

Quantitative genetic variability maintained by mutation-stabilizing selection balance: sampling variation and response to subsequent directional selection

PETER D. KEIGHTLEY† AND WILLIAM G. HILL

Department of Genetics, University of Edinburgh, West Mains Road, Edinburgh EH9 3JN

(Received 10 October 1988 and in revised form 22 December 1988)

Summary

A model of genetic variation of a quantitative character subject to the simultaneous effects of mutation, selection and drift is investigated. Predictions are obtained for the variance of the genetic variance among independent lines at equilibrium with stabilizing selection. These indicate that the coefficient of variation of the genetic variance among lines is relatively insensitive to the strength of stabilizing selection on the character. The effects on the genetic variance of a change of mode of selection from stabilizing to directional selection are investigated. This is intended to model directional selection of a character in a sample of individuals from a natural or long-established cage population. The pattern of change of variance from directional selection is strongly influenced by the strengths of selection at individual loci in relation to effective population size before and after the change of regime. Patterns of change of variance and selection responses from Monte Carlo simulation are compared to selection responses observed in experiments. These indicate that changes in variance with directional selection are not very different from those due to drift alone in the experiments, and do not necessarily give information on the presence of stabilizing selection or its strength.

1. Introduction

Many selection experiments use as their source material samples from caged insect populations (often *Drosophila*) which have been in the cage environment for many generations. Samples are taken from the cage, separate selection lines started, and response patterns of individual lines and the variation in response among lines are obtained. Similarly, selection of artificial populations is usually with species which have been under domestication for many generations and there are often independent selection lines. Such experiments should provide information on the underlying genetic basis of quantitative variation.

Many characters in natural (and perhaps in artificial) populations are thought to have intermediate optima. A popular model of selection with an intermediate optimum is stabilizing selection. This has intuitive appeal, there is some evidence for its operation in nature and the model is amenable to analysis. Previous analyses of the consequences of stabilizing selection in natural populations have been concerned with the genetic variance maintained when

the character is at or near the optimum, in which case it is necessary to invoke mutation to maintain genetic variation. Kimura (1965) analysed a 'continuum of alleles' model, which involves loci at which the effect of mutant alleles differs only slightly from the previous allelic state. Kimura derived a formula for the equilibrium genetic variance of a locus at a mutation-stabilizing selection equilibrium, and showed that the equilibrium distribution of allelic effects segregating at the locus is normal. The model was further analysed by Lande (1976) who argued that it predicts that substantial variation can be maintained even with strong stabilizing selection.

The second type of analysis differs from the first because the effects of mutant alleles can be large. Turelli (1984) analysed a 'House of Cards' approximation of the continuum of alleles model, which was originally proposed by Kingman (1978). The critical assumption which differed from that in Kimura's (1965) analysis above was that the effect of a mutant allele swamps existing variation at a locus controlling genetic variation in the trait. This gives different qualitative predictions of the variance maintained at the locus at equilibrium and agrees with two allele analyses of Latter (1960) and Bulmer (1972). Turelli

† Corresponding author.

argued that a 'House of Cards' approximation of the mutation process is appropriate for per-locus mutation rates likely to be found in nature.

With the exception of Bulmer (1972) the analyses described above have been of infinite populations. This has an important consequence for the probability distribution at equilibrium of allele frequencies at a locus influencing the trait. With stabilizing selection in populations near equilibrium, mutations are unconditionally deleterious (Robertson, 1956). In an infinite population, if there are few alleles segregating at the locus the equilibrium probability distribution of allele frequencies is therefore highly leptokurtic, i.e. mutant alleles are almost always very rare, and intermediate allele frequencies are absent. Such a distribution is very 'U-shaped'. This affects the consequences of a shift in the optimum. Barton & Turelli (1987) analysed the dynamics of the population mean and variance after a change of the optimum and showed an accelerating rise in the mean, slowing down as it approached the new optimum, and a rise in the genetic variance because some previously deleterious mutant alleles became advantageous and were selected to intermediate frequencies where they contributed more substantially to the genetic variance. In some cases, the variance fell again close to its original value (i.e. alleles became fixed) and in others a new equilibrium was reached with a higher variance. The existence of such multiple equilibria was predicted by Barton (1986). A change of optimum is similar to imposing directional selection on a character previously subject to stabilizing selection.

In a finite population, alleles are able to drift in frequency, so the equilibrium probability distribution of allele frequencies becomes less leptokurtic than described above. Bulmer (1972) derived an expression for the probability distribution of allele frequency at a mutation-stabilizing selection-drift balance for the case of up to two alleles per locus. This was used by Keightley & Hill (1988) and Burger, Wagner & Stettinger (1988) to investigate finite population models of genetic variance maintained at a mutation-stabilizing selection balance for various distributions of mutant gene effects.

Here, expressions are derived for the variance of the genetic variance among independent lines at a mutation-selection-drift balance. The results of previous investigations of the problem using neutral models (Bulmer, 1976; Avery & Hill, 1977; Lynch & Hill, 1986) are compared. The consequences of a change from stabilizing selection to directional selection on the genetic variance of a character in a finite population are investigated. These results together with patterns of response to directional selection from Monte Carlo simulation of populations previously under stabilizing selection are compared to response patterns obtained from experiments published in the literature.

2. Model

The assumptions of the model are essentially the same as described by Keightley & Hill (1988). The population consists of N random mating diploid individuals with constant population size and non-overlapping generations. The phenotypic value of an individual is X , the sum of contributions from the alleles affecting the trait plus an independent normally distributed environmental effect of variance V_E . New mutations appear independently at random and effects are added to the value at the locus at which they occur. A mutant effect, a , is the difference in value between the homozygotes and q is the frequency of the higher valued allele. Mutant effects are sampled from a gamma distribution,

$$f(a) = \alpha^\beta e^{-\alpha a} a^{\beta-1} / \Gamma(\beta) \quad (0 < a < \infty).$$

The distribution was reflected about zero by randomly assigning the sign of each mutant effect with probability $\frac{1}{2}$ of being positive. The resulting distribution is termed a 'reflected gamma distribution'. The parameter β determines the shape of the distribution ($\beta \rightarrow 0$ gives a highly leptokurtic (geometric) distribution; $\beta \rightarrow \infty$ is the limiting case of equal absolute values of effects), and the parameter α is the scale of the distribution. The root mean square of the distribution of mutant effects is defined as

$$\epsilon = E(a^2/V_E)^{\frac{1}{2}} = [\beta(\beta+1)/(\alpha^2 V_E)]^{\frac{1}{2}}.$$

The parameter $N_e \mu$, where μ is the mutation rate at the locus, is assumed to be sufficiently small, or the parameter $|N_e s|$, where s is the selection coefficient of the mutant allele, is assumed to be sufficiently large that no more than two alleles segregate at any time at each locus. The genomic mutation rate is $\lambda = n\mu$, where n is the total number of loci affecting the trait. The expected genetic variation arising per generation in the population is $V_M = \lambda E(a^2)/2$ (Hill, 1982a).

The *genic* variance of the character is the sum of variance contributions from each locus, $\Sigma a^2 q(1-q)/2$. The *additive* variance is the variance of genotypic values and includes terms due to departures from Hardy-Weinberg and linkage equilibrium.

Where stabilizing selection is operating, it is assumed that it has acted for sufficient time that the population is close to equilibrium. The fitness under stabilizing selection is defined by a 'non-optimal' model with optimum fixed at zero. The relative fitness is

$$W(X) = \exp(-X^2/2w^2), \quad (1)$$

where w is an inverse measure of the strength of selection. The change in gene frequency is approximated by

$$\Delta q = a^2(q - \frac{1}{2})q(1-q)/[4(w^2 + \sigma^2)] \quad (2)$$

(Robertson, 1956), where σ is the phenotypic standard deviation. New mutants are unconditionally deleterious and there is a meta-stable equilibrium at $q = \frac{1}{2}$. Monte Carlo simulation has shown that this approximation is accurate over a wide range of parameters (Keightley & Hill, 1988). The term $w^2 + \sigma^2$ is often referred to as V_s , the strength of stabilizing selection.

The model is used to investigate the effect on the genetic variance of a change from stabilizing selection to directional selection. There are therefore two phases of selection, a stabilizing selection phase followed by a directional selection phase. From diffusion theory, in the first phase the steady state probability distribution of allele frequency with recurrent mutation is a function of $N_1 s_1$, where N_1 is the effective population number in the first phase and s_1 is the selective value of the mutant allele. With stabilizing selection s_1 is frequency dependent (Robertson, 1956) and is approximated by

$$s_1 = (q - \frac{1}{2}) a^2 / [4(w^2 + \sigma^2)] \quad \text{or} \\ s_1^* = -a^2 / [8(w^2 + \sigma^2)],$$

when the mutant allele is rare. The steady state distribution of allele frequency with recurrent mutation and stabilizing selection is therefore a function of the parameter

$$N_1 E(s_1^*) = -N_1 \epsilon^2 V_E / [8(w^2 + \sigma^2)].$$

In the second phase of directional selection, the distribution of gene frequency at generation t is a function of t/N_2 and $N_2 s_2$, where N_2 is the effective population size in this phase and s_2 is the selection coefficient of the favourable allele. For example, the pattern of change of allele frequency over 10 generations in a population of 100 individuals with a selection coefficient of 0.1 is the same as that over 20 generations in a population of 200 individuals with a selection coefficient of 0.05. With truncation selection the selection coefficient is approximately $s_2 = ia/\sigma$, where i is the intensity of selection. Since the distribution of a is symmetrical in the stabilizing selection phase, the distribution of initial gene frequencies in the directional selection phase is also symmetrical, so the directional selection phase is parameterized by $N_2 E(s_2) = N_2 iE(|a|)/\sigma$.

3. Methods

(i) Transition matrix iteration

Using this method, the expectation and variance of heterozygosity maintained at a locus with recurrent mutation and the expected steady state allele frequency distribution with recurrent mutation were computed. Details of the method are given by Keightley & Hill (1988). To model a change from stabilizing to directional selection, the expected steady state allele frequency distribution with recurrent mutation in the

stabilizing selection phase was computed using the transition matrix with the expected change of gene frequency from (2). The expected variance was then computed for each generation of directional selection by iterating a transition matrix with change of gene frequency given by

$$\Delta q = s_2 q(1 - q) / [2(1 + s_2 q)].$$

The transition matrices were of dimension $N = 80$. To generate expected responses for a gamma distribution of mutant effects, numerical integration was used (see Keightley & Hill, 1988 for details).

(ii) Monte Carlo simulation

Using this method, the effects of simultaneously segregating mutants are assessed. The simulation procedure has been described in detail elsewhere (Keightley & Hill, 1983). Here, only the case of free recombination was considered. Essentially, there was an infinite number of freely recombining sites and new mutants, with effects sampled from a reflected gamma distribution, arose at unique sites. Selection was performed by assigning fitnesses or probabilities of producing progeny (i.e. fertilities) to each individual. With stabilizing selection, the phenotypic value, X , of an individual was the sum of genotypic contributions plus an environmental effect of variance $V_E = 1$, and the relative fitness given by (1). Where the equilibrium behaviour was needed, the population was allowed to approach equilibrium by initiating the population from an isogenic state and allowing mutations to accumulate for $6N$ generations. The variances of genic and additive variances at equilibrium were computed from independent runs. After a change of selection mode to directional selection, the fitness was simply the genotypic value, because in this case it was assumed that $i/\sigma = 1$, so the selection coefficient of an allele was the same as its effect.

4. Results

The first part of the results considers the variation among independent lines of the genetic variance maintained with mutation and stabilizing selection. Subsequently, the consequences of a change from stabilizing selection to directional selection are considered.

(i) Single locus analysis

Using diffusion theory (Kimura, 1969), Bulmer (1972) derived the density function of gene frequency, q_i , at a locus (i) under stabilizing selection with equal forward and backward mutation rates, μ , between two possible alleles

$$f(q_i) = C \exp [(-4\Phi q_i(1 - q_i))[q_i(1 - q_i)]^{\theta-1}. \quad (3)$$

where $\Theta = 4N_e\mu$, $\Phi = N_e a^2/[8(w^2 + \sigma^2)]$, N_e is the effective population size, and C is a normalizing constant. It follows that the expected heterozygosity at the locus is

$$E[q_i(1 - q_i)] = \left\{ \int_0^1 \exp[-4\Phi q_i(1 - q_i)][q_i(1 - q_i)]^\Theta dq \right\} / \left\{ \int_0^1 \exp[-4\Phi q_i(1 - q_i)][q_i(1 - q_i)]^{\Theta-1} dq \right\}. \quad (4)$$

This was shown to reduce to

$$E[q_i(1 - q_i)] = I(4\Phi, \Theta + 1)/I(4\Phi, \Theta), \quad (5)$$

where $I(x, y)$ is a function of the complete beta function and the confluent hypergeometric function (Bulmer, 1972). Assuming Θ is small (i.e. ignoring back-mutation) (5) simplifies to

$$E[q_i(1 - q_i)] = \left[\sum_{i=0}^{\infty} \Phi^i / ((2i + 1) i!) \right] / e^\Phi,$$

(Keightley & Hill, 1988). Integrating over the distribution of gene effects, $f(a)$, gives the expected genic variance

$$E(V_g) = N_e \lambda \int_{-\infty}^{+\infty} \left\{ \sum_{i=0}^{\infty} \Phi^i / ((2i + 1) i!) \right\} / e^\Phi a^2 f(a) da. \quad (6)$$

Bulmer's (1972) analysis can be extended to derive a formula, with similar assumptions, for the variance of the genic variance at a locus with recurrent mutation among independent lines. It follows from (3), (4) and (5) that

$$E[q_i^2(1 - q_i)^2] = I(4\Phi, \Theta + 2)/I(4\Phi, \Theta). \quad (7)$$

Assuming $\Theta \rightarrow 0$, by similar analysis the variance of the genic variance among lines is

$$V(V_g) = (N_e \lambda / 12) \int_{-\infty}^{+\infty} \left\{ \sum_{i=0}^{\infty} 3\Phi^i / (i!(4(i + 1)^2 - 1)) \right\} / e^\Phi a^4 f(a) da. \quad (8)$$

Equations (6) and (8) can be evaluated easily by iteration on a computer and converge readily. They were checked against results obtained from a transition matrix and were found to agree almost exactly. Equation (8) also agrees with results from the Monte Carlo simulation (Table 1). Two limiting cases are of particular interest.

Neutrality, $\Phi \rightarrow 0$. From (6) the expected genic variance is

$$E(V_g) = N_e \lambda E(a^2) = 2N_e V_M.$$

The variance of the genic variance from (8) is

$$V(V_g) = N_e \lambda E(a^4) / 12, \quad (9)$$

in agreement with Lynch & Hill (1986) who used a different derivation. The coefficient of variation of V_g is therefore

$$CV(V_g) = [(E(a^4)/E^2(a^2)) / (12N_e \lambda)]^{1/2}. \quad (10)$$

Table 1. Comparison of predictions of variance of genic variance derived from (8) (diffusion theory of independent genes) and Monte Carlo simulation

N	Theory $V(V_g) \times 10^4$	Simulation $V(V_g) \pm 1 \text{ s.e.} \times 10^4$
10	1.94	1.95 \pm 0.08
15	2.92	2.85 \pm 0.17
20	3.89	3.68 \pm 0.13
30	5.83	5.65 \pm 0.18

The parameters of the simulation were $\lambda = 0.2$, $\epsilon = 0.1$, and a reflected gamma distribution of mutant effects with shape parameter $\beta = \frac{1}{2}$. There was no selection.

Strong selection, $\Phi \rightarrow \infty$. By simplifying (6), Keightley & Hill (1988) show that the expected genic variance is $E(V_g) = 4\lambda(w^2 + \sigma^2)$ which equals that obtained by Latter (1960) and Turelli (1984) which assumed infinite population size. The variance of the genic variance among lines by similar analysis from (8) is

$$V(V_g) = 4\lambda(w^2 + \sigma^2)^2 / N_e. \quad (11)$$

Both the expectation and the variance of the genic variance with strong stabilizing selection (large $N_e s$) are therefore independent of the magnitude of the effects of mutant alleles. The coefficient of variation of V_g is

$$CV(V_g) = 1 / (4N_e \lambda)^{1/2}. \quad (12)$$

Comparison of (10) and (12) shows that for new mutant alleles of equal effect, the coefficient of variation varies by only a factor of $1/\sqrt{3}$ between cases of weak and strong selection. The shape of the

distribution of effects of mutant alleles becomes important as selection becomes weak. Fig. 1 shows the coefficient of variation of the genic variance among lines expressed as a fraction of that expected for strong stabilizing selection ($\Phi \rightarrow \infty$) as a function of $\Phi = N_e \epsilon^2 V_M / [8(w^2 + \sigma^2)]$. The curves are for a range of values of β , the shape parameter of the gamma distribution. All curves converge with increasing Φ as the CV becomes independent of the shape of the distribution but the shape parameter has increasing influence as $\Phi \rightarrow 0$.

(ii) *Disequilibrium*

The above analysis applies to the genic variance and its variance among independent lines. Such quantities cannot easily be measured. The additive variance and its variance among lines, which can be estimated, is influenced by departures from Hardy-Weinberg and

linkage equilibrium at different loci affecting the trait. Bulmer (1976) and Avery & Hill (1977) showed that variation in disequilibrium can be an important contributor to the variation in the additive variation among lines. Here, Monte Carlo simulation is used to examine previous results on a neutral model and examine the effects of selection.

Neutrality. Using results obtained by Avery & Hill (1979), Lynch & Hill (1986) derived an expression for the coefficient of variation of the additive variance

among independent lines at an equilibrium between drift and mutation in the absence of linkage,

$$CV(V_A) = [(E(a^4)/E^2(a^2))/(12N_e\lambda) + 2/(3N_e)]^{1/2} \quad (13)$$

(note, the additional n^{-1} term given by Lynch & Hill (1986) is inappropriate). The first term is the variance of the genic variance [cf. (10)] and the second is the variance of disequilibrium (note, although there is no selection and no net disequilibrium, the disequilibrium in each line varies stochastically about zero). The variance of the additive variance is therefore

$$V(V_A) = N_e\lambda E(a^4)/12 + 8N_e V_M^2/3. \quad (14)$$

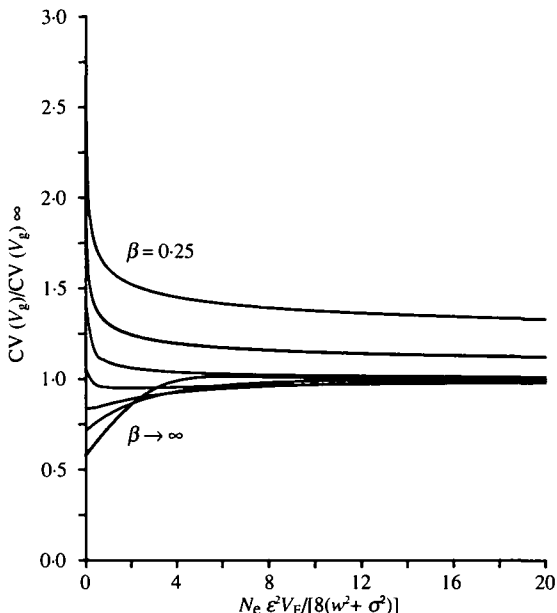


Fig. 1. The coefficient of variation of the genic variance among independent lines at equilibrium as a proportion of that predicted for very strong stabilizing selection, namely $[1/(4N_e\lambda)]^{1/2}$, as a function of $N_e E(s^*) = \Phi = N_e \epsilon^2 V_g / [8(w^2 + \sigma^2)]$. The curves were generated by numerical integration of (6) and (8) using gamma distributions of mutant effects. The shape parameter ranges from $\beta \rightarrow \infty$ to $\beta = \frac{1}{4}$ with intermediate values of $\beta = 8, 4, 2, 1, \frac{1}{2}$.

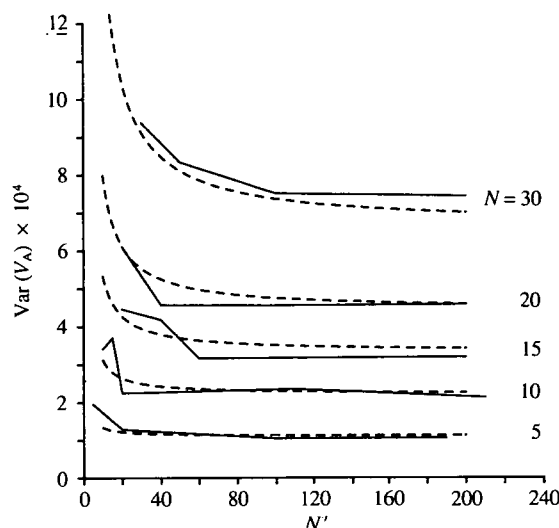


Fig. 2. The variance of the additive variance among independent lines for various parental population sizes (N) with variation in N' , the number of individuals used to estimate V_A . The curves were generated by (a) Monte Carlo simulation (—) in the absence of selection and mutation parameters such that many alleles segregate ($\lambda = 0.2, \epsilon = 0.1$ sampled from a reflected gamma distribution with $\beta = \frac{1}{2}$); (b) equation (15) (----) with $E(a^4) = (\beta + 2)(\beta + 3)\epsilon^4 V_g^2 / [\beta(\beta + 1)]$ and $\beta = \frac{1}{2}$.

Table 2. Comparison of predictions of variances of genic and additive variances among lines from Monte Carlo simulation and theory

N_1	N_c	Theory $V(V_g) \times 10^4$	Simulation $V(V_g) \pm 1 \text{ S.E.} \times 10^4$	Theory $V(V_A) \times 10^4$	Simulation $V(V_A) \pm 1 \text{ S.E.} \times 10^4$
5	4.1	0.75	0.78 ± 0.04	0.86	0.86 ± 0.02
10	8.5	1.45	1.35 ± 0.06	1.66	1.53 ± 0.03
20	17.0	2.58	2.81 ± 0.13	2.97	3.03 ± 0.06
30	26.0	3.54	3.62 ± 0.20	4.09	4.07 ± 0.07

The mutation parameters of the Monte Carlo simulation were the same as Table 1 ($\lambda = 0.2, \epsilon = 0.1$, and a gamma distribution of mutant effects with shape parameter $\beta = \frac{1}{2}$ reflected about zero). The character was under stabilizing selection with $w^2 + \sigma^2 = 2$. The effective population size was measured in the simulation by following the fates of independent neutral alleles. The effective population size is less than the actual population size because of selection. The value of N_c computed by the computer program was used to compute the theoretical values in the Table. The theoretical value of $V(V_g)$ is from (8). The theoretical value of $V(V_A)$ is $V(V_g) + 2E^2(V_g)/(3N_e)$, i.e. includes the disequilibrium component from the neutral model [cf. (14)] and $E(V_g)$ was computed from (6).

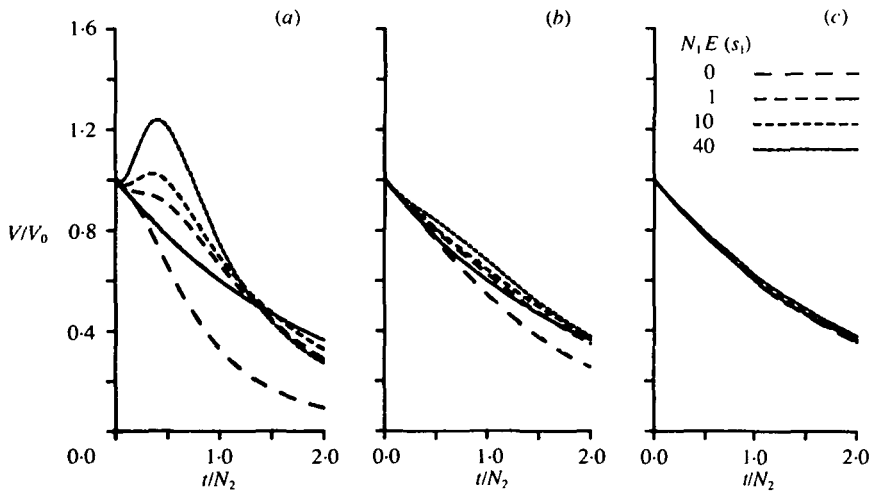


Fig. 3. Strength of directional selection, $N_2 E(s_2) = 1$.

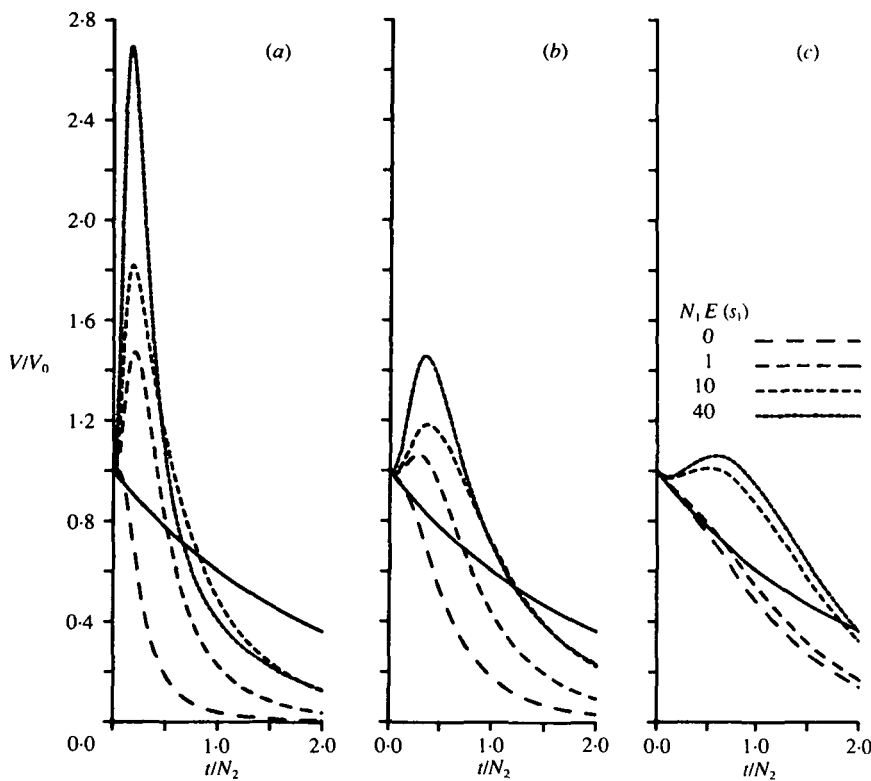


Fig. 4. $N_2 E(s_2) = 5$

Figs. 3–5. The expected genic variance as a proportion of that at $t = 0$ plotted against t/N_2 for four strengths of stabilizing selection, $N_1 E(s_1^*)$. The solid line on each figure is the expected variance for any value of $N_1 E(s_1^*)$ for no directional selection [$N_2 E(s_2) = 0$]. Curves for three reflected gamma distributions of mutant effects are shown (a) $\beta = \frac{1}{3}$; (b) $\beta = 1$; (c) $\beta \rightarrow \infty$.

Bulmer (1980, ch. 12) points out that there is an ambiguity in the interpretation of (13) and (14) because the variance of the additive variance depends on the number of individuals measured to estimate the additive variance within each line, and this may be different from the effective population size, N_e . As a

starting point for resolving this difficulty, let N_e be the effective number of parents in each line (as before) and N' be the number of individuals per line used to estimate the additive variance. Assuming a normal distribution of observations, the variance in the estimate of the variance among lines due to sampling

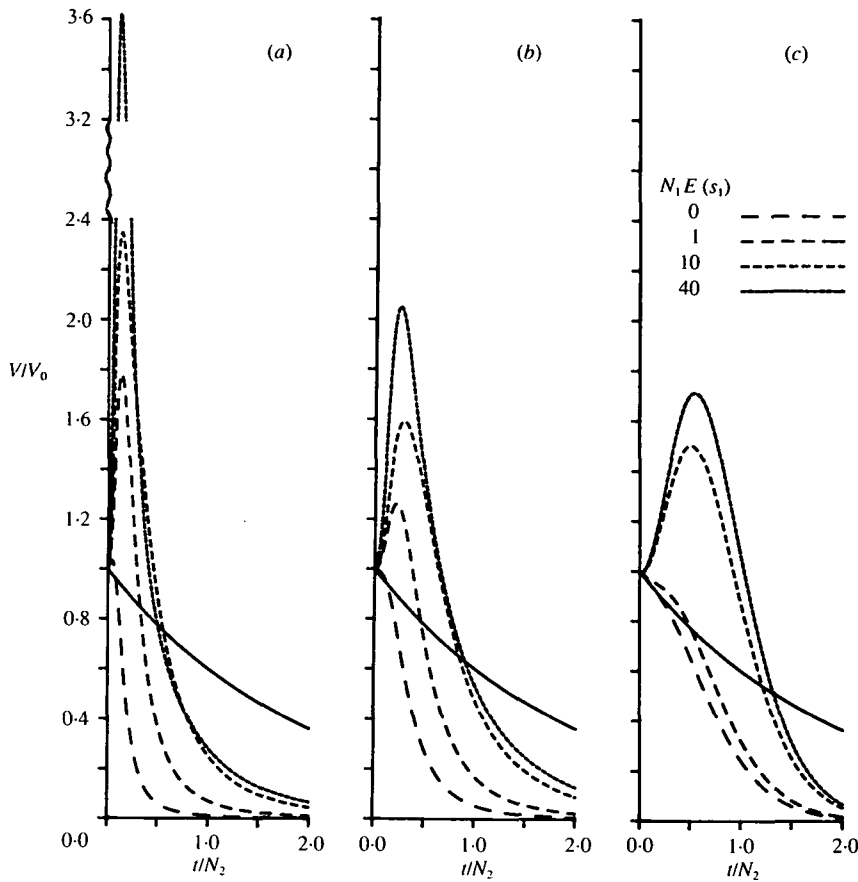


Fig. 5. $N_2 E(s_2) = 10$.

is approximately $2V_A^2/(N' - 1) \approx 8N_c^2 V_M^2/N'$. This additional source of variation affects $V(V_A)$ so (14) can be rewritten taking this source into account,

$$V(V_A) = N_c \lambda E(a^4)/12 + (2N_c V_M)^2 [2/(3N_c) + 2/N']. \tag{15}$$

This reduces to (14) as N' becomes large, in which case the additional variance from estimating V_A becomes small. The results of Monte Carlo simulation of a neutral model with many alleles segregating are compared to evaluation of (15) in Fig. 2. Lines of various sizes were allowed to reach steady state and the additive variance and its variance among lines were computed using various numbers of progeny. The agreement between the models is very close.

There would be an additional source of variation in estimating $V(V_A)$ caused by error in estimating V_A within lines, using, for example, offspring-parent regression, correlation of sibs or other appropriate method. For example, if V_A were estimated within lines from the covariance of half-sibs, it can be shown (cf. Robertson, 1959) that the variance of the estimate of V_A is approximately $2V_A^2[1 + 4/(\nu h^2)]^2/N'$, where ν is the number of progeny per half-sib family (assumed constant), h^2 is the heritability and N' in this case means the number of sires. Adding this additional

term to (14), $V(V_A)$ (where the estimation of V_A is done by half-sib covariance) becomes,

$$V(V_A) = N_c \lambda E(a^4)/12 + (2N_c V_M)^2 [2/(3N_c) + (1 + 8/(\nu h^2))^2/N']. \tag{16}$$

For small numbers of progeny and sires and traits of low heritability at equilibrium, the last term in (16) can dominate. Equation (16) gives the variance of the additive variance in the absence of linkage. With linkage, the term $2/(3N_c)$ is inflated, although not by much for species with many chromosomes (Lynch & Hill, 1986).

Stabilizing selection. No formulae are available to predict the variance among sublines of the disequilibrium component at steady state with stabilizing selection, although formulae for the expected disequilibrium with selection and an 'infinitesimal model' have been derived elsewhere (Keightley & Hill, 1987). However, with free recombination, disequilibrium in a population has a very short 'memory', with on average half the previous disequilibrium lost due to recombination each generation. It is likely therefore that the additional term in (14) for the neutral case is a good predictor of the variation in disequilibrium for the case of stabilizing selection. Simulation runs (Table 2) show good agreement with eqn (8) for $V(V_R)$ and (14) for $V(V_A)$.

Table 3. The mean, coefficient of variation (CV), and range of cumulative response to selection to generations $t = 10$ and $t = 20$ from 10 independent populations initially at a mutation-stabilizing selection-drift equilibrium

$N_1 E(s_1^*)$	ϵ	β	t	Mean	CV	Range
0	0.2	∞	10	2.21	0.19	1.27–2.90
			20	3.62	0.17	2.74–4.91
0	0.2	$\frac{1}{2}$	10	2.15	0.36	1.06–3.63
			20	3.35	0.34	1.77–5.12
0	0.8	∞	10	1.69	0.44	0.52–2.82
			20	2.13	0.41	0.88–3.34
0	0.8	$\frac{1}{2}$	10	1.78	0.51	0.39–3.38
			20	1.49	0.66	0.39–3.45
$\frac{1}{2}$	0.2	∞	10	1.36	0.32	0.64–2.29
			20	2.46	0.22	1.85–3.82
$\frac{1}{2}$	0.2	$\frac{1}{2}$	10	0.66	0.36	0.42–1.18
			20	1.16	0.41	0.58–1.95
8	0.8	∞	10	0.42	1.14	–0.06–1.50
			20	0.64	1.06	0.02–2.26
8	0.8	$\frac{1}{2}$	10	0.25	1.64	–0.02–1.31
			20	0.36	1.67	–0.02–1.94

Mutations occurred in the stabilizing selection phase only and were sampled from a reflected gamma distribution with shape parameter β and scale parameter ϵ given in the Table. The mutation rate λ was such that $V_M/V_E = 10^{-3}$. The population size in the stabilizing selection phase was $N_1 = 160$ and in the directional selection phase was $N_2 = 20$. Response in units of i/σ .

(iii) Effect of change of selection mode on variance

The results of the previous section show that the tendency for stabilizing selection to generate an extremely U-shaped distribution of allele frequencies influences the variance of the genetic variance between sublines. This effect also influences the pattern of response and change of variance of a character under stabilizing selection subsequently subjected to directional selection.

For a gamma distribution of mutant effects, the effect of selection on the genetic variance of the character is a function of three parameters: (1) $N_1 E(s_1^*)$, the expected selective value in the stabilizing selection phase; (2) $N_2 E(s_2)$, the expected selective value in the directional selection phase; (3) β , the shape parameter of the gamma distribution. Figs. 3–5 show expected genic variances (ignoring disequilibrium) in the generations after a change in mode of selection for a range of $N_1 E(s_1^*)$ and $N_2 E(s_2)$. Mutants occurred only in the stabilizing selection phase and effects were sampled from reflected gamma distributions. Curves for three different values of the shape parameter, β , are shown: (1) $\beta = \frac{1}{4}$, a highly leptokurtic distribution (see Keightley & Hill, 1988); (2) $\beta = 1$, an exponential distribution; (3) $\beta \rightarrow \infty$, all mutant effects have equal absolute values. In all cases, the probability of a mutant of positive or negative effect was assumed to be the same. An implicit assumption of the analysis is that there is no stabilizing selection operating in the directional selection phase or, equivalently, that directional selection is strong relative to stabilizing

selection. Figs. 3–5 show a wide range of values of the parameters. The selective values in the stabilizing selection phase range from $N_1 E(s_1^*) = 0$ (neutrality) to $N_1 E(s_1^*) = 40$. The latter case would pertain, for example, if $[E(a^2)]^{\frac{1}{2}} = 0.1$, $w^2 + \sigma^2 = 20$ and $N_1 = 6.4 \times 10^5$. The range of selective values during the directional selection phase is from $N_2 E(s_2) = 0$ (neutrality) to $N_2 E(s_2) = 10$ [e.g. $E(|a|)/\sigma = 0.1$, $i = 1$, and $N_2 = 100$].

With values of $N_2 E(s_2)$ at the high end of the range shown, the pattern of change of variance departs substantially from that observed with neutrality. In some cases, there is a marked rise in variance followed by a rapid fall. The rise in variance occurs during the fixation of beneficial alleles segregating initially at low frequency. Such a pattern is therefore observed when the following conditions pertain: (i) strong directional selection, so such alleles have a high probability of fixation; (ii) strong stabilizing selection because the probability distribution of allele frequencies becomes increasingly U-shaped with increasingly strong stabilizing selection, so the expected initial frequency of beneficial mutants of large effect is low. The pattern of change of variance becomes most extreme with a leptokurtic distribution of mutant effects (e.g. $\beta = \frac{1}{4}$) because as the mutational distribution becomes more leptokurtic, a higher proportion of the mutational variance is contributed by mutations of large effect.

The pattern of rapid rise followed by rapid fall in variance depends on the presence of beneficial mutants in the directional selection phase, i.e. mutants of positive effect. In Figs. 3–5 there are equal probabilities

of mutants of positive and negative effect, but the pattern of rapid rise followed by a rapid fall in variance becomes more extreme with a higher proportion of mutants of positive effect (results not shown). In other cases, the variance falls off more quickly than for neutral genes. This occurs with weak stabilizing selection, but strong directional selection, in which case alleles initially segregating at intermediate frequencies become fixed at a high rate and the genetic variance falls rapidly.

The curves show the expected genic variance and ignore the consequences of disequilibrium. Selection generates a negative disequilibrium component of variance which increases with increasingly tight linkage (see Bulmer, 1980, ch. 9; Keightley & Hill, 1987). In such circumstances, the additive variance is less than the genic variance and the pattern of increase in additive variance would be less extreme than shown.

(iv) Variation in response

Using similar methods to the above, Hill & Rasbash (1986) analysed the variation in response to directional selection. Higher variation in response was noted with increasingly leptokurtic distributions of effects of segregating alleles and with increasingly U-shaped probability distributions of allele frequency. The variation in the genic variance or response can be easily computed with a transition matrix. Using Monte Carlo simulation, however, it is possible to generate replicates of responses for different $N_1 E(s_1^*)$ and $N_2 E(s_2)$ parameter combinations and the general pattern of the response is perhaps easier to visualise (and compare to the results of experiments).

Table 3 shows cumulative responses and CVs of cumulative responses to generations 10 and 20 among 10 independent replicates sampled from independent populations initially at a mutation-stabilizing selection balance with a range of strengths of stabilizing selection and sizes of gene effects (examples of responses are plotted in Figs. 6 and 7). The Table compares results from a reflected gamma distribution of mutant effects with shape parameter $\beta = \frac{1}{2}$ and equal probabilities of positive and negative effects ($\beta \rightarrow \infty$). The population size in the stabilizing selection phase was $N_1 = 160$ and in the directional selection phase was $N_2 = 20$, and V_M was 10^{-3} . The main points to note from the table are: (i) In theory, the average initial response rate is equal to the standing additive variance in the stabilizing selection phase. With no stabilizing selection [$N_1 E(s_1^*) = 0$], the theoretical initial rate is therefore $2N_1 V_M = 0.32 V_E$, but the cumulative response to generation 10 was less than ten times this because of the presence of disequilibrium generated by directional selection and a loss of genetic variance due to changes in gene frequency; the average response with stabilizing selection is lower than the average response from initially unselected populations

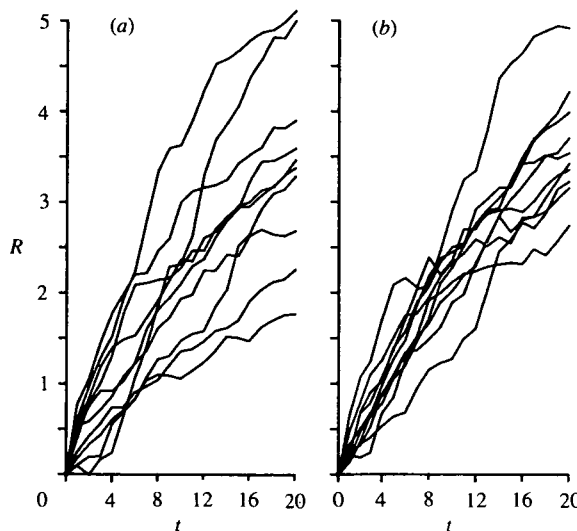


Fig. 6. There was no selection in the stabilizing phase. The sizes of gene effects were given by $\epsilon = 0.2$ and V_M was 10^{-3} .

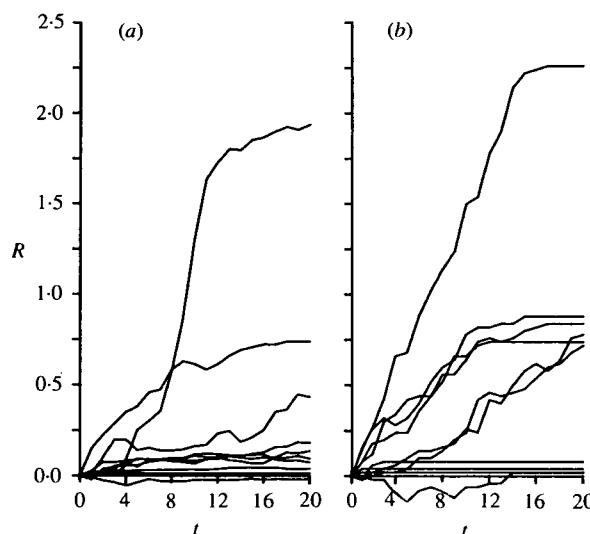


Fig. 7. The strength of stabilizing selection was such that $N_1 E(s_2) = 8$, and the value of ϵ was 0.8 with $V_M = 10^{-3}$.

Figs. 6–7. Examples of selection responses generated by Monte Carlo simulation from directional selection of samples of $N_2 = 20$ individuals from independent populations of size $N_1 = 160$ individuals at mutation-stabilizing selection-drift equilibrium (a) Reflected gamma distribution with $\beta = \frac{1}{2}$. (b) Equal probabilities of positive and negative effects. Response in units of i/σ .

because the expected steady state variance is lowered; (ii) the more leptokurtic distribution of mutant effects ($\beta = \frac{1}{2}$) gives a higher coefficient of variation of response than that for equal absolute values of mutant effects; (iii) with stronger stabilizing selection [$N_1 E(s_1^*) = 8$], because there is little standing variance, average response is generally small, but occasionally an allele of large effect segregating at low frequency gives a rapid early response so the range of response is large relative to the mean; (iv) many response patterns give similar means, variances and ranges of

response and in practice would not be distinguishable from one another. Some examples of responses, the results of which are summarized in Table 3, are shown in Figs. 6 and 7. Fig. 6(a, b) shows results for a reflected gamma distribution with shape parameters $\beta \rightarrow \infty$ and $\beta = \frac{1}{2}$ respectively, for the case of no stabilizing selection and small gene effects ($\epsilon = 0.2$). These response patterns are similar to those commonly observed in selection experiments, although it should be emphasised that weak stabilizing selection would not lead to much change in pattern. In contrast, Fig. 7(a, b) which again is for cases of reflected gamma distributions with shape parameter $\beta = \infty$ and $\beta = \frac{1}{2}$ respectively show much higher variation in response. In these cases, stabilizing selection was relatively strong [$N_1 E(s_1^*) = 8$] and gene effects were large ($\epsilon = 0.8$). Such responses are not typical of selection experiments.

The responses in Figs. 6 and 7 show the change in mean genotype, so there would be more variation in mean phenotype than shown because of the presence of an environmental component of variation. However, this would contribute little to the variation of response unless N_e is very small or the heritability of the character is low. The simulated selection responses were generated from independent populations. In practice, a caged population is often used to initiate independent lines, so variation in genes segregating initially and hence variation in response would be due to sampling from the base population rather than to different genes segregating in independent populations. Simulation showed that responses generated from sub-populations of a single population rather than a set of independent populations tend to vary less than the responses discussed above because of reduced variance between lines in the initial genetic variance and alleles of large effect which are not rare are likely to be fixed in all replicate lines. If, however, the mutational variance is generated by few mutants of large effects (e.g. $\epsilon = 0.8$, cf. Fig. 7), under stabilizing selection, because genes segregate at very low frequency at equilibrium, the variation in response is very similar to that obtained by sampling from independent populations.

5. Discussion

(i) Mutation-stabilizing selection balance

The two allele models of Latter (1960) and Bulmer (1972) and Turelli's (1984) 'House of Cards' approximation of the continuum of alleles model showed that the expected genic variance in an infinite population is independent of the effects of alleles at the loci controlling the trait. Similarly, our results show that with strong stabilizing selection the variance of the genic variance among independent lines is also independent of the size of gene effects, and is only a function of the effective population size, strength of

stabilizing selection, and the genomic mutation rate, λ . As drift becomes more important relative to selection, the coefficient of variation of the genic variance becomes increasingly dependent on the shape of the distribution of effects of new mutants, but we have little if any information about this parameter and can only conjecture that distributions of mutant effects are very leptokurtic (Robertson, 1967; Keightley & Hill, 1988; Shrimpton & Robertson, 1988). Variation in estimates of genetic variance from different populations therefore does not necessarily tell us much about the selective forces operating in the population.

Turelli's (1984) 'House of Cards' analysis of mutation-stabilizing selection balance is multi-allele, but the formula for the expected genic variance at equilibrium is the same as obtained from the two allele analyses of Latter (1960) and Bulmer (1972). Why is this so? In these models, the population size is assumed to be infinite or, equivalently, selection is assumed to be very strong. The probability distribution of allele frequencies is therefore of extremely U-shaped form with alleles at intermediate frequencies absent. Each new mutant allele almost always occurs at a locus previously carrying the 'wild-type' allele. The fates of new mutant alleles are therefore almost independent of any other mutant alleles segregating at the same locus in the population, and a two allele analysis with the parameter $n\mu$ replaced with genomic mutation rate, λ , is sufficient. Similarly, with weak selection the fates of different alleles at the same locus are essentially independent of one another, they can be considered as occurring at separate loci, and the two allele treatment also applies. As shown previously (Keightley & Hill, 1988), the model of stabilizing selection is very similar to a model of unconditionally deleterious genes with selection coefficient s proportional to a^2 and independent of gene frequency, q .

The tendency for stabilizing selection to generate an extremely U-shaped probability distribution of allele frequencies has important consequences for subsequent changes of variance and hence responses with directional selection. Under certain circumstances, namely strong directional and strong stabilizing selection, large increases in variance occur in early generations due to the fixation of alleles initially at low frequency; in other circumstances, namely strong directional selection and weak stabilizing selection, a rapid fall in variance (much faster than expected from drift alone) can occur due to the rapid fixation of genes initially at intermediate frequencies. Seldom, if ever, are such response patterns seen in selection experiments. For example, the replicated *Drosophila* abdominal bristle selection experiments of Clayton, Morris & Robertson (1957), Frankham, Jones & Barker (1968) and Yoo (1980), which were initiated from cage populations, showed little sign of early accelerated or rapidly falling responses. Similar regular patterns were observed in a *Tribolium* egg

production selection experiment (i.e. a character closely related to fitness) initiated from a cage population (Ruano, Orozco & Lopez-Fanjul, 1975) and in a selection experiment of cannon-bone length in Scottish blackface sheep (Atkins & Thompson, 1986). In the latter case, Atkins & Thompson showed that the response closely matched the predicted response of an 'infinitesimal model' which incorporated the effect of disequilibrium on the additive variance (Bulmer, 1980 ch. 9). The results of the experiment of Frankham *et al.* (1968) are of particular interest because selection on bristle score was performed using various population sizes and a range of selection intensities. The expected response can be estimated if the 'infinitesimal model' is assumed and the initial genetic variance is obtained from: (i) the heritability estimated from the base population; or (ii) the realised heritability estimated from the selection response in the first one or two generations which in theory is almost independent of the magnitude of gene effects. Such analysis shows substantially lower predicted responses using the base population genetic variance estimate than observed in the experiment. Using the genetic variance obtained from the realised heritability in the first two generations, the agreement between the experimental results and the infinitesimal model is closer. With the strongest selection strength (10%) and the biggest population ($N = 80$), some hint of an accelerating response was observed, but unfortunately this line was not replicated. The responses from Yoo's (1980) long-term experiment fit closely an infinitesimal model if parameters derived from the initial generations are assumed although the response continued longer than predicted by the infinitesimal model. The results of Falconer's (1973) replicated mouse body-weight selection experiment also give a reasonable fit to the infinitesimal model if the genetic variance derived from the response in the initial generations is assumed, but this realised heritability is rather higher than Falconer's estimate of the heritability in the base population. In this case the lines were derived from crosses of inbred lines and presumably some alleles were initially at intermediate frequencies.

The experimental selection response patterns do not tell us very much about the strength of stabilizing selection which might affect the characters in the base population because, as $N_2 s_2 \rightarrow 0$, the expected variance each generation of directional selection becomes the same irrespective of the distribution of allele frequencies. There is, however, a general absence of observations of either rapid rises or rapid falls in variance compared to those expected for drift alone. At least two explanations are possible for these observations, although they are not mutually exclusive. (i) Selective values of directional selection [$N_2 E(s_2^*)$] are not high, say greater than one, so drift dominates. This would also imply a very large number of loci of small effect controlling variation in the trait. (ii) The

initial distribution of allele frequencies is not extreme (implying weak stabilizing selection or some other mechanism generating such a distribution). Variation from initially segregating alleles falls due to changes in gene frequency from directional selection, but mutation contributes sufficiently to variation to maintain responses (Hill, 1982*a, b*).

Many of the responses generated by Monte Carlo simulation, using a wide range of parameters, are very similar to selection responses obtained experimentally. Some types of response patterns in Table 3 and Figs. 6–7 are not, however, observed experimentally. For example, with $\epsilon = 0.8$ and a gamma distribution with shape parameter $\beta = \frac{1}{2}$, responses are very variable because there are few genes segregating. This pattern becomes more extreme with strong stabilizing selection $N_1 E(s_1^*) = 8$, Fig. 7*b*) with some lines giving a rapid early burst of response. Bursts of response have been seen in selection lines, generally in long-term experiments (Thoday, Gibson & Spickett, 1964; Yoo, 1980), and are most likely the result of fixation of mutants appearing since the start of the experiment (Hill, 1982*b*). The alternative hypothesis of segregation of rare recessive alleles is also possible, though unlikely if the burst occurs late in the experiment as in the cases cited above (Robertson, 1978). Breakdown of linkage disequilibrium is also an unlikely explanation (Keightley & Hill, 1983).

6. Concluding remarks

The validity of the stabilizing selection model of natural selection has been discussed extensively elsewhere (Robertson, 1973; Turelli, 1984, 1985; Keightley & Hill, 1988; Hill, 1989). The most important weakness is that the pure stabilizing selection model ignores selection which might be acting at the locus through pleiotropic effects on characters directly related to fitness. Other models where the mutant allele is at a selective disadvantage (e.g. Hill & Keightley, 1988) have similar qualitative effects on the probability distribution of allele frequencies. The essential problem in explaining quantitative genetic variation is not whether mutation is an adequate force to explain observed variation, for in the absence of selection it is a more than adequate force. The problem is the mode of action of natural selection and the selective values of the genes affecting the character.

The analysis here is purely additive and ignoring dominance is a serious limitation. As shown by Kacser & Burns (1981), the larger the absolute effect of a mutant allele, the more likely it is to behave as nearly recessive. The consequences of this could be deduced with specific models of the relationship between mutant effect and dominance. It is likely that the tendency to give bursts or rapid falls in response would be reduced, however, because alleles of large additive effect would contribute little to the variance

in the stabilizing selection phase and would have little chance of fixation from directional selection.

Other models of the mutation process might also be considered. For example, Cockerham & Tachida's (1987) model differs from the present step-wise model because the effect of a new mutation replaces the current value at the locus, not as in this case adding to the value. This additional constraint does not affect the equilibrium behaviour with stabilizing selection as the model is formally the same as the 'House of Cards'. It can lead, however, to limits in the case of directional selection. The present results, therefore, would only be applicable in the short term.

We are grateful to Michael Turelli for helpful discussions. We wish to thank the Agricultural and Food Research Council for financial support.

References

- Atkins, K. D. & Thompson, R. (1986). Predicted and realised responses to selection for an index of bone length and body weight in Scottish Blackface sheep. *Animal Production* **43**, 421–435.
- Avery, P. J. & Hill, W. G. (1977). Variability in genetic parameters among small populations. *Genetical Research* **29**, 193–213.
- Avery, P. J. & Hill, W. G. (1979). Variance in quantitative traits due to linked dominant genes and variance in heterozygosity in small populations. *Genetics* **91**, 817–844.
- Barton, N. H. (1986). The maintenance of polygenic variation through a balance between mutation and stabilizing selection. *Genetical Research* **47**, 209–216.
- Barton, N. H. & Turelli, M. (1987). Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genetical Research* **49**, 157–173.
- Bulmer, M. G. (1972). The genetic variability of polygenic characters under optimizing selection, mutation and drift. *Genetical Research* **19**, 17–25.
- Bulmer, M. G. (1976). The effect of selection on genetic variability: a simulation study. *Genetical Research* **28**, 101–117.
- Bulmer, M. G. (1980). *The Mathematical Theory of Quantitative Genetics*. Oxford: Clarendon Press.
- Burger, R. E., Wagner, G. P. & Stettinger, F. (1988). How much heritable variation can be maintained in finite populations by a mutation selection balance? *Evolution* (in the press).
- Clayton, G. A., Morris, J. A. & Robertson, A. (1957). An experimental check on quantitative genetical theory. I. Short-term responses to selection. *Journal of Genetics* **55**, 131–151.
- Cockerham, C. C. & Tachida, H. (1987). Evolution and maintenance of quantitative genetic variation by mutations. *Proceedings of the National Academy of Sciences, USA* **84**, 6205–6209.
- Falconer, D. S. (1973). Replicated selection for body weight in mice. *Genetical Research* **22**, 291–321.
- Frankham, R., Jones, L. P. & Barker, J. S. F. (1968). The effects of population size and selection intensity in selection for a quantitative character in *Drosophila*. I. Short term response to selection. *Genetical Research* **12**, 237–248.
- Hill, W. G. (1982a). Rates of change in quantitative traits from fixation of new mutations. *Proceedings of the National Academy of Sciences, USA* **79**, 142–145.
- Hill, W. G. (1982b). Predictions of response to artificial selection from new mutations. *Genetical Research* **40**, 255–278.
- Hill, W. G. (1989). Mutation and maintenance of quantitative genetic variation. In: *Evolution and Animal Breeding* (ed. W. G. Hill and T. F. C. Mackay). Wallingford, Oxford: C.A.B. International (in the press).
- Hill, W. G. & Keightley, P. D. (1988). Interrelations of mutation, population size, artificial and natural selection. In *Proceedings of the Second International Conference on Quantitative Genetics* (ed. B. S. Weir, E. J. Eisen, M. M. Goodman and G. Namkoong), pp. 57–70. Sunderland, Massachusetts: Sinauer.
- Hill, W. G. & Rasbash, J. (1986). Models of long term artificial selection in finite populations. *Genetical Research* **48**, 41–50.
- Kacser, H. & Burns, J. A. (1981). The molecular basis of dominance. *Genetics* **97**, 639–666.
- Keightley, P. D. & Hill, W. G. (1983). Effects of linkage on response to directional selection from new mutations. *Genetical Research* **42**, 193–206.
- Keightley, P. D. & Hill, W. G. (1987). Directional selection and variation in finite populations. *Genetics* **117**, 573–582.
- Keightley, P. D. & Hill, W. G. (1988). Quantitative genetic variability maintained by mutation-stabilizing selection balance in finite populations. *Genetical Research* **52**, 33–43.
- Kingman, J. F. C. (1978). A simple model for the balance between selection and mutation. *Journal of Applied Probability* **15**, 1–12.
- Kimura, M. (1965). A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proceedings of the National Academy of Sciences, USA* **54**, 731–736.
- Kimura, M. (1969). The number of heterozygous nucleotide sites maintained in a finite population due to a steady flux of mutations. *Genetics* **61**, 893–903.
- Lande, R. (1976). The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genetical Research* **26**, 221–235.
- Lande, R. (1988). Quantitative genetics and evolutionary theory. In *Proceedings of the Second International Conference on Quantitative Genetics* (ed. B. S. Weir, E. J. Eisen, M. M. Goodman and G. Namkoong), pp. 71–85. Sunderland, Massachusetts: Sinauer.
- Latter, B. D. H. (1960). Natural selection for an intermediate optimum. *Australian Journal of Biological Science* **13**, 30–35.
- Lynch, M. & Hill, W. G. (1986). Phenotypic evolution by neutral mutation. *Evolution* **40**, 915–935.
- Robertson, A. (1956). The effect of selection against extreme deviants based on deviation or on homozygosity. *Journal of Genetics* **54**, 236–248.
- Robertson, A. (1959). The sampling variance of the genetic correlation coefficient. *Biometrics* **15**, 219–226.
- Robertson, A. (1967). The nature of quantitative genetic variation. In *Heritage from Mendel* (ed. R. A. Brink), pp. 265–280. University of Wisconsin Press, Madison, Milwaukee and London.
- Robertson, A. (1973). The validity of the optimum model. *Advances in Applied Probability* **6**, 17–18.
- Robertson, A. (1978). The time of detection of recessive visible genes in small populations. *Genetical Research* **31**, 255–264.
- Ruano, R. G., Orozco, F. & Lopez-Fanjul, C. (1975). The effect of different selection intensities on selection responses in egg-laying of *Tribolium castaneum*. *Genetical Research* **25**, 17–27.
- Shrimpton, A. E. & Robertson, A. (1988). The isolation of polygenic factors controlling bristle score in *Drosophila*

- melanogaster*. II. Distribution of third chromosome bristle effects within chromosome sections. *Genetics* **118**, 445–459.
- Thoday, J. M., Gibson, J. B. & Spickett, S. G. (1964). Regular responses to selection. 2. Recombination and accelerated response. *Genetical Research* **5**, 1–19.
- Turelli, M. (1984). Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. *Theoretical Population Biology* **25**, 138–193.
- Turelli, M. (1985). Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits. *Genetics* **111**, 165–195.
- Yoo, B. H. (1980). Long-term selection for a quantitative character in large replicate populations of *Drosophila melanogaster*. *Genetical Research* **35**, 1–17.

