Directional Selection and Variation in Finite Populations

Peter D. Keightley and William G. Hill

Institute of Animal Genetics, University of Edinburgh, West Mains Road, Edinburgh EH9 3JN, Scotland

Manuscript received April 12, 1987
Revised copy accepted July 16, 1987

ABSTRACT

Predictions are made of the equilibrium genetic variances and responses in a metric trait under the joint effects of directional selection, mutation and linkage in a finite population. The "infinitesimal model" is analyzed as the limiting case of many mutants of very small effect, otherwise Monte Carlo simulation is used. If the effects of mutant genes on the trait are symmetrically distributed and they are unlinked, the variance of mutant effects is not an important parameter. If the distribution is skewed, unless effects or the population size is small, the proportion of mutants that have increasing effect is the critical parameter. With linkage the distribution of genotypic values in the population becomes skewed downward and the equilibrium genetic variance and response are smaller as disequilibrium becomes important. Linkage effects are greater when the mutational variance is contributed by many genes of small effect than few of large effect, and are greater when the majority of mutants increase rather than decrease the trait because genes that are of large effect or are deleterious do not segregate for long. The most likely conditions for "Muller's ratchet" are investigated.

IN recent years there has been much interest in the production and maintenance of variation in populations by mutation, stimulated by the presence of abundant variation in natural and artificial populations at the protein and DNA levels. Also, the genome is now seen as a fluid entity with transposition a particularly potent force in generating molecular variability. Variation at the phenotypic level must also originate from mutation, but the rate at which such variation is generated has been thought to be slow. This belief was derived mainly from observations of experimental populations of Drosophila. For example, the gain from new mutations in bristle score variation is of the order of one thousandth of the environmental variation per generation (discussed by LANDE 1976; HILL 1982b), and mutagenesis experiments have failed to produce large amounts of new variation in such quantitative traits (CLAYTON and ROBERTSON 1964; KITAGAWA 1967; HOLLINGDALE and BARKER 1971).

Despite the apparent slowness of accumulation of new mutational variance, theoretical analyses of the interaction of mutation and natural selection in the absence of drift have shown that mutation may be a powerful force in maintaining variation in natural populations (LANDE 1976), although the extent predicted depends on assumptions in the model (TURELLI 1984). Theoretical studies in finite populations have concentrated on the combined effect of mutation and directional selection in influencing quantitative variability and selection response rates (HILL 1982a,b). The equilibrium variance of a quantitative character is attained more quickly in the presence of selection than in its absence, and is highly dependent on population size. Thus in the early generations of a selection experiment or breeding program the response from variation generated by new mutations is expected to be small. In later generations, however, the contribution to the total variation present and hence to the response can be very important, especially in large populations. The results of long-term selection experiments can be interpreted in light of these analyses. The continued response after 120 generations of directional selection for increased pupal weight in Tribolium (ENFIELD 1980), after at least 75 generations of selection for increased bristle score of Drosophila (YOO 1980), and after 76 generations for increased oil content in maize (DUDLEY 1977) were likely to have been strongly influenced by variation arising while the experiment was proceeding.

In a previous paper (KEIGHTLEY and HILL 1983) Monte Carlo simulation was used to investigate the effect of linkage on asymptotic selection responses in small populations with new mutations and it was concluded that the asymptotic response rate is little affected by linkage, especially for species with many chromosomes. Moreover, the variance of effects of mutants on the trait was not an important parameter of the model for asymptotic selection responses were little affected by whether the new mutational variance arose from a few genes of large effect or many genes of small effect. Mutant effects on the trait were, however, assumed to be symmetrically distributed about zero.

In this paper we attempt to develop a theoretical framework to predict the amount of quantitative variation maintained and selection responses in finite populations from the simultaneous segregation of newly arising linked mutations. The analysis is based on the 'infinitesimal model' first used by FISHER (1918), which provides a reference point to allow us to better understand the results of the more complex Monte Carlo simulation.

As an extension of the previous work, we shall also investigate cases where the effects of new mutations come from a skewed distribution, i.e., mutations of increasing or decreasing effect on the character are not equally likely. While it may be reasonable to assume that mutant effects are symmetrically distributed for bristle number, it is likely that mutations affecting characters close to fitness are mainly detrimental (MUKAI *et al.* 1972). We also discuss the implication of the behavior of the model in theories of the evolutionary advantages of recombination, and in particular, investigate the conditions necessary for the operation of "Muller's ratchet" (MULLER 1964; FELSENSTEIN 1974), where the population mean can decline due to the fixation of recurrent mutations of negative effect.

ANALYSIS OF THE INFINITESIMAL MODEL

Definitions

The quantitative character is assumed to be determined by the sum of an infinite number of genes of infinitely small effects. Dominance and epistasis are assumed to be absent and expression are given for the case of two alleles per locus. At the ith locus, a_i is defined as the difference between homozygotes and q_i is the frequency of the higher valued allele.

We utilize the following symbols for three fundamental quantities which are essentially the same as defined by BULMER (1976):

$V_g = \Sigma a_i^2 q_i(1 - q_i)/2$ is the genic variance in the population. It is the genetic variance which would be present if the frequencies of each individual gene were measured and it were assumed that both correlations between loci due to linkage and deviations from Hardy-Weinberg equilibrium were zero.

V_A is the additive genetic variance in the population. It is the variance of breeding values among individuals.

$$\Sigma D = \sum\sum_{i \neq j} a_i a_j D_{ij}/2, \quad (1)$$

is the disequilibrium covariance in the population, where $D_{ij} = f_{ij} - q_i q_j$, and f_{ij} is the frequency of the corresponding gamete. It is the sum of the covariances between the values of pairs of loci in the population, and can be either negative or positive. At any generation, t, the above three quantities are related by

$$V_{A,t} = V_{g,t} + \Sigma D_t. \quad (2)$$

There is also a covariance component due to deviations from Hardy-Weinberg equilibrium (BULMER 1976). We ignore this in the model since it is transient and disappears after one generation irrespective of linkage.

Change of variance

In an isogenic population, all three quantities V_g, V_A and ΣD are zero. With constant forces of mutation, selection and drift, they approach equilibria when the rate of loss of variation due to selection and drift is balanced by the rate of gain from new mutations. In a finite population the variances will drift stochastically about the equilibrium. A prediction of the infinitesimal model is that the genic variance (V_g) is unaffected by selection (see CROW and KIMURA 1970, pp. 236–239; BULMER 1980, Ch.9) and therefore in an infinite population with mutation V_g will never reach an equilibrium. Designating V_M the expected increment in variance in the population per generation from new mutations, the expected change in genic variance in one generation is given by

$$V_{g,t+1} = V_{g,t}(1 - 1/2N) + V_M, \quad (3)$$

where N is the effective population number. In a finite population, the expected equilibrium value of V_g is $2NV_M$ (*cf.* CLAYTON and ROBERTSON 1955).

The equilibrium value of the additive genetic variance is affected by selection, mutation, drift and recombination, and is less simply derived than that of the genic variance. As a starting point, we consider the effects of selection in an infinite population with free recombination in the absence of mutation.

Selection: The effect of selection on genetic variability in the infinitesimal model has been discussed by BULMER (1971, 1980, pp. 153–154), and by FALCONER (1981, pp. 179–189). After one generation of selection of parents and breeding of progeny, the total genetic variance in the population can be divided into fractions between and within full-sib families. With random mating, the between family component is one-half of the genetic variance among parents. Selection by truncation reduces the variance in the parents by a factor of $1 - h^2k$, where h^2 is the heritability ($h^2 = V_A/V_P$, the squared correlation between phenotype and genotype; with selection on an arbitrary index h^2 is replaced by the squared accuracy of the index), and k is a constant factor describing the strength of selection. For truncation selection of a normally distributed population, $k = i(i - x)$, where i is the intensity of selection (standardized selection differential) and x is the standardized deviation of the truncation point. Thus after one generation, the between family component is given by $(1 - h^2k)V_A/2$.

Selection leads to a reduction in the genetic variance between family means, which appears as a negative disequilibrium covariance component within families. With free recombination the within family variance component is simply given by $V_g/2$ because recombination completely eliminates disequilibrium within families, but only half of the total genetic variance is initially present within full-sib families.

The total additive variance in the population after one generation of selection is obtained by adding the between and within family components,

$$V_{A,t+1} = (1 - h_t^2 k)V_{A,t}/2 + V_{g,t}/2. \quad (4)$$

The recurrence relation (4) corresponds to Equation 9.30 of BULMER (1980). Its validity depends on a normal distribution of genotypic values in the progeny, since skew in the distribution can affect the amount of variation removed by selection, but the results of BULMER (1980, Ch. 9) and ZENG (1987) indicate that, in many cases, skewness effects are small and can be ignored. Here, we use simulation to investigate possible effects of such skewness.

Mutation: As with the genic variance, the additive variance increases by V_M units each generation from mutation. Equation 4 becomes

$$V_{A,t+1} = (1 - h_t^2 k)V_{A,t}/2 + V_{g,t}/2 + V_M. \quad (5)$$

Finite population size: With a poisson distribution of family size the expected reduction in the additive genetic variance in the population is by a proportion $1/2N$ in the absence of selection. With truncation selection, the within family variance is independent of the population size, but the expected reduction in the between family component is by a proportion $1/N$ due to sampling of parents with replacement. Equation 5 becomes

$$V_{A,t+1} = (1 - 1/N)(1 - h_t^2 k)V_{A,t}/2 + V_{g,t}/2 + V_M. \quad (6)$$

Linkage: Linkage does not affect the variance between family means after one generation of selection, but affects the within family variance by reducing the amount of variation recovered from the disequilibrium covariance component due to recombination between loci. If c_{ij} is the recombination fraction between loci i and j, the disequilibrium remaining in the within family component is given by

$$2 \sum_{i \neq j} \sum D_{ij,t} (1 - 2c_{ij})a_i a_j = \Sigma D_t - 2\Sigma(D_t c),$$

where ΣD_t is defined by (1) and similarly $2\Sigma(D_t c) = \Sigma\Sigma_{i \neq j} D_{ij} a_i a_j c_{ij}$. More generally (6) becomes

$$V_{A,t+1} = (1 - 1/N)(1 - h_t^2 k)V_{A,t}/2 + [V_{g,t} + \Sigma D_t - 2\Sigma(D_t c)]/2 + V_M. \quad (7)$$

The recurrence relation for the disequilibrium component is obtained by combining equations (2) and (7) to give

$$\Sigma D_{t+1} = \Sigma D_t - \Sigma D_t/(2N) - \Sigma(D_t c) - h_t^2 k(1 - 1/N)(V_{g,t} - \Sigma D_t)/2. \quad (8)$$

Asymptotic variance

As $t \rightarrow \infty$ for finite N, the variances reach expected equilibrium values about which they fluctuate stochastically due to sampling. For free recombination (6) can be reexpressed as a quadratic

$$\tilde{V}_A^2(1 + k + 1/N - k/N) + \tilde{V}_A(V_E - 2NV_M) \cdot(1 + 1/N) - 2V_M V_E(N + 1) = 0, \quad (9)$$

where $\tilde{V}_A$ is the equilibrium value of V_A. Ignoring second order terms, (9) is approximated by

$$\tilde{V}_A^2(1 + k) + \tilde{V}_A(V_E - 2NV_M) - 2NV_M V_E = 0. \quad (10)$$

Thus, $\tilde{V}_A$ is a function of mutation rate and population size as their product NV_M.

For complete linkage ($c = 0$ for all possible pairs of loci) a quadratic in $\tilde{V}_A$ is obtained by combining (2) and (7),

$$\tilde{V}_A^2(Nk + 1 - k) + \tilde{V}_A(V_E - 2NV_M) - 2NV_M V_E = 0. \quad (11)$$

Here, the relationship of $\tilde{V}_A$ to V_M and N is not as simple as in (10), but is a function of NV_M and Nk.

When recombination is finite, the simultaneous recurrence relations (3), (7) and (8) do not appear to have a simple solution. Their properties were investigated by iterating until steady state was achieved with initial values of V_g, V_A and ΣD set at zero, as would be the case in an isogenic population. The effect of a finite chromosome length was modeled by dividing the chromosome into a large number of equivalent segments (typically 100) and calculating the recombination fraction and hence the disequilibrium contribution from each possible pair of segments. This method exactly models the case of infinitesimally small effects as the number of segments approaches infinity, but increasing the number of segments beyond 100 made almost no difference to the results. The total amount of recombination was specified by L, the map length of the chromosome, and HALDANE's (1919) mapping function was used to related recombination fraction to map distance (l_{ij}) between pairs of loci: $c_{ij} = [1 - \exp(-2l_{ij})]/2$. Previous analyses (AVERY and HILL 1979) indicate that other models relating recombination fraction to map length (e.g. with crossover interference) make little difference in this type of model.

THE SIMULATION MODEL

The model is similar to that described by KEIGHTLEY and HILL (1983), except here we have simulated

truncation selection rather than the previous fertility selection model.

Population structure and truncation selection regime: The parental population consisted of N monoecious diploid individuals with random family size and mating, but selfing excluded. The present simulations were restricted to one chromosome pair per individual, with the number of crossovers per chromosome pair per generation sampled from a poisson distribution with parameter L. The positions of crossovers were uniformly distributed on the chromosome. Each generation, T progeny were bred from the parents and to the value of each progeny, an independent environmental deviate of mean zero and variance $V_E = 1$ was added. The N individuals of highest value were selected for breeding in the following generation so N/T is the fraction selected.

Mutation effects: The expected number of new mutations per haploid genome (constant each generation) was λ, and their effects a, the difference between homozygotes, were sampled from a time-invariant distribution. The expected increment in genotypic variance, V_M, each generation is given by

$$V_M = \lambda E(a^2)/2 \qquad (12)$$

(HILL 1982a). Importantly, V_M is independent of the sign of the mutations (*i.e.*, whether they have a negative or positive effect on the value of the character).

Mutations occurred in the T offspring and were assumed to have an immediate effect on their phenotype. Mutations had two basic attributes of a position on the chromosome and a value. They were assumed to be equally likely to occur at any position on the chromosome, and the number of mutations per chromosome was sampled from a poisson distribution with parameter λ. The choice of distribution for mutational effects is essentially arbitrary, but since it is likely that mutations of small effect are much more common than those of large effect (MUKAI *et al.* 1972) the effects of both positive and negative mutations were assumed to have a gamma distribution with shape parameter one-half. For example, for positive mutants,

$$f(a) = \alpha^{1/2} e^{-\alpha a} a^{-1/2}/\Gamma(1/2), \qquad 0 < a < \infty \qquad (13)$$

with moments $E(a) = 1/(2\alpha)$, $E(a^2) = 3/(4\alpha^2)$, and the root mean square is defined as $\epsilon = [E(a^2)/V_E]^{1/2}$.

This model was used to investigate the effect on genetic variability and selection responses of different proportions of positive and negative mutational values. In practice, therefore, mutational effects were sampled from a gamma distribution of parameter ϵ with sign randomly allocated, and probability P of being positive. For the resulting distribution, referred to as the 'reflected double gamma,' $E(a) = (2P - 1)/(2\alpha)$, and $E(a^2)$ and thus ϵ are unchanged.

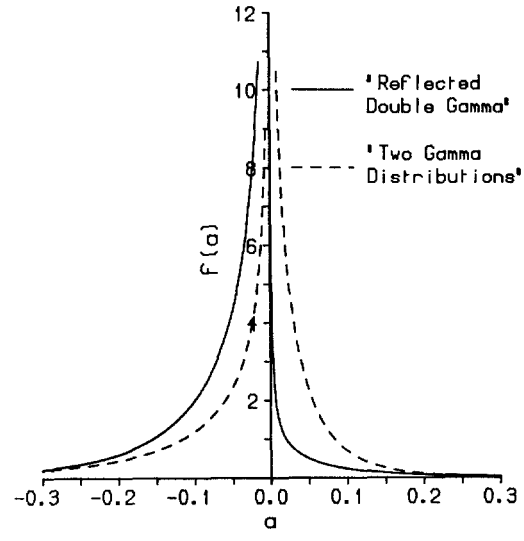

FIGURE 1.—Comparison of 'reflected double gamma' and 'two-gamma distributions' for the same values of $E(a)$ and $E(a^2)$. In the 'reflected double gamma', $[E(a^2)]^{1/2} = \epsilon = 0.1$ and P (proportion positive) is 0.1. In the 'two-gamma distributions,' $\epsilon_1 = 0.0517$, $\epsilon_2 = 0.132$ and one-half are positive.

As an alternative, to compare the 'reflected double gamma' with a different skewed distribution we used a scheme referred to as 'two-gamma distributions.' The expected number of positive and negative effects were equal ($P = 0.5$), but the values of positive mutations were sampled from a gamma distribution with parameter ϵ_1 and negative mutations were sampled from a gamma distribution with parameter ϵ_2, and allocated a negative sign. An example of a 'reflected double gamma' along with two-gamma distributions together giving the same $E(a)$ and $E(a^2)$ is given in Figure 1. Values of ϵ_1 and ϵ_2 were chosen so that the mean and variance were the same as for the reflected double gamma with specified P.

The case of an infinite number of infinitesimally small effects was simulated by adding a random normal deviate of mean zero and variance $V_M/2$ to the value of each chromosome. This was only possible for the simulation of zero recombination where it was not necessary to store individual effects.

Computer simulation programs: The simulation programs were essentially the same as described by KEIGHTLEY and HILL (1983) except that selection was by truncation rather than through fertility differences. In order to check the validity of the programs we used two main techniques: (1) running of 'marginal cases' which produce known results, *e.g.*, $N = T$ gives $\tilde{V}_A = \tilde{V}_g = 2NV_M$, $\Sigma D = 0$; and (2) the programs for free recombination, no recombination and specified L were separate and used slightly different algorithms, so could independently check each other at the margins.

Computation of results: The population was initialized in an isogenic state and the simulation started. To allow the system to reach steady state, the early

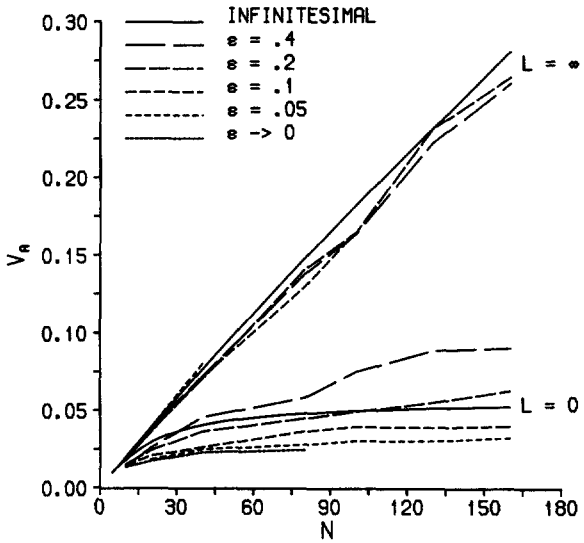

FIGURE 2.—The equilibrium variance is shown for the infinitesimal and simulation models for various population sizes and 50% truncation selection. $V_M/V_E = 10^{-3}$. In the simulation model, a range of values of ϵ and corresponding λ were used, with mutations coming from a symmetrical 'reflected double gamma' distribution ($P = 0.5$).

generations (200 for the populations simulated) were ignored. Thereafter the asymptotic response rate was calculated from the difference in mean value every other 10 generations and the mean genotypic variance, skew and kurtosis were computed every 10 generations. For a given V_M the computing time was approximately proportional to N^2 and inversely proportional to ϵ^2. So results for small ϵ (*e.g.*, 0.05) were only obtainable for N of 40 or less.

RESULTS

Comparison of the simulation and the infinitesimal model: Predictions of $\tilde{V}_A$ from both the Monte Carlo simulation and from the infinitesimal model for varying population size are shown in Figure 2. A value of V_M of $10^{-3}V_E$ was used, but in the simulation a range of sizes of effects was compared with corresponding values for the number of mutants per generation to satisfy (12).

With free recombination, the infinitesimal and simulation models are in good agreement. Surprisingly, the agreement is close even with relatively large effects and few mutants ($\epsilon = 0.4$). The disequilibrium present in the populations simulated can be estimated by subtracting the observed $\tilde{V}_A$ from the genic variance (given by $2NV_M$ in the infinitesimal case). As expected, with free recombination the amount of disequilibrium is small.

With complete linkage, the curves for different ϵ values differ substantially, larger values of ϵ giving higher predictions of $\tilde{V}_A$. The infinitesimal model is a poor predictor for complete linkage especially when effects are large, but it also overestimates $\tilde{V}_A$ when $\epsilon \to 0$. The overestimation can be explained by the presence of negative skew in the distribution of genotypic values of individuals (Table 1). Negative skew leads to a greater loss of variance each generation than predicted by the constant factor $i(i - x)$, and hence a lower $\tilde{V}_A$.

TABLE 1

Equilibrium skewness of genotypic progeny values, computed as $g_1 = [\Sigma(X - \bar{X})^3/N]/\tilde{V}_A$, given for the case of no recombination ($L = 0$)

	Population size (N)			
	10	20	40	80
ϵ	Equilibrium skewness (g_1) among progeny			
$\to 0$	−0.0794	−0.147	−0.180	−0.184
0.05	−0.0592	−0.105	−0.149	−0.145
0.1	−0.0533	−0.0540	−0.0717	−0.0381
0.2	0.0142	0.0220	0.0465	0.0852
0.4	0.0289	0.0568	0.138	0.0735

$V_M/V_E = 10^{-3}$ and mutants come from a symmetrical 'reflected double gamma' distribution. Fifty percent truncation selection was simulated.

The effects of a finite amount of recombination are shown in Figure 3. The simulation and infinitesimal models agree at the free recombination limit but there is an increasing discrepancy at low recombination fractions. At the population sizes simulated, most of the effect of linkage is eliminated by one or two crossovers per chromosome per generation. The results are in agreement with those of KEIGHTLEY and HILL (1983) which used a fertility model of selection rather than the present viability model.

Asymmetrical distribution of mutational effects: Previous analyses (HILL 1982b; KEIGHTLEY and HILL 1983) have indicated that if the distribution of mutational effects is symmetrical (*i.e.*, the mutational variance contributed by negative and positive mutations is equal), then the shape of the density function of effects does not have much influence on selection responses and variation maintained.

Predictions of $\tilde{V}_A$ from simulations of different population sizes using the 'reflected double gamma' are plotted in Figure 4 for free recombination and values of P representing cases where mutants are mostly negative ($P = 0.1$), positive ($P = 0.9$) or symmetrically distributed ($P = 0.5$). The results show that $\tilde{V}_A$ is higher than the infinitesimal prediction (also shown in the figure) when mutants have predominantly positive effects and lower when most are negative. As the expected value of mutational effects approaches zero, however, the results approach the infinitesimal prediction. In the limit all the effects become infinitely small and the models must coincide. With finite effects, there are two reasons for the discrepancy from the infinitesimal prediction. Firstly, most negative mutations are lost almost immediately and contribute

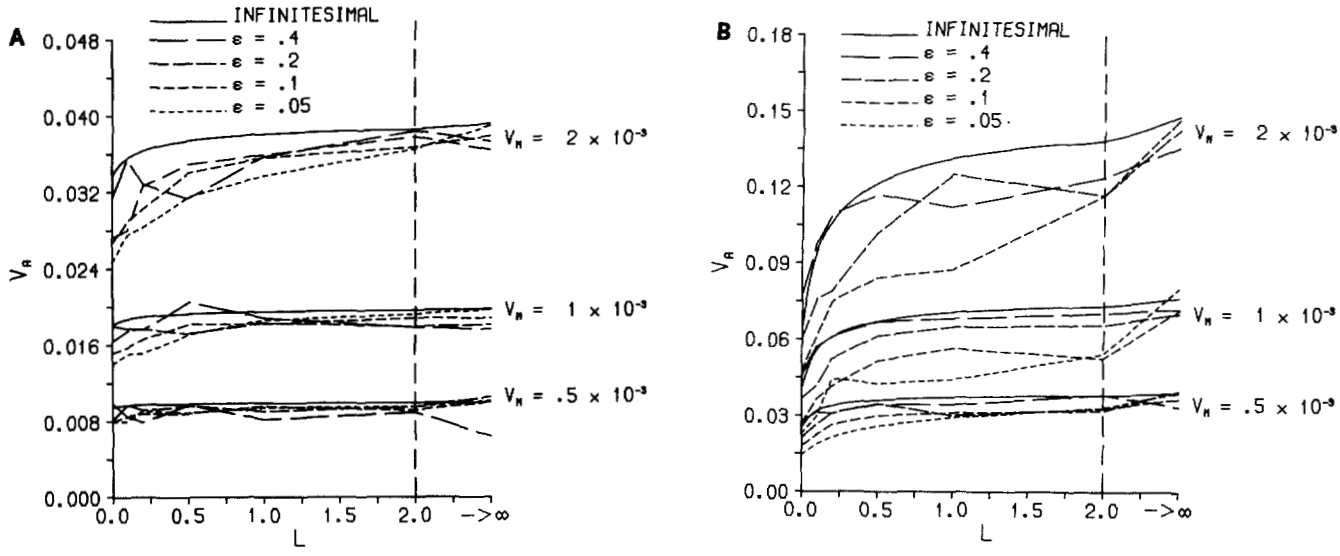

FIGURE 3.—Equilibrium variance ($\tilde{V}_A$) predicted from the simulation and infinitesimal models for three different values of mutational variance are plotted for different chromosome lengths (L) with 50% truncation selection. (A) $N = 10$; (B) $N = 40$.

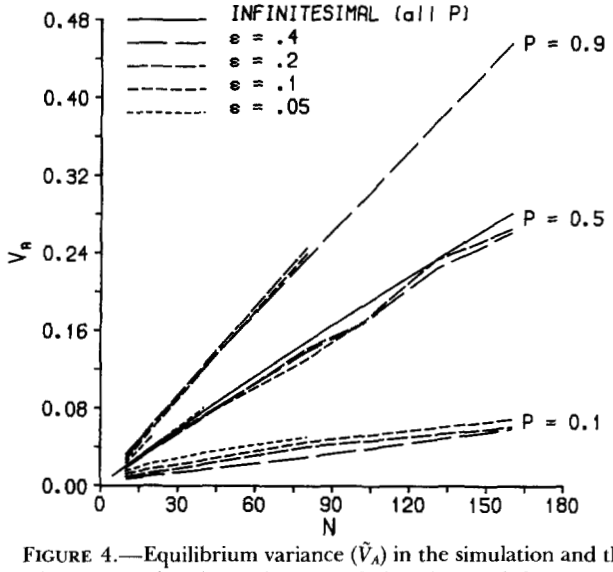

FIGURE 4.—Equilibrium variance ($\tilde{V}_A$) in the simulation and the infinitesimal models for various population sizes and three values of P (proportion of mutants positive). Free recombination, otherwise parameters as in Figure 2.

TABLE 2

Equilibrium skewness of progeny genotypic values computed as in Table 1 for the reflected double gamma distribution of mutants, $V_M/V_E = 10^{-3}$ and free recombination

	Proportion of mutants positive (P)		
	0.1	0.5	0.9
ϵ	Equilibrium skewness (g_1) among progeny		
0.05	−0.118	0.0312	0.0785
0.1	−0.285	−0.0121	0.0539
0.2	−0.699	0.0514	0.134
0.4	−1.67	0.0773	0.385

The population size (N) = 20 and 50% truncation simulation was simulated.

little to the variance maintained (HILL 1982b), while many positive mutations get fixed and contribute substantially to variance especially when at intermediate frequencies. The variance maintained is therefore proportional to the fraction of the mutational variance contributed by positive effects, $E^+(a^2)$ (HILL 1982a). Secondly, the mutation pressure generates skew in the distribution of genotypes (negative or positive depending on the sign of the mutations) since the density function of mutational effects is itself skewed (Table 2).

Where the skewness in the distribution of mutations is generated by two gamma distributions of different scale we obtain a similar pattern. In Table 3, we compare the results from the 'reflected double gamma' with results from two such distributions together giving the same mutational variance and the same mean effect. With small effects ($\epsilon = 0.1$) there is little difference in $\tilde{V}_A$, but where effects are large ($\epsilon = 0.4$), the dominating influence of the positive mutants, which are much more likely to get fixed and hence contribute to $\tilde{V}_A$, leads to noticeable differences between the models.

Response to selection: The response is given by $R = iV_A/\sigma_P$ where σ_P is the phenotypic s.d. If the distribution of genotypes and environmental effects are normal and independent of one another, the regression of A on P is linear (*e.g.*, FALCONER 1981). When the mean value of mutational effects is non-zero, there is an additional change in mean, $\Delta_m = 2\lambda E(a)$, due to the mutational pressure. In the 'reflected double-gamma' distribution $E(a) = \epsilon(2P - 1)/\sqrt{3}$, so

$$\Delta_m = 2\lambda\epsilon(2P - 1)/\sqrt{3}. \quad (14)$$

The responses to selection for various population sizes with both free and zero recombination and three

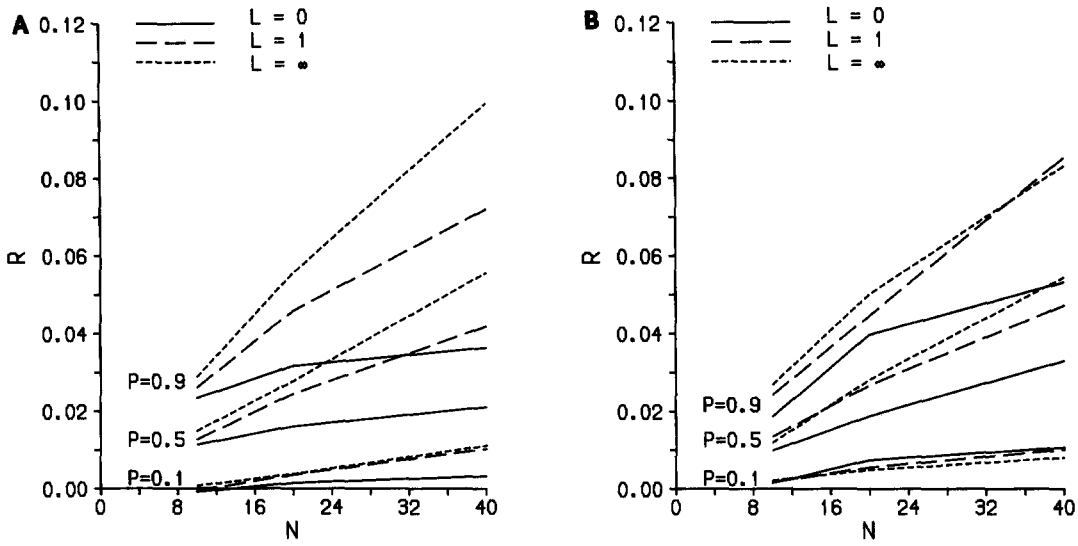

FIGURE 5.—Response rates where mutant effects have a reflected double gamma distribution with three values of P (proportion positive) in various population sizes. $V_M/V_E = 10^{-3}$ and 50% truncation selection. (A) $\epsilon = 0.1$; (B) $\epsilon = 0.4$.

values of P are plotted in Figure 5. Since Δ_m is independent of population size but $\tilde{V}_A$ is highly dependent, net responses become negative in small populations if most mutations are deleterious ($P = 0.1$).

Restating (12) and (14), $V_M \propto \lambda\epsilon^2$ and $\Delta_m \propto \lambda\epsilon$. It is clear that for a given V_M, as the expected magnitude of effects decreases ($\epsilon \rightarrow 0$) and hence the number of mutations increases, Δ_m must increase. Thus, if the new mutational variance is due to a large number of small negative effects, the mean value of a population will decline faster than if the mutational variance is due to a small number of large effects.

Paradoxically, the effects of linkage, *i.e.*, the difference between response rates for free recombination and complete linkage, are most severe when most mutations are positive ($P = 0.9$). Both positive and negative mutants interfere with each other's fixation probabilities (HILL and ROBERTSON 1966), but when most mutations are of positive value, there are more segregating so linkage effects are more important.

The response to selection from the fixation of freely recombining mutants can be approximated analytically if we assume that disequilibrium effects are small, so that the fate of each mutant is independent. In this case the response is given by

$$R = 2N\lambda \int_{-\infty}^{\infty} au(a)f(a)da \qquad (15)$$

where $f(a)$ is the density function of mutant effects and $u(a)$ is the fixation probability of mutants of effect a (HILL 1982a). For a gamma-distribution of mutational effects an approximation for R can be obtained by replacing $u(a)$ by the diffusion approximation of KIMURA (1957) for the fixation probability of additive genes [see HILL and RASBASH (1986) and Appendix].

Predicted response rates from simulation and Equation 15 are compared in Table 4. In general, the simulation agrees quite closely with the model of independent mutants. Comparing the results from the 'two-gamma' mutational distribution with those from the 'reflected double gamma' substantial differences in response rates can occur. The differences in response are consistent with the differences in variance maintained (Table 3). Also the differences are most extreme in small populations when the proportion of positive mutants (P) is 0.1. In this case, response rates are near zero so any difference is magnified.

DISCUSSION

Models: The computer simulation model is in itself of interest because it has been set up as far as possible in terms of known or measurable parameters, parental and progeny population sizes, map length of the chromosome, new mutational variance and distribution of mutational effects (assumed to be gamma form). The number of genes in the model is not fixed as in other models (*e.g.*, LANDE 1976; TURELLI 1984), but more naturally the number of loci with alleles segregating varies while the simulation is running. Furthermore, mutations which occur very close together on the chromosome can be considered either as alleles at separate loci or multiple alleles at the same locus. The model therefore connects and concurs simultaneously with the infinite locus models of BULMER (1971, 1976) and the 'infinite alleles' model of KIMURA (1965), and the possibility of intragenic recombination is accounted for. There are also similarities to a 'stepwise mutation' model (*e.g.*, TURELLI 1984). Any model of the mutational process, however, needs to be justified in terms of the effect on series of mutations on a gene (for say an enzyme) which in turn affects a quantitative character (say a flux), for which models have been

TABLE 3

Additive genetic variation maintained ($\tilde{V}_A$) for 'reflected double gamma' (RDG) and 'two-gamma distributions' (TG)

					Population size (N)			
P	ϵ	ϵ_1	ϵ_2	Model	10	20	40	80
					Maintained Additive Variation ($\tilde{V}_A$)			
0.1	0.1			RDG	0.0119	0.0173	0.0262	0.0451
0.5		0.0517	0.132	TG	0.0105	0.0165	0.0259	0.0474
0.9	0.1			RDG	0.0264	0.0579	0.120	0.239
0.5		0.132	0.0517	TG	0.0280	0.0587	0.114	0.214
0.1	0.4			RDG	0.00646	0.0103	0.0165	0.0294
0.5		0.207	0.527	TG	0.00660	0.0116	0.0228	0.0399
0.9	0.4			RDG	0.0324	0.0617	0.121	0.234
0.5		0.527	0.207	TG	0.0256	0.0554	0.110	0.217

$V_M/V_E = 10^{-3}$ and 50% truncation selection was simulated.

TABLE 4

Response rate in the character predicted by the analytical model of independent mutants (IM) and simulation where mutational effects are sampled from a 'reflected double gamma' (RDG) distribution and 'two-gamma' distributions (TG)

					Population size (N)		
					10	20	40
P	ϵ	ϵ_1	ϵ_2	Model	Response rate in character		
0.1	0.1			RDG	−0.00072	0.00427	0.0197
0.1	0.1			IM	0.00009	0.00428	0.0108
		0.0517	0.132	TG	0.00311	0.00883	0.0130
		0.0517	0.132	IM	0.00376	0.00811	0.0162
0.5	0.1			RDG	0.0149	0.0278	0.0554
0.5	0.1			IM	0.0143	0.0268	0.0572
0.9	0.1			RDG	0.0304	0.0512	0.152
0.9	0.1			IM	0.0304	0.0529	0.104
		0.132	0.0517	TG	0.0258	0.0484	0.0890
		0.132	0.0517	IM	0.0242	0.0479	0.0959
0.1	0.4			RDG	0.00245	0.00551	0.0111
0.1	0.4			IM	0.00208	0.00435	0.0088
		0.207	0.527	TG	0.00377	0.00849	0.0190
		0.207	0.527	IM	0.00347	0.00688	0.0137
0.5	0.4			RDG	0.0120	0.0281	0.0543
0.5	0.4			IM	0.0110	0.0220	0.0440
0.9	0.4			RDG	0.0240	0.0452	0.0890
0.9	0.4			IM	0.0199	0.0396	0.0791
		0.527	0.207	TG	0.0256	0.0412	0.0835
		0.527	0.207	IM	0.0179	0.0359	0.0719

$V_M/V_E = 10^{-3}$, and 50% truncation selection was simulated.

developed and explored by KACSER and BURNS (1973, 1981). It also needs to be justified in terms of the possible effects of mutations on the enzyme activities themselves.

This study has extended work done previously (KEIGHTLEY and HILL 1983), and our understanding of the behavior of the model has been greatly improved. The results agree with the previous paper, where a fertility model of selection was investigated, although in that case the predicted value of V_A maintained was slightly lower since selection has more impact on the effective population size.

With V_M/V_E of the order of 10^{-3}, it is clear that with directional selection and free recombination, disequilibrium will be a minor part of the total variation present, and a small number of crossovers goes most of the way to free recombination. This is not to say that disequilibrium between alleles at closely linked loci will be absent, only that the disequilibrium variance which depends on the effect of the alleles on the character will be low.

Using the 'infinitesimal model' we have developed an analytical solution for the equilibrium genetic variance under the joint effects of mutation, linkage and

selection in a finite population. The model agree well with the Monte Carlo simulations both where effects are small ($\epsilon \rightarrow 0$), and also where mutational effects are relatively large. This behavior is consistent with the results of HILL (1982a) where independent genes give $\tilde{V}_A$ of $2NV_M$, irrespective of the mutational distribution. The higher fixation probability of mutants of large effect and their higher contribution to the variance in the character is nearly exactly balanced by their shorter fixation times and fewer number when compared to genes of small effect.

Asymmetry of mutant effects: When mutants come from an asymmetrical distribution the behavior is not as simple. As $\epsilon \rightarrow 0$ for any population size, the equilibrium additive variance will be essentially the same as predicted by the infinitesimal model. This will be true irrespective of the selection regime, selection only generating disequilibrium. With larger effects, $\tilde{V}_A$ becomes dependent on the proportion of mutants of positive effect (P). When the value of most mutational effects exceeds σ/Ni, $\tilde{V}_A$ is approximated by $4PNV_M$. This is so because the fixation probability of such mutants (and therefore the probability that they will reach intermediate frequencies and contribute substantially to the population variance) is proportional to a and independent of N. The number of mutants appearing in the population is, however, proportional to N. For small effects or in small populations, terms for which $|a| < \sigma/Ni$ become more important and $\tilde{V}_A$ will approach the infinitesimal prediction of $2NV_M$. Figure 4 shows, however, that effects must become very small or N very small before $\tilde{V}_A$ will be much different from $4PNV_M$.

With an asymmetrical mutational distribution, the distribution of genotypic values becomes skewed in the same direction as the mutational skew. A skewed distribution will lose more or less variance from directional selection depending on the sign of the skewness, so equilibrium variances are affected by such skewness. The simulations show that skewness is more important when mutational effects are large. At this point, we should mention that directional selection also generates skewness (positive) in the genotypic distribution (BULMER 1980, Ch. 9) so predicting the asymptotic distribution of genotypes when mutational effects are skewed is a difficult task.

The behavior of the system where genes are linked is also strongly affected by skewness in the genotypic distribution. When effects are small (*e.g.*, $\epsilon = 0.05$), the infinitesimal model overestimates the equilibrium genetic variance. The most likely explanation is a negatively skewed mutational distribution generated due to the loss of all but the best 'haplotype' and the presence of a 'tail' of individuals of lower value from mutation. This tendency to generate negative skewness is partially opposed by truncation selection generating positive skewness as mentioned earlier. The effect of linkage in generating skewness has been noted in earlier two locus studies (HILL and ROBERTSON 1966). Where effects are large, and therefore fewer mutations are occurring per generation, genes behave more as if they were independent and therefore higher V_A is maintained.

As a consequence of a negatively skewed mutational distribution, the rate of fixation of deleterious genes can exceed the rate for beneficial mutants and the population mean can decline; an effect corresponding to "Muller's ratchet" (MULLER 1964). We have identified a number of conditions necessary for the operation of the ratchet: (1) small population size since the fixation of deleterious mutants depends on chance; (2) many mutants of small effect (as opposed to a few of larger effect) since the 'mutation pressure' on the population mean is greater in this case; and (3) tight linkage since less standing variation will be available to oppose the mutation pressure. Linkage is also more important with small effects (*cf.* Figure 2).

Somewhat surprisingly the simulations show that linkage has most influence where most mutants are of positive value, and linkage effects can all but disappear when most are negative (*cf.* Figure 5). The explanation, however, is simple: deleterious genes almost never get fixed, while positive mutants get fixed with probability proportional (if independent) to a. In this latter case, however, linked positive mutants present simultaneously in the population can form unfavorable repulsion combinations leading to a reduction in fixation probabilities.

Directional *vs.* stabilizing selection: We should now point out that the free recombination results differ markedly from models of the maintenance of heritable variation in quantitative characters under mutation-stabilizing selection balance (*e.g.*, LANDE 1976; TURELLI 1984). In these models, the predicted equilibrium genetic variance $\tilde{V}_A$ is finite in an infinite population. With directional selection, $\tilde{V}_A$ will become infinite in an infinite population. The underlying cause of this discrepancy is the presence of mutants of positive effect on the character and hence on reproductive success with directional selection, but mutants of both positive and negative effect on the trait are deleterious for fitness with stabilizing selection (ROBERTSON 1956). Further consequences of the models are discussed elsewhere (HILL and KEIGHTLEY 1987).

We wish to thank the Agricultural and Food Research Council for financial support.

LITERATURE CITED

AVERY, P. J. & W. G. HILL, 1979 Variance in quantitative traits due to linked dominant genes in finite populations. Genetics **91:** 817–844.

BULMER, M. G., 1971 The effect of selection on genetic variability. Am. Nat. **105:** 201–211.

BULMER, M. G., 1976 The effect of selection on genetic variability: a simulation study. Genet. Res. **28:** 101–117.

BULMER, M. G., 1980 *The Mathematical Theory of Quantitative Genetics.* Oxford University Press, Oxford.

CLAYTON, G. A. and A. ROBERTSON, 1955 Mutation and quantitative genetic variation. Am. Nat. **89:** 151–158.

CLAYTON, G. A. and A. ROBERTSON, 1964 The effects of X-rays on quantitative characters. Genet. Res. **5:** 410–422.

CROW, J. F., and M. KIMURA, 1970 *An Introduction to Population Genetics Theory.* Harper & Row, New York.

DUDLEY, J. W., 1977 76 generations of selection for oil and protein percentage in maize. pp. 459–473. In *Proceedings of the International Conference on Quantitative Genetics,* Edited by E. POLLAK, O. KEMPTHORNE and T. B. BAILEY, JR. Iowa State University Press, Ames.

ENFIELD, F. D., 1980 Long term effects of selection; the limits to response. pp. 69–86. In *Selection Experiments in Laboratory and Domestic Animals,* Edited by A. ROBERTSON. Commonwealth Agricultural Bureaux, Slough.

FALCONER, D. S., 1981 *Introduction to Quantitative Genetics,* Ed. 2. Longman, London.

FELSENSTEIN, J., 1974 The evolutionary advantage of recombination. Genetics **78:** 737–756.

FISHER, R. A., 1918 The correlation between relatives on the supposition of Mendelian inheritance. Trans. R. Soc. Edinburgh **52:** 399–433.

HALDANE, J. B. S., 1919 The combination of linkage values and the calculation of distance between loci of linked factors. J. Genet. **8:** 299–309.

HAMMING, R. W., 1962 *Numerical Methods for Scientists and Engineers,* McGraw-Hill, New York.

HILL, W. G., 1982a Rates of change in quantitative traits from fixation of new mutations. Proc. Natl. Acad. Sci. USA **79:** 142–145.

HILL, W. G., 1982b Predictions of response to artificial selection from new mutations. Genet. Res. **40:** 255–278.

HILL, W. G., and P. D. KEIGHTLEY, 1987 Interrelations of mutation, population size, artificial and natural selection. In: *Proceedings of the Second International Conference on Quantitative Genetics.* Edited by E. J. EISEN, M. M. GOODMAN, G. NAMKOONG and B. S. WEIR. in press.

HILL, W. G., and J. RASBASH, 1986 Models of long term artificial selection in finite populations. Genet. Res. **48:** 45–50.

HILL, W. G., and A ROBERTSON, 1966 The effects of linkage on limits to artificial selection. Genet. Res. **8:** 269–294.

HOLLINGDALE, B., and J. S. F. BARKER, 1971 Selection for increased abdominal bristle number in *Drosophila melanogaster* with concurrent irradiation. Theor. Appl. Genet. **41:** 208–215.

KACSER, H., and J. A. BURNS, 1973 The control of flux. Symp. Soc. Exp. Biol. **27:** 65–104.

KACSER, H., and J. A. BURNS, 1981 The molecular basis of dominance. Genetics **97:** 639–666.

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

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

KIMURA, M., 1965 A stochastic model concerning the maintenance of genetic variability in quantitative characters. Proc. Natl. Acad. Sci. USA **54:** 731–736.

KITAGAWA, O., 1967 The effects of x-ray irradiation on selection response in *Drosophila melanogaster.* Jpn. J. Genet. **42:** 121–137.

LANDE, R., 1976 The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genet. Res. **26:** 221–235.

MUKAI, T., S. J. CHIGUSA, L. E. METTLER and J. F. CROW, 1972 Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster.* Genetics **72:** 339–355.

MULLER, H. J., 1964 The relation of recombination to mutational advance. Mutat. Res. **1:** 2–9.

ROBERTSON, A., 1956 The effect of selection against extreme deviants based on deviation or on homozygosis. J. Genet. **54:** 236–248.

TURELLI, M., 1984 Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominance bristle. Theor. Popul. Biol. **25:** 138–193.

YOO, B. H., 1980 Long term selection for a quantitative character in large replicate populations of *Drosophila melanogaster.* Genet. Res. **35:** 1–17.

ZENG, Z.-B., 1987 Genotypic distribution at the limits to natural and artificial selection with mutation. *Theor. Popul. Biol.* **32:** 90–113

Communicating editor: M. TURELLI

APPENDIX

Equation 15 can be restated in terms of $s = ia/\sigma_P$, the selective advantage of an allele of effect a:

$$R = (2N\lambda\sigma_P/i)\int_{-\infty}^{\infty} u(s)sf(s)ds. \qquad \text{(A1)}$$

For positive mutants, an expression for the response, $R(+)$, is obtained by replacing $u(s)$ by KIMURA's diffusion approximation for the fixation probability of a mutant of effect s and $f(s)$ by the density function for the gamma distribution (13):

$$R(+) = (2N\lambda\sigma_P/i)\int_0^{\infty} \frac{1-e^{-s}}{1-e^{-2Ns}}\, s\, \frac{\alpha^{1/2}e^{-\alpha s}s^{-1/2}}{\Gamma(1/2)}\, ds. \qquad \text{(A2)}$$

Expanding (A2), we obtain

$$\begin{aligned} R(+) = {} & (2N\lambda\sigma_P/i)(\alpha^{1/2}/\Gamma(1/2)) \\ & \cdot\int_0^{\infty} [(e^{-\alpha s}s^{3/2} + e^{-2Ns-\alpha s}s^{3/2} + \cdots) \qquad \text{(A3)} \\ & - (1/2!)(e^{-\alpha s}s^{5/2} + e^{-2Ns-\alpha s}s^{5/2} + \cdots) \\ & + (1/3!)(e^{-\alpha s}s^{7/2} + \cdots) + \ldots]ds. \end{aligned}$$

Integrating (A3):

$$R(+) = (2N\lambda\sigma_P/i)\left\{\frac{1/2(1/2+1)}{\alpha^2}[1 + 1/(2N/\alpha+1)^{5/2} + \cdots]\right.$$

$$\left. - \frac{1/2(1/2+1)(1/2+2)}{2!\alpha^3}[1 + 1/(2N/\alpha+1)^{7/2} + \ldots] + \cdots\right\}.$$

Convergence was accelerated using the Gregory integration formula (HAMMING 1962, p. 138).

For negative mutants, a similar series expansion is obtained, and the total response is given by

$$R = PR(+) - (1-P)R(-).$$