 0.25, strict additivity for fitness maintains less equilibrium genetic variance than in the three-allele model. In the case of mutants recessive for fitness, $\hat{V}_G$ will fall between the predictions of the neutral and three-allele models. Overdominance with $\hat{q} > 0.15$ (Robertson, 1962) would maintain more genetic variance than the neutral model when N is small.

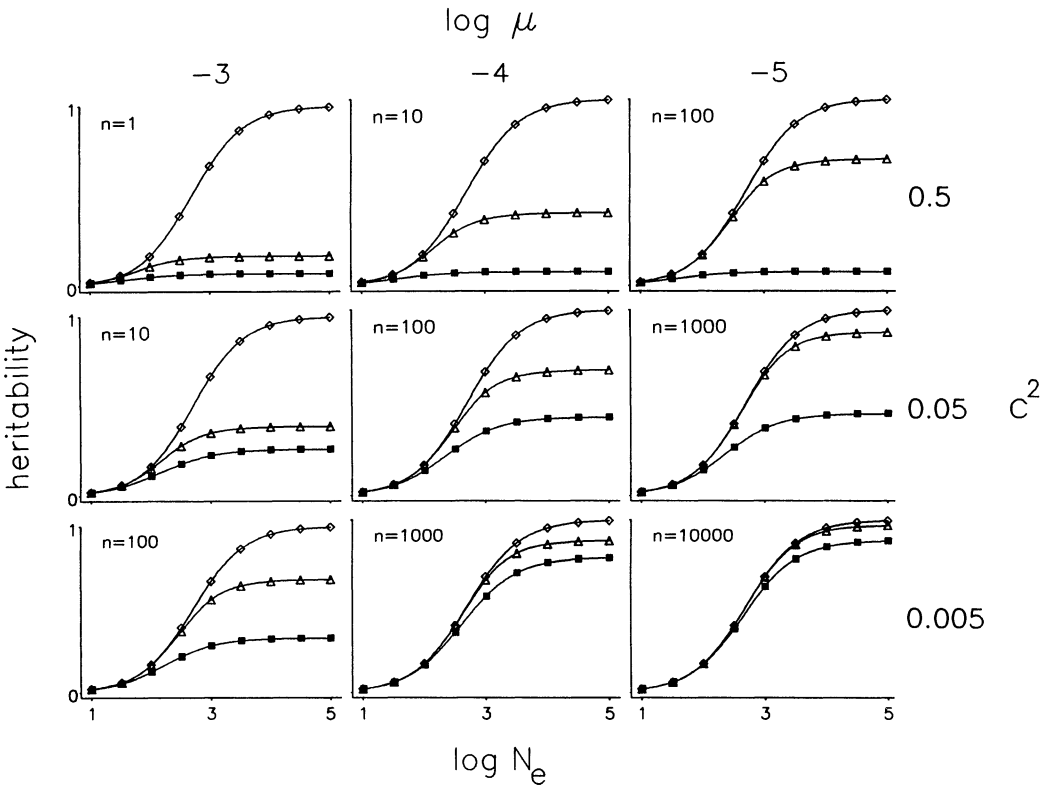

FIG. 3. Heritability when $V_M = 0.001$. The solid squares correspond to the three-allele model, triangles to the Gaussian model, and diamonds to the neutral model. For the three-allele and Gaussian models $V_s = 20$.

Relative Genetic Variance at Different Population Sizes.—Since we are ignorant of such parameters as numbers of loci, selection coefficients, and effective mutation rates at loci underlying quantitative characters, the genetic variance seen in any one population can be explained by any of the models presented here. If, however, we compare related populations of different effective population sizes, N, they will share the same underlying parameters, providing some potential to discriminate among selection models and parameter values. To present comparisons among populations of different sizes, I divide $\hat{V}_G$ for each approximation by $\hat{V}_G$ at $N = 1{,}000$, calculated from the same approximation. This value of N may be the proper order of magnitude for many vertebrates and plants. In Figures 4 and 5, the slope of the normalized $\hat{V}_G$ curve around $N = 1{,}000$ is a measure of the sensitivity of $\hat{V}_G$ to drift at that population size.

In Figure 4, I have graphed this normalized $\hat{V}_G$ for the same models and mutational parameters as in Figure 3. For the largest average effect used, $c = 0.5$, the sensitivity to drift in the three-allele model is slight in the neighborhood of $N = 1{,}000$. However, this average effect is about twice the maximum experimental estimates of the average effect of new mutations (Hoi-Sen, 1972; Russell et al., 1963), and the experimental data can only overestimate the true average effect. For all other parameter combinations, the sensitivity is high, especially when population size is reduced from 1,000.

In Figure 5, I explore the sensitivity of $\hat{V}_G$ to drift when dominance for fitness, h, is varied while the dominance for phenotype is held constant at 0.5, the purely additive case. To explore different degrees of dominance, I assume that selection is no longer solely a function of the phenotype of the character considered, as there is no way to obtain stable overdominance by weak selection on a quantitative character (Rob-

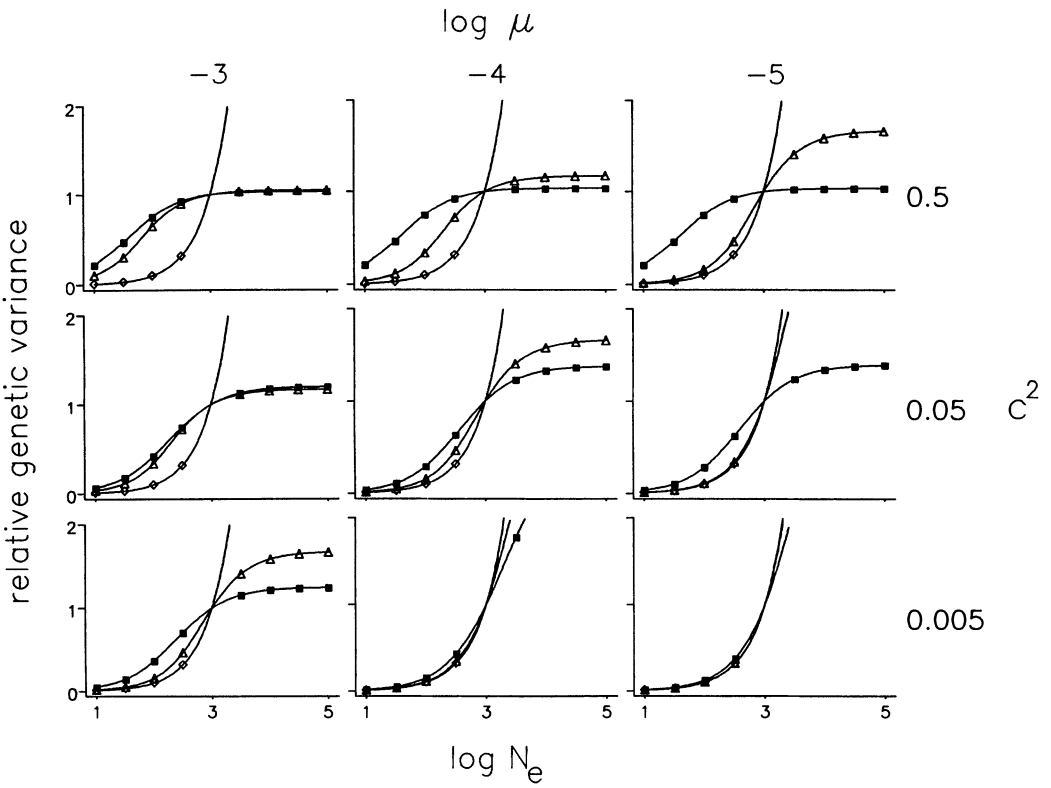

FIG. 4. Susceptibility of genetic variance to drift around $N = 1{,}000$ for the three-allele, Gaussian, and neutral models. Solid squares correspond to the three-allele model, triangles to the Gaussian model, and diamonds to the neutral model. Genetic variance was generated as in Figure 3, then normalized to 1.0 at $N = 1{,}000$ for each model.

ertson, 1956; Lewontin, 1964). Under such an assumption, there is no longer any necessary connection between s and c, so I separate them entirely and present the relative equilibrium heterozygosity, $\hat{H}$, as a function of s, rather than $\hat{V}_G$ as a function of c. Recall that $\hat{V}_G = nc^2\hat{H}$.

For comparison, I present data for the nor-optimal three-allele model to represent $h = 0.25$. To model other dominance relationships, I use the biallelic Wright model. I consider two cases: complete recessivity of nonoptimal alleles ($h = 0$) and overdominance with an equilibrium frequency of the lower-fitness allele of 0.3. In this case, I used an alternate parameterization, assuming that the fitness of the heterozygote is 1 and that the two homozygotes have fitnesses $1 - 0.429s$ and $1 - s$. For the overdominant case, there are no experimental guidelines for appropriate parameter values. It is clear that at most a small proportion of new mutations can interact overdominantly (Mukai et al., 1972). Since such alleles are almost never observed, we can be confident that their average effects on fitness must be small, probably less than a few percent. Therefore, the presentation of comparisons at particular combinations of s and μ is arbitrary. The appropriate μ and s for each kind of dominance relationship may differ by orders of magnitude. It is most unlikely that μ for overdominant alleles can be as high 10^{-3}; consequently, I present results for $\mu = 10^{-6}$ instead.

The results shown in Figure 5 demonstrate that populations in which genetic variance is maintained by mutant alleles with less dominance will be substantially more susceptible to drift when selection is relatively strong, and equally susceptible when selection is weak. For the overdominant case with the largest selection coefficient, increasing the population size from N

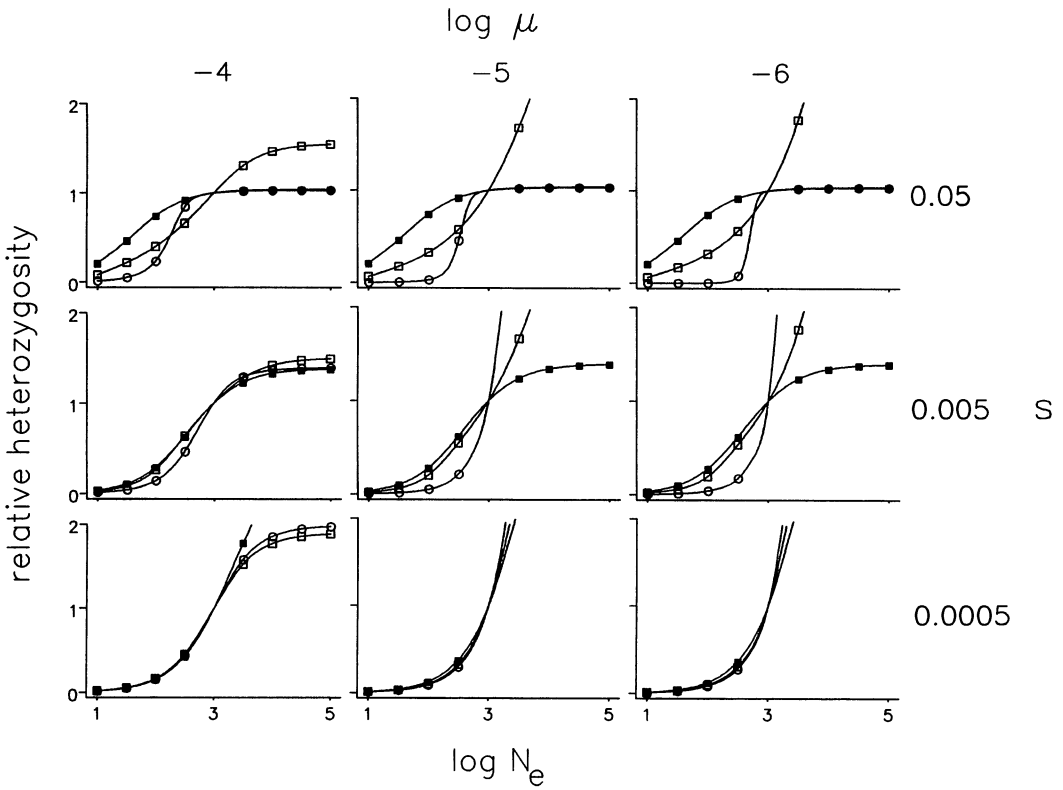

FIG. 5. Susceptibility of genetic variance to drift around N = 1,000 with arbitrary dominance for fitness. Solid squares correspond to the three-allele model, open squares to complete recessivity, and open circles to overdominance with $\hat{q}$ = 0.3. Genetic variance was normalized to 1.0 at N = 1,000, as in Figure 4.

= 1,000 is predicted to have no effect on $\hat{V}_G$, while reducing N leads to a drastic reduction in $\hat{V}_G$. However this value of s, a 5% fitness difference, is unrealistically large for overdominant alleles. This sensitivity of low dominance allelic combinations to drift is somewhat counterintuitive, as far more absolute equilibrium genetic variance is maintained by overdominance than in more additive cases.

DISCUSSION

In this paper, I have explored the consequences of a drift-mutation-selection balance using a variety of simple single-locus models. All of these models, including one of balancing selection, prove very susceptible to drift at effective population sizes (N) less than about 10^4, unless the average selection coefficient (s) is unrealistically large. While high s values would reduce sensitivity to drift, they also reduce the equilibrium genetic variance ($\hat{V}_G$) when Ns is not much less than one. This would make it very difficult for any of these models, except one involving overdominance, to explain observed levels of V_G.

I have ignored the effects of linkage in the models presented here. Previous results indicate that this is justified in some cases, such as effectively infinite populations (Lande, 1980*b*; Turelli, 1984). For effectively neutral characters, the expected genetic variance is not affected by linkage, although the variance in V_G among populations will be increased by it (Lynch and Hill, 1986). The simulation results of Keightley and Hill (1983, 1987, 1988) show that tight linkage will generally cause $\hat{V}_G$ to be overestimated by the models above in small populations when Ns is of order 1. This effect peaks at intermediate population sizes, rather than in small populations, because in small populations fewer loci will be segregating. This effectively increases the

amount of recombination between segregating sites. However, in Keightley and Hill's simulations, the discrepancy between models incorporating linkage and those ignoring it disappears if there is even one crossover per genome per generation. The negligible importance of linkage in infinite populations (Lande, 1980*b*; Turelli, 1984) suggests that whatever underestimate of $\hat{V}_G$ there is due to linkage when $Ns \approx 1$ must disappear as Ns increases. This would have the effect of decreasing the rate of increase in $\hat{V}_G$ with N. Thus, $\hat{V}_G$ would be susceptible to drift at higher values of N, while the susceptibility in smaller populations would be less severe.

Effective population size is extremely difficult to estimate for most populations. The only method for estimating it that seems at all general is sampling third-base-position heterozygosity (Kreitman, 1987). This is very expensive and at present has only been applied to *Drosophila*. Even with such data, the estimate one is really interested in is not for neutral alleles but for variation that experiences selection pressures similar to those acting on alleles that influence the studied character. Selection determines the time it takes to achieve genetic equilibrium and, therefore, the importance of migration. Unfortunately, without data on effective population size with selection, we are limited to qualitative predictions about $\hat{V}_G$. If effective population sizes are frequently less than 10^4, large and small populations should differ greatly in their genetic variance. Is there any evidence for this? The data on this point are quite equivocal. While small populations clearly seem to have less genetic variance than larger ones, nothing suggests the kind of dramatic differences predicted by the models in Figures 4 and 5.

There is a great deal of data on V_G for domesticated and artificial populations; however, such populations are unlikely to be near genetic equilibrium. The expectation that small populations will have little genetic variance often leads investigators to synthesize a base population by combining inbred lines or small populations, obscuring any potential effects of population size. The problems with using nonequilibrium populations are strikingly illustrated by Bryant et al.'s (1986) finding that genetic variance for morphological traits in houseflies went up when isolates from a large, natural population were taken through bottlenecks of four, eight, or 16 flies. Bryant et al. ascribed this result to epistasis in the base population. Such increases will not, however, persist at equilibrium (Goodnight, 1988). It is well known that severe inbreeding quickly reduces genetic variance within populations, but this will be true irrespective of the mechanism by which variance was maintained in the base population. Experiments have shown that long-term selection in small populations leads to somewhat less response than is obtained in large populations (e.g., Jones et al., 1968); however, since such experiments draw from a common base population, itself not at equilibrium, no conclusions about the quantitative agreement with models can be made. Laboratory populations could, however, be used to shed light on this issue. The time required to eliminate 95% of the difference between equilibrium heterozygosity and either lower or higher heterozygosity for neutral loci is approximately $6N$ generations in small populations ($1/N \gg \mu$) (Malécot, 1969). This time is shorter for variation acted on by selection. It is possible to maintain populations of organisms such as *Drosophila* this long under constant conditions.

Two kinds of data from natural populations have relevance to this issue. First, heritability data suggest little correspondence between N and V_G. For example, many mammals, such as mice, probably have effective population sizes on the order of a few hundred (Chepko-Sade et al., 1987), yet there is no obvious trend for heritabilities or genetic variances to be small in such organisms. High heritabilities for morphological characters are found in the mouse species *Peromyscus maniculatus* (Sumner, 1918), the red-backed salamander *Plethodon cinereus* (Highton, 1960), and Darwin's finches (Grant, 1986). Such results seem inconsistent with the general results of the models considered here. Such examples are weak evidence, however, as there is certainly a bias against the publication of insignificant results. In addition, the implicit comparisons that one makes based on such data are between different characters in very different organisms. Such comparisons would be

invalid if, for example, V_M/V_E were different for the characters compared. Lynch (1988) has recently shown that estimates of V_M/V_E vary over two orders of magnitude.

Data that allowed comparisons among populations of the same or closely related species that differ in N would be much more convincing. I do not know of any. A few studies have indicated a correlation between N and phenotypic variance. In the case of characters that are highly heritable, differences in phenotypic variance may reflect genetic variance. Soulé et al. (1973) showed a highly significant positive correlation between allozyme heterozygosity and the coefficient of variation of scale counts among eight *Anolis* species and a weakly significant correlation among island populations of another lizard, *Uta stansburiana.* Using the data from *Uta,* Soulé (1972) also showed that phenotypic variance is correlated positively with island area and negatively with distance from the mainland. Soulé considered and rejected the hypothesis that these correlations were due to effective population size and proposed instead that they reflected differences in environmental variance and niche width. On balance, a simple dependence on population size seems a more likely explanation for the data than this elaborate alternative. Qualitatively similar studies of bird populations on islands give conflicting results. Van Valen (1965) found more variation in island populations, while Grant (1967) found less. Fisher (Fisher and Ford, 1928; Fisher, 1937) studied phenotypic variance in moths and in egg dimensions of birds and found that species that were classified as "abundant" tended to have more phenotypic variance than those that were more rare. However, Fisher did not control for geographic variation and drew samples from all over the British Isles.

All of these studies are limited in their relevance to the issue, as only a qualitative prediction concerning relative variance can be made. All mechanisms that maintain genetic variation will show some sensitivity to population size. Only in exceptional circumstances will it be possible to accumulate sufficient information about population size to make a more quantitative prediction. Such a comparison might be possible in a species whose habitat requirements were patchy but temporally stable, or among populations isolated by a dispersal barrier. Comparisons could also be made between a common species and a rare one. Any such study would prove extremely valuable. At the very least, the magnitude of any dependence on $\hat{V}_G$ on N sets limits on the strength of mutation and selection, assuming that the mutation-selection-drift balance model is correct. A major problem with any such tests would be the variance around $\hat{V}_G$. In the neutral case, Lynch and Hill (1986) show that this variance among populations will be large. The simulation results of Bürger et al. (1989) also show a very high variance of V_G, and strong temporal autocorrelation when there is stabilizing selection. Thus, replicate study populations that have been independent for some time will be required in such studies.

If it is true that genetic variance is not as susceptible to drift as predicted here, then some of the assumptions of these models must be false. All of the models applied here are equilibrium ones, and it is still very unclear whether one should expect quantitative characters to achieve genetic equilibrium very often, particularly if the selection regime often changes on an ecological time scale. If populations are not at equilibrium, this will cause genetic variances from large and small populations to be more similar that they would be at equilibrium. Because equilibration time is of order N generations, currently large populations will be more likely to have less heterozygosity than the equilibrium level, making large and small populations more similar in heterozygosity and genetic variance.

A related problem is the sensitivity of equilibria to historical events. Barton (1986) showed that, in an infinite population, a two-allele, multiple-locus model predicts that there will be multiple genetic equilibria, differing very slightly in their means but differing substantially in their variances. Therefore the system will have an important historical component, depending on which equilibrium a population achieves. More recently, Barton has investigated the effect of finite population size, and found that this effect disappears for most parameter values when $N < 20{,}000$ (Barton, pers.

comm.). The models considered here suggest that $\hat{V}_G$ will covary with N for populations smaller than this.

The presence of epistatic interactions among loci could also lead to a strong historical component for the phenotypic means that populations achieve (Wright, 1931). If populations can achieve different selected optima, which in turn vary in the strength of selection around the optima (i.e., the flatness of adaptive peaks), then this could affect $\hat{V}_G$. This could result in a weaker correlation between N and $\hat{V}_G$, although the expected slope of the relationship should remain the same. Since we know virtually nothing about epistasis for fitness, we cannot say whether epistatic interactions would tend to increase or decrease the average amount of selection against alleles which are currently nonoptimal. Therefore, there is no reason to suspect that, on average, epistasis would alter the sensitivity of $\hat{V}_G$ to drift.

Another important assumption of all of these models is that there is a single panmictic population. If selection acts to differentiate the character of interest over space, migration does more than just alter N. Slatkin (1978), in an analysis based on the Gaussian allelic approximation to the infinite-alleles model, showed that the effect of migration in a cline on $\hat{V}_G$ within local populations will be small unless the cline is quite steep relative to dispersal distances. Use of a rare-alleles approximation in this case might show more susceptibility of $\hat{V}_G$ to migration, as heterozygosity is very low in this approximation. A potential case of this was provided by Grant (1986), who hypothesized that the high heritabilities in Darwin's finches are maintained in the face of strong selection by gene flow from other species and differentiated populations on other islands. If populations also achieve different stable equilibrium states with respect to particular loci (Barton, 1986), then migrants could have a disproportionate effect on $\hat{V}_G$ without any differentiation of phenotypic means.

Turelli (1985) has taken the ubiquity of pleiotropy as a particular challenge to experimental verification of the mechanisms involved in the maintenance of V_G. The existence of pleiotropy does not affect the basic conclusions of this paper. While the fitness of an allele may be a complex function of its effects on many phenotypes, its effect on only one phenotype, fitness, determines its evolutionary dynamics. Efforts to infer the strength of selection from changes in phenotypic variance may be inaccurate because of pleiotropy, but this leaves unaffected the basic dilemma that assuming weaker selection yields more genetic variance but a greater sensitivity to drift, while stronger selection yields the converse.

This poses a particularly difficult challenge to the rare-allele approximation to the infinite-allele model (Barton and Turelli, 1987) and related finite-allele models. In an infinite population, these models require hundreds of loci to explain heritabilities as high as 0.50 when typical parameter estimates are used (Turelli, 1984). If the amount of selection is assumed to be high enough to minimize the susceptibility of $\hat{V}_G$ to drift, as in the top set of graphs in Figures 3–5, then the number of loci necessary to account for high heritabilities becomes enormous, as is evident from Figure 3.

In addition, previous analyses have pointed out that the rare-allele assumptions predict that heritability should increase rapidly in the short term when selection is applied to a character (Barton and Turelli, 1987; Slatkin, 1987). This is because one allele is expected to be at high frequency at each locus at equilibrium, and therefore, selection will initially tend to increase heterozygosity and V_G. This is clearly not in accordance with the experimental evidence.

While this suggests that the rare-alleles approximation and finite-allele analogues have shortcomings, other models also fare poorly when compared to the real world. The Gaussian allelic approximation requires the assumption that mutation rates are much higher than empirical evidence suggests (Turelli, 1984). This difficulty is intensified in finite populations, where the assumption of a Gaussian distribution of allelic effects should require even higher mutation rates than proposed by Turelli. The pleiotropic-overdominance model is plausible, but there is no empirical evidence that alleles interact this way, other than a few peculiar cases such as hemoglobin *S*. This stands in sharp contrast to the volumes of evidence for additive and dominance in-

teractions. Furthermore, all of these models share the susceptibility to drift that seems incompatible with empirical evidence. This suggests that alternatives to these simple single-locus models must be developed. One possibility would be models involving frequency-dependent selection on either the genotypic or phenotypic level. This might reduce the impact of drift (Hedrick, 1972). Another would be a stepping-stone, or isolation-by-distance model, with occasional long-distance migration, as outlined above.

Turelli (1985) has predicted that it will be no easier to determine the mechanisms that maintain genetic variance than to discriminate, in general, between neutralist and selectionist explanations for allozyme variation. I agree that this task will surely prove difficult, but the addition of finite population size to the picture provides another axis on which to evaluate the evidence. This cannot fail to be helpful.

Acknowledgments

I am indebted to M. Turelli for suggesting the rare-alleles derivation for Equation (10). I thank W. F. Eanes for his patience, W. F. Eanes, D. J. Futuyma, M. Slatkin, D. Stratton, M. Turelli, G. P. Wagner, and two anonymous reviewers for their comments on earlier versions of this manuscript, and L. Ginzburg for helpful discussions. This work was supported by NSF grants BSR-8402967 and BSR-8500461 to W. F. Eanes. This is contribution number 698 from the graduate program in Ecology and Evolution, State University of New York at Stony Brook.

Literature Cited

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

Barton, N. H., and M. Turelli. 1987. Adaptive landscapes, genetic distance and the evolution of quantitative characters. Genet. Res. 49:157–173.

Bryant, E. H., S. A. McCommas, and L. M. Combs. 1986. The effect of an experimental bottleneck upon quantitative genetic variation in the housefly. Genetics 114:1191–1211.

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

———. 1973. The maintenance of the genetic variability of polygenic characters by heterozygous advantage. Genet. Res. 22:9–12.

———. 1980. The Mathematical Theory of Quantitative Genetics. Oxford Univ. Press, Oxford, U.K.

Bürger, R., G. P. Wagner, and F. Stettinger. 1989. How much heritable variation can be maintained in finite populations by a mutation-selection balance? Evolution 43:1748–1766.

Chakraborty, R., and M. Nei. 1982. Genetic differentiation of quantitative characters between populations or species. I. Mutation and random genetic drift. Genet. Res. 39:303–314.

Chepko-Sade, B. D., W. M. Shields, J. Berger, Z. T. Halpin, W. T. Jones, L. L. Rogers, J. P. Rood, and A. T. Smith. 1987. The effects of dispersal and social structure on effective population size, pp. 287–321. *In* B. D. Chepko-Sade and Z. T. Halpin (eds.), Mammalian Dispersal Patterns: The Effects of Social Structure on Population Genetics. Univ. Chicago Press, Chicago, IL.

Clayton, G. A., and A. Robertson. 1955. Mutation and quantitative genetic variation. Amer. Natur. 89:151–158.

Fisher, R. A. 1937. The relation between variability and abundance shown by the measurements of the eggs of British nesting birds. Proc. Roy. Soc. Lond. B 122:1–26.

Fisher, R. A., and E. B. Ford. 1928. The variability of species in the Lepidoptera, with reference to abundance and sex. Trans. Entomol. Soc. (London) 76:367–384.

Fleming, W. H. 1979. Equilibrium distributions of continuous polygenic traits. J. Appl. Mathemat. 36: 148–168.

Foley, P. 1987. Molecular clock rates at loci under stabilizing selection. Proc. Nat. Acad. Sci. USA 84: 7996–8000.

Gillespie, J. H. 1984. Pleiotropic overdominance and the maintenance of genetic variation in polygenic characters. Genetics 107:321–330.

Goodnight, C. J. 1988. Epistasis and the effect of founder events on the additive genetic variance. Evolution 42:441–454.

Grant, P. R. 1967. Bill length variability in birds of the Tres Marias Islands, Mexico. Can. J. Zool. 45: 805–815.

———. 1986. Ecology and Evolution of Darwin's Finches. Princeton Univ. Press, Princeton, NJ.

Hedrick, P. W. 1972. Maintenance of genetic variation with a frequency-dependent selection model as compared to the overdominant model. Genetics 72:771–775.

Highton, R. 1960. Heritability of geographic variation in trunk segmentation in the red-backed salamander, *Plethodon cinereus*. Evolution 14:351–360.

Hoi-Sen, Y. 1972. Is subline differentiation a continuing process in inbred strains of mice? Genet. Res. 19:53–59.

Jones, L. P., R. Frankham, and J. S. F. Barker. 1968. The effects of population size and selection intensity in selection for a quantitative character in *Drosophila*. II. Long term response to selection. Genet. Res. 12:249–266.

Keightley, P. D., and W. G. Hill. 1983. Effects of linkage on response to directional selection from new mutations. Genet. Res. 42:193–206.

———. 1987. Directional selection and variation in finite populations. Genetics 117:573–582.
———. 1988. Quantitative genetic variability maintained by mutation-stabilizing selection balance in finite populations. Genet. Res. 52:33–43.
KIMURA, M. 1965. A stochastic model concerning the maintenance of genetic variability in quantitative characters. Proc. Nat. Acad. Sci. USA 54: 731–736.
———. 1981. Possibility of extensive neutral evolution under stabilizing selection with special reference to nonrandom usage of synonymous codons. Proc. Nat. Acad. Sci. USA 78:5773–5777.
KIMURA, M., T. MARUYAMA, AND J. F. CROW. 1963. The mutation load in small populations. Genetics 48:1303–1312.
KREITMAN, M. 1987. Molecular population genetics. Oxford Surv. Evol. Biol. 4:38–60.
LANDE, R. 1976. The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genet. Res. 26:221–235.
———. 1979. Quantitative genetic analysis of multivariate evolution applied to brain : body size allometry. Evolution 33:402–416.
———. 1980*a*. Genetic variation and phenotypic evolution during allopatric speciation. Amer. Natur. 116:463–479.
———. 1980*b*. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215.
LATTER, B. D. H. 1960. Natural selection for an intermediate optimum. Aust. J. Biol. Sci. 13:30–35.
———. 1970. Selection in finite populations with multiple alleles. II. Centripetal selection, mutation, and isoallelic variation. Genetics 66:165–186.
LEWONTIN, R. C. 1964. The interaction of selection and linkage. II. Optimum models. Genetics 50:757–782.
LYNCH, M. 1988. The rate of polygenic mutation. Genet. Res. 51:137–148.
LYNCH, M., AND W. G. HILL. 1986. Phenotypic evolution by neutral mutation. Evolution 40:915–935.
MALÉCOT, G. 1969. The Mathematics of Heredity. Freeman, San Francisco, CA.
MUKAI, T., S. I. CHIGUSA, L. E. METTLER, AND J. F. CROW. 1972. Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. Genetics 72:335–355.
NEI, M., AND Y. IMAIZUMI. 1966. Effects of restricted population size and increase in mutation rate on the genetic variation of quantitative characters. Genetics 54:763–782.
ROBERTSON, A. 1956. The effect of selection against extreme deviants based on deviation or on homozygosis. J. Genet. 54:236–248.
———. 1962. Selection for heterozygotes in small populations. Genetics 47:1291–1300.
RUSSELL, W. A., G. F. SPRAGUE, AND L. H. PENNY. 1963. Mutations affecting quantitative characters in long-time inbred lines of maize. Crop Sci. 3:175–178.
SLATKIN, M. 1978. Spatial patterns in the distribution of polygenic characters. J. Theoret. Biol. 70:213–228.
———. 1987. Heritable variation and heterozygosity under a balance between mutation and stabilizing selection. Genet. Res. 50:53–62.
SOULÉ, M. 1972. Phenetics of natural populations. III. Variation in insular populations of a lizard. Amer. Natur. 106:429–446.
SOULÉ, M. E., S. H. YANG, M. G. W. WEILER, AND G. C. GORMAN. 1973. Island lizards: The genetic-phenetic variation correlation. Nature 242:191–193.
SUMNER, F. B. 1918. Continuous and discontinuous variation and their inheritance in *Peromyscus*. Amer. Natur. 52:177–208, 290–301, 439–454.
TURELLI, M. 1984. Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. Theoret. Popul. Biol. 25:138–193.
———. 1985. Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits. Genetics 111:165–195.
VAN VALEN, L. 1965. Morphological variation and the width of the ecological niche. Amer. Natur. 99: 377–389.
WRIGHT, S. 1931. Evolution in Mendelian populations. Genetics 16:97–159.
———. 1937. The distribution of gene frequencies in populations. Proc. Nat. Acad. Sci. USA 23:307–320.
———. 1969. Evolution and the Genetics of Populations, Vol. 2. The Theory of Gene Frequencies. Univ. Chicago Press, Chicago, IL.

Corresponding Editor: P. W. Hedrick