Pleiotropic Models of Polygenic Variation, Stabilizing Selection, and Epistasis

Sergey Gavrilets*,†,[1] and G. de Jong‡

**N. I. Vavilov Institute of General Genetics, 117809 GSP-1, Moscow B-333, Russia, †Laboratoire de Génétique Cellulaire, INRA Centre de Toulouse, Auzeville B. P. 27, F-31326 Castanet Tolosan, France, and ‡Department of Plant Ecology and Evolutionary Biology, University of Utrecht, 3584 CH Utrecht, The Netherlands*

Manuscript received January 2, 1992
Accepted for publication February 11, 1993

ABSTRACT

We show that in polymorphic populations many polygenic traits pleiotropically related to fitness are expected to be under apparent "stabilizing selection" independently of the real selection acting on the population. This occurs, for example, if the genetic system is at a stable polymorphic equilibrium determined by selection and the nonadditive contributions of the loci to the trait value either are absent, or are random and independent of those to fitness. Stabilizing selection is also observed if the polygenic system is at an equilibrium determined by a balance between selection and mutation (or migration) when both additive and nonadditive contributions of the loci to the trait value are random and independent of those to fitness. We also compare different viability models that can maintain genetic variability at many loci with respect to their ability to account for the *strong* stabilizing selection on an additive trait. Let V_m be the genetic variance supplied by mutation (or migration) each generation, V_g be the genotypic variance maintained in the population, and n be the number of the loci influencing fitness. We demonstrate that in mutation (migration)-selection balance models the strength of apparent stabilizing selection is order V_m/V_g. In the overdominant model and in the symmetric viability model the strength of apparent stabilizing selection is approximately $1/(2n)$ that of total selection on the whole phenotype. We show that a selection system that involves pairwise additive by additive epistasis in maintaining variability can lead to a lower genetic load and genetic variance in fitness (approximately $1/(2n)$ times) than an equivalent selection system that involves overdominance. We show that, in the epistatic model, the apparent stabilizing selection on an additive trait can be as strong as the total selection on the whole phenotype.

BOTH in nature and in artificial selection experiments individuals with a phenotype that deviates from the mean prove to have reduced fitness. Two extreme explanations of this observation have been proposed (ROBERTSON 1967). According to the first, this is evidence of "stabilizing" selection working directly on quantitative traits. In keeping with this interpretation, most mathematical models describing the evolution of quantitative traits have included stabilizing selection as a basic part; *i.e.*, fitness takes the form of a quadratic or Gaussian function of the phenotypic value (*e.g.*, LANDE 1975, 1976; GIMELFARB 1986, 1989; NAGYLAKI 1989; HASTINGS and HOM 1990; TURELLI and BARTON 1990; BARTON and TURELLI, 1991). Practical measures of the mode and intensity of natural selection also are based on the beliefs that direct stabilizing selection acts on most quantitative traits and that quantitative traits are decisive in establishing fitness (LANDE and ARNOLD 1983; ARNOLD and WADE 1984a,b; ENDLER 1986; MITCHELL-OLDS and SHAW 1987; SCHLUTER 1988).

This "direct" explanation has, however, some weak points. First, the negative correlation between fitness and the squared deviation of the trait value from the mean does not necessarily imply that an individual has reduced fitness just because its quantitative trait deviates from the "optimum." The difference between statistical and functional relations in the context of measuring natural selection acting on quantitative traits was recently emphasized by WADE and KALISZ (1990). Further, observed relationships between fitness and a quantitative trait may be rather misleading. For example, one may observe "stabilizing" selection on a trait that is neutral by definition, provided there are certain pleiotropic relations between this trait and fitness (ROBERTSON 1956; BARTON 1990; KEIGHTLEY and HILL 1990). Another example is a model in which natural selection moves the mean of a quantitative trait away from the observed optimum (ROBERTSON 1967). The view that quantitative variation can be understood in terms of direct stabilizing selection on the traits also can be criticized using load arguments, ideas about widespread pleiotropy, and the results of selection experiments (ROBERTSON 1967; FALCONER 1989; BARTON 1990; KEIGHTLEY and HILL 1990).

1 Present address: Division of Environmental Studies, University of California, Davis, California 95616.

An alternative extreme view is that the observed variability of quantitative traits is a side effect of polymorphism maintained for some other reasons, and that observed differences in fitness of individuals with different values of a quantitative trait have nothing to do with selection on that trait. ROBERTSON (1956) proposed a model in which genetic variation was maintained by overdominance at each of n loci. The loci also control pleiotropically an additive neutral quantitative trait that will be under apparent stabilizing selection provided the population is at a polymorphic equilibrium. This balancing selection model was reanalyzed by BARTON (1990), who also considered a similar model in which genetic variability was maintained by mutation [see also KEIGHTLEY and HILL (1990)].

In this paper, we consider a similar class of pleiotropic models. First, we show that for any dependence of fitness on genotype, one expects to observe "stabilizing selection" on an additive polygenic trait that is pleiotropically related to fitness. One set of conditions is that the genetic system is at a stable polymorphic equilibrium determined by selection while (1) the nonadditive (*i.e.*, dominant and epistatic) contributions of loci to the trait value are absent, or (2) the nonadditive contributions of loci to the trait value are random and independent of those to fitness. Stabilizing selection also is expected to be observed if the polygenic system is at an equilibrium determined by a balance between selection and mutation (or migration), and both additive and nonadditive contributions of the loci to the trait value are random and independent of those to fitness. On the other hand, one expects to observe directional selection on the trait if the genetic system is at a stable polymorphic equilibrium and the contributions of the loci to the trait value are related to the contributions of the loci to fitness. Apparent disruptive selection may arise, if, for example, conditions (1) and/or (2) are satisfied, but the population is at an unstable polymorphic equilibrium. The above conclusions are based on an approximation of the apparent fitness function. This approximation assumes that the number of loci underlying the trait is large and that linkage disequilibrium can be neglected, while the "real" fitness function and the distribution of the trait can be arbitrary.

In natural populations, both abundant polygenic variation and strong stabilizing selection are found. This forms rather a conundrum, as stabilizing selection should rapidly eliminate that variation. As a consequence, the analysis of possible mechanisms of the maintenance of polygenic variability under stabilizing selection was stimulated (*e.g.*, BULMER 1972, 1973; LANDE 1975; GILLESPIE and TURELLI 1989; GIMELFARB 1986, 1989; HASTINGS and HOM 1990; NAGYLAKI 1989; ZHIVOTOVSKY and GAVRILETS 1992). However, abundant data on reduced fitness of individuals with extreme phenotypes (*e.g.*, ENDLER 1986) cannot be interpreted exclusively as evidence of ecological stabilizing selection. The problem: "How can quantitative variability be maintained under strong stabilizing selection?" has been intensively analyzed during the last 15 years, but if we accept that stabilizing selection is only or predominantly "apparent," then this problem will not be of much importance. The questions that become important now are "How can polygenic variability be maintained under selection, and how can this lead to strong apparent stabilizing selection?"

BARTON (1990) formulated this problem and made an attempt to solve it. He considered two models with additive fitness. In the first, polygenic variability was maintained by recurrent mutation to deleterious alleles. In the second, proposed by ROBERTSON (1956), polygenic variability was maintained by overdominance at each locus. In both models the loci also control a neutral additive quantitative trait which is under apparent stabilizing selection. The general conclusion of his analysis was, however, that "neither mutation/selection balance nor balancing selection alone can easily account for both the high heritability and the strong stabilizing selection which are commonly observed" (p. 779). The mutation-selection model also was analyzed by KEIGHTLEY and HILL (1990) using numerical simulations. The conclusion of these authors was that the model "can be made to fit the observations, but it is easy to construct examples where it does not, particularly in which predicted stabilizing selection is too weak" (p. 99).

In the second part of this paper, we apply our results on apparent fitness functions to a general mutation (or migration)-selection balance model and to three specific viability models. These are: the overdominant model (ROBERTSON 1956; BARTON 1990), a symmetric model (KARLIN and AVNI 1982) and a model with additive by additive epistatic interactions between pairs of loci (GAVRILETS 1993; ZHIVOTOVSKY and GAVRILETS 1992). For all these models, the existence of stable multilocus polymorphism has been proven; therefore, they represent a possible solution to the first part of the problem. We shall consider the second part of the problem, *i.e.*, whether it is possible to generate strong apparent stabilizing selection on quantitative traits in each of these four models. Our results show that epistasis may solve the problem stated by BARTON. The main advantage of the epistatic models is the possibility of the maintenance of high levels of polymorphism while the genetic load and genetic variance in fitness associated with this polymorphism remain very low.

THE FORM OF APPARENT SELECTION ON AN ADDITIVE POLYGENIC TRAIT

To characterize apparent selection on a polygenic trait, ROBERTSON (1956) considered the fitness con-

ditioned on the trait value, while BARTON (1990) and KEIGHTLEY and HILL (1990) calculated the covariance of fitness with the squared value of the trait. In this section we shall use ROBERTSON's approach as more informative.

Consider a diploid monoecious randomly mating population with distinct nonoverlapping generations. We assume viability selection; let $\bar{w}$ be the mean fitness of the population. Let there be n loci with two alleles each: A_i and a_i ($i = 1,\ldots,n$), and let p_i and $q_i = 1 - p_i$ be the frequencies of allele A_i and a_i respectively. Throughout the paper we shall use the indicator variables l_i (l_i') which equal to 1, if the allele at the ith locus of the paternal (maternal) gamete is A_i, and 0, if this allele is a_i. Assume that linkage disequilibrium can be neglected. This assumption is reasonable if, for example, selection is much weaker than recombination.

Each individual is characterized by the value of an additive quantitative trait, X. Let $f(x)$ be the distribution of X in the population with mean $E\{x\} = \bar{x}$ and variance $\mathrm{var}\{x\} = P$. Denote by $(\delta\bar{x})_i$ and $(\delta P)_i$ the average effects of allele A_i on the mean value $\bar{x}$ and phenotypic variance P. Thus,

$$\begin{aligned}(\delta\bar{x})_i &= E\{x \mid l_i = 1\} - \bar{x},\\ (\delta P)_i &= P - \mathrm{var}\{x \mid l_i = 1\},\end{aligned} \tag{1a}$$

where we use the notions of the conditional mean and conditional variance. Denote by $(\delta\bar{x})_{ii}$ and $(\delta P)_{ii}$ the corresponding effects of the one-locus marginal genotype A_iA_i

$$\begin{aligned}(\delta\bar{x})_{ii} &= E\{x \mid l_i = l_i' = 1\} - \bar{x},\\ (\delta P)_{ii} &= P - \mathrm{var}\{x \mid l_i = l_i' = 1\}.\end{aligned} \tag{1b}$$

Without loss of generality, we shall assume that $\bar{x} = 0$. Note that if there is no dominance, $(\delta\bar{x})_{ii} = 2(\delta\bar{x})_i$, $(\delta P)_{ii} = 2(\delta P)_i$. We shall assume that the trait is controlled by a large number of loci with small effects, so that all $(\delta\bar{x})_i$ and $(\delta\bar{x})_{ii}$ values are small and all $(\delta P)_i$ and $(\delta P)_{ii}$ values are second order in $(\delta\bar{x})$-values.

Now let us assume that one has measurements of the fitnesses of different individuals with the same phenotype $X = x$. According to general practice [*e.g.*, SCHLUTER (1988)] the corresponding mean fitness, $E\{w \mid X = x\}$, will be considered as the "real" fitness of phenotype x, the deviations of the fitnesses from $E\{w \mid X = x\}$ as a random "noise," and the function $w(x) \equiv E\{w \mid X = x\}$ as the phenotypic fitness function. In the Appendix we show that for arbitrary "real" fitness function, w, and for arbitrary phenotypic distribution, $f(x)$, the "apparent" fitness function, $w(x)$, can be approximated as

$$w(x) = \bar{w} + \sum_i \frac{\partial\bar{w}}{\partial p_i}\left[-(\delta\bar{x})_i p_i \frac{f'(x)}{f(x)} + \frac{1}{2}\left((\delta\bar{x})_i^2 - (\delta P)_i\right)p_i\frac{f''(x)}{f(x)} + \cdots\right] \tag{2a}$$

$$+ \sum_i \frac{1}{2}\frac{\partial^2\bar{w}}{\partial p_i^2}\left[-\left((\delta\bar{x})_{ii} - 2(\delta\bar{x})_i\right)p_i^2\,\frac{f'(x)}{f(x)}\right] \tag{2b}$$

$$+ \frac{1}{2}\left((\delta\bar{x})_{ii}^2 - 4(\delta\bar{x})_i^2 - (\delta P)_{ii} + 2(\delta P)_i\right)p_i^2\frac{f''(x)}{f(x)}\Bigg] \tag{2c}$$

$$+ \sum_{i,j}\frac{1}{2}\frac{\partial^2\bar{w}}{\partial p_i\partial p_j}(\delta\bar{x})_i(\delta\bar{x})_j p_i p_j \frac{f''(x)}{f(x)} + \cdots. \tag{2d}$$

Here $f'(x) = df(x)/dx$, $f''(x) = d^2f(x)/dx^2$, and the partial derivatives of the mean fitness are evaluated at the point $(p_1,\ldots,p_n)$; the error is third order in $(\delta\bar{x})$-values.

First note that both terms in (2a) and (2b) that are first order in $(\delta\bar{x})$ are proportional to $f'(x)/f(x)$. Since the covariances $\mathrm{cov}((f'(x)/f(x)), x) = -1$ and $\mathrm{cov}((f'(x)/f(x)), x^2) = 0$, we can interpret the corresponding sums in (2) as components of directional selection in the apparent fitness function (LANDE and ARNOLD 1983). If, for example, the phenotypic distribution is normal,

$$\frac{f'(x)}{f(x)} = -\frac{x - \bar{x}}{P},$$

and the first order terms give linear dependence of fitness on phenotype. The terms in (2a, 2c, 2d) that are second order are proportional to $(f'')(x))/(f(x))$. Since the covariances $\mathrm{cov}((f''(x))/(f(x)), x) = 0$ and $\mathrm{cov}((f''(x))/(f(x)), x^2) = 2$, we can interpret the corresponding sums in (2) as components of stabilizing (disruptive) selection in the apparent fitness function provided they are negative (positive) (LANDE and ARNOLD, 1983). For example, if $f(x)$ is normal, then

$$\frac{f''(x)}{f(x)} = -\frac{1}{P} + \frac{(x - \bar{x})^2}{P^2}.$$

Thus, the second order terms give a quadratic dependence of $w(x)$ on x. In general, the directional component of apparent selection (which is first order) dominates the components of stabilizing and disruptive selection (which are second order). Below we consider how the relative order of these components depends on the genetic structure of the population.

Additive trait without dominance: Assume that the quantitative trait is additive both between and within loci. Such a trait can be described as

$$x = \Sigma\,\alpha_i(l_i + l_i' - 1) + e, \tag{3}$$

where α_i is the additive contribution of the i-th locus, and e stands for random microenvironmental deviation. In this case, using the expressions for the average effects (1) given in the Appendix, the apparent fitness function (2) can be rewritten as

$$w(x) = \bar{w} + \sum_i \frac{\partial \bar{w}}{\partial p_i}\left[-\alpha_i p_i q_i \frac{f'(x)}{f(x)} + \frac{1}{2}\alpha_i^2 p_i q_i (q_i - p_i)\frac{f''(x)}{f(x)} + \cdots\right] \quad (4a)$$

$$+ \sum_{i,j} \frac{1}{2}\frac{\partial^2 \bar{w}}{\partial p_i \partial p_j}\alpha_i p_i q_i \alpha_j p_j q_j \frac{f''(x)}{f(x)} + \cdots. \quad (4b)$$

Let us assume now that the genetic system is at a stable polymorphic equilibrium determined by selection. If linkage disequilibrium can be neglected, the dynamics of the allele frequencies under selection are approximated by the general relation $\Delta p_i = (p_i q_i/2)(\partial \ln \bar{w}/\partial p_i)$ [*e.g.*, WRIGHT (1935) and BARTON and TURELLI (1987)]. At a polymorphic equilibrium, $\partial \bar{w}/\partial p_i = 0$, so that the apparent fitness function becomes

$$w(x) = \bar{w} + \left(\sum_{i,j} \frac{1}{2}\frac{\partial^2 \bar{w}}{\partial p_i \partial p_j}\alpha_i p_i q_i \alpha_j p_j q_j\right)\frac{f''(x)}{f(x)} + \cdots. \quad (5)$$

At a stable equilibrium, the matrix of second order derivatives $\partial^2 \bar{w}/\partial p_i \partial p_j$ is negative definite, therefore the term in brackets is negative. We observe "pure" stabilizing selection on an additive trait in an equilibrium population, and for a normally distributed trait, quadratic stabilizing selection. Note that if the population is at an unstable equilibrium where the matrix of second order derivatives $\partial^2 \bar{w}/\partial p_i \partial p_j$ is positive definite, then we would observe "pure" disruptive selection.

In the situation just considered, the first order term in (4a) disappeared due to the assumption that the genetic system was at a stable polymorphic equilibrium determined by selection. Another possibility is to assume that the contributions of the loci to the trait, α_i, are drawn from a probability distribution which is independent of fitness and that the α_i have a mean value of zero. A natural interpretation of such a situation is that the trait is "neutral." In this case the terms in (4a) proportional to $f'(x)/f(x)$ will not dominate the terms proportional to $f''(x)/f(x)$. We expect to observe "stabilizing" (or "disruptive") selection on the trait in a population that is not necessarily at an equilibrium determined by selection, but, for example, at an equilibrium determined by a balance of mutation (or migration) and selection. Let the polygenic system be at a polymorphic equilibrium where a low level of variability is maintained by mutation or migration. Assume, without loss of generality, that it is allele A_i that has a low frequency, say order $\epsilon \ll 1$. In this case all $p_i q_i$ values will be order ϵ. Therefore the sum in (4d), which is now second order in ϵ, can be neglected, and the apparent fitness function is described by the terms in (4a) which are first order in ϵ:

$$w(x) = \bar{w} + \sum_i \frac{\partial \bar{w}}{\partial p_i}\left[-\alpha_i p_i q_i \frac{f'(x)}{f(x)} + \frac{1}{2}\alpha_i^2 p_i q_i (q_i - p_i)\frac{f''(x)}{f(x)} + \cdots\right]. \quad (6)$$

At a polymorphic equilibrium where selection tends to eliminate the allele having low frequency (*i.e.* where $\partial \bar{w}/\partial p_i < 0$ and $p_i \ll q_i$), $(\partial \bar{w}/\partial p_i)(q_i - p_i) < 0$, and, hence, the terms is (6) that are proportional to $f''(x)/f(x)$ correspond to "stabilizing" selection. Therefore, we expect to observe "stabilizing" selection on a "neutral" trait in a population that is at an equilibrium determined by a balance of mutation (or migration) and selection.

THE STRENGTH OF APPARENT SELECTION ON AN ADDITIVE POLYGENIC TRAIT

The results from the preceding section show that one can expect to observe stabilizing selection in many situations. However, nothing was said about the strength of this apparent selection. In the following sections, we shall use expression (4) for calculating strength of apparent stabilizing selection in different models. The component of relative apparent fitness $w(x)/\bar{w}$ that stands for stabilizing selection can be rewritten as $-sP^2\,(f''(x)/f(x))$, or, if we assume that the phenotypic distribution $f(x)$ is approximately normal, as $-s(x - \bar{x})^2$, where parameter $s > 0$ depends on the model under consideration. The parameter s is a practical measure of the intensity of stabilizing selection (LANDE and ARNOLD 1983). Alternative dimensionless measures are the genetic load, L_{app}, and the genetic variance in the relative fitness, $\mathrm{var}(w/\bar{w})_{app}$, associated with apparent stabilizing selection. In the case of a normal distribution of x, these values are approximated by

$$L_{app} = sP,\ \mathrm{var}(w/\bar{w})_{app} = 2\,s^2P^2. \quad (7)$$

In the following sections, we shall calculate these characteristics of the apparent fitness function in different models.

Mutation (migration)-selection balance models

Maintenance of variability: Let us assume that the polygenic system is at a polymorphic equilibrium where a low level of variability is maintained by mutation or migration. In this case, the dynamics of the allele frequencies near the equilibrium are approximated by

$$\Delta p_i = \frac{p_i q_i}{2}\frac{\partial \ln \bar{w}}{\partial p_i} + \Delta p_i^m, \quad (8)$$

where Δp_i^m is the change in p_i caused by mutation or migration, $i = 1, \ldots, n$. If variability is maintained by recurrent mutation that occurs at an equal rate in each direction, $\Delta p_i^m = \mu_i(p_i - q_i)$, where $\mu_i \ll 1$ is the mutation rate at the ith locus (*e.g.*, BARTON 1986). If variability is maintained by gene flow from another "source" population, $\Delta p_i^m = \mu(p_i - p_{i,0})$, where $\mu \ll 1$ is the migration rate and $p_{i,0}$ is the frequency of allele A_i in the "source" population. Let us assume without loss of generality that it is allele A_i that has a low frequency. Multiplying (8) by $2\alpha_i^2$ and summing over all loci controlling the trait, we find that at an equilibrium (with $\Delta p_i = 0$) the genotypic variance is approximately

$$V_g = V_m/\hat{L}, \tag{9}$$

where $V_m = \sum_x \mu_i 2\alpha_i^2(q_i - p_i) \cong \sum_x \mu_i 2\alpha_i^2$ for the mutation-selection balance case, and $V_m = \sum_x \mu 2\alpha_i^2(p_{i,0} - p_i) \cong \sum_x \mu 2\alpha_i^2 p_{i,0}$ for the migration-selection balance case. Here the sum $\sum_x$ is over n_x loci influencing both fitness and trait ($n_x \leqslant n$). In both cases, V_m is the new genetic variance supplied in the population each generation. Parameter $\hat{L}$ is the mean value of $L_i = -(\frac{1}{2})\partial \ln \bar{w}/\partial p_i$ weighted according to the genotypic variance contributed by that gene, $\hat{L} = \sum_x L_i 2\alpha_i^2 p_i q_i / \sum_x 2\alpha_i^2 p_i q_i$. Parameter $\hat{L}$ can be interpreted as the mean intensity of selection against the rare allele having pleiotropic effect on the trait. Expression (9) generalizes the formula for the genetic variance in an additive trait maintained under direct gaussian (and quadratic) selection by mutation (BULMER 1972) for the case of arbitrary fitness functions and for the case of migration-selection balance. The question of whether it is possible to maintain high levels of genetic variability by mutation remains controversial (BARTON 1990; KEIGHTLEY and HILL 1990).

Apparent stabilizing selection: For equilibria with a low level of variability, the apparent fitness function is approximated by expression (6). In this case the "intensity" of stabilizing selection is

$$s_{\text{balance}} = \sum_i \frac{\partial \bar{w}}{\partial p_i} \frac{1}{2} \alpha_i^2 p_i q_i (q_i - p_i)/\bar{w} \cong \frac{1}{2} \hat{L}\, V_g/P^2. \tag{10}$$

Assuming approximate normality of the phenotypic distribution, the apparent load and the genetic variance in relative apparent fitness (7) become

$$L_{\text{app}} = \frac{1}{2} \hat{L}(V_g/P),\ \text{var}(w/\bar{w})_{\text{app}} = \frac{1}{2} (\hat{L})^2 (V_g/P)^2. \tag{11}$$

Note that V_g/P is the heritability. Thus, the apparent genetic load in this model is maximum half the intensity of selection against the rare allele having pleiotropic effect on the trait, $\hat{L}$. As follows from expression (9), the parameter $\hat{L}$ equals the ratio of the new genetic variance supplied to the population by mutation (migration) in one generation to that one maintained in the population, $\hat{L} = V_m/V_g$. The typical experimental estimate of V_m/V_g is about 10^{-3} [*e.g.*, LYNCH (1988)]. This makes it impossible to explain any strong stabilizing selection observed in terms of this mutation-selection balance model. For more discussion of the level of genetic variability and intensity of apparent stabilizing selection in similar mutation-selection balance models, see BARTON (1990), KEIGHTLEY and HILL (1990), and KONDRASHOV and TURELLI (1992).

Overdominant viability model

Let viability be characterized by dominance "within" loci. Then fitness can be described as

$$w = m + \sum^{n} [a_i(l_i + l_i') + 2b_i l_i l_i']. \tag{12}$$

Maintenance of variability: If $a_i > 0$, $a_i + 2b_i < 0$ (*i.e.*, if there is overdominance at each locus), then an equilibrium with allele frequencies $p_i^* = -a_i/(2b_i)$ is globally stable. At this equilibrium, the mean fitness of the population, the segregation load associated with polymorphism, and the genetic variance in relative fitness are

$$\bar{w} = \mu + \sum a_i p_i^*, \tag{13a}$$

$$L = (\bar{w}/w_{\max})\sum - 2p_i^* q_i^* b_i/\bar{w} \cong \sum L_i = n\bar{L}, \tag{13b}$$

$$\text{var}(w/\bar{w}) = \sum 4(p_i^* q_i^*)^2 b_i^2/\bar{w}^2 \cong \sum L_i^2 = n\bar{L^2}, \tag{13c}$$

respectively. Here $L_i = -2b_i p_i^* q_i^*/\bar{w}$ is the segregation load at the ith locus (provided that the overall load is not large, *i.e.*, $\bar{w}/w_{\max} \cong 1$); $\bar{L}$ and $\bar{L^2}$ are the arithmetic means of L_i and L_i^2.

Apparent stabilizing selection: Assume that the genetic system is at a polymorphic equilibrium. In this case, $\partial \bar{w}/\partial p_i = 0$, $\partial^2 \bar{w}/\partial p_i \partial p_j = 0$, $i \neq j$, $\partial^2 \bar{w}/\partial p_i^2 = 4b_i$, and "intensity" of stabilizing selection s can for this overdominant model be represented as

$$s_{\text{over}} = \sum_x (L_i/2) 2\alpha_i^2 p_i q_i / P^2 = -b(b\hat{p}q/\bar{w})V_g/P^2 = (\hat{L}/2)V_g/P^2, \tag{14}$$

where $\hat{L}$ is the weighted mean segregation load at a single locus, *i.e.*, the mean value of $L_i = -2b_i p_i q_i/\bar{w}$, weighted according to the variance contributed by that gene, $\hat{L} = \sum_x L_i 2\alpha_i^2 p_i q_i / \sum_x 2\alpha_i^2 p_i q_i$. Again the sum $\sum_x$ is over the n_x loci influencing both fitness and trait ($n_x \leqslant n$). In this case, expressions (6) become

$$L_{\text{app}} = \frac{1}{2} \hat{L}(V_g/P),\ \text{var}(w/\bar{w})_{\text{app}} = \frac{1}{2} (\hat{L})^2 (V_g/P)^2, \tag{15}$$

The expression for L_{app} can also be derived from ROBERTSON's (1956) estimates (see also BARTON

(1990) for an alternative derivation). Obviously, apparent load and variance in fitness depend on the real load and variance in fitness. Measures that depend only on the structure of the pleiotropic model are the ratios

$$L_{app}/L = \frac{1}{2n}(\hat{L}/\bar{L})(V_g/P),$$

$$\mathrm{var}(w/\bar{w})_{app}/\mathrm{var}(w/\bar{w}) = \frac{1}{2n}[(\hat{L})^2/\overline{L^2}](V_g/P)^2, \quad (16)$$

where $\bar{L}$ and $\overline{L^2}$ are the arithmetic means of L_i and L_i^2. Note that these ratios can be interpreted as a part of the genetic load (or genetic variance in relative fitness) "explained" by apparent stabilizing selection. Right-hand sides of (16) behave approximately as $(1/2n)$. It is widely supposed that the number of the loci influencing fitness is very large; therefore, it is impossible to explain any strong stabilizing selection observed in terms of this overdominant model. BARTON (1990) made a similar conclusion.

Symmetric viability models

KARLIN and AVNI (1981) analyzed a class of symmetric viability models where fitness, w, depended on the proportion of heterozygous loci, h. In our notation, this model can be described as

$$w = w(h),\ h = \frac{1}{n}\sum(l_i + l_i' - 2l_il_i'). \quad (17)$$

Maintenance of variability: In this model the central polymorphic equilibrium (exhibiting equal frequencies for each gametotype) always exists and can be stable provided function $w(h)$ and the recombination rates satisfy certain conditions (KARLIN and AVNI 1981). The mean value and the variance of h in the population are

$$\bar{h} = \frac{1}{n}\sum 2p_iq_i,\ \mathrm{var}(h) = \left(\frac{1}{n}\right)^2 \sum 2p_iq_i(1 - 2p_iq_i), \quad (18)$$

where the latter expression assumes linkage equilibrium. At the central equilibrium where all allele frequencies equal one half, $\bar{h} = 1/2$, $\mathrm{var}(h) = 1/(4n)$. The fact that $\mathrm{var}(h)$ is small if the number of loci is large suggests that we can use a linear approximation $w(h) \cong w(\bar{h}) + (dw/dh)(h - \bar{h})$. This gives the mean fitness, the genetic load, and the genetic variance in relative fitness as

$$\bar{w} \cong w(\bar{h}),\ L \cong \frac{1}{2}|d\ \ln\ \bar{w}/dh|,$$
$$\mathrm{var}(\bar{w}/w) \cong \frac{1}{4n}(d\ \ln\ \bar{w}/dh)^2. \quad (19)$$

Apparent stabilizing selection: To calculate the apparent fitness function we need to know the values of the second order partial derivatives $\partial^2\bar{w}/\partial p_i\partial p_j$ and $\partial^2\bar{w}/\partial p_i^2$ at equilibrium. Approximating these partial derivatives to the leading order in $1/n$, we get the intensity of stabilizing selection

$$s_{sym} = \frac{1}{4}\left|\frac{1}{n}\frac{1}{\bar{w}}\frac{dw(h)}{dh}\right|(V_g/P^2). \quad (20)$$

Accordingly, provided the phenotypic distribution is normal, the apparent genetic load and genetic variance in fitness are approximated as

$$L_{app} = \frac{1}{4}\left|\frac{1}{n}\frac{1}{\bar{w}}\frac{dw(h)}{dh}\right|(V_g/P),$$
$$\mathrm{var}(w/\bar{w})_{app} = \frac{1}{8}\left(\frac{1}{n}\frac{1}{\bar{w}}\frac{dw(h)}{dh}\right)^2(V_g/P)^2. \quad (21)$$

Calculating the ratios of the apparent to the real values, we find that

$$L_{app}/L = \frac{1}{2n}(V_g/P),$$

$$\mathrm{var}(w/\bar{w})_{app}/\mathrm{var}(w/\bar{w}) = \frac{1}{2n}(V_g/P)^2. \quad (22)$$

This shows that in the symmetric viability model, as in the overdominant model, the strength of apparent stabilizing selection on an additive trait is approximately $1/(2n)$ of the strength of the overall selection on the phenotype. The symmetric model also cannot explain observed strong stabilizing selection.

EPISTATIC VIABILITY MODELS

In all models just considered, we observe "stabilizing selection" on an additive trait. However, none of them can account for high heritability and strong stabilizing selection occurring simultaneously. This suggests that we need a more complex model. A possible candidate for the analysis is a viability model that accounts for additive, dominant, and epistatic pairwise additive by additive effects (GAVRILETS 1993; ZHIVOTOVSKY and GAVRILETS 1992).

Model with equivalent loci

Let us consider the completely symmetric case where fitness is described by

$$w = m + \sum_i^n \left[a(l_i + l_i') + 2bl_il_i'\right] + \sum_{i\neq j}^n c(l_i + l_i')(l_j + l_j'). \quad (23)$$

Maintenance of variation: If a multilocus system is under viability selection (23), and selection is much weaker than recombination, there exists a single globally stable equilibrium with allele frequencies

$$p_i = p^* = -\frac{a}{2(b + 2c(n-1))}, \quad (24)$$

provided $a > 0$, $b < -a/2n$, $b/2 < c < -(a + 2b)/(4(n-1))$ (ZHIVOTOVSKY and GAVRILETS 1992). At this equilibrium the mean fitness and the genetic variance in fitness are

$$\bar{w} = \mu + \text{nap}^*,$$

$$\text{var}(w) = 4n(p^*q^*)^2[b^2 + 2c^2(n-1)]. \quad (25)$$

We do not present an expression for the genetic load, which is much more complex than expressions (25) and takes different forms for different configurations of the parameters.

It is interesting to compare characteristics of this polymorphic equilibrium with those of the overdominant model. For this reason, let us assume that the parameters of the overdominant model (for which we shall use the subscript "over") and the parameters of the epistatic model are connected by the following equations:

$$\mu_{\text{over}} = \mu, \; a_{\text{over}} = a, \; b_{\text{over}} = b + 2c(n-1). \quad (26)$$

Under these conditions the equilibrium allele frequencies and the equilibrium mean fitness are equal in both models. Let us introduce the parameter ω measuring the "strength" of epistasis:

$$b = (1-\omega)b_{\text{over}}, \; 2c(n-1) = \omega b_{\text{over}}. \quad (27)$$

If $\omega = 0$, we have exactly the overdominant model with $c = 0$, and if $\omega \neq 0$, then the sign of c is opposite to that of ω (since $b_{\text{over}} < 0$). Using (25, 27) one can show that

$$\text{var}(w)/\text{var}(w)_{\text{over}} = (1-\omega)^2 + \omega^2/(2n-2). \quad (28)$$

This is a quadratic in ω; it is larger than one if $\omega < 0$ (*i.e.*, if $c > 0$) with a minimum at $\omega = 1/(2n-2)$. The condition for existence and stability of the polymorphic equilibrium $b < 2c$ imposes a restriction on the ω values: $\omega < 1 - 1/n$. As ω tends to $1 - 1/n$, the ratio of the variances in fitness tends to $1/(2n-2)$ provided that the number of loci, n, is large. This means that in the epistatic model, the genetic variance in fitness at equilibrium can be approximately $2n$ times lower than in the corresponding overdominant model. In other words, epistasis can maintain the same level of polymorphism as overdominance, while the overall genetic variance in fitness (genetic load) associated with this polymorphism equals approximately half the genetic variance in fitness (genetic load) for a single locus in the corresponding overdominant model. It is interesting that this conclusion is valid only for negative values of the epistatic parameter c.

Apparent phenotypic selection: Let the trait be additive, without any microenvironmental deviation:

$$x = \mu + \sum^{n_x} \alpha(l_i + l_i').$$

where α is the locus contribution to the trait. The number of loci affecting both fitness and the trait, n_x, is not necessary equal to the number of loci affecting fitness, n. At a polymorphic equilibrium, the mean and variance for the quantitative trait are $\bar{x} = \mu + 2n_x\alpha p^*$, and $V_g = 2n_x\alpha^2 p^*q^*$. The mean fitness conditioned on the trait value can now be calculated exactly (see the Appendix):

$$w(x) = \bar{w} - \frac{2n_x}{2n_x - 1} pq[b + 2(n_x - 1)c] \quad (29a)$$

$$+ t\sqrt{2n_x pq}\,[a + 2bp + 4c(n-1)p] \quad (29b)$$

$$+ t\frac{2n_x}{2n_x - 1}\sqrt{pq/(2n_x)}(p - q) \quad (29c)$$

$$\cdot [b + 2(n_x - 1)c]$$

$$+ t^2\frac{2n_x}{2n_x - 1} pq[b + 2(n_x - 1)c], \quad (29d)$$

where $t = (x - \bar{x})/\sqrt{V_g}$ is the normalized deviation of x from the mean value $\bar{x}$ of the trait. Expression (29) is exactly a quadratic in x. The term (29b) shall dominate if $a + 2bp + 4c(n-1)p \neq 0$. This means that when the genetic system is not at equilibrium ($p_i \neq p^*$), we will observe "directional" selection. If the genetic system is at equilibrium, then term (29b) disappears, and we have apparent "stabilizing" selection on x. In this case the normalized deviation of the mean value of the trait from the observed "optimum" (at which $w(x)$ has its maximum) is $(\bar{x} - x_{\text{opt}})/\sqrt{V_g} = (p - q)/\sqrt{8n_x pq}$. This tends to zero if n_x is large.

Note that the "intensity" of stabilizing selection in this model can be approximated by

$$s_{\text{epi}} = = -[b + 2c(n_x - 1)](pq/\bar{w})/V_g, \quad (30)$$

and that the apparent genetic load and the "apparent" genetic variance in fitness are approximately

$$L_{\text{app}} = -[b + 2(n_x - 1)c]p^*q^*/\bar{w},$$

$$\text{var}(w/\bar{w})_{\text{app}} = 2[(b + 2(n_x - 1)c)p^*q^*/\bar{w}]^2. \quad (31)$$

All these values increase with the number of the pleiotropic loci, n_x. From (26) and (31) it follows that

$$\frac{\text{var}(w)_{\text{app}}}{\text{var}(w)} = \frac{[1 + (n_x - n)\omega/(n-1))]^2}{n[(1-\omega)^2 + \omega^2/(2n-2)]}, \quad (32)$$

where we again use the parameter ω measuring the "strength" of epistasis (see Eq. 27). As ω tends to $1 - 1/n$, this ratio tends to n_x/n provided the number of loci, n, is large. This means that the genetic variance in the apparent fitness can be as large as the genetic variance in "real" fitness (if $n_x \cong n$). Again this conclu-

sion is valid only for negative values of the epistasis parameter c.

Asymmetric model

The completely symmetric model assumes in implicit form that the effects of alleles on fitness and on the quantitative trait are strongly related; resulting, in particular, in very strong apparent selection. In this section we consider a nonsymmetric model without such a restriction. Let fitness be represented by

$$w_{ll'} = m + \sum_{i}^{n} [a_i(l_i + l_i') + 2b_i l_i l_i'] + \sum_{i \neq j}^{n} c_{ij}(l_i + l_i')(l_i + l_i'). \quad (33)$$

The maintenance of variability: Let us define the $(n \times n)$-matrix $\mathbf{S}$ with the components $(\mathbf{S})_{ij} = S_{ij}$, $i, j = 1, \ldots, n$, where $S_{ij} = 4c_{ij}$ $(i \neq j)$, $S_{ii} = 2b_i$, and the vector $\mathbf{A}$ with components $(\mathbf{A})_i = a_i + \sum_j S_{ij}$, $i = 1, \ldots, n$. The single equilibrium with n polymorphic allele frequencies $p_i^* = (-\mathbf{S}^{-1}\mathbf{A})_i$ exists and is stable if for all components of the vector $-\mathbf{S}^{-1}\mathbf{A}$, $0 < (-\mathbf{S}^{-1}\mathbf{A})_i < 1$, and the matrix $\mathbf{S}$ is negative definite (ZHIVOTOVSKY and GAVRILETS 1992).

Apparent phenotypic selection on an additive trait: In this model, $(\delta\bar{x})_i = \alpha_i q_i$, $(\delta P)_i = \alpha_i^2 p_i q_i$, $\partial^2\bar{w}/\partial p_i \partial p_j = 8c_{ij}$, $i \neq j$, $\partial^2\bar{w}/\partial p_i^2 = 4b_i$, and the intensity of stabilizing selection becomes

$$s_{\text{ep}} = -\left[\sum_i 2b_i\alpha_i^2 p_i^2 q_i^2 + \sum_{i \neq j} 4c_{ij}\alpha_i p_i q_i \alpha_j p_j q_j\right] / (\bar{w}P^2). \quad (34)$$

Let the values of the contributions of the loci to the trait, α_i, be drawn from a probability distribution with mean $\bar{\alpha}$ and the variance σ_α^2. The expectation of the terms in the squared bracket in (34) approximately becomes

$$\bar{\alpha}^2\left[\sum_i 2b_i p_i^2 q_i^2 + \sum_{i \neq j} 4c_{ij} p_i q_i p_i q_j\right] + \sigma_\alpha^2 \sum_i 2b_i p_i^2 q_i^2. \quad (35)$$

Let us introduce the weighted means

$$\hat{b} = \frac{\sum(b_i p_i q_i) p_i q_i}{\sum p_i q_i (\bar{pq})}, \quad \hat{c} = \frac{\sum(c_{ij} p_j q_j) p_i q_i}{(n-1)\sum p_i q_i (\bar{pq})},$$

where $(\bar{pq}) = \sum p_i q_i / n$ is the mean of the $p_i q_i$ values. Note that if all allele frequencies are equal, $\hat{b}$ and $\hat{c}$ are exactly the arithmetic means of b_i and c_{ij}. Using these weighted means, we can rewrite expression (35) as

$$[\hat{b} + 2(n-1)\hat{c}](\bar{pq})(\sum 2\bar{\alpha}^2 p_i q_i) + \hat{b}(\bar{pq})(\sum 2\sigma_\alpha^2 p_i q_i). \quad (36)$$

The last expression shows that in this epistatic model the value of parameter s_{ep} crucially depends on the relative order of the mean and the variance of the contributions of the loci to the trait.

Let us first consider the situation when $\bar{\alpha}^2 \ll \sigma_\alpha^2$. In this case, the sign of the effects of the loci on the trait is random with respect to their effects on fitness. The "independence" of the allele effects of these two types can naturally be interpreted as "neutrality" of the trait. If $\bar{\alpha}^2 \ll \sigma_\alpha^2$, the parameter s_{ep} can be approximated by

$$s_{\text{epi}}^{\text{neut}} \cong -\left[\hat{b}(\bar{pq})/\bar{w}\right] V_g/P^2 \quad (37)$$

and does not depend on the epistatic terms. If $\bar{\alpha}^2 \gg \sigma_\alpha^2$, the sign of the effects of the loci on the trait is fixed with respect to their effects on fitness. In this case, the trait can be naturally considered as "nonneutral" or "adaptive." If $\bar{\alpha}^2 \gg \sigma_\alpha^2$, the strength of apparent selection is characterized by

$$s_{\text{epi}}^{\text{adap}} \cong -[\hat{b} + 2(n-1)\hat{c}](\bar{pq})/\bar{w}\, V_g/P^2. \quad (38)$$

Comparison of (37) with (14) suggests that for "neutral" traits, apparent selection is approximately $\frac{1}{2}n$ times weaker than real selection, just as in the overdominant model. In contrast, for "adaptive traits, comparison of (38) with expression (30) suggests that the apparent selection can be as strong as real selection. The relationships between apparent and real selection on traits with different degrees of adaptivity are investigated numerically in the next section.

NUMERICAL RESULTS

In order to get a clearer view of the relationships between trait values, fitness, and apparent selection, we computed numerical examples for the additive model, for a symmetric model, and for the model with additive by additive epistatic interactions between pairs of loci. For each we used $n = 20$ loci. One of the assumptions underlying our analysis was that linkage disequilibrium could be neglected. In additive models, polymorphic equilibria are always in linkage equilibrium. In symmetric models this is true with respect to the central equilibria (where all allele frequencies equal one half), while in models with additive by additive epistatic interactions between pairs of loci, linkage disequilibrium can be neglected if selection is weak relative to recombination.

In additive models fitness was determined by expression (12) where the contributions of loci to fitness a_i and b_i were drawn independently from normal distributions with means $\bar{a}$, $\bar{b}$, and variances σ_a^2 and σ_b^2. In the symmetric model, fitness was determined as

$$w = w(h) = (h + \zeta)^\xi,$$

where h is the proportion of heterozygous loci, and ξ and ζ are parameters [KARLIN and AVNI (1981),

expression 6.2b]. In the models with additive by additive epistatic interaction between pairs of loci, fitness was determined by expression (33) where the contributions of loci to fitness a_i, b_i, and c_{ij} were drawn independently from normal distributions with means $\bar{a}$, $\bar{b}$, $\bar{c}$ and variances σ_a^2, σ_b^2, and σ_c^2, respectively. Mean values $\bar{a}$, $\bar{b}$, and $\bar{c}$ were chosen such that the corresponding epistatic model with $a_i = \bar{a}$, $b_i = \bar{b}$, $c_{ij} = \bar{c}$ has a stable polymorphic equilibrium, at $p^* = -\bar{a}/(2\bar{b} + 4\bar{c}(n-1))$.

Individuals were sampled from the equilibrium population. In the paternal gamete, locus i was allotted the allele A_i with probability p_i^* and the allele a_i with probability $q_i^* = 1 - p_i^*$, and similarly for the maternal gamete. For an individual with known genotype, fitness was computed according to a corresponding model. A pleiotropic trait was determined by the same loci. The microenvironmental deviation was absent. Both additive traits and traits with dominance and additive by additive epistasis were considered. The general formula used for calculating the trait value was

$$x = \mu + \sum_i [\alpha_i(l_i + l_i') + 2\beta_i l_i l_i'] + \sum_{i \neq j} \gamma_{ij}(l_i + l_i')(l_j + l_j').$$

Here if $\alpha_i \neq 0$, $\beta_i = \gamma_{ij} = 0$, the trait is additive both "between" and "within" loci; if $\alpha_i \neq 0$, $\beta_i \neq 0$, $\gamma_{ij} = 0$, the trait is additive "between" loci with dominance "within" loci, and the case with $\alpha_i \neq 0$, $\beta_i \neq 0$, $\gamma_{ij} \neq 0$ corresponds to the trait with additive, dominant, and additive by additive epistatic interactions between pairs of loci. The contributions of loci to the trait α_i, β_i, γ_{ij} were drawn independently from normal distributions with means $\bar{\alpha}$, $\bar{\beta}$, and $\bar{\gamma}$, and variances σ_α^2, σ_β^2 and σ_γ^2.

In order to investigate the relations between traits and fitness, 5000 individuals were sampled. The maximum fitness observed amongst those 5000 individuals was used to convert absolute fitness w to relative fitness $w' = w/w_{\max}$. Using the sample trait mean, $\hat{x}$, and the sample trait variance, $\hat{P}$, the trait value for each individual was normalized to $x' = (x - \hat{x})/\sqrt{\hat{P}}$. The individuals were collected according to x' into 50 classes of a width of about 0.1 standard deviation; for the individuals in each class, relative fitness was averaged to arrive at the apparent fitness $w_{\text{obs}}(x')$.

Form of apparent selection on a quantitative trait

In a nonequilibrium population, we expect $w_{\text{obs}}(x')$ for an additive trait to show directional selection; Figure 1A gives $w_{\text{obs}}(x')$ for an additive trait, at random allele frequencies. We expect to observe stabilizing selection on an additive trait if the population is at a stable polymorphic equilibrium determined by selection. Figure 1, B–D, present the form of apparent selection on a purely additive trait for the three viability models considered. Figure 2A represents $w_{\text{obs}}(x')$ for a traits with random dominance. Figure 2B describes apparent selection on a trait with both random dominance and epistasis. This figure suggests that the conclusion about apparent stabilizing selection on additive traits with random dominance may be generalized for nonadditive traits with random epistasis.

In these figures, the average relative fitness per class $w_{\text{obs}}(x')$ is almost perfectly quadratic, with its highest value at or very near $\bar{x}$. This changes if the average of the dominance contributions is not zero, but now the fitness model as well as the trait composition becomes important (Figure 2C). We observe directional, almost linear, selection if the coefficients of the trait are constant multipliers of the coefficients of fitness (Figure 2D). In other situations (G. DE JONG and S. GAVRILETS, unpublished results) we can observe very complex fitness functions with clearly asymmetric selection at low epistasis in the trait; at higher epistasis the maximum is displaced from $\bar{x}$.

Strength of apparent selection on a quantitative trait: In Table 1, different characteristics of the real and apparent strength of selection in the additive model and in the epistatic model are given. In this table quantities from theoretical expectations arrived by summation over loci as in (10, 11) and (34) are listed together under the heading "Theoretical," and quantities computed from a sample of 5000 individuals are listed together under the heading "Observed." The observed coefficient s was estimated by the regression of individual fitness w on the squared deviation of an individual's trait value from the sample mean. The apparent load is estimated by sV_g where V_g is the phenotypic variance. This apparent load corresponds to $V_g/2V_s$ in (BARTON 1990; KEIGHTLEY and HILL 1990). The observed fitness deviation for any individual can be divided into two independent parts, the deviation of $w(x)$ from $\bar{w}$ and the deviation of w from the apparent fitness $w(x)$: $w - \bar{w} = [w(x) - \bar{w}] + [w - w(x)]$. The total observed variance in fitness $\text{var}(w)_{\text{obs}}$ becomes the sum of $\text{var}(w(x) - \bar{w})_{\text{obs}}$ and $\text{var}(w - w(x))_{\text{obs}}$. The quantity $\text{var}(w)_{\text{obs}}$ estimates the variance in the real fitness, $\text{var}(w)$, while the quantity $\text{var}(w(x) - \bar{w})_{\text{obs}}$ estimates $\text{var}(w)_{\text{app}}$.

In Table 1, two cases for the overdominant fitness model and two cases for the epistatic fitness model are presented. The overdominant case 1 has the same average fitness and equilibrium allele frequencies as the epistatic case 3, while the overdominant case 2 has about the same total load as the epistatic case 3. Case 4 represents epistasis at a higher number of loci, with the same value of $b + 2c(n-1)$ as cases 1 and 3. Table 1 shows that, as expected, the genetic load in the epistatic model is much lower than in the correspond-

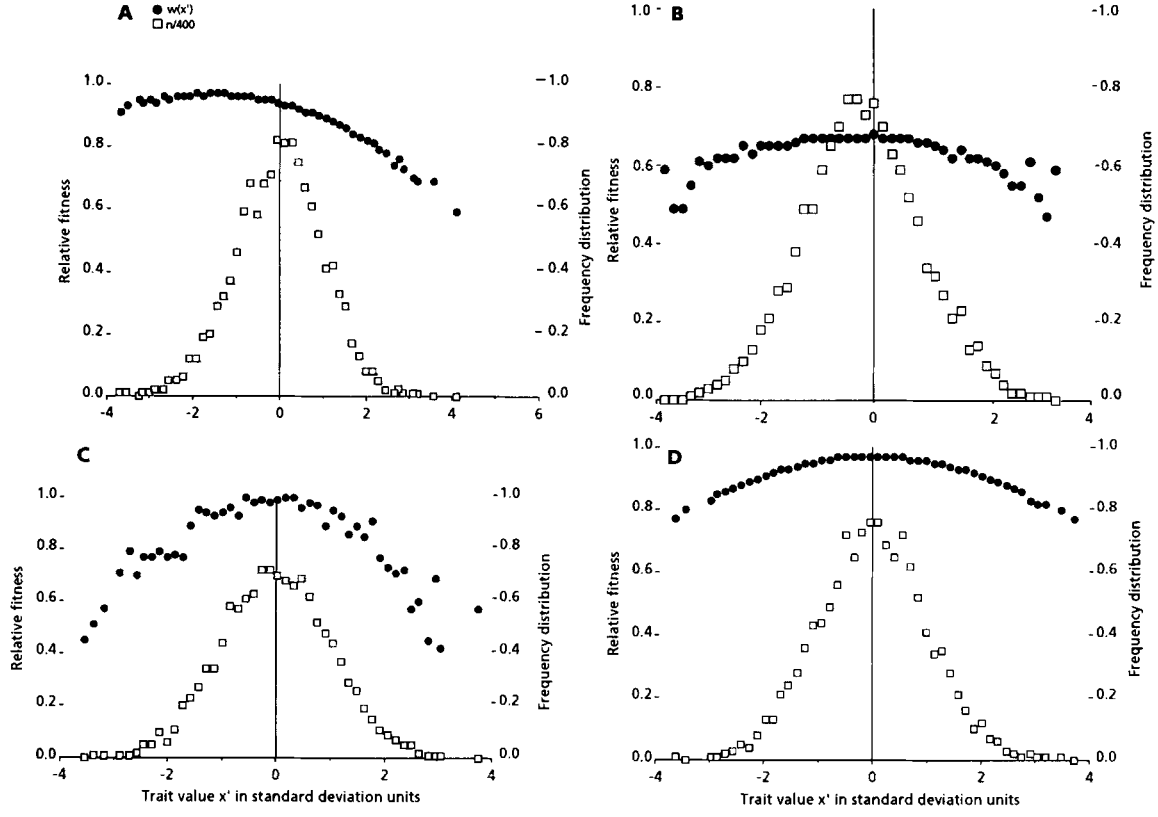

FIGURE 1.—Apparent selection on an additive trait pleiotropic to fitness. Shown are the apparent fitness as a function of trait value and the frequency distribution of trait values. The trait is determined according to expression (40), with: $\mu = 0$, $\bar{\alpha} = 10.0$, $\sigma_\alpha^2 = 4.0$, $\bar{\beta} = 0$, $\sigma_\beta^2 = 0$, $\bar{\gamma} = 0$, $\sigma_\gamma^2 = 0$. (A) Nonequilibrium population, random allele frequencies. Epistatic asymmetric fitness model, expression (33), with $m = 100$, $\bar{a} = 20.0$, $\sigma_a^2 = 0.0625$, $\bar{b} = -2.0$, $\sigma_b^2 = 0.01$, $\bar{c} = -0.5$, $\sigma_c^2 = 0.000625$. (B) Equilibrium population. Overdominant fitness model, expression 12, with $m = 100$, $\bar{a} = 20.0$, $\sigma_a^2 = 0.0625$, $\bar{b} = -21.0$, $\sigma_b^2 = 0.0225$. (C) Equilibrium population. Symmetric fitness model, expression (39), $\zeta = 0.5$; $\xi = 4.5$. (D) Equilibrium population. Epistatic asymmetric fitness model, expression (33), with $m = 100$, $\bar{a} = 20.0$, $\sigma_a^2 = 0.0625$, $\bar{b} = -2.0$, $\sigma_b^2 = 0.01$, $\bar{c} = -0.5$ and $\sigma_c^2 = 0.000625$.

ing overdominant model. The ratio of the strength of apparent selection to the strength of overall selection is much higher in the epistatic model.

In the case of overdominance, the characteristics of the apparent selection do not depend upon the degree of "adaptivity," *i.e.*, on the relation between the mean and the variance of α_i. In Figure 3, the ratio of apparent load to total load is given. In Figure 3, A and B, the equilibrium allele frequencies and the mean fitness are the same. If epistasis is absent, as in Figure 3A, no effect of the mean and the variance of the contributions of the loci to the trait is found at all. Without epistasis we find that both the explained part of the genetic load and the explained part of the genetic variance in fitness are very small (Table 1, case 1). In the epistatic model these values are small if $\bar{\alpha} \ll \sigma_\alpha$, but become very large if $\bar{\alpha} \gg \sigma_\alpha$ (Figure 3B). In Figure 3C, the explained part of the genetic load is given for different strength of epistasis, for both $\bar{\alpha} \ll \sigma_a$ and $\bar{\alpha} \gg \sigma_\alpha$. The apparent load at $\bar{\alpha} \gg \sigma_a$ is independent of the parameter c, only dependent upon $b + 2c(n - 1)$, as shown in expression (35). At $\bar{\alpha}/\sigma_a = 0$, the intensity of selection is demonstrated by expression (37). The part of the genetic load explained by the apparent stabilizing selection only appreciably rises with the strength of epistasis when the absolute value of $2c(n - 1)$ is larger than the absolute value of b, at around $c = -0.3$.

This is illustrated in Figure 4, A and B, by plotting the actual fitness of 100 individuals as well as their apparent fitness value on the basis of their trait value. As can be seen in Figure 4A, the overdominant model does not lead to a clustering of fitness values around $w(x)$. In Figure 4B, the degree of epistasis is high, and individual fitnesses follow the apparent fitness function almost perfectly. This is a visual example of the difference between the overdominant and epistatic models in explaining the genetic variance in fitness. The part of the genetic variance in fitness that is explained increases with the number of loci in the presence of high epistasis (Figure 4C). The epistatic model performs well in both the ratio of explained fitness variance to the total fitness variance, and the ratio of apparent load to total genetic load. The overdominant model cannot distinguish between "neutral" and "adaptive" traits.

DISCUSSION

The surprising result here is how few assumptions are necessary to arrive at apparent stabilizing selection

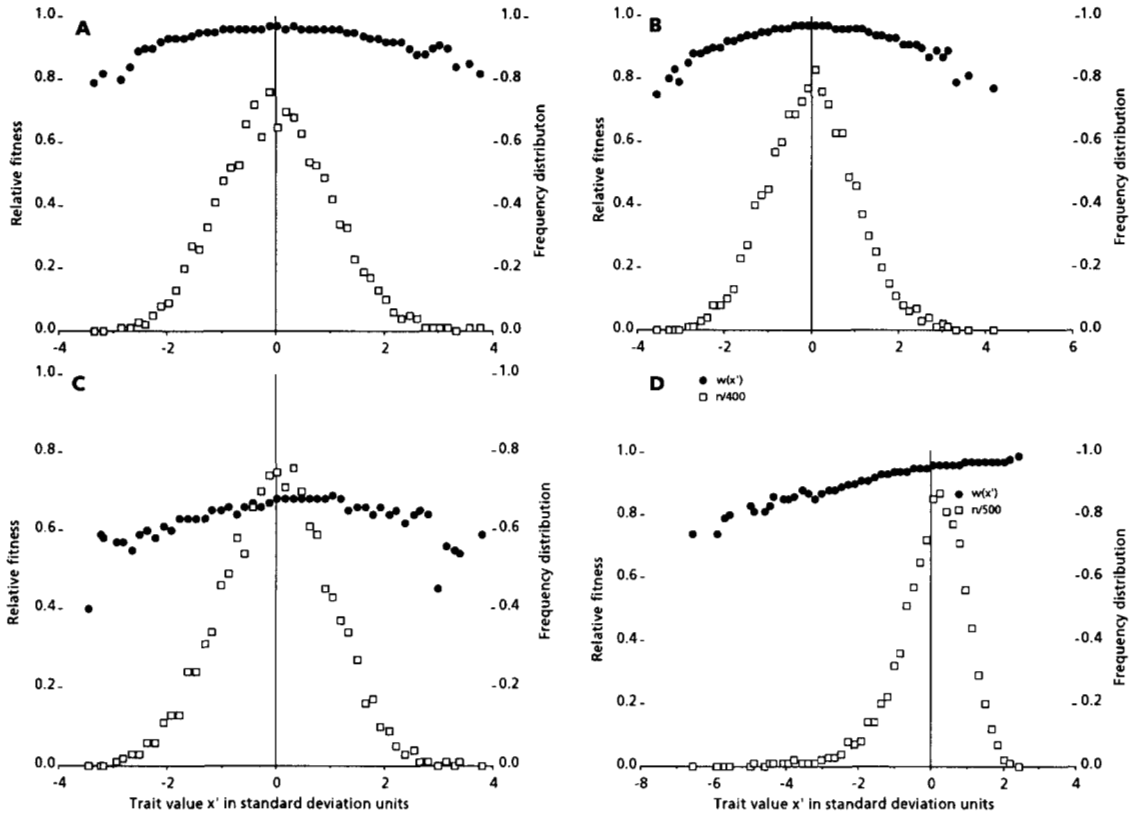

FIGURE 2.—Apparent selection on traits pleiotropic to fitness and influenced by dominance and epistasis. Shown are the apparent fitness as a function of trait value and the frequency distribution of trait values. Fitness in the equilibrium populations is according to the epistatic asymmetric fitness model, expression (33), with $m = 100$, $\bar{a} = 20.0$, $\sigma_a^2 = 0.0625$, $\bar{b} = -2.0$, $\sigma_b^2 = 0.01$, $\bar{c} = -0.5$ and $\sigma_c^2 = 0.000625$. (A) Random dominance, no epistasis in the trait: $\mu = 0$, $\bar{\alpha} = 10.0$, $\sigma_\alpha^2 = 4.0$, $\overline{\beta} = 0$, $\sigma_\beta^2 = 9.0$, $\bar{\gamma} = 0$ and $\sigma_\gamma^2 = 0$. (B) Random dominance, random epistasis in the trait: $\mu = 0$, $\bar{\alpha} = 10.0$, $\sigma_\alpha^2 = 4.0$, $\overline{\beta} = 0$, $\sigma_\beta^2 = 9.0$, $\bar{\gamma} = 0$ and $\sigma_\gamma^2 = 0.25$. (C) Partial dominance, no epistasis in the trait: $\mu = 0$, $\bar{\alpha} = 10.0$, $\sigma_\alpha^2 = 4.0$, $\overline{\beta} = -4.0$, $\sigma_\beta^2 = 0.0$, $\bar{\gamma} = 0$ and $\sigma_\gamma^2 = 0$. (D) Dominance and epistasis in the trait, averages of the trait coefficients a constant multiple of the averages of the fitness coefficients: $\mu = 0$, $\bar{\alpha} = 10.0$, $\sigma_\alpha^2 = 16.0$, $\overline{\beta} = -4.0$, $\sigma_\beta^2 = 0.16$, $\bar{\gamma} = -1.0$ and $\sigma_\gamma^2 = 0.025$.

on a quantitative trait pleiotropically connected with fitness. For any polygenic system in which variability is maintained by selection, all we need is either that the trait be additive, or that dominance and epistatic effects of loci on the trait to be considered as independent random variables with mean zero. Such a consideration implies that the sign of the dominance and epistatic effects of the loci on the trait is random with respect to their effect on fitness. This "independence" of the allele effects of these two types can be naturally interpreted as "neutrality" of the allele effects on the trait. For a polygenic system in which a low level of variability is maintained by mutation or migration, an additional condition is that the additive effects are "neutral" too. On the other hand, in order to observe "directional" selection on a quantitative trait, either the genetic system has to be at a "transient" state, or the effects of the loci on the trait have to be systematically related to those on fitness. The natural interpretation of the latter case is "adaptivity" of the corresponding contributions of the loci to the trait. Assuming that the genetic system is at a stable equilibrium determined by selection, we therefore can say that any quantitative trait for which nonadditive contributions are absent or random and independent of those to fitness will exhibit "stabilizing selection," while traits for which nonadditive contributions are related to those to fitness will exhibit "directional selection." For genetic systems in which variability is maintained by mutation or migration, any trait for which both additive and nonadditive contributions are random and independent of those to fitness will be under apparent stabilizing selection and any trait with contributions related to those to fitness will be under apparent directional selection.

Two pleiotropic models with additive fitness were recently analyzed with respect to their utility to account for both the high heritability and the strong apparent stabilizing selection which are commonly observed: mutation-selection balance model (BARTON 1990; KEIGHTLEY and HILL 1990) and the overdominant model (BARTON 1990). The results of these analyses show, however, that neither model can be easily accepted as a possible explanation. The difficulties with the mutation-selection balance model lie both in the low levels of maintained variability (BARTON, 1990) and in the weak stabilizing selection predicted (BARTON 1990; KEIGHTLEY and HILL 1990). The overdominant model fails due to load arguments (BARTON 1990).

TABLE 1

Comparison of overdominant and epistatic fitness model: effect of "neutral" and "adaptive" traits on apparent selection, apparent load and explained load, apparent fitness variance and explained fitness variance

	Overdominance				Epistasis			
Fitness model	Case 1[a]		Case 2[b]		Case 3[c]		Case 4[d]	
A. Contributions of loci to fitness equal								
$\tilde{\alpha}/\sigma_\alpha$[e]	0	5	0	5	0	5	0	5
Parameters								
μ	100		100		100		100	
a	20.0		1.0		20.0		20.0	
b	−21.0		−1.05		−2.0		−1.008	
c	0.0		0.0		−0.5		−0.2564	
$b + 2c(n-1)$	−21.0		−1.05		−21		−21	
p^*	0.47619		0.47619		0.47619		0.47619	
$b - 2c$	−21.0		−1.05		−1.0		−0.488	
n	20		20		20		40	
Theoretically predicted values								
$\bar{w}$	290.47	290.47	109.52	109.52	290.47	290.47	480.95	480.95
L	0.4190	0.4190	0.0637	0.0637	0.0506	0.0506	0.0262	0.0262
L_{apparent}[f]	0.0177	0.0177	0.00239	0.00239	0.0033	0.0173	0.000732	0.0104
L_{apparent}/L[f]	0.0422	0.0422	0.0274	0.0274	0.0648	0.3413	0.0278	0.3978
var$\{w\}$	2195.01	2195.01	5.4875	5.4875	67.194	67.194	61.156	61.156
var$\{w\}_{\text{apparent}}$[f]	54.87	54.87	0.1372	0.1372	1.8353	52.308	0.2481	51.405
var$\{w\}_{\text{app}}$/var$\{w\}$[f]	0.0250	0.0250	0.0250	0.0250	0.0273	0.7785	0.0041	0.8405
Observed values over 5000 simulated individuals								
$\bar{w}$	290.75	290.75	109.54	109.54	290.49	290.49	481.06	481.06
L	0.3392	0.3392	0.0638	0.0638	0.0507	0.0507	0.0262	0.0262
L_{apparent}[f]	0.0166	0.0177	0.00240	0.0241	0.0030	0.0177	0.000870	0.0104
L_{apparent}/L[f]	0.0491	0.0522	0.0376	0.0377	0.0590	0.3499	0.0332	0.3970
var$\{w\}$	2186.92	2186.92	5.4668	5.4668	66.29	66.29	60.12	60.12
var$\{w\}_{\text{apparent}}$[f]	53.45	53.51	0.1336	0.1337	1.49	52.96	0.3609	51.47
var$\{w\}_{\text{app}}$/var$\{w\}$[f]	0.0245	0.0245	0.0244	0.0244	0.0225	0.7988	0.0060	0.8562
B. Contributions of loci to fitness normally distributed[g]								
Parameters								
μ	100		100		100		100	
$\bar{a}$	20.0		1.0		20.0		20.0	
σ_a	0.25		0.0125		0.25		0.25	
$\bar{b}$	−21.0		−1.05		−2.0		−1.008	
σ_b	0.15		0.0075		0.1		0.05	
$\bar{c}$	0.0		0.0		−0.5		−0.2564	
σ_c	0.0		0.0		0.025		0.0125	
$\bar{b} + 2\bar{c}(n-1)$	−21.0		−1.05		−21		−21	
p^*	0.4785		0.47789		0.4790		0.4769	
$\sigma\{p\}$	0.0056		0.01015		0.1599		0.2574	
$\bar{b} - 2\bar{c}$	−21.0		−1.05		−1.0		−0.488	
n	20		20		20		40	
Theoretically predicted values								
$\bar{w}$	291.12	291.12	109.60	109.60	293.01	293.01	485.63	480.95
L	0.4183	0.4183	0.0872	0.0872	0.0449	0.0449	0.0207	0.0207
L_{apparent}[f]	0.0176	0.0176	0.0024	0.0024	0.0032	0.0154	0.00056	0.0076
L_{apparent}/L[f]	0.0422	0.0427	0.0373	0.0377	0.0723	0.3431	0.0271	0.3657
var$\{w\}$	2192.03	2192.03	5.4790	5.4790	53.4526	53.4526	33.9919	33.9919
var$\{w\}_{\text{apparent}}$[f]	54.9948	54.8226	0.1371	0.1371	1.6808	41.1845	0.1522	27.6821
var$\{w\}_{\text{app}}$/var$\{w\}$[f]	0.0251	0.0250	0.0250	0.0250	0.0315	0.7705	0.0045	0.8144
Observed values over 5000 simulated individuals								
$\bar{w}$	292.34	292.34	109.62	109.62	292.92	292.92	485.62	485.62
L	0.3383	0.3383	0.0638	0.0638	0.0452	0.0452	0.0208	0.0208
L_{apparent}[f]	0.0175	0.0191	0.0024	0.0026	0.0034	0.0162	0.00055	0.0078
L_{apparent}/L[f]	0.0518	0.0564	0.0379	0.0408	0.0748	0.3593	0.0266	0.3746
var$\{w\}$	2137.25	2137.25	5.3761	5.3761	58.025	58.025	34.1307	34.1307
var$\{w\}_{\text{apparent}}$[f]	7.168	7.959	0.3631	0.3999	1.3298	46.6110	0.1426	28.2695
var$\{w\}_{\text{app}}$/var$\{w\}$[f]	0.0033	0.0037	0.0675	0.0744	0.0315	0.8034	0.0042	0.8283

[a] Case 1 and case 3 have similar mean fitness $\bar{w}$ but differ in epistasis c.
[b] Case 2 and case 3 have approximately similar genetic load L but differ in epistasis c.
[c] Cases 1, 3 and 4 have identical nonadditivity $b + 2c(n-1)$.
[d] Case 3 and case 4 differ in number of loci n.
[e] $\tilde{\alpha}/\sigma_\alpha = 0$ stands for a "neutral" trait; $\tilde{\alpha}/\sigma_\alpha = 5$ stands for an "adaptive" trait.
[f] "Neutral" *vs.* "adaptive" has an influence only under the epistatic fitness model.
[g] The variance over loci in equilibrium allele frequency due to the normal distribution of the locus parameters has little influence.

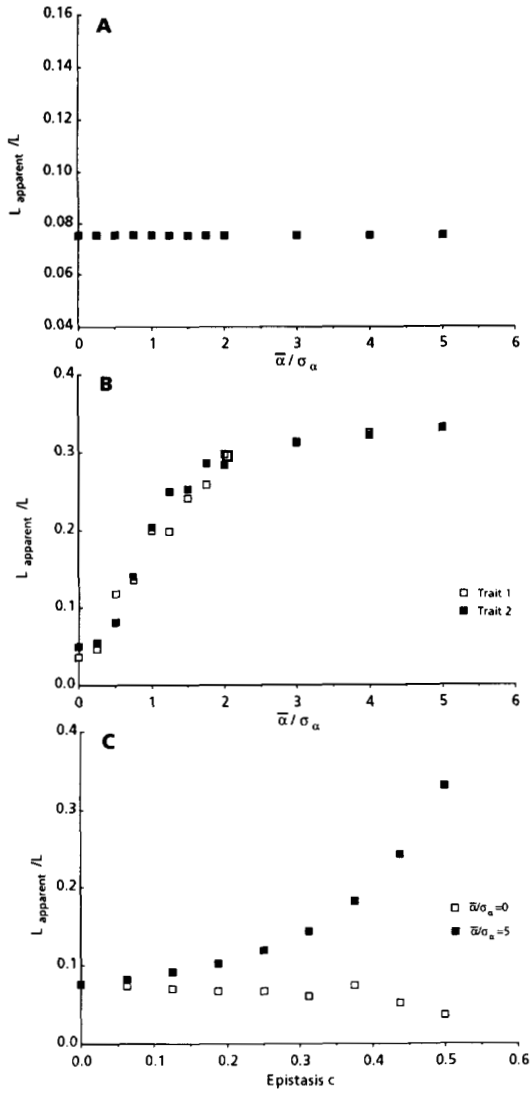

FIGURE 3.—Part of the genetic load that is explained by the apparent fitness function, depending upon the fitness model and the ratio of mean to standard deviation for the trait coefficients. Additive traits: trait 1: $\sigma_\alpha^2 = 4.0$; trait 2: $\sigma_\alpha^2 = 225.0$. All loci identical values for fitness. (A) Overdominant fitness model: $m = 100$, $a = 20.0$, $b = -21.0$, $c = 0$: $\omega = 0$. Without epistasis, the genetic load explained by the apparent fitness function does not depend on the ratio $\bar{\alpha}/\sigma_\alpha$. (B) Epistatic fitness model: $m = 100$, $a = 20.0$, $b = -2.0$, $c = -0.5$: $\omega = 0.9$. With epistasis, the genetic load explained by the apparent fitness function depends on the ratio $\bar{\alpha}/\sigma_\alpha$, *i.e.*, on the ordering of the trait coefficients α with respect to the fitness coefficients a. (C) The influence of the degree of epistasis in fitness on the part of the genetic load explained by the apparent fitness function: $m = 100$, $a = 20$, $b + 2c(n-1) = -21.0$, $-c$ as indicated. If $\bar{\alpha} = 0$, and the sign of the trait coefficient α is random with respect to the fitness coefficient a, the part of the genetic load explained by the apparent fitness function decreases with increasing level of epistasis. If $\bar{\alpha}/\sigma_\alpha = 5$, and the sign of the trait coefficient α is identical to that of the fitness coefficient a, the part of the genetic load explained by the apparent fitness function increases with increasing level of epistasis.

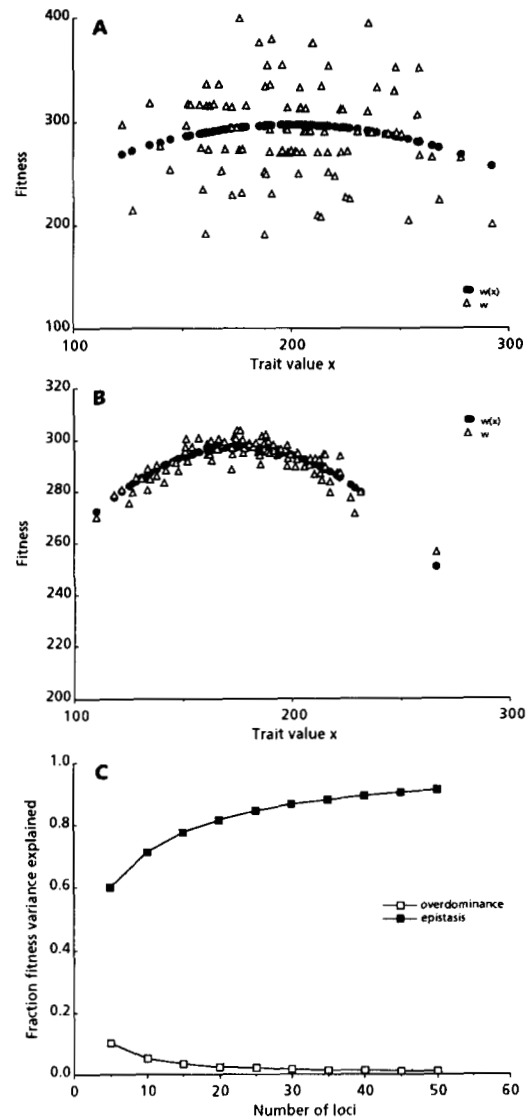

FIGURE 4.—Genetic variance in fitness as explained by the apparent fitness function. Additive trait, $\bar{\alpha} = 10.0$, $\sigma_\alpha^2 = 4.0$, $\bar{\alpha}/\sigma_\alpha = 5.0$. (A) Fitness and apparent fitness for 100 individuals under the overdominant fitness model; $n = 20$ loci, $m = 100$, $\bar{a} = 20.0$, $\sigma_a^2 = 0.0625$, $\bar{b} = -21.0$, $\sigma_b^2 = 0.0225$, $\omega = 0$. The apparent fitness function explains ½n of the genetic variance in fitness, see in C under $n = 20$ loci. (B) Fitness and apparent fitness for 100 individuals under the epistatic fitness model; $n = 20$ loci, $m = 100$, $\bar{a} = 20.0$, $\sigma_a^2 = 0.0625$, $\bar{b} = -2.0$, $\sigma_b^2 = 0.01$, $\bar{c} = -0.5$, $\sigma_c = 0.000625$, $\omega = 0.9$. The apparent fitness function explains the greater part of the genetic variance in fitness, see in C under $n = 20$ loci. (C) The part of the genetic variance in fitness that is explained by the overdominant and by the epistatic fitness model as a function of the number of loci n. For the epistatic fitness model, the highest degree of epistasis that is compatible with n polymorphic loci is chosen. Overdominant fitness model: $m = 100$, $a = 20.0$, $b = -21.0$: The part of the genetic variance in fitness that is explained equals (½n). Epistatic fitness model: $m = 100$, $a = 20.0$, $b + c(n-1) = -21.0$, $b = 4c$. The part of the genetic variance in fitness that is explained equals (½n). $[b + 2c(n-1)]^2/[b^2 + 2c^2(n-1)]$.

In this paper, we have applied our general results on apparent fitness functions both to the additive viability models analyzed by BARTON and KEIGHTLEY and HILL, and to a general mutation (migration)-selection balance model and to two non-additive viability models. These are: a symmetric viability model, in which fitness depends on the proportion of heterozygous loci, and a model with additive by additive

epistatic interactions between pairs of loci. We demonstrated that the strength of apparent stabilizing selection on an additive trait in the population, where a low level of polygenic variability is maintained by mutation or migration, is order V_m/V_g independently of the real selection. The typical experimental estimate of V_m/V_g for mutation is about 10^{-3} [*e.g.*, LYNCH (1988)]. This makes it impossible to explain strong stabilizing selection observed in terms of mutation-selection balance models. We showed that both in the overdominant model and in the symmetric viability model the strength of apparent stabilizing selection on an additive quantitative trait is approximately $1/(2n)$ times the strength of the overall selection on whole phenotype. On the general supposition that the number of loci, n, influencing fitness is very large, it is impossible to explain strong stabilizing selection in terms of these two pleiotropic models.

Epistasis can change all this. The epistatic fitness model can maintain the same (very high) level of genetic polymorphism as the overdominant model. Any model that attempts to explain the maintenance of high levels of genetic variability at many loci meets the problem of genetic load. The overdominance per locus model can, in principle, maintain very high levels of variability. However, in this model the genetic load associated with polymorphism increases very quickly (linearly or exponentially) with the number of loci. This means that models involving overdominance per locus cannot be taken as a general case of selectively maintained polymorphism. Nor is per locus overdominance indicated as a general case of selection by experimental results. Contrary to the overdominant model, genetic load and the genetic variance in fitness in the epistatic model can be very low. For fitnesses involving epistasis, the genetic variance in fitness (and genetic load) can be as small as $1/(2n)$ that of the overdominant model (Equation 19, Figure 1). In other words, these values in the epistatic model are approximately half the corresponding values at a single locus of the overdominant model. The results of numerical simulations show that this is also true in more general models with "unequivalent" loci. Thus, from a theoretical point of view, epistasis seems to be able to solve the problem of genetic load. From general considerations epistasis in fitness is likely to occur in real organisms. For example, epistasis has to occur in any organism where both the acquisition of a resource and the allocation of that resource to competing traits are under genetic control (DE JONG and VAN NOORDWIJK 1992), or when both viability and fecundity are genetically correlated and total fitness is their product.

If we accept that stabilizing selection is only or predominantly "apparent," then we have to be able to explain the fact that this apparent stabilizing selection is sometimes very strong (BARTON 1990). Neither the general mutation-selection balance model nor the overdominant viability model and the symmetric viability model can produce strong apparent stabilizing selection on an additive trait. This conclusion does not depend on whether we measure the strength of selection by the genetic load or by the genetic variance in relative fitness. In the epistatic model with "equivalent" loci, the strength of apparent stabilizing selection (measured by genetic load or by genetic variance in fitness) may be as large as the strength of "real" selection. This model, however, assumes in an implicit form that effects of alleles on fitness and on the quantitative trait are strongly related. In the case where these two types of allele effects are completely independent, the strength of apparent phenotypic selection is weak (approximately $1/(2n)$ times less than the total strength of selection). What we really have seems to lie somewhere between two extremes, the interdependence of these two types of allele effects exists, but is not absolute (MACKAY 1990; KEIGHTLEY and HILL 1990). In this paper, we used the ratio $\bar{\alpha}/\sigma_\alpha$ as a measure of the interdependence. If $\bar{\alpha} \ll \sigma_\alpha$, then the sign of the effects of the loci on the trait is random with respect to their effects on fitness. Such a situation can be naturally interpreted as "neutrality" of the trait. A trait for which $\bar{\alpha} \gg \sigma_\alpha$, can be naturally interpreted as "non-neutral" or "adaptive." Results of numerical simulations show that in the epistatic model the ratio of the strength of apparent to real selection can take any value between $1/(2n)$ and 1 depending on $\bar{\alpha}/\sigma t\bar{\alpha}$. The effect of this ratio on the strength of apparent stabilizing selection resembles that of the correlation of ρ of KEIGHTLEY and HILL (1990). An intuitive explanation of this similarity can be given if one considers the situation where both fitness and the trait is additive. In this case the covariance $\mathrm{cov}(w, x)$ depends on the expected value $E\{\alpha_i a_i\} = \mathrm{cov}(\alpha_i, a_i) + \bar{\alpha}\bar{a}$. This shows that we can change $\mathrm{cov}(w, x)$ either by changing covariance of α_i and a_i, or changing the corresponding mean values.

In conclusion we stress two points. First, our results on apparent stabilizing selection do not necessarily lead to the suggestion that the traits are not subject to direct selection. There are many mechanisms which can generate apparent stabilizing selection on the trait. Pleiotropy may be one of them. Second, the epistatic models considered in this paper are not intended to be an ultimate explanation of the maintenance of polygenic variability and apparent stabilizing selection. More analysis of these and more complex models are needed. Models and results presented here should point out possible points of attention and sorts of observations that should be appropriate. One of the essential points is whether selection can be fully described as an environmental influence on an already fully developed phenotype, as any model of the form

$w = w(x)$ tends to suggest, or whether apparent selection comes about during development due to pleiotropy between fitness and the trait. As with most conceptual dichotomies, this is overstating the difference. What we need is to know how development and environment connect fitness and the trait.

We thank N. BARTON, C. CHEVALET, W. GABRIEL, T. VAN DER HAVE, A. KONDRASHOV and W. SCHARLOO for helpful discussions and R. BÜRGER, C. HOM, S. SCHEINER, M. TURELLI and two anonymous reviewers for comments on earlier drafts. S.G. was supported by a visitor's grant from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) and by a postdoctoral fellowship from the Ministère Française de la Recherche et de la Technologie.

LITERATURE CITED

ARNOLD, S. J., and M. J. WADE, 1984a On the measurement of natural and sexual selection: theory. Evolution **38:** 709–719.

ARNOLD, S. J., and M. J. WADE, 1984b On the measurement of natural and sexual selection: applications. Evolution **38:** 720–734.

BARTON, N. H., 1983 Multilocus clines. Evolution **37:** 454–471.

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

BARTON, N. H., 1990 Pleiotropic models of quantitative variation. Genetics **124:** 773–782.

BARTON, N. H., and M. TURELLI, 1987 Adaptive landscapes, genetic distances and the evolution of quantitative characters. Genet. Res. **49:** 157–174.

BARTON, N. H., and M. TURELLI, 1991 Natural and sexual selection on many loci. Genetics **127:** 229–255.

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

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

CAVENER, D. R., and M. T. CLEGG, 1981 Evidence for biochemical and physiological differences between enzyme genotypes in Drosophila melanogaster. Proc. Natl. Acad. Sci. USA **78:** 4444–4447.

DE JONG, G., and A. J. VAN NOORDWIJK, 1992 Acquisition and allocation of resources: genetic (co)variances, selection, and life histories. Am. Nat. **139:** 749–770.

ENDLER, J. A., 1986 *Natural Selection in the Wild*. Princeton University Press, Princeton, N.J.

EWENS, W., and G. THOMSON, 1977 Properties of equilibria in multilocus genetic systems. Genetics **87:**807–819.

FALCONER, D., 1989 *Introduction to Quantitative Genetics*, Ed. 3. Longman, Harlow, Essex.

GAVRILETS, S., 1993 Arbitrary strength of selection in a model of epistatic selection. J. Math. Biol. (in press).

GILLESPIE, J. H., and M. TURELLI, 1989 Genotype-environment interaction and the maintenance of polygenic variability. Genetics **121:** 129–138.

GIMELFARB, A., 1986 Additive variation maintained under stabilizing selection: a two-locus model of pleiotropy for two quantitative traits. Genetics **112:** 717–725.

GIMELFARB, A., 1989 Genotypic variation for a quantitative character maintained two-locus model of pleiotropy for two quantitative traits. Genetics **123:** 217–227.

HASTINGS, A., 1990 Second-order approximations for selection coefficients at polygenic loci. J. Math. Biol. **28:** 475–483.

HASTINGS, A., and C. L. HOM, 1990 Multiple equilibria and maintenance of additive genetic variance in a model of pleiotropy. Evolution **44:** 1153–1163.

KARLIN, S., and H. AVNI, 1981 Analysis of central equilibria in multilocus systems: a generalized symmetric viability regime. Theor. Popul. Biol. **20:** 241–280.

KEIGHTLEY, P. D., and W. G. HILL, 1990 Variation maintained in quantitative traits with mutation-selection balance: pleiotropic side-effects on fitness traits. Proc. R. Soc. Lond. B. **242:** 95–100.

KONDRASHOV, A. S., and M. TURELLI, 1992 Deleterious mutations, apparent stabilizing selection, and the maintenance of quantitative variability. Genetics **132:** 603–618.

LANDE, R., 1975 The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genet. Res. **26:** 221–235.

LANDE, R., 1976 Natural selection and random genetic drift in phenotypic evolution. Evolution **30:** 314–334.

LANDE, R., and S. J. ARNOLD, 1983 The measurement of selection on correlated characters. Evolution **36:** 1210–1226.

LYNCH, M., 1988 The rate of polygenic mutation. Genet. Res. **51:** 137–148.

MACKAY, T. F. G., 1990 Distribution of effects of new mutations affecting quantitative traits, pp. 219–228 in *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production*, edited by W. G. HILL, R. THOMSON and J. A. WOOLIAMS. AWCGALP, Edinburgh.

MITCHELL-OLDS, T., and R. G. SHAW, 1987 Regression analysis of natural selection: Statistical inference and biological interpretation. Evolution **41:** 1149–1161.

NAGYLAKI, T., 1989 The maintenance of genetic variability in two-locus models of stabilizing selection. Genetics **122:** 235–248.

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

ROBERTSON, A., 1967 The nature of quantitative genetic variation, pp. 265–280 in *Heritage from Mendel*, edited by R. A. BRINK. University of Wisconsin Press, Madison.

SCHLUTER, D., 1988 Estimating the form of natural selection on a quantitative trait. Evolution **42:** 849–861.

SLATKIN, M., 1972 On treating the chromosome as the unit of selection. Genetics **72:** 157–168.

TURELLI, M., and N. H. BARTON, 1990 Dynamics of polygenic characters under selection. Theor. Popul. Biol. **38:** 1–57.

WADE, M. J., and S. KALISZ, 1990 The cases of natural selection. Evolution **44:** 1947–1955.

ZHIVOTOVSKY, L., and S. GAVRILETS, 1992 Quantitative variability and multilocus polymorphism under epistatic selection. Theor. Popul. Biol. **42:** 254–283.

Communicating editor: B. S. WEIR

APPENDIX

Here we calculate the apparent fitness function (2). Let there be n loci with two alleles each: A_i and a_i ($i = 1,\ldots, n$). Let the indicator variable l_i (l_i') be equal to 1, if the allele at the ith locus of the paternal (maternal) gamete is A_i, and to 0, if this allele is a_i. Let us introduce the ($2n \times 1$) vector L with the components, $L_i = l_i$, $L_{n+i} = l_i'$, $i = 1,\ldots, n$. This vector defines the genotype. Any dependence of fitness, w,

on the genotype, L, can be represented as a sum in the form

$$w(L) = m + \sum_{i}^{2n} \mathfrak{A}_i L_i + \sum_{i\neq j}^{2n} \mathfrak{B}_{ij} L_i L_j + \sum_{i\neq j\neq k}^{2n} \mathfrak{C}_{ijk} L_i L_j L_k + \ldots, \quad \text{(A1)}$$

where m, $\mathfrak{A}_i$, $\mathfrak{B}_{ij}$, $\mathfrak{C}_{ijk}$, . . . are parameters of the fitness function. We shall assume that the relative contribution of each sum in (A1) to w does not depend on the number of loci. The mean fitness of the population is

$$E\{w\} = m + \sum_{i}^{2n} \mathfrak{A}_i \bar{L}_i + \sum_{i\neq j}^{2n} \mathfrak{B}_{ij} E\{L_i L_j\} + \sum_{i\neq j\neq k}^{2n} \mathfrak{C}_{ijk} E\{L_i L_j L_k\} + \ldots, \quad \text{(A2)}$$

where E denotes the average over the population and $\bar{L}_i = E\{L_i\}$. Note that $\bar{L}_i$ ($\bar{L}_{i+n}$), $i = 1,\ldots, n$, is the frequency of allele A_i at the ith locus of the paternal (maternal) gamete, and that the expectations in (A2) represent the frequencies of the corresponding marginal gametes and genotypes (SLATKIN 1972; EWENS and THOMSON 1977). For example, $E\{L_1L_2\}$ is the frequency of the two-locus marginal gamete A_1A_2, while $E\{L_1L_{n+1}\}$ is the frequency of the one-locus marginal genotype A_1A_1. Let

$$C_{i\ldots j} = E\{(L_i - \bar{L}_i)\ldots(L_j - \bar{L}_j)\}$$

denote the covariance of the indicator variables $L_i,\ldots, L_j$. For example, if both $i \leqslant n$ and $j \leqslant n$, then $D_{ij} \equiv C_{ij}$ is the standard linkage disequilibrium between the ith and the jth loci, while $D_{i,i} \equiv C_{i(i+n)}$ characterizes the deviation from Hardy-Weinberg proportions"within" the ith locus. The expectations in (A2) (*i.e.*, the frequencies of marginal gametes and genotypes) can be expressed as functions linear in C, for example

$$E\{L_1L_2\ldots L_N\} = \Pi + \sum_{i<j}^{N} (\partial^2\Pi/\partial\bar{L}_i\partial\bar{L}_j)C_{ij} + \sum_{i<j<k}^{N} (\partial^3\Pi/\partial\bar{L}_i\partial\bar{L}_j\partial\bar{L}_k)C_{ijk} + \ldots, \quad \text{(A3)}$$

where $\Pi = \bar{L}_1\bar{L}_2\ldots\bar{L}_N$ (SLATKIN 1972; BARTON 1983). Substituting the expectation in (A2) by the corresponding terms in the form similar to (A3) and gathering terms, we get

$$E\{w\} = \bar{w} + \sum_{i\neq j}^{2n} \frac{1}{2} (\partial^2\bar{w}/\partial\bar{L}_i\partial\bar{L}_j)C_{ij} + \sum_{i\neq j\neq k}^{2n} \frac{1}{6} (\partial^3\bar{w}/\partial\bar{L}_i\partial\bar{L}_j\partial\bar{L}_k)C_{ijk} + \ldots, \quad \text{(A4)}$$

where $\bar{w}$ incorporates all terms that depend only on $\bar{L}_i$:

$$\bar{w} = m + \sum_{i}^{2n} \mathfrak{A}_i \bar{L}_i + \sum_{i\neq j}^{2n} \mathfrak{B}_{ij} \bar{L}_i \bar{L}_j + \sum_{i\neq j\neq k}^{2n} \mathfrak{C}_{ijk} \bar{L}_i \bar{L}_j \bar{L}_k + \ldots. \quad \text{(A5)}$$

This is just the mean fitness of the population both at linkage and Hardy-Weinberg equilibrium. Assume now that the distributions of paternal and maternal gametes are equal. This implies that for $i \leqslant n$, $j \leqslant n$, $\bar{L}_i = \bar{L}_{i+n} = p_i$, $C_{ij} = C_{(i+n)(j+n)} = D_{ij}$, $C_{i(j+n)} = C_{(i+n)j} = D_{i,j}$, and so on. Note also that $\partial^2\bar{w}/\partial\bar{L}_i\partial\bar{L}_{i+n} = (1/2)\partial^2\bar{w}/\partial p_i^2$, $\partial^2\bar{w}/\partial\bar{L}_i\partial\bar{L}_j = (1/4)\partial^2\bar{w}/\partial p_i\partial p_j$, and so on. Expression (A4) can be rewritten as

$$\begin{aligned} E\{w\} = \bar{w}(p_l, \ldots, p_n) &+ \sum_{i} \frac{1}{2} (\partial^2\bar{w}/\partial p_i^2)D_{i,i} \\ &+ \sum_{i\neq j} \frac{1}{4} (\partial^2\bar{w}/\partial p_i\partial p_j)(D_{ij} + D_{i,j}) \\ &+ \sum_{i\neq j} \frac{3}{8} (\partial^3\bar{w}/\partial p_i^2\partial p_j)D_{i,ij} \\ &+ \sum_{i\neq j\neq k} \frac{1}{24} (\partial^3\bar{w}/\partial p_i\partial p_j\partial p_k)(D_{ijk} + 3D_{i,jk}) + \ldots, \end{aligned} \quad \text{(A6)}$$

where $p_i = \bar{L}_i = \bar{L}_{i+n}$ is the frequency of allele A_i in the population, and $D_{i\ldots j,k\ldots l}$ is the "linkage disequilibrium" among loci $i,\ldots, j$ at the paternal gamete and loci $k,\ldots, l$ at the maternal gamete. The indexes in (A6) run from 1 to n.

The mean fitness of the subpopulation having the value of x of a quantitative trait, $w(x) \equiv E\{w|x\}$, is described by

$$\begin{aligned} w(x) = \bar{w}(p_1^x, \ldots, p_n^x) &+ \sum_{i} \frac{1}{2} (\partial^2\bar{w}/\partial p_i^2)D_{i,i}^x \\ &+ \sum_{i\neq j} \frac{1}{4} (\partial^2\bar{w}/\partial p_i\partial p_j)(D_{ij}^x + D_{i,j}^x) \\ &+ \sum_{i\neq j} \frac{3}{8} (\partial^3\bar{w}/\partial p_i^2\partial p_j)D_{i,ij}^x \\ &+ \sum_{i\neq j\neq k} \frac{1}{24} (\partial^3\bar{w}/\partial p_i\partial p_j\partial p_k)(D_{ijk}^x + 3D_{i,jk}^x) + \ldots, \end{aligned} \quad \text{(A7)}$$

where p_i^x, D_{ij}^x, $D_{i,j}^x$, . . . are the corresponding characteristics of the subpopulation, and the partial derivatives are evaluated at $(p_1^x,\ldots, p_n^x)$. Note that even if the population is at linkage and Hardy-Weinberg equilibrium, the subpopulation with a fixed value of the trait is not. Expanding $\bar{w}$ and partial derivatives in a Taylor series at the point $(p_1,\ldots, p_n)$, we have

$$w(x) = \bar{w} + \sum_i \frac{\partial \bar{w}}{\partial p_i}(p_i^x - p_i) + \sum_{i,j} \frac{1}{2}\frac{\partial^2 \bar{w}}{\partial p_i \partial p_j}(p_i^x - p_i)(p_j^x - p_j) + \sum_i \frac{1}{2}(\partial^2 \bar{w}/\partial p_i^2) D_{i,i}^x + \sum_{i \neq j} \frac{1}{4}(\partial^2 \bar{w}/\partial p_i \partial p_j)(D_{ij}^x + D_{i,j}^x) + \ldots, \quad \text{(A8)}$$

where now the partial derivatives are evaluated at $(p_1, \ldots, p_n)$. The last formula shows that to calculate the "apparent" fitness function one has to know the conditional frequencies p_i^x, the conditional disequilibria $D_{i,i}^x$, D_{ij}^x, $D_{i,j}^x$, and so on.

Let us first consider the conditional mean $p_i^x = E\{l_i | X = x\}$, which can be represented as

$$p_i^x = E\{l_i | X = x\} = p_i \frac{\Pr\{X \in (x, x + dx | l_i = 1\}}{\Pr\{X \in (x, x + dx\}} = p_i \frac{f(x | l_i = 1)}{f(x)}, \quad \text{(A9)}$$

where $f(x)$ is the phenotypic distribution of x in the population, and $f(x | l_i = 1)$ is the phenotypic distribution conditioned on the presence of allele A_i. HASTINGS (1990) has shown that if epistasis in the trait and linkage disequilibrium are absent, the phenotypic distribution conditional on the presence of the specified allele, A_i, at the specified locus, i, can be approximated as

$$f(x | l_i = 1) = f(x) - (\delta\bar{x})_i f'(x) + \frac{1}{2}((\delta\bar{x})_i^2 - (\delta P)_i) f''(x), \quad \text{(A10)}$$

where $(\delta x)_i = E\{x | l_i = 1\} - \bar{x}$ and $(\delta P)_i = P - \text{var}\{x | l_i = 1\}$ are the mean effects of allele A_i on the mean value and phenotypic variance, and the error is third order in the mean effects $(\delta\bar{x})$. Substituting (A10) into (A9),

$$p_i^x - p_i = -(\delta\bar{x})_i p_i \frac{f'(x)}{f(x)} + \frac{1}{2}((\delta\bar{x})_i^2 - (\delta P)_i p_i \frac{f''(x)}{f(x)} + \ldots. \quad \text{(A11)}$$

HASTINGS (1990) has also shown that an expression analogous to (A10) is valid for the phenotypic distribution conditioned on the presence of two specified genes at the ith locus. Applying HASTING's approach for the phenotypic distribution conditioned on the presence of two alleles A_i at the ith locus, we get

$$E\{l_i l_i' | X = x\} - p_i^2 = -(\delta\bar{x})_{ii} p_i^2 \frac{f'(x)}{f(x)} + \frac{1}{2}((\delta\bar{x})_{ii}^2 - (\delta P)_{ii}) p_i^2 \frac{f''(x)}{f(x)} + \ldots, \quad \text{(A12)}$$

where $(\delta\bar{x})_{ii} = E\{x | l_i = l_i' = 1\} - \bar{x}$, and $(\delta P)_{ii} = P - \text{var}\{x | l_i = l_i' = 1\}$. Hence, $D_{i,i}^x \equiv E\{l_i l_i' | X = x\} - E\{l_i | X = x\} E\{l_i' | X = x\}$ can be represented as

$$D_{i,i}^x = -((\delta\bar{x})_{ii} - 2(\delta\bar{x})_i) p_i^2 \frac{f'(x)}{f(x)} + \frac{1}{2}((\delta\bar{x})_{ii}^2 - 4(\delta\bar{x})_i^2 - (\delta P)_{ii} + 2(\delta P)_i) p_i^2 \frac{f''(x)}{f(x)} + (\delta\bar{x})_i^2 p_i^2 \frac{d}{dx}\left(\frac{f'(x)}{f(x)}\right). \quad \text{(A13)}$$

Using HASTINGS' formulas for the phenotypic distribution conditioned on the presence of allele A_i at the ith locus and allele A_j at the jth locus, we have

$$E\{l_i l_j | X = x\} - p_i p_j = -(\delta\bar{x})_{ij} p_i p_j \frac{f'(x)}{f(x)} + \frac{1}{2}((\delta\bar{x})_{ij}^2 - (\delta P)_{ij}) p_i p_j \frac{f''(x)}{f(x)} + \ldots, \quad \text{(A14)}$$

where $(\delta\bar{x})_{ij} = E\{x | l_i = l_j = 1\} - \bar{x}$, and $(\delta P)_{ij} = P - \text{var}\{x | l_i = l_j = 1\}$. Formulas (A11, A14) allow us to approximate the conditional disequilibria. Since the trait is additive between loci, $(\delta\bar{x})_{ij} = (\delta\bar{x})_i + (\delta\bar{x})_j$, $(\delta P)_{ij} = (\delta P)_i + (\delta P)_j$, and

$$D_{ij}^x = (\delta\bar{x})_i (\delta\bar{x})_j p_i p_j \left[\frac{f''(x)}{f(x)} - \left(\frac{f'(x)}{f(x)}\right)^2\right] = (\delta\bar{x})_i (\delta\bar{x})_j p_i p_j \frac{d}{dx}\left(\frac{f'(x)}{f(x)}\right). \quad \text{(A15)}$$

The same formula is valid for $D_{i,j}^x$. Substituting (A11, A13, A15) into (A8) and slightly rearranging the terms, we get Equation 2 of the main text.

Average effects on the mean and the variance: An additive trait with arbitrary degree of dominance can be described as

$$x = \mu + \sum[\alpha_i(l_i + l_i') + 2\beta_i l_i l_i'] + e,$$

For this trait the average effects of alleles on the mean value and on the variance of the trait are

$$(\delta\bar{x})_i = \alpha_i q_i + 2\beta_i p_i q_i, \quad (\delta P)_i = \alpha_i^2 p_i q_i + 4\alpha_i \beta_i p_i q_i (p_i - q_i) + 4\beta_i^2 p_i q_i (p_i^2 + p_i - 1).$$

$$(\delta\bar{x})_{ii} = 2\alpha_i q_i + 2\beta_i (p_i + 1) q_i, \quad (\delta P)_{ii} = 2\alpha_i^2 p_i q_i + 8\alpha_i \beta_i p_i^2 q_i + 4\beta_i^2 p_i^2 q_i (p + 1).$$