Genetic Variation Maintained in Multilocus Models of Additive Quantitative Traits Under Stabilizing Selection

Reinhard Bürger* and Alexander Gimelfarb†

**Institut für Mathematik, Universität Wien, A-1090 Vienna, Austria and †Department of Zoology, University of Maryland, College Park, Maryland 20742-4415*

Manuscript received October 29, 1998
Accepted for publication March 15, 1999

ABSTRACT

Stabilizing selection for an intermediate optimum is generally considered to deplete genetic variation in quantitative traits. However, conflicting results from various types of models have been obtained. While classical analyses assuming a large number of independent additive loci with individually small effects indicated that no genetic variation is preserved under stabilizing selection, several analyses of two-locus models showed the contrary. We perform a complete analysis of a generalization of Wright's two-locus quadratic-optimum model and investigate numerically the ability of quadratic stabilizing selection to maintain genetic variation in additive quantitative traits controlled by up to five loci. A statistical approach is employed by choosing randomly 4000 parameter sets (allelic effects, recombination rates, and strength of selection) for a given number of loci. For each parameter set we iterate the recursion equations that describe the dynamics of gamete frequencies starting from 20 randomly chosen initial conditions until an equilibrium is reached, record the quantities of interest, and calculate their corresponding mean values. As the number of loci increases from two to five, the fraction of the genome expected to be polymorphic declines surprisingly rapidly, and the loci that are polymorphic increasingly are those with small effects on the trait. As a result, the genetic variance expected to be maintained under stabilizing selection decreases very rapidly with increased number of loci. The equilibrium structure expected under stabilizing selection on an additive trait differs markedly from that expected under selection with no constraints on genotypic fitness values. The expected genetic variance, the expected polymorphic fraction of the genome, as well as other quantities of interest, are only weakly dependent on the selection intensity and the level of recombination.

MANY quantitative characters in natural populations are apparently subject to stabilizing selection toward an intermediate optimum (*e.g.*, Endler 1986). This means that extreme phenotypes have lower fitness than those near the population mean. Therefore, stabilizing selection is expected to exhaust genetic variation. This view has been substantiated by classical analyses based on the assumption that loci are independent and each allele contributes only an extremely small amount to the total genetic variance (Fisher 1930; Haldane 1932; Robertson 1956). Further support came from Wright's (1935) study of the so-called quadratic optimum model in which two diallelic loci contribute additively to the character whose fitness deviates in a quadratic way from its maximum value. By contrast, most quantitative traits exhibit relatively high levels of genetic variability in nature.

This apparent contradiction has been a fundamental problem in evolutionary genetics, and several mechanisms that can potentially contribute genetic variability under stabilizing selection have been proposed and investigated. In principle, variation can be maintained either by mechanisms acting directly on the considered trait or as a side effect of polymorphisms that are independent of the observed character. Among the direct mechanisms are migration, mutation, frequency-dependent selection, genotype-environment interaction, and epistasis. Considerable progress has been made in elucidating the potential and the limitations of these mechanisms to maintain genetic variation (*cf.*, Lewontin 1974; Lande 1975; Barton and Turelli 1989; Gimelfarb 1989; Falconer and Mackay 1995; Maynard Smith 1998).

However, relatively little is known about the ability of stabilizing selection *per se* to maintain polymorphisms and genetic variability in quantitative traits controlled by more than one locus acting additively. Conflicting results have been obtained, depending on model assumptions about the number of involved loci, magnitude of allelic effects, and linkage equilibrium.

All the classical analyses of additive quantitative traits assumed many loci in linkage equilibrium, with each locus having a very small effect on the trait. We review these first following Bulmer's (1971, 1980) generalization and formalization of the approaches of Fisher (1930), Haldane (1932), and Robertson (1956). The

Corresponding author: Reinhard Bürger, Institut für Mathematik, Universität Wien, Strudlhofgasse 4, A-1090 Vienna, Austria.
E-mail: reinhard.buerger@univie.ac.at

model assumes that at each locus contributing to the trait, there are two alleles with sufficiently small effects on the phenotype, such that the density function of the trait in the subpopulation with genotype A_iA_j is simply shifted by the amount g_{ij}, where g_{ij} is the deviation of the genotypic contribution of A_iA_j from the mean (*cf.* Nagylaki 1984 for proof that, in the absence of linkage disequilibrium, this approximation is correct to first order). Assuming additionally that the phenotypic distribution is normal, Bulmer (1971) showed that the polymorphic equilibria are unstable under stabilizing selection with allelic frequencies converging to either zero or one. Therefore, no genetic variability can be maintained under these assumptions.

Further support for the view that stabilizing selection alone cannot maintain genetic variation in additive traits comes from analyses of mutation-selection-balance models that assume n loci in linkage equilibrium and with mutational effects drawn from a continuous probability distribution (*cf.* Crow and Kimura 1964; Lande 1975; Fleming 1979; Turelli 1984; Bürger 1986; Turelli and Barton 1990). Bürger and Hofbauer (1994) proved for such a model under quite general assumptions that the genetic variance at any equilibrium decreases to zero as the per-locus mutation rates decrease to zero.

Wright (1935, 1952) studied a model in which a quantitative trait is determined additively by two diallelic loci with equal and symmetric effects with respect to the optimum. The fitness of the trait was assumed to decrease quadratically as the phenotype deviates from its optimum value. He showed that no stable polymorphic equilibrium could exist under these assumptions, thus supporting the view that stabilizing selection depletes genetic variation. Barton (1986) generalized Wright's quadratic optimum model to many unlinked equivalent loci and omitted the assumption that allelic effects are symmetric with respect to the optimum. He showed that for half of the possible positions of the optimum (within the range of genotypic values) selection maintains no genetic variation at all, while for the other half of the positions variation is maintained at exactly one locus.

In general, it might be expected that in a genetic system with very many loci contributing additively to a trait and the effects of alleles varying among loci, the optimum phenotype can be matched closely by the genotypic value of some homozygote and, hence, either no genetic variation will be maintained by such a system or at most a polymorphism will be maintained at just one locus.

In contrast to the classical view and to the results discussed above, several analyses of two-locus systems that allowed for linkage or nonequivalent loci arrived at very different conclusions. Gale and Kearsey (1968) and Kearsey and Gale (1968) observed that in a two-locus diallelic additive model with a triangular fitness function, both loci can be stably polymorphic if their effects are sufficiently diverse. They also noted that the amount of diversity required to maintain the polymorphism decreases as linkage becomes tighter. Nagylaki (1989) investigated two-locus diallelic models of stabilizing selection under the assumption of linkage equilibrium, but with rather general fitness functions that have their optimum at the value of the double heterozygote and decrease monotonically and symmetrically from the optimum. For a large class of such fitness functions, both loci may be stably polymorphic if the ratio of the effect of the "major" locus to that of the "minor" locus exceeds a critical value. Gavrilets and Hastings (1993, 1994) investigated an extended version of Wright's (1935) quadratic optimum model that includes linkage and disposes of the assumption of equivalent loci. They showed that for sufficiently tight linkage and unequal allelic effects, stable two-locus polymorphisms may exist. Thus, several analyses of two-locus systems suggest that substantial amounts of genetic variation can be maintained under stabilizing selection.

A number of interesting and unresolved questions emerge from the above results: How many loci have to contribute to an additive quantitative trait in order for no (or almost no) genetic variance to be maintained at equilibrium by stabilizing selection *per se*? How does this depend on the distribution of allelic effects and the strength of stabilizing selection? What is the role of linkage? More generally, what can be said about the equilibrium structure of multilocus systems under stabilizing selection on an additive quantitative trait?

In this article, we address these questions by studying diallelic two-, three-, four-, and five-locus models with arbitrary recombination and allelic effects varying between loci. We assume the quadratic optimum model in which the optimum coincides with the totally heterozygous genotype. For two loci, the equilibrium structure can be determined explicitly, while for three, four, and five loci, we follow the approach of Gimelfarb (1998) and perform numerical iterations of the recursion equations for large sets of randomly chosen recombination rates and allelic effects. In this way, we obtain information on how the expected equilibrium properties under stabilizing selection depend on the number of loci and can compare the properties of multilocus systems under stabilizing selection with the properties of systems having the same number of loci but randomly assigned genotypic fitness, *i.e.*, under selection without constraints on the genotypic fitnesses.

THE GENERAL MODEL

In an infinite, randomly mating diploid population, a quantitative character that is controlled additively by n diallelic loci is considered. The contribution of one allele at each locus ℓ is zero, whereas the contribution, β_ℓ, of the other allele is a random number between zero and one. It is assumed that the minimum and maximum genotypic values are always zero and one. Therefore,

the actual contribution by the second allele at locus ℓ is scaled to be $\alpha_\ell = \frac{1}{2}\beta_\ell / \Sigma_{k=1}^{n}\beta_k$. This implies that the genotypic value of the total heterozygote is always $\frac{1}{2}$, and the average allelic effect among the n loci controlling the trait is $\bar{\alpha} = 1/(2n)$. Environmental variance is ignored, so that genotypic values and phenotypic values are identical. We assume the quadratic optimum model; *i.e.*, the fitness of an individual with genotypic value G is

$$W(G) = 1 - s(G - \tfrac{1}{2})^2. \qquad (1)$$

The strongest possible selection occurs for $s = 4$, when the minimum and maximum phenotypes are rendered lethal. This normalization has the advantage that the strength of selection on genotypes can be compared for different numbers of contributing loci.

Let gametes be designated by i, their frequencies among zygotes in consecutive generations by p_i and p_i', and the fitness of a zygote consisting of gametes j and k by W_{jk}. Further, let $R(i|j,k)$ denote the probability that a randomly chosen gamete produced by a (j,k) individual is i. The function R is determined by the pattern of recombination between loci. With these ingredients, the well-known system of recurrence equations describing the dynamics of the distribution of gametes under viability selection followed by recombination is given by

$$p_i' = \frac{1}{\bar{W}} \sum_j \sum_k p_j p_k W_{jk} R(i|j,k), \qquad (2)$$

where $\bar{W} = \Sigma_{j,k} p_j p_k W_{jk}$ denotes mean fitness.

ANALYTIC THEORY FOR TWO LOCI

For two loci, fairly complete analytical results of the above general model can be obtained. If the alleles at the first and second locus are denoted by A_1, A_2, and B_1, B_2, respectively, the genotypic values of the four possible gametes, A_1B_1, A_1B_2, A_2B_1, A_2B_2 are 0, α_2, α_1, and $\alpha_1 + \alpha_2 = \frac{1}{2}$, respectively. For definiteness, we assume $\alpha_1 \geq \alpha_2$ and refer to these loci as major and minor, respectively. With the fitness assignments as above, the fitness values of the nine possible genotypes are given by

$$\begin{array}{c} \\ A_1A_1 \\ A_1A_2 \\ A_2A_2 \end{array} \begin{array}{c} \begin{array}{ccc} B_1B_1 & B_1B_2 & B_2B_2 \end{array} \\ \begin{pmatrix} 1-d & 1-b & 1-a \\ 1-c & 1 & 1-c \\ 1-a & 1-b & 1-d \end{pmatrix} \end{array}, \qquad (3)$$

where $a = s(\alpha_1 - \alpha_2)^2$, $b = s\alpha_1^2$, $c = s\alpha_2^2$, $d = \frac{1}{4}s$. This is the symmetric viability model that has been intensively studied (see Karlin and Feldman 1970 for an extensive equilibrium analysis and further references). The parameters satisfy the additional relation $a + d = 2(b + c)$, which is of technical importance. Wright (1935, 1952) was the first to investigate this model of stabilizing selection. More recently, Gavrilets and Hastings (1993, 1994) calculated all equilibria and explored the transient dynamics of the mean genotypic value and the genetic variance.

Our aim here is to study not only the equilibrium structure of this model, but in particular the amount of genetic variance that can be maintained through selection, as well as the magnitude of further quantities, like mean fitness, deviation of the mean phenotype from the optimum, and linkage disequilibrium. Proofs of the results presented below are given in the appendix, and a graphical representation is shown in Figure 1.

There are four types of equilibria: (i) a fully polymorphic equilibrium that is symmetric with all allele frequencies equal to $\frac{1}{2}$ (in the appendix, this is denoted by F1); (ii) a pair of fully polymorphic equilibria that, in the language of the symmetric viability model, are called unsymmetric because their coordinates satisfy no simple symmetry relations (they are denoted by F2 and F3); (iii) a pair of equilibria (denoted F4 and F5) with the major locus polymorphic and the minor locus fixed; (iv) two monomorphic equilibria corresponding to fixation of A_1B_2 or A_2B_1 (these are denoted by F6 and F7).

The existence and stability of these equilibria depend on the relation between the parameters α_2/α_1 and r/s, where r is the recombination rate. For consistency with the multilocus results, we use the standard deviation of allelic effects, $\sigma_\alpha = \frac{1}{2}(\alpha_1 - \alpha_2) = \frac{1}{4} - \alpha_2$, as a measure for the disparity of effects. An important value is $\sigma_\alpha = \frac{1}{12}$, which is equivalent to $\alpha_1 = 2\alpha_2$. The ranges of stability are displayed in the bottom of Figure 1 for $r/s \leq \frac{1}{8}$, which covers the full range of parameters if selection is as strong as possible ($s = 4$). For weaker selection, larger values of r/s can occur. Fully polymorphic equilibria can be stable only if $r/s \leq \frac{1}{12}$. Only one class of equilibria can be stable for a given parameter combination. The unsymmetric equilibria are stable whenever they exist, while the other equilibria always exist but may be unstable.

For each equilibrium, mean fitness, absolute deviation of the mean phenotype from the optimum, genetic variance, and linkage disequilibrium can be calculated (see appendix). For a range of recombination rates, the top and middle parts of Figure 1 display the equilibrium genetic variance and the deviation of the mean phenotype from the optimum as function of σ_α.

The following is a summary of the main findings:

1. Substantial genetic variance is maintained for any recombination rate if the effects of the alleles at the two loci are sufficiently different ($\sigma_\alpha > \frac{1}{12}$), in which case one or both loci, depending on the recombination rate, are stably polymorphic. The reason is that no homozygote genotype is close to the optimum and, from double heterozygotes, a high recombination rate would generate many complete homozygotes with extreme phenotype and low fitness. Therefore, if recombination is strong relative to selection, the minor locus is fixed and the major locus is kept

polymorphic by overdominance. With tight linkage, both gametes A_1B_2 and A_2B_1 are maintained at high frequency.

2. If the effects of the two loci are not very different ($\sigma_\alpha < 1/12$), very tight linkage ($r < 4s\sigma_\alpha^2$) is necessary to maintain genetic variation. In this case, a homozygote ($A_1A_1B_2B_2$ or $A_2A_2B_1B_1$) is close to the optimum and, therefore, fixed unless linkage is extremely tight.
3. For any value of r, the equilibrium genetic variance decreases from its maximum value $\frac{1}{2}(\alpha_1^2 + \alpha_2^2) = 1/8$ to 0 as σ_α decreases from its maximum value $1/4$ to 0.
4. For fixed σ_α, the genetic variance has a maximum value for intermediate recombination rates. If the effects are not very different ($\sigma_\alpha < 1/12$), then the variance is lower under loose linkage than under very tight linkage, while the reverse is true if the effects are sufficiently different. Thus, tighter linkage may increase the degree of polymorphism at equilibrium but at the same time reduce the genetic variance.
5. The mean phenotype evolves rapidly to its equilibrium value (Hastings 1987). However, for unequal effects the equilibrium does not coincide with the optimum unless linkage is extremely tight. Convergence of the genetic variance and of the genotype frequencies may be very slow (Gavrilets and Hastings 1993, 1994). Therefore, if there is substantial initial genetic variability and the selection regime is such that it eventually depletes genetic variability, this process may be very slow even for relatively strong selection.

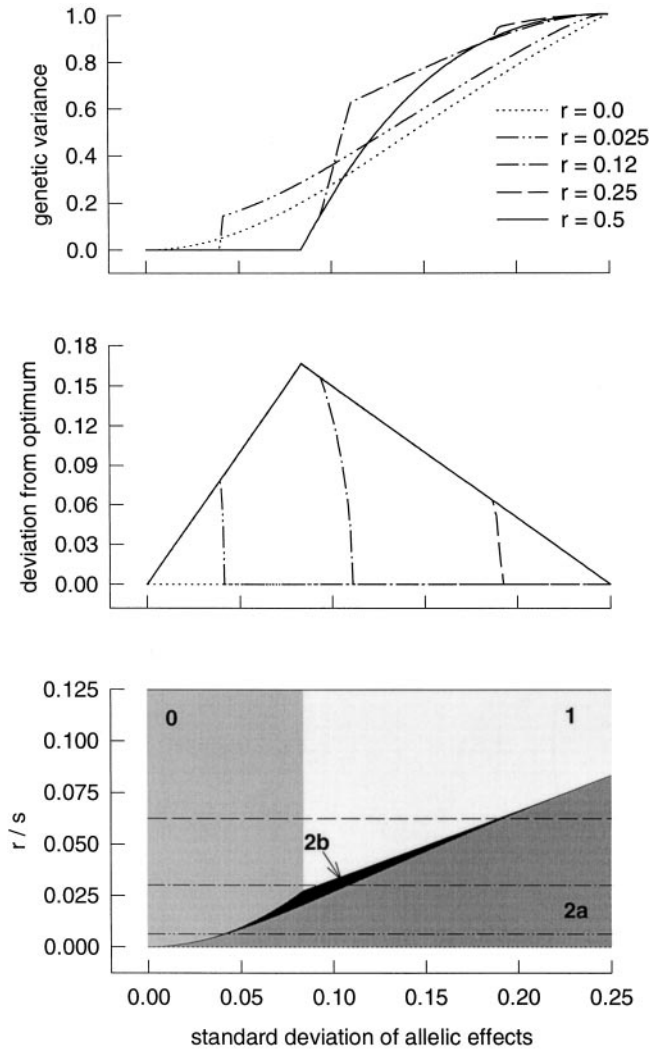

Figure 1.—Equilibrium properties of two-locus two-allele systems under stabilizing selection. (Top) The ratio of the equilibrium genetic variance, $\hat{\sigma}_G^2$ to the maximum possible variance, $V_{max} = \frac{1}{2}(\alpha_1^2 + \alpha_2^2)$, as a function of the standard deviation of allelic effects, $\sigma_\alpha = 1/4 - \alpha_2$, for five different rates of recombination and $s = 4$. (Middle) The corresponding absolute deviation of the mean from the optimum. (Bottom) The regions of stability of the four types of equilibria. The following abbreviations are used for the regions of stability. 0, two monomorphic equilibria; 1, two equilibria with one (the major) locus polymorphic; 2a, the symmetric equilibrium with both loci polymorphic; 2b, two asymmetric equilibria with both loci polymorphic. The upper boundary of region 2a is given by r_1/s (A9) and the upper boundary of 2b by r_2/s (A10). At σ_α, we have $r_1/s = r_2/s = 1/12$. The boundary between the regions 0 and 1 is at $\sigma_\alpha = 1/12$ and intersects with region 2b at $r/s = 1/36$. The horizontal lines correspond to the curves in the top and middle parts. It should be noted that the areas indicating the different types of equilibria are not proportional to their respective probabilities, because the horizontal scale is σ_α. It may be noted that $0 \le r/s \le 1/8$ covers the whole range only if $s = 4$. For smaller s (weaker selection), the maximum value of r/s is larger than $1/8$; hence the proportion of parameter values yielding stable two-locus polymorphisms becomes correspondingly smaller.

THE STATISTICAL APPROACH

Usually, parameters of genetic systems controlling quantitative traits are unknown or can be inferred only indirectly. In addition, because the dimensionality of the parameter space and of the space of gamete frequencies increases rapidly as the number of loci increases, an explicit and analytical characterization of the equilibrium properties of multilocus models in terms of all parameters and initial conditions would be of limited value, even if the necessary analytical methods were available. Therefore, we used a different approach to evaluate the quantities of interest for randomly chosen parameter sets and initial conditions and, consequently, to obtain statistical results.

For a genetic system with a given number of loci ($n = 2, 3, 4, 5$), we constructed 4000 parameter sets (allelic effects of loci, coefficients of recombination between adjacent loci, the strength of stabilizing selection). For each parameter set, allelic effects were obtained by generating values β_ℓ ($= 1, 2, \ldots, n$) as independent random variables uniformly distributed

between 0 and 1 and transforming them into the actual allelic effects $\alpha_\ell = \frac{1}{2}\beta_\ell / \Sigma_k \beta_k$. The strength of stabilizing selection, s, was obtained as a random variable uniformly distributed between 1 and 4, with $s = 4$ corresponding to the strongest possible quadratic selection (the fitness of extreme phenotypes is zero), while $s = 1$ represents "weak" selection (the fitness of extreme phenotypes is 75% of the fitness of the best fit phenotype). On the basis of such obtained values of allelic effects and the strength of selection, genotypic fitnesses, W_{ij}, were calculated and substituted into the recursion equations (2). Recombination coefficients between adjacent loci, $r_{\ell,\ell+1}$ ($\ell = 1, \ldots, n-1$), were obtained as independent random variables (no interference) uniformly distributed between 0 and 0.5. For four-locus genetic systems, we also constructed 4000 parameter sets with random allelic effects and random recombination, but with "strong" ($s = 4$) and weak ($s = 1$) selection, as well as with random allelic effects and random selection but with free recombination, $r_{\ell,\ell+1} = \frac{1}{2}$. In addition, we constructed 4000 parameter sets for two-, three-, four-, and five-locus systems with random recombination between adjacent loci and with genotypic fitness values, W_{ij}, chosen as independent random variables uniformly distributed between 0 and 1, as described by Gimelfarb (1998).

For each of the 4000 parameter sets, the recursion equations (2) were numerically iterated starting from 20 random initial distributions of gametes. To make the initial distributions more evenly distributed in the gametic space, they were chosen such that the (Euclidean) distance between any two of them was no less than a predetermined value (0.25, 0.30, 0.35, and 0.30 for 2, 3, 4, and 5 loci, respectively). Starting from an initial distribution, the recursion equations (2) were iterated until either equilibrium was reached or the number of iterations exceeded 300,000. In the latter case, the parameter set was excluded from the analysis. For each parameter set, the number of different equilibria, the gametic frequencies at each equilibrium, and the number of trajectories (initial distributions) converging to each equilibrium were recorded. Using this database, the equilibrium properties of multilocus genetic systems were analyzed.

For two-locus genetic systems, results can be obtained not only by the statistical approach, but also by numeric integration of equations in the appendix, as well as by placing a fine grid over the parameter space and averaging over the appropriate quantities whose equations are given in the appendix. All three methods yielded, within statistical and numeric accuracy, identical results.

STATISTICAL RESULTS

Results obtained by the statistical approach are summarized in Tables 1–3 and in Figures 2 and 3.

Table 1 shows the proportion of parameter sets for which there was at least one equilibrium with a given number of polymorphic loci (*e.g.*, among 4000 of four-locus systems under stabilizing selection, 68% yielded at least one monomorphic equilibrium, while only 1% of such sets yielded at least one equilibrium with two polymorphic loci). Also shown in Table 1 are the proportions of trajectories that, starting from a random initial distribution, converge to an equilibrium with a given number of polymorphisms (*e.g.*, 58% of trajectories for four-locus genetic systems under stabilizing selection converged to a monomorphic equilibrium).

Table 2 presents some of the parameters characterizing stable equilibria that are expected for genetic systems with a given number of loci: the number of different stable equilibria per parameter set; the polymorphic fraction of the genome (the probability for a locus to be polymorphic); and the mean fitness. For stabilizing selection, Table 2 also presents the expected deviation of the mean from the optimum; the genetic variance at equilibrium; the ratio of the genetic variance at equilibrium to the maximum genetic variance that can be maintained by the given genetic system; and the ratio of the allelic effect at a polymorphic locus to the average allelic effect among all loci in the system. The maximum genetic variance that can be maintained in linkage equilibrium by an additive trait controlled by n loci with allelic effects $\{\alpha_\ell\}$ is $V_{\max} = \frac{1}{2}\Sigma_\ell \alpha_\ell^2$, while the average allelic effect among loci is $\bar{\alpha} = 1/(2n)$.

Table 3 compares parameters of stable equilibria for four-locus genetic systems with different recombination rates and under stabilizing selection of different strength.

Equilibrium structure: The data in Table 1 demonstrate that, while the probability for a genetic system under stabilizing selection to maintain a monomorphism or a one-locus polymorphism is higher for systems with more loci, the probability of maintaining more than one polymorphic locus drops rapidly for systems with more loci. In fact, for a five-locus genetic system, the probability of maintaining a polymorphism in two or more loci is negligible. The same is not true under selection with randomly assigned fitnesses, in which case the probability of maintaining more than one locus polymorphic is higher for genetic systems with more loci. In addition, Table 2 shows that with increasing number of loci, the polymorphic fraction of the genome decreases at a much higher rate under stabilizing selection than under selection with random fitness. Indeed, as a function of n, the polymorphic fraction of the genome under stabilizing selection is very closely approximated by $0.99n^{-1.58}$, while for random fitnesses it decreases proportionally to $n^{-0.45}$.

For two loci, we proved in the appendix that different types of equilibria cannot stably coexist (except for $r = 0$), and the maximum number of coexisting stable equilibria is two. It can be noted that the sums of the entries

TABLE 1

Equilibrium structure

Number of loci	Number of polymorphisms	Stabilizing selection		Random fitnesses	
		Sets	Trajectories	Sets	Trajectories
2	0	0.48	0.48	0.70	0.50
	1	0.35	0.35	0.56	0.40
	2	0.17	0.17	0.11	0.10
3	0	0.52	0.52	0.75	0.43
	1	0.47	0.45	0.71	0.40
	2	0.03	0.03	0.29	0.15
	3	0.00	0.00	0.03	0.02
4	0	0.68	0.58	0.77	0.38
	1	0.64	0.41	0.80	0.41
	2	0.01	0.01	0.46	0.18
	3	0.00	0.00	0.09	0.03
	4	—	—	0.01	0.00
5	0	0.83	0.59	0.75	0.34
	1	0.75	0.41	0.83	0.41
	2	0.01	0.00	0.54	0.20
	3	—	—	0.15	0.05
	4	—	—	0.01	0.00
	5	—	—	—	—

Proportion of parameter sets yielding a stable equilibrium maintaining a given number of polymorphisms, and the proportion of trajectories converging to such an equilibrium under stabilizing selection of random strength, $1.0 \leq s \leq 4.0$, and under selection with random genotypic fitnesses (recombination is random; 0.00 entry indicates that the corresponding proportion is <0.005, whereas a dash indicates that the corresponding number of polymorphisms was not observed).

in the column "sets" for stabilizing selection in Table 1 are 1.00, 1.02, 1.33, and 1.59 for 2, 3, 4, and 5 loci, respectively. This indicates that the probability of several simultaneously stable equilibria with different degrees of polymorphism, while zero for 2 loci and very low for 3 loci, becomes substantial if the number of loci increases. Also, Table 2 shows that with increasing number of loci in the genetic system, the expected number of simultaneously stable equilibria increases faster under stabilizing selection than with randomly assigned fitnesses, and for 4 and 5 loci, it is actually higher under stabilizing selection. The maximum numbers of stably

TABLE 2

Stable equilibria and their genetic variability

	Stabilizing selection (Number of loci)				Random fitnesses (Number of loci)			
	2	3	4	5	2	3	4	5
Number of stable equilibria[a]	1.85	2.35	3.87	5.52	1.90	2.72	3.53	3.88
Polymorphic fraction of genome	0.34	0.17	0.11	0.08	0.30	0.26	0.22	0.20
Mean fitness	0.89	0.98	0.99	0.99	0.87	0.87	0.87	0.87
Deviation from optimum	0.07	0.05	0.03	0.02				
Genetic variance	0.031	0.007	0.002	0.001				
Genetic variance / maximum[b]	0.27	0.11	0.05	0.03				
Effect of polymorphic locus[c]	1.48	1.06	0.99	0.85				

Parameters of stable equilibria expected for genetic systems with a given number of loci under stabilizing selection of random strength, $1.0 \leq s \leq 4.0$, and under selection with random genotypic fitnesses (recombination is random). Entries are averages over all trajectories unless indicated otherwise.

[a] Average over all parameter sets.

[b] Ratio of equilibrium genetic variance to the maximum genetic variance for a given genetic system.

[c] Ratio of the average allelic effect among polymorphic loci to the average allelic effect among all loci.

TABLE 3

The role of selection intensity and recombination

	s = weak r = random	s = strong r = random	s = random r = random	s = random r = 0.5
Number of stable equilibria[a]	3.94	3.76	3.87	3.97
Polymorphic fraction of genome	0.11	0.11	0.11	0.11
Mean fitness	0.99	0.98	0.99	0.98
Deviation from optimum	0.03	0.03	0.03	0.03
Genetic variance/maximum[b]	0.05	0.05	0.05	0.05
Effect of polymorphic locus[c]	1.00	0.97	0.99	1.00

Effect of strength of stabilizing selection, s, and of recombination, r, on the expected parameters of equilibria for four-locus genetic systems (weak selection, $s = 1.0$; strong selection, $s = 4.0$; random selection, $1.0 \leq s \leq 4.0$). Entries are averages over all trajectories unless indicated otherwise (no equilibria with more than one polymorphism were observed under random selection and $r = 0.5$).

[a] Average over all parameter sets.

[b] Ratio of equilibrium genetic variance to the maximum genetic variance for a given genetic system.

[c] Ratio of the average allelic effect among polymorphic loci to the average allelic effect among all loci.

coexisting equilibria detected by numerical iteration starting with 20 initial distributions were 6, 12, and 15 for $n = 3$, 4, and 5, respectively, and there was exactly 1 polymorphic locus at all these equilibria. These results indicate that historic effects may be of paramount importance in the evolution of populations under stabilizing selection.

Because recombination rates between adjacent loci in our parameter sets were generated as independently and uniformly distributed random variables, the probability of obtaining an n-locus system with all loci tightly linked decreases exponentially as n increases. Therefore, we also performed some iterations for four-locus genetic systems with completely linked loci. All trajectories converged to fully polymorphic equilibria. However, only two complementary types of gametes, *e.g.*, 0110 and 1001, were present at these equilibria and at equal frequency. Such equilibria correspond to the symmetric equilibrium F1 in the two-locus case. Sometimes convergence to such equilibria was extremely slow. By perturbation arguments (Karlin and McGregor 1972), these findings should extend to sufficiently small recombination rates.

Linkage disequilibrium: The average amount of linkage disequilibrium was calculated for equilibria with at least two polymorphic loci and was found to be very small (data not shown). Given that the probability of a stable equilibrium with two or more polymorphic loci is very small if more than two loci control the trait, linkage disequilibrium can be neglected in multilocus systems under stabilizing selection.

Strength of selection and recombination: As is well known and shown by Figure 1, for two loci the equilibrium structure depends on the relative magnitude, r/s, of recombination rate and selection intensity. For a given standard deviation of allelic effects, σ_α, the equilibrium genetic variance is not very sensitive to changes in the recombination rate (relative to the selection intensity) except for small σ_α and very small r. Furthermore, for any particular value of s, the equilibrium genetic variance, V_G, averaged over all trajectories and all parameter sets $\{r, \alpha_1, \alpha_2\}$ is almost independent of s. It decreases from $V_G = 0.0311$ for $s = 4$ to $V_G = \ln 3/2 - 3/8 \approx 0.0305$ as $s \to 0$. However, as shown by Figure 1 and discussed in analytic theory for two loci, the number of polymorphic loci at a stable equilibrium is strongly effected by r/s.

Employing the statistical approach, we investigated the effect of the strength of stabilizing selection and of recombination in the case of four loci. The results are summarized in Table 3 and demonstrate that, indeed, many expected equilibrium properties, notably the average number of equilibria, the average degree of polymorphism, and the average genetic variance at equilibrium, are almost independent of selection intensity and recombination rate. This is further supported by additional iterations with $s = 4$ and $r = 0.5$ (not shown). Presumably, this is so because most genetic variance comes from single-locus polymorphisms that are stable because of overdominance. In this case, allele frequencies are independent of s but depend only on the relative disadvantage of the two homozygotes. Different results are observed only if all loci are completely (or very tightly) linked, in which case, as mentioned above, all trajectories converge to fully polymorphic equilibria with gametes of only two complementary types present.

The explored values of selection intensities ($1 \leq s \leq 4$) cover the range from weak to the strongest possible selection but not extremely weak or no selection. Inclusion of values $s < 1$ would change our results very little because it would not much increase our parameter range and because all our results indicate that the equilibrium structure is nearly unaffected by changes in s unless selection is very strong relative to recombination. In addition, Tables 2 and 3 show that even strong selection imposes only a small genetic load on the popula-

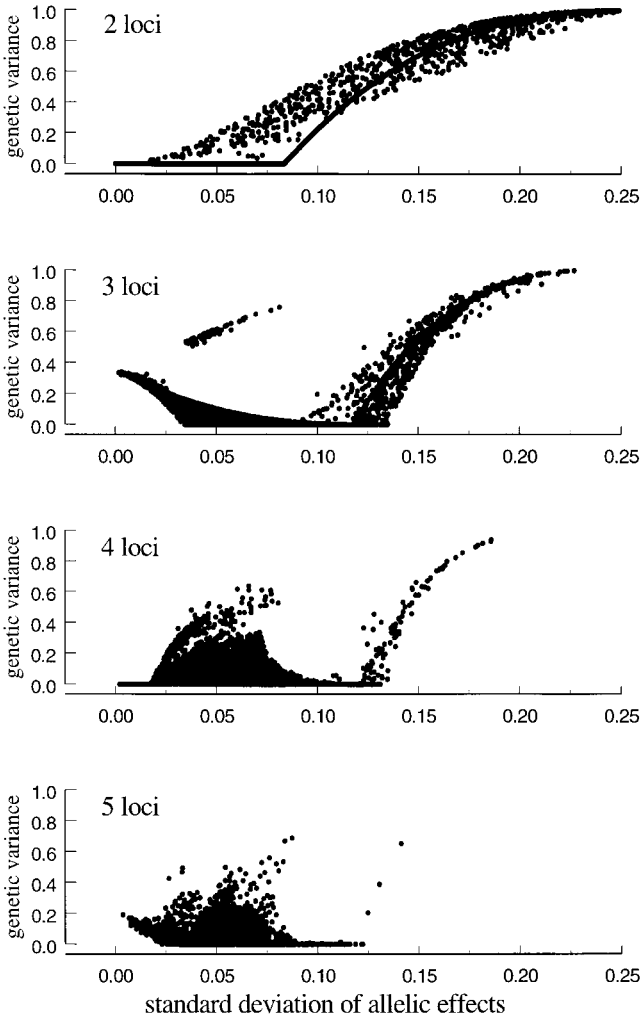

Figure 2.—Ratio of the expected equilibrium genetic variance, V_G, to the maximum genetic variance under linkage equilibrium, V_{max}, as a function of the standard deviation of allelic effects, σ_α. Each data point represents the average over all trajectories for the given parameter set. Allelic effects, recombination rates between adjacent loci, and strength of stabilizing selection are independently drawn from uniform distributions as described in the statistical approach.

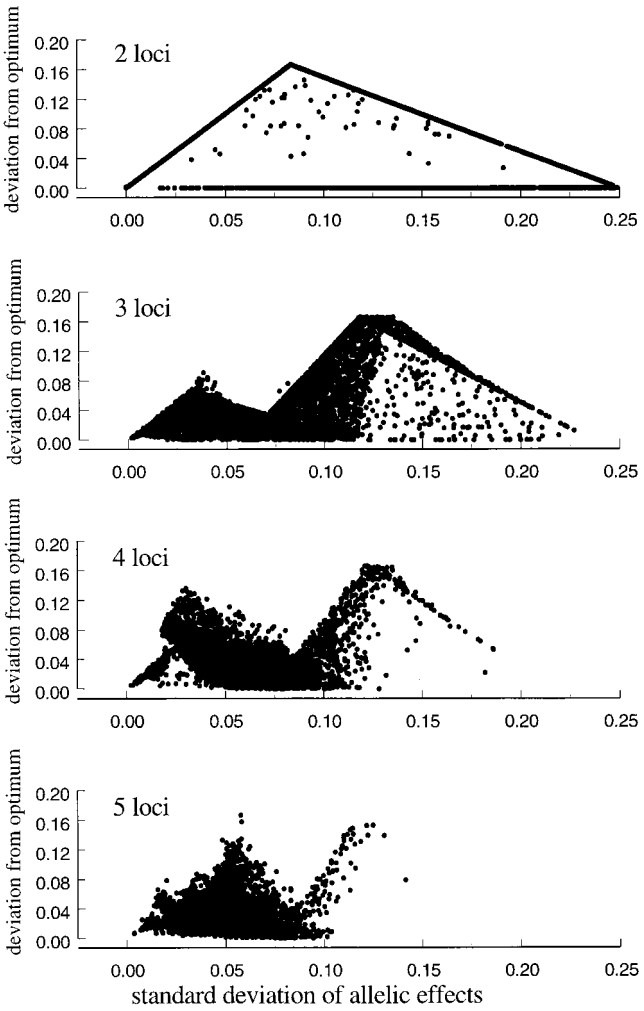

Figure 3.—Deviation of the expected mean genotypic value from the optimum as a function of the standard deviation of allelic effects, σ_α.

tion. Thus, purely on the basis of observation, one might conclude that these populations are under weak selection even if, in fact, selection is strong.

Deviation from the optimum: In general, the mean genotypic value deviates from the optimum. This deviation is displayed in Figure 3. It may be substantial for certain sets of parameters (see also below). However, as shown in Table 2, the average deviation from the optimum decreases as the number of loci increases. This is due to the fact that the optimum can be matched more closely in systems with more loci of varying effects. The present data quantify this statement and show that the deviation from the optimum declines slightly faster than the average allelic effect $\bar{\alpha}$. The ratio of the expected deviation from the optimum to the average allelic effect is 0.29, 0.31, 0.26, and 0.22 for $n = 2$, 3, 4, and 5, respectively.

Genetic variance: Given that our scaling is such that the average allelic effect, $\bar{\alpha} = 1/(2n)$, is smaller for systems with more loci controlling a trait, it is expected that the genetic variance at equilibrium, $\bar{V}_G(n)$, averaged over all trajectories and parameter sets (s, $\{r_\ell\}$, $\{\alpha_\ell\}$), must decline with increasing number of loci, n. Table 2 shows that this indeed is the case. However, the decline occurs at a rate much faster than expected. To a close approximation, we have $\bar{V}_G(n) = 0.52n^{-4.0}$. If the genetic variance is scaled relative to the maximum genetic variance that can be maintained for a given genetic system in linkage equilibrium, $V_{max} = \frac{1}{2}\Sigma_\ell\alpha_\ell^2$, the resulting overall average value of $V_G(n)/V_{max}$ [with $V_G(n)$ being the variance for a given parameter set, averaged over all trajectories] still decreases in proportion to $n^{-2.4}$.

What is the reason for such a fast decline? First, we have already seen that under stabilizing selection the polymorphic fraction of the genome decreases with increasing number of loci at a rate of approximately $n^{-1.58}$, which is much higher than under selection with random fitnesses. Second, if a locus maintains a polymorphism,

i.e., contributes to the genetic variance, the allelic effect of such a locus (as compared to $\bar{\alpha}$, the average effect among all loci) is smaller for systems with more loci. If a trait is controlled by two loci then, as we have seen above, it is either the major locus or both loci that are segregating at a polymorphic equilibrium. If $n = 3$ or 4, Table 2 shows that the effect of a polymorphic locus is expected to be approximately equal to the average allelic effect, $\bar{\alpha}$, whereas if $n = 5$, it is expected to be smaller than the average ($0.85\bar{\alpha}$). For comparison, if $n = 5$, the expected minimum allelic effect among polymorphic loci is $0.4\bar{\alpha}$, while the expected maximum effect is $1.6\bar{\alpha}$. It is also interesting to note that, in contrast to the two-locus case, for $n \geq 3$ the expected genetic variance at an equilibrium is higher if this equilibrium maintains several loci polymorphic. For example, with $n = 4$, the expected genetic variance at equilibria with two polymorphic loci is almost four times as large as at equilibria with one polymorphic locus (data not shown).

Figure 2 displays the ratio V_G/V_{max} averaged over all trajectories for a given parameter set as a function of the standard deviation, σ_α, of allelic effects in the given set. The maximum genetic variance is attained if σ_α is maximal, which is the case if one locus has maximum effect ($= \frac{1}{2}$), while the others have zero effect. The maximum value of σ_α is $\sigma_{\alpha,max} = \sqrt{(1/4n)(1 - 1/n)}$, which is 0.25, 0.236, 0.217, 0.2 for $n = 2, 3, 4, 5$, respectively. If σ_α is maximum, then both alleles at the major locus are segregating at equal frequency and the mean phenotypic value coincides with the optimum.

Let us next discuss the properties of stable equilibria, in particular the equilibrium genetic variance and the deviation of the mean from the optimum as affected by the diversity of allelic effects measured by σ_α. First, we concentrate on two-locus systems with loci not very tightly linked ($r/s > \frac{1}{36}$, which is the value at which the regions 0, 1, and 2b in the bottom of Figure 1 intersect). In such a system, the major locus or both loci are segregating if $\frac{1}{12} \leq \sigma_\alpha \leq \frac{1}{4}$. The genetic variance decreases from its maximum value to zero, as σ_α decreases from $\frac{1}{4}$ to $\frac{1}{12}$, while the deviation from the optimum increases from zero and to its maximum value at $\sigma_\alpha = \frac{1}{12}$ (Figure 1). If $\sigma_\alpha < \frac{1}{12}$, then only monomorphic equilibria are stable unless linkage is very tight. The smaller σ_α is, *i.e.*, the more similar effects of loci are, the more closely the optimum can be matched by a completely homozygous genotype (the one that gives rise to the stable equilibrium). Therefore, the deviation from the optimum converges to 0 as $\sigma_\alpha \to 0$.

On the basis of this explanation, the two-locus panel of Figure 2 is easily interpreted. The thick line of densely packed points represents parameter sets with loose linkage because, for given σ_α, all two-locus systems with sufficiently large r/s have the same stable equilibria. Tightly linked loci are represented by the scattered dots.

With more than two loci, the complexity increases. Figures 2 and 3 show that there is no monotone relation between the standard deviation of allelic effects and the genetic variance. However, they also indicate that for sufficiently large σ_α, the equilibrium behavior is similar to the two-locus case. As σ_α decreases from $\sigma_{\alpha,max}$, the effect of the major locus decreases from its maximum value of $\frac{1}{2}$, and alleles with smaller effects are present at other loci. Therefore, the genetic variance declines. In analogy with the two-locus case, we suspect that if the loci are loosely linked, then it is the major locus that is polymorphic while the other loci are monomorphic. In the parameter range below $\sigma_{\alpha,max}$, the mean increasingly deviates from the optimum as is illustrated by the almost straight lines ascending from $\sigma_{\alpha,max}$ to the left in the two-, three-, and four-locus panels of Figure 3. The deviation is caused by the asymmetries introduced by the monomorphic loci, because for such values of σ_α, no homozygote genotype will closely match the optimum. The scattered points below these lines represent parameter sets with linked loci. For sufficiently tight linkage, two or more loci may be polymorphic and, as in the two-locus case, at such equilibria the mean genotypic value may be close to or coincide with the optimum. This, however, is rarely the case if the number of loci is larger than three (see Table 1 and Figure 3).

If the number of loci is greater than two, a problem arises in that the same value of σ_α can be caused by different sets of allelic effects, and the smaller σ_α is, *i.e.*, the more similar the effects of loci are, the more important this problem becomes. As σ_α decreases further, the genetic variance declines and eventually becomes 0. As shown in Figure 2, there is an interval of values of σ_α where this happens (for $n = 3, 4, 5$, this interval is contained in the range $0.11 < \sigma_\alpha < 0.14$). It is the same interval where the maximum deviation of the mean from the optimum occurs (*cf.* Figure 3). For a range of smaller σ_α values, the figures indicate that most stable equilibria are monomorphic and that the deviation of the mean now decreases. The few data points in this range of σ_α values that have a positive genetic variance come from tightly linked loci; their number decreases as n increases. As σ_α decreases further (below ≈ 0.1), the situation becomes more complicated as can be seen from the figures: obviously, for given σ_α, there are now many different types of stable equilibria; some maintain high levels of genetic variance (because they are polymorphic in more than one locus or because an allele of large effect is segregating), while most maintain relatively little variance. There also emerges another peak of the mean's deviation from the optimum.

Finally, if all loci have very similar effects, *i.e.*, if σ_α is close to zero, the equilibrium structure is simple again. If the number of loci is even, then all loci are fixed because there exists a homozygote genotype very close to the optimum, and V_G/V_{max} converges to 0. If, on the other hand, the number of loci is odd, then one of the (almost equivalent) loci is segregating, and the genetic variance converges to $1/(8n)^2$, *i.e.*, V_G/V_{max} converges to $1/n$, as observed in Figure 2. In both cases, the mean

coincides with the optimum because of the symmetries of the model.

DISCUSSION

While previous analyses, assuming a quantitative trait controlled additively by a large number of loci of individually very small effects, indicated that stabilizing selection on such a character depletes genetic variation, several two-locus analyses showed the contrary (see Introduction). Because analyses of the first kind are based on assumptions such as linkage equilibrium and equivalent loci, it has not been obvious if their conclusions would still be valid without these assumptions. We demonstrated that under our model assumptions, which include linkage and loci with different effects, the expected genetic variance maintained under stabilizing selection declines very rapidly from a high value in two-locus systems to an extremely low value in five-locus systems, thus providing quantitative support for the classical results. In particular, already four- and five-locus systems exhibit equilibrium properties as expected under the infinitesimal model. In addition, with more than three loci, the probability of equilibria involving at least two polymorphic loci is almost negligible, implying that in such systems no linkage disequilibrium is to be expected. A further interesting finding is that many quantities, *e.g.*, the average number of stable equilibria, the average polymorphic fraction of the genome, the average deviation of the mean phenotype from the optimum, and the average genetic variance are virtually independent of the strength of stabilizing selection and the level of recombination.

What are the implications of these findings for the classical problem of how genetic variation is maintained in traits that apparently experience stabilizing selection? Obviously, our results suggest that it is very unlikely, although not impossible, that appreciable levels of genetic variation are maintained at equilibrium by the selective forces resulting from direct stabilizing selection if the loci controlling the trait act additively. From this, however, no support can be deduced for any of the theories about the maintenance of genetic variation, be it by mutation-selection balance or by pleiotropic effects of polymorphisms maintained independently of the observed character.

First, our model provides a kind of null hypothesis in assuming that allelic effects at loci as well as recombination rates between "adjacent" loci are drawn from a uniform distribution. There is empirical evidence that allelic effects have a highly leptokurtic distribution (Mackay *et al.* 1992; López and López-Fanjul 1993), and it has been suggested that a (reflected) gamma distribution may be adequate to approximate distributions of mutational effects (Hill and Rasbash 1986). How would this change our results? Probably just a little. With such a distribution, the expected standard deviation of allelic effects is higher than under the present uniform distribution, while the mean, $\bar{\alpha}$, remains unchanged. [For instance, with $n = 5$, the expected σ_α for uniformly distributed β_ℓ is 0.057, while for exponentially distributed β_ℓ, it is 0.082. For $n \to \infty$, it can be shown that the expected σ_α converges to $\sigma_\beta/(2n\bar{\beta})$, where σ_β is the standard deviation of the distribution of the β_ℓ, and $\bar{\beta}$ is its mean.] Therefore, there would be more data points in Figures 2 and 3 with larger σ_α; in particular, it would be more likely to obtain a system with one major and many minor loci. As indicated by Figure 2, this might increase the expected genetic variance at equilibrium by increasing the average effect of a polymorphic locus (Table 2), but it could not offset the decline caused by the rapidly decreasing polymorphic fraction of the genome.

Little is known about linkage relations between loci affecting quantitative traits. If quantitative trait loci tended to be tightly linked, polymorphisms at several loci could be maintained. Except for the two-locus case, this would lead to higher levels of genetic variance.

Our analysis, as well as most previous analyses, is based on the assumption of additive loci. However, recent experimental results indicate that typical quantitative traits may be under control of loci with substantial epistatic effects (Mackay and Fry 1996; Routman and Cheverud 1997). Theoretical analyses of multilocus models under stabilizing selection have shown that with epistasis, high degrees of genetic variation can be maintained (Gimelfarb 1989). Future theoretical investigations of multilocus systems underlying quantitative traits should definitely pay more attention to epistatic effects. Because to date, only little is known about the nature of epistatic effects, such investigations had to be based on ad hoc assumptions.

Following the tradition in population genetics, we have explored the equilibrium properties of our model. But are natural populations ever in or close to equilibrium? The more loci are affecting a trait, the more likely it is that transient polymorphisms are contributing to genetic variation. Such transient polymorphisms can be present for a variety of reasons, for instance, new mutations sweeping through the population or balancing selection caused by forces that may be independent of the character under consideration. We have no quantitative information about the transient properties of our model populations, but we have observed that in some cases, particularly for tightly linked loci, convergence to equilibrium may be exceedingly slow, on the order of tens of millions of generations. Therefore, it might prove useful to devote more consideration to evolving populations and to study nonequilibrium properties.

Keeping all these reservations in mind, let us compare the amount of genetic variation that can be maintained under stabilizing selection alone with that maintained under mutation-selection balance. It is well known that the equilibrium genetic variance that can be maintained under mutation-selection balance in a large population

is approximately $2n\mu/s$, provided the loci are not very tightly linked and the expected squared effect, $\overline{\alpha^2} = (1/n)\Sigma_i\alpha_i^2$, is larger than $10\mu/s$ (*cf.* Latter 1960; Bulmer 1972; Turelli 1984; Bürger and Lande 1994; *cf.* Bürger 1998, for a review). In our simulations, we have $\overline{\alpha^2} = 0.0132$ if $n = 5$, and $\overline{\alpha^2}$ is larger if n is smaller. If there are only two loci and if the character is not nearly neutral, then for all realistic mutation rates, the equilibrium genetic variance maintained by mutation-selection balance is much lower than the average value, $\overline{V}_G = 0.031$, that is maintained under pure stabilizing selection. If there are five loci, however, then the average genetic variance maintained in our model is only approximately $\overline{V}_G \approx 0.001$. Such an amount of variance can be easily maintained by mutation-selection balance, namely whenever $\mu/s > 0.0001$. This argument is not intended to imply that mutation-selection balance is the primary mechanism for maintaining genetic variation in quantitative traits (see Barton and Turelli 1989 and Bürger and Lande 1994 for discussions). It shows, however, that pure stabilizing selection can be excluded as an important mechanism, unless genetic systems are very different from those assumed in this article.

In accordance with previous theoretical investigations (*cf.* Turelli and Barton 1990; Bürger and Lande 1994), our results show that the average level of linkage disequilibrium is exceedingly low. The main reason is that with more than three loci affecting the trait and independent, uniformly distributed recombination rates, polymorphisms involving two or more loci are so rare.

Finally, we draw attention to the multitude of stable equilibria that may exist under stabilizing selection. We not only found that the expected number of coexisting stable equilibria is higher under stabilizing selection than under selection with randomly assigned fitnesses (see Table 2), but we also found that in a five-locus system, up to 15 of 20 randomly chosen initial conditions converged to different equilibria. This implies that historic effects may play an important role for populations under stabilizing selection. Therefore, it cannot be expected that closely related (but isolated) populations that have experienced similar selective regimes are genetically identical at all or most loci contributing to traits under stabilizing selection, even if the populations are large so that effects of random drift are negligible.

We thank Ellen Baake and an anonymous reviewer for their comments. R.B. was supported by the Austrian Science Foundation (FWF), Project P12865-MAT, and by the Adaptive Dynamics Network of the International Institute of Applied Systems Analysis (IIASA) in Laxenburg, Austria. A.G. was supported by National Science Foundation Research Training Grant DBI9602266.

LITERATURE CITED

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

Barton, N. H., and M. Turelli, 1989 Evolutionary quantitative genetics: How little do we know? Annu. Rev. Genet. **23:** 337-370.

Bulmer, M. G., 1971 The stability of equilibria under selection. Heredity **27:** 157–162.

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

Bulmer, M. G., 1980 *The Mathematical Theory of Quantitative Genetics*. Clarendon Press, Oxford.

Bürger, R., 1986 On the maintenance of genetic variation: global analysis of Kimura's continuum-of-alleles model. J. Math. Biol. **24:** 341–351.

Bürger, R., 1998 Mathematical properties of mutation-selection models. Genetica **102/103:** 279–298.

Bürger, R., and J. Hofbauer, 1994 Mutation load and mutation-selection-balance in quantitative genetic traits. J. Math. Biol. **32:** 193-218.

Bürger, R., and R. Lande, 1994 On the distribution of the mean and variance of a quantitative trait under mutation-selection-drift balance. Genetics **138:** 901–912.

Crow, J. F., and M. Kimura, 1964 The theory of genetic loads, pp. 495–505 in *Proceedings of the XI International Congress of Genetics*, edited by S. J. Geerts. Pergamon, Oxford.

Endler, J. A., 1986 *Natural Selection in the Wild*. Princeton University Press, Princeton, NJ.

Falconer, D. S., and T. F. C. Mackay, 1995 *Introduction to Quantitative Genetics*, Ed. 4. Addison Wesley Longman, New York.

Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

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

Gale, J. S., and M. J. Kearsey, 1968 Stable equilibria under stabilising selection in the absence of dominance. Heredity **23:** 553–561.

Gavrilets, S., and A. Hastings, 1993 Maintenance of genetic variability under strong stabilizing selection: a two-locus model. Genetics **134:** 377–386.

Gavrilets, S., and A. Hastings, 1994 Dynamics of genetic variability in two-locus models of stabilizing selection. Genetics **138:** 519–532.

Gimelfarb, A., 1989 Genotypic variation for a quantitative character maintained under stabilizing selection without mutations: epistasis. Genetics **123:** 217–227.

Gimelfarb, A., 1998 Stable equilibria in multilocus genetic systems: a statistical investigation. Theor. Popul. Biol. **54:** 133–145.

Haldane, J. B. S., 1932 *The Causes of Evolution*. Longmans, Green, London (reprinted with a new introduction and afterword by E. G. Leigh, Jr., by Princeton University Press, 1992).

Hastings, A., 1987 Monotonic change of the mean phenotype in two-locus models. Genetics **117:** 583–585.

Hill, W. G., and J. Rasbash, 1986 Models of long-term artificial selection in finite population. Genet. Res. **48:** 41–50.

Karlin, S., and W. M. Feldman, 1970 Linkage and selection: two locus symmetric viability model. Theor. Popul. Biol. **1:** 39–71.

Karlin, S., and J. L. McGregor, 1972 Application of method of small parameters to multi-niche population genetic models. Theor. Popul Biol. **3:** 186–208.

Kearsey, M. J., and J. S. Gale, 1968 Stabilising selection in the absence of dominance: an additional note. Heredity **23:** 617–620.

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

Latter, B. D. H., 1960 Natural selection for an intermediate optimum. Aust. J. Biol. Sci. **13:** 30-35.

Lewontin, R. C., 1974 *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.

López, M.A., and C. López-Fanjul, 1993 Spontaneous mutation for a quantitative trait in *Drosophila melanogaster*. II. Distribution of mutant effects on the trait and fitness. Genet. Res. **61:** 117–126.

Mackay, T. F. C., and J. D. Fry, 1996 Polygenic mutation in *Drosophila melanogaster*: Genetic interactions between selection lines and candidate quantitative trait loci. Genetics **144:** 671–688.

Mackay, T. F. C., R. F. Lyman and M. S. Jackson, 1992 Effects of P element insertions on quantitative traits in *Drosophila melanogaster*. Genetics **130:** 315–332.

Maynard Smith, J., 1998 *Evolutionary Genetics*, Ed. 2. Oxford University Press, Oxford.

Nagylaki, T., 1984 Selection on a quantitative character, pp. 275–

306 in *HumanPopulation Genetics: The Pittsburgh Symposium*, edited by A. Chakravarti. Van Nostrand, New York.
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 homozygosis. J. Genet. **54:** 236–248.
Routman, E. J., and J. M. Cheverud, 1997 Gene effects on a quantitative trait: two-locus epistatic effects measured at microsatellite markers and at estimated QTL. Evolution **51:** 1654–1662.
Turelli, M., 1984 Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. Theor. Popul. Biol. **25:** 138-193.
Turelli, M., and N. H. Barton, 1990 Dynamics of polygenic characters under selection. Theor. Popul. Biol. **38:** 1-57.
Wolfram, S., 1991 *Mathematica*, Ed. 2. Addison Wesley, Redwood City, CA.
Wright, S., 1935 Evolution in populations in approximate equilibrium. J. Genet. **30:** 257–266.
Wright, S., 1952 The genetics of quantitative variability, pp. 5–47 in *Quantitative Inheritance*, edited by E. C. R. Reeve and C. H. Waddington. Her Majesty's Stationary Office, London.

Communicating editor: W. Stephan

APPENDIX

Let x_1, x_2, x_3, x_4 denote the frequencies of the four gametes, A_1B_1, A_1B_2, A_2B_1, A_2B_2, respectively. Collectively, all quadruples (x_1, x_2, x_3, x_4) satisfying $0 \le x_i \le 1$ and $\Sigma_i x_i = 1$ form the 4-dimensional simplex S_4. This is the state space for the two-locus two-allele dynamics. Let $D = x_1x_4 - x_2x_3$ be the linkage disequilibrium measure and let W_i denote the marginal fitness of gamete i. Then, the well-known recursion equations can be written as

$$\overline{W}x_i' = x_iW_i - \eta_i rD, \quad i = 1, \ldots, 4, \tag{A1}$$

where $\eta_1 = \eta_4 = -\eta_2 = -\eta_3$. The equations in this appendix are written without the constraint $\alpha_1 + \alpha_2 = \frac{1}{2}$. However, it is assumed that the fitness optimum coincides with the value of the double heterozygote, *i.e.*, $W(\mathcal{G}) = 1 - s[\mathcal{G} - (\alpha_1 + \alpha_2)]^2$.

Equilibria: We assume $r > 0$. Then there exist up to nine equilibria, seven of which may be stable (but not simultaneously). We denote the possibly stable ones by F1–F7.

There always exists a symmetric equilibrium, F1, which is calculated to be given by

$$\text{F1: } \hat{x}_1 = \hat{x}_4 = \tfrac{1}{4} + \hat{D}_1, \quad \hat{x}_2 = \hat{x}_3 = \tfrac{1}{4} - \hat{D}_1, \tag{A2a}$$

where

$$\hat{D}_1 = \frac{1}{4s\alpha_1\alpha_2}\left[r - \sqrt{s^2\alpha_1^2\alpha_2^2 + r^2}\right]. \tag{A2b}$$

Two further interior equilibria, F2 and F3, may exist that are unsymmetric. They can be calculated explicitly using the results of Karlin and Feldman (1970). Let us introduce the coordinates

$$x = x_1 - x_4, \quad y = x_2 - x_3, \quad z = x_1 + x_4 - x_2 - x_3. \tag{A3}$$

Then we have

$$x_1 = \frac{1 + x + 2x}{4}, \quad x_2 = \frac{1 - z + 2y}{4},$$
$$x_3 = \frac{1 - z - 2y}{4}, \quad x_4 = \frac{1 + z - 2x}{4}. \tag{A4}$$

Because of the specific form of the selection parameters a, b, c, and d in (3), the quadratic polynomial in Equation A7 of Karlin and Feldman simplifies dramatically (with a little help from *Mathematica*; Wolfram 1991) and becomes a linear expression in z. Its root, $\hat{z}$, is given in (A6c) below. Implementing this value into Equation A2 of Karlin and Feldman yields two possible values for the equilibrium mean fitness at the prospective unsymmetric equilibria. It turns out that only the value

$$\overline{W}_e = 1 - \tfrac{9}{8}r + \tfrac{1}{4}rs(\alpha_1^2 + \alpha_2^2) - \tfrac{1}{8}s^2(\alpha_1^2 - \alpha_2^2)^2 \tag{A5}$$

leads to equilibria in S_4. Inserting these expressions in (A3) and (A4) of Karlin and Feldman (1970) yields the desired solutions. These are

$$\text{F2: } \hat{x} = \frac{(\alpha_1 + \alpha_2)[r - s(\alpha_1 - \alpha_2)^2]\sqrt{R}}{8s^{1/2}r^{3/2}\alpha_1\alpha_2} \tag{A6a}$$

$$\hat{y} = \frac{(\alpha_1 + \alpha_2)[-r + s(\alpha_1 + \alpha_2)^2]\sqrt{R}}{8s^{1/2}r^{3/2}\alpha_1\alpha_2} \tag{A6b}$$

$$\hat{z} = -\left(\frac{\alpha_1}{\alpha_2} + \frac{\alpha_2}{\alpha_1}\right) + \frac{3r}{4s\alpha_1\alpha_2} + \frac{s\alpha_1\alpha_2}{4r}\left(\frac{\alpha_1}{\alpha_2} - \frac{\alpha_2}{\alpha_1}\right)^2, \tag{A6c}$$

where

$$R = 3r^2 + 2rs(\alpha_1^2 + \alpha_2^2) - s^2(\alpha_1^2 - \alpha_2^2)^2. \tag{A7}$$

Substituting $\hat{x}$, $\hat{y}$, and $\hat{z}$ into (A4) gives the coordinates of the unsymmetric equilibria in terms of the chromosome frequencies x_1, x_2, x_3, and x_4. The equilibrium F3 is symmetric to F2 upon interchanging $\hat{x}_1$ with $\hat{x}_4$ and $\hat{x}_2$ with $\hat{x}_3$ (or, equivalently, taking $-\hat{x}$ and $-\hat{y}$). It is not difficult to show that F2 and F3 exist if and only if

$$r_1 < r < r_2, \tag{A8}$$

where

$$r_1 = -\tfrac{1}{3}s(\alpha_1^2 + \alpha_2^2) + \tfrac{2}{3}s\sqrt{\alpha_1^4 - \alpha_1^2\alpha_2^2 + \alpha_2^4} \tag{A9}$$

is the positive root of the equation $R = 0$ and

$$r_2 = \min\{s(\alpha_1 - \alpha_2)^2, \tfrac{1}{3}s(\alpha_1^2 - \alpha_2^2)\}. \tag{A10}$$

The reader may observe that

$$s(\alpha_1 - \alpha_2)^2 \ge \tfrac{1}{3}s(\alpha_1^2 - \alpha_2^2)$$

if and only if

$$\alpha_1 \ge 2\alpha_2$$

and

$$\tfrac{1}{3}s(\alpha_1 - \alpha_2)^2 < r_1 < \tfrac{1}{3}s(\alpha_1^2 - \alpha_2^2). \tag{A11}$$

We also note that the linkage disequilibrium at the unsymmetric equilibria is negative and given by $\hat{D} = \frac{1}{4}(\hat{z} +$

$\hat{y}^2 - \hat{x}^2$). This expression can be factorized and is negative between r_1 and r_2.

Next, there may exist two edge equilibria, F4 and F5, with the major locus polymorphic:

$$\text{F4: } \hat{x}_1 = \hat{x}_3 = 0, \quad \hat{x}_2 = \tfrac{1}{2} + \frac{\alpha_2}{\alpha_1}, \quad \hat{x}_4 = \tfrac{1}{2} - \frac{\alpha_2}{\alpha_1}, \qquad \text{(A12)}$$

$$\text{F5: } \hat{x}_2 = \hat{x}_4 = 0, \quad \hat{x}_1 = \tfrac{1}{2} - \frac{\alpha_2}{\alpha_1}, \quad \hat{x}_3 = \tfrac{1}{2} + \frac{\alpha_2}{\alpha_1}, \qquad \text{(A13)}$$

These exist (*i.e.*, are in S_4) if and only if $\alpha_1 > 2\alpha_2$.

Finally, there are four corner (vertex) equilibria corresponding to fixation of one of the chromosomes. These always exist but only the vertices corresponding to fixation of A_1B_2 or A_2B_1 can be stable. These are denoted by F6 and F7, respectively, and their coordinates are

$$\text{F6: } \hat{x}_2 = 1, \quad \hat{x}_1 = \hat{x}_3 = \hat{x}_4 = 0, \qquad \text{(A14)}$$

$$\text{F7: } \hat{x}_3 = 1, \quad \hat{x}_1 = \hat{x}_2 = \hat{x}_4 = 0. \qquad \text{(A15)}$$

If $r = r_1$ then F1 = F2 = F3 and if $r = r_2$ then F2 and F3 coincide either with F4 and F5 (if $\alpha_1 > 2\alpha_2$) or with F6 and F7 (otherwise). All equilibria exhibit either negative linkage disequilibrium (the three interior equilibria) or are in linkage equilibrium (the boundary equilibria). Indeed, more can be proved, namely that all orbits eventually enter the region where $D \le 0$. We consider the function $Z = x_2x_3/x_1x_4$ and show that $Z' > Z$ if $D > 0$ [note that $D = x_1x_4(1 - Z)$]. This follows from

$$Z' = \frac{x_2'x_3'}{x_1'x_4'} = \frac{(x_2W_2 + rD)(x_3W_3 + rD)}{(x_1W_1 - rD)(x_4W_4 - rD)}$$
$$> \frac{x_2x_3}{x_1x_4}\cdot\frac{W_2W_3}{W_1W_4} = Z\,\frac{W_2W_3}{W_1W_4}$$

and a straightforward calculation that reveals that $W_2W_3 \ge W_1W_4$.

Next, we investigate the stability properties of these equilibria.

Stability: The precise condition for local stability of the symmetric equilibrium in the case $a + d = 2(b + c)$ was derived by Karlin and Feldman (1970). Their condition (4.6) implies that F1 is locally stable if and only if

$$r \le r_1. \qquad \text{(A16)}$$

The eigenvalues of F6 and F7 are readily calculated and are

$$\frac{1 - s\alpha_1^2}{1 - s(\alpha_1 - \alpha_2)^2}, \quad \frac{1 - s\alpha_2^2}{1 - s(\alpha_1 - \alpha_2)^2}, \quad \frac{1 - r}{1 - s(\alpha_1 - \alpha_2)^2}.$$

Local asymptotic stability requires that these eigenvalues are less in modulus than one. Therefore, the equilibria F6 and F7 are locally stable if and only if

$$\alpha_1 \le 2\alpha_2 \quad \text{and} \quad r > s(\alpha_1 - \alpha_2)^2. \qquad \text{(A17)}$$

Also, the eigenvalues of the boundary equilibria F4 and F5 can be calculated explicitly and are given by

$$\frac{1 - s(\alpha_1^2 - 3\alpha_2^2)}{1 - \tfrac{1}{2}s(\alpha_1^2 - 2\alpha_2^2)}$$

and

$$\frac{1}{[1 - \tfrac{1}{2}s(\alpha_1^2 - 2\alpha_2^2)]}\left[1 - \tfrac{1}{2}r - \tfrac{1}{2}s\alpha_1^2 \pm \sqrt{r^2 - 8\alpha_2^2rs + 4\alpha_1^2\alpha_2^2s^2}\right].$$

Again, all these eigenvalues must be less in modulus than one. The first one satisfies this condition if and only if $\alpha_1 > 2\alpha_2$, *i.e.*, if and only if the equilibria exist (in S_4). The expression under the square root of the second and third eigenvalue is always $\ge (r - 4s\alpha_2^2)^2$ if $\alpha_1 \ge 2\alpha_2$. Therefore, these are real and it follows easily that F4 and F5 are locally stable if and only if they exist ($\alpha_1 > 2\alpha_2$) and

$$r \ge \tfrac{1}{3}s(\alpha_1^2 - \alpha_2^2). \qquad \text{(A18)}$$

The stability analysis of the unsymmetric equilibria is rather complicated because, in general, the eigenvalues cannot be determined explicitly. However, the determinant of the Jacobian matrix (which is the constant term of the characteristic polynomial) is of relatively simple structure. It is a polynomial in r of degree six and decomposes into six linear factors. Among the zeroes of this polynomial are r_1, $s(\alpha_1 - \alpha_2)^2$, and $\tfrac{1}{3}s(\alpha_1^2 - \alpha_2^2)$. Hence, if $r = r_1$ or $r = r_2$, the eigenvalues can be determined explicitly. The largest eigenvalue is one and the others are less in absolute value than one (and agree either with those of F1 or those of F4, F5 or F6, F7). Taking the derivative of the determinant of the Jacobian with respect to r and evaluating at r_1 and r_2 shows that all three eigenvalues are less in modulus than unity if r is slightly larger than r_1 or slightly smaller than r_2. Numerical calculations suggest that the unsymmetric equilibria are stable whenever they exist, *i.e.*, if $r_1 < r < r_2$.

Thus, if $0 < r \le r_1$ then F1 is locally stable (and apparently globally stable). When $r = r_1$, F1 loses its stability and the unsymmetric equilibria F2 and F3 arise by a pitchfork bifurcation. Then these are stable and converge to F6 and F7 as $r \to s(\alpha_1 - \alpha_2)^2$ if $\alpha_1 \le 2\alpha_2$, while they converge to F4 and F5 as $r \to \tfrac{1}{3}s(\alpha_1^2 - \alpha_2^2)$ if $\alpha_1 > 2\alpha_2$. In the first case the equilibria F4 and F5 do not exist. When the pair of unsymmetric equilibria hits a pair of boundary equilibria (either F4 and F5 or F6 and F7), a transcritical bifurcation occurs and the boundary equilibria become stable and remain stable for all larger values of r, while the unsymmetric equilibria leave the simplex.

Therefore, in this model no more than two equilibria can be simultaneously stable. As long as linkage is sufficiently tight, *i.e.*, if

$$r < r_2, \qquad \text{(A19)}$$

there exist one or two stable polymorphisms in the interior and all boundary equilibria are unstable. If $r > r_2$

then no two-locus polymorphism can be maintained. Because $r_2 < \frac{1}{3}$ always holds ($s\alpha_1^2 \leq 1$), no stable two-locus polymorphism exists if $r > \frac{1}{3}$. If $r > r_2$ and the disparity between the allelic effects of the two loci is sufficiently large ($\alpha_1 > 2\alpha_2$), then the major locus is kept stably polymorphic by marginal overdominance while the other is monomorphic. If the disparity of effects is not large enough ($\alpha_1 \leq 2\alpha_2$), no polymorphism can be maintained and either chromosome A_1B_2 or A_2B_1 becomes fixed (*cf.* Gavrilets and Hastings 1993).

Genetic variance: The genetic variance and all other quantities are measured before selection, so that the population is in Hardy-Weinberg equilibrium. Therefore, it is sufficient to calculate the genetic variance among gametes and multiply it by two. The mean haploid genotypic value, $\overline{\mathcal{G}}_h$, is simply

$$\overline{\mathcal{G}}_h = \alpha_2 x_2 + \alpha_1 x_3 + (\alpha_1 + \alpha_2) x_4. \qquad \text{(A20)}$$

From this, the deviation of the mean genotypic value at equilibrium from the optimum is calculated to be 0, $\sqrt{R}/(2\sqrt{rs})\,\alpha_2$, and $\alpha_1 - \alpha_2$ for F1, F2 and F3, F4 and F5, and F6 and F7, respectively.

The (additive) genetic variance is given by

$$\sigma_G^2 = 2(\alpha_2^2 x_2 + \alpha_1^2 x_3 + (\alpha_1 + \alpha_2)^2 x_4 - \overline{\mathcal{G}}_h^2). \qquad \text{(A21)}$$

This yields the equilibrium genetic variances

$$\hat{\sigma}_G^2(\text{F1}) = \tfrac{1}{2}(\alpha_1^2 + \alpha_2^2) + 4\alpha_1\sigma_2\hat{D}_1, \qquad \text{(A22)}$$

$$\hat{\sigma}_G^2(\text{F2, F3}) = \frac{3}{8rs}[r^2 - 2rs(\alpha_1^2 + \alpha_2^2) + s^2(a_1^2 - a_2^2)^2], \qquad \text{(A23)}$$

$$\hat{\sigma}_G^2(\text{F4, F5}) = \tfrac{1}{2}(\alpha_1^2 - 4\alpha_2^2), \qquad \text{(A24)}$$

and

$$\hat{\sigma}_G^2(\text{F6, F7}) = 0. \qquad \text{(A25)}$$

Although two equilibria may be simultaneously stable, it is important to note that alternative stable equilibria have the same mean phenotype and the same genetic variance.