# Intraspecific Competitive Divergence and Convergence under Assortative Mating

**Reinhard Bürger**[1,2,*] **and Kristan A. Schneider**[1,†]

1. Department of Mathematics, University of Vienna, Nordbergstrasse 15, 1090 Vienna, Austria;
2. Program for Evolutionary Dynamics, Harvard University, Cambridge, Massachusetts 02138

*Submitted March 24, 2005; Accepted September 23, 2005; Electronically published January 9, 2006*

*Online enhancement:* appendix.

ABSTRACT: Ecologically driven sympatric speciation has received much attention recently. We investigate a multilocus model of a quantitative trait that is under frequency-dependent selection caused by intraspecific competition and acts as mating character for assortment. We identify the conditions that lead to the establishment of reproductively isolated clusters. This may be interpreted as evolutionary splitting or sympatric speciation. In our model, there are parameters that independently determine the strength of assortment, the costs for being choosy, and the strength of frequency-dependent natural selection. Sufficiently strong frequency dependence leads to disruptive selection on the phenotypes. The population consists of (sexual) haploid individuals. If frequency dependence is strong enough to induce disruptive selection and costs are absent or low, the result of evolution depends in a distinctive nonlinear way on the strength of assortment: under moderately strong assortment, less genetic variation is maintained than under weak or strong assortment, and sometimes there is none at all. Evolutionary splitting occurs only if frequency dependence and assortment are both strong enough and costs are low. Even then, the evolutionary outcome depends on the genetics and the initial conditions. The roles of the number of loci, of linkage, and of asymmetric selection are also explored.

*Keywords:* assortative mating, frequency-dependent selection, density-dependent selection, disruptive selection, recombination, sympatric speciation.

* Corresponding author; e-mail: reinhard.buerger@univie.ac.at.

† E-mail: kristan.schneider@univie.ac.at.

Am. Nat. 2006. Vol. 167, pp. 190–205.  0003-0147/2006/16702-40971$15.00. 

The study of the mechanisms of speciation has gained much momentum during the past decade, and the recent publication of three books reflects these developments. Coyne and Orr (2004) focus on empirical and comparative evidence. The volume edited by Dieckmann et al. (2004) is primarily dedicated to the ecological mechanisms and the adaptive dynamics approach to speciation. Gavrilets (2004) provides a comprehensive review of the various theories. In particular, theories of sympatric speciation have received much attention. Speciation in sympatry may be driven by a variety of mechanisms. Among them are sexual selection (van Doorn et al. 1998; Higashi et al. 1999), sexual conflict (Parker and Partridge 1998; Gavrilets and Waxman 2002), and habitat or resource specialization (Kawecki 1997), to cite only some of the recent work. An important class of mechanisms derives from ecological interactions that induce disruptive selection. This is the topic we are concerned with. Previous work has demonstrated that the joint action of disruptive selection and assortative mating can cause sympatric speciation (Doebeli 1996; Dieckmann and Doebeli 1999; Kondrashov and Kondrashov 1999; Doebeli and Dieckmann 2000; Drossel and McKane 2000; Geritz and Kisdi 2000). But since then various limitations have been indicated (Matessi et al. 2001; Bolnick 2004*b*; Gavrilets 2004; Gourbiere 2004; Kirkpatrick and Nuismer 2004).

One of the problems with studies of ecologically driven sympatric speciation, as well as of (sympatric) speciation in general, is that they are primarily based on numerical work, and general analytical or systematic results are largely missing (Turelli et al. 2001; Gavrilets 2003). In particular, many of the previous studies aim at demonstrating a phenomenon rather than exploring the conditions for speciation. A notable exception is work by Gavrilets (2003, 2004, pp. 374–382), who derived explicit conditions for sympatric speciation in five simple two-locus models. A substantial part of this work is also computational but much more comprehensive and systematic than previous studies. The purpose of this work is the identification of the conditions that lead to competitively

driven divergence and the establishment of reproductively isolated clusters in a population.

In our model, there is a single trait that is under frequency-dependent natural selection and acts as a mating character; that is, the probability that two individuals mate decreases with increasing phenotypic difference. The trait is determined by an arbitrary number of recombining diallelic loci that can have arbitrary effects. The population is sufficiently large to ignore random genetic drift, is density regulated, and consists of sexual haploid individuals (a concession to limited computer resources). We tackle a simpler question than, for instance, the ambitious approach of Dieckmann and Doebeli (1999)—first, by using a "magic trait" (sensu Gavrilets 2004, chap. 10.3), and second, by using a parameter that tunes the strength of assortment instead of letting it evolve. In fact, there is a second mating parameter that specifies the costs for being choosy.

We pursue a numerical and statistical approach, and for each combination of ecological and mating parameters, we investigate the evolutionary behavior for 1,000 randomly chosen genetic systems, each from 10 (random) initial conditions. Also, linkage is included. This is computationally much more demanding and time-consuming than any previous approach, but it is rewarding because a more complete picture is obtained, one that does not depend on assumptions such as loci of equal effects and free recombination. We complement the numerical results by analytical ones. We derive the eigenvalues of all monomorphic equilibria and give simple conditions guaranteeing their stability as a function of the strength of natural selection, assortative mating, costs of assortment, and the underlying genetics. The monomorphic equilibria play a prominent role in models of sympatric speciation, even though this has rarely been acknowledged. These results generalize those of Gavrilets (2004) on one of his four single-trait two-locus models. In addition, we characterize the equilibrium structure under the assumptions of random mating and linkage equilibrium. This yields an important guide to interpreting the results for assortative mating.

Another problem is that studies of (sympatric) speciation are based on a zoo of different assumptions (Kirkpatrick and Ravigné 2002). Recently, Bürger (2005) has studied the maintenance of polygenic variation in an ecological model that is the weak-selection limit of, to the best of our knowledge, all models of intraspecific competition for a continuous unimodal resource that have been used in the literature. Although natural selection is weak overall, frequency dependence can be very strong. Here, we adopt this ecological model and include assortative mating. Therefore, our results apply to most, if not all, such models as long as natural selection is not too strong.

## The Model

We consider a sexually reproducing population of haploid individuals with discrete generations and equivalent sexes that is sufficiently large to ignore random genetic drift. Natural selection acts through differential viabilities on an additive polygenic trait. Individual fitness is assumed to be determined by two components: stabilizing selection on this trait and frequency-dependent competition among individuals of similar phenotypes.

### *Ecological Assumptions*

The first fitness component is frequency independent and may reflect some sort of direct selection on the trait, for example, by differential supply of a resource whose utilization efficiency is phenotype dependent. We ignore environmental variation and deal directly with the fitnesses of genotypic values. In the absence of genotype-environment interaction, this is no restriction because in our model, the only effect of including environmental variance would be a deflation of the selection intensity on genotypes. For simplicity, we use the words genotypic value and phenotype synonymously.

Stabilizing selection is modeled by the quadratic function

$$S(g) = 1 - s(g - \theta)^2, \tag{1}$$

where $s \geq 0$ measures the strength of stabilizing selection and $\theta$ is the position of the optimum. Of course, $S(g)$ is assumed positive on the range of possible phenotypes, which is scaled to $[-\Gamma, \Gamma]$, where $\Gamma$ is a positive constant. Thus, we have the restriction $0 \leq s \leq (\Gamma + |\theta|)^{-2}$. We exclude pure directional selection by assuming $-\Gamma < \theta < \Gamma$.

The second component of fitness is frequency dependent. We assume that competition between phenotypes $g$ and $h$ can be described by

$$\alpha(g, h) = 1 - c(g - h)^2, \tag{2}$$

where $0 \leq c \leq 1/(4\Gamma^2)$. This implies that competition between individuals of similar phenotypes is stronger than between individuals of very different phenotypes, as will be the case if different phenotypes preferentially utilize different food resources. Large $c$ implies a strong frequency-dependent effect of competition, whereas in the limit $c \to 0$, frequency dependence vanishes. Let $\mathcal{P}(h)$ denote the relative frequency of individuals with phenotype $h$. Then the intraspecific competition function $\bar{\alpha}(g)$, which

measures the strength of competition experienced by phenotype $g$ if the population distribution is $\mathcal{P}$, is given by

$$\bar{\alpha}(g) = \sum_h \alpha(g, h)\mathcal{P}(h)$$

and calculated to be

$$\bar{\alpha}(g) = 1 - c[(g - \bar{g})^2 + V_A]. \tag{3}$$

Here, $\bar{g}$ and $V_A$ denote the mean and (additive genetic) variance, respectively, of the distribution $\mathcal{P}$ of genotypic values. In the following, it will be convenient to measure the strength of frequency-dependent competition relative to the strength of stabilizing selection. Therefore, we define

$$f = \frac{c}{s}. \tag{4}$$

We shall treat $f$ and $s$ as independent parameters.

Our model includes density-dependent population growth, which, in the absence of genetic variation, occurs according to

$$N' = NF(N). \tag{5}$$

Here, $N$ and $N'$ are the population sizes in consecutive generations and $F: [0, \infty) \to [0, \infty)$ is a strictly decreasing function of $N$ (on the interval of admissible values) so that $F(N) = 1$ has a unique positive solution $K$, the carrying capacity. The function $F$ and the parameters that determine it are assumed such that they ensure a simple demographic dynamics; that is, convergence to $K$ occurs for all (admissible) initial conditions (see Thieme 2003, chap. 9, for general conditions on $F$). We shall primarily be concerned with discrete logistic growth, that is,

$$F(N) = \rho - \frac{N}{\kappa},$$
$$0 \le N < \rho\kappa. \tag{6}$$

The carrying capacity in this model is $K = (\rho - 1)\kappa$. Monotone convergence to $K$ occurs for all $N$ with $0 < N < \rho\kappa$ if $1 < \rho \le 2$ and oscillatory convergence (at a geometric rate) if $2 < \rho < 3$.

One way of defining fitness caused by natural selection is to assume that the two fitness components are multiplicative, for instance, because they act in different phases of the life cycle. This yields

$$W_*(g) = F(N\bar{\alpha}(g))S(g) \tag{7}$$

(Bulmer 1974, 1980; Bürger 2002; Bürger and Gimelfarb 2004). Assuming that stabilizing selection is weak, that is, ignoring terms of order $s^2$, we can approximate $W_*(g)$ by

$$W(g) = F(N)\{1 - s(g - \theta)^2 + s\eta(N)[(g - \bar{g})^2 + V_A]\}, \tag{8}$$

where the dependence of $W(g)$ on $N$ and $\mathcal{P}$ is omitted and

$$\eta(N) = \frac{-NF'(N)}{F(N)}f. \tag{9}$$

For discrete logistic growth, we obtain

$$\eta(N) = f\left(\frac{\rho\kappa}{N} - 1\right)^{-1}. \tag{10}$$

We interpret $\eta(N)$ as a compound measure of the strength of frequency and density dependence relative to stabilizing selection. We note that $f$ can assume any value $\ge 0$; thus frequency dependence may be arbitrarily strong.

Because $W(g)$ is quadratic in $g$, selection is disruptive if $W(g)$ is convex and the minimum is within the range of phenotypic values $(-\Gamma, \Gamma)$. For a given population distribution $\mathcal{P}$, $\bar{g}$ and $V_A$ are constants and $W(g)$ is twice differentiable. By straightforward calculation, we obtain that $W(g)$ is convex if and only if $\eta(N) > 1$, and the minimum lies in the interior of the phenotypic range if and only if

$$\eta(N)(\Gamma + \bar{g}) > \Gamma + \theta,$$
$$\eta(N)(\Gamma - \bar{g}) > \Gamma - \theta. \tag{11}$$

If both conditions are satisfied, $\eta(N) > 1$ holds. Therefore, the conditions (inequalities [11]) are necessary and sufficient for selection to be disruptive for every population distribution with mean $\bar{g}$. Hence, if $\eta(N) > 1$, selection is disruptive if $\bar{g}$ is sufficiently close to $\theta$; otherwise it is directional. If $\bar{g} = \theta$, then $\eta(N) > 1$ is necessary and sufficient for selection to be disruptive.

Bürger (2005) showed that $W(g)$ is the weak-selection approximation of fitness (i.e., to first order in $s$) in most models of intraspecific competition for a continuum of resources, for example, in those of Bulmer (1974, 1980), Slatkin (1979), Christiansen and Loeschcke (1980), Loeschcke and Christiansen (1984), Dieckmann and Doebeli (1999), Bürger (2002), Bürger and Gimelfarb (2004), Gourbiere (2004), Kirkpatrick and Nuismer (2004), Schneider (2005), and Schneider and Bürger (2005). If, instead of discrete logistic population growth, the Hassell or the Beverton-Holt model is assumed, then for the corresponding $F$, equation (8) yields the weak-selection ap-

proximation to the models of Doebeli (1996), Bolnick and Doebeli (2003), and Bolnick (2004*b*), respectively. Therefore, our results are representative for a large class of models as long as selection is not too strong.

### *Assortative Mating*

We assume that individuals mate assortatively with respect to the character under natural selection. We adopt the model of Matessi et al. (2001), which is a specification of that of Gavrilets and Boake (1998), and assume that females express preferences based on the similarity of their phenotypic value to that of their prospective mating partner. To be explicit, we choose this preference function as

$$\pi(g - h) = \exp[-a(g - h)^2]. \tag{12}$$

This is the probability that an encounter of a female with a male results in mating. If $a = 0$, then females have no preferences and mating is random. The larger $a$ is, the stronger is assortative mating.

Females mate only once, whereas males may participate in multiple matings. If an encounter was not successful, in which case she remains unmated, she may try again until the total number of encounters has reached a maximum number $M$. This reflects the idea that choosiness has costs, for instance, because of a limited mating period. If $M$ is infinity, there are no costs because every female is sure to find a mating partner; if $M = 1$, the costs for being choosy are very high. The probability that an encounter of a female of type $g$ with a random male results in mating is

$$\bar{\pi}(g) = \sum_h \pi(g - h)\mathcal{P}(h), \tag{13}$$

and the probability that she eventually mates with a male of type $h$ is $Q(g, h)\mathcal{P}(h)$, where

$$Q(g, h) = \sum_{m=0}^{M-1} [1 - \bar{\pi}(g)]^m \pi(g - h). \tag{14}$$

Matessi et al. (2001) call $Q$ the mating rate. They observed that if $M = 1$, whence

$$Q(g, h) = \pi(g - h), \tag{15}$$

this model can be conceived as a model of fertility selection (Bodmer 1965; Hadeler and Liberman 1975) or as a model of parental selection (Gavrilets 1998). The case $M = 1$ can also be interpreted as both sexes being choosy because then the probability of a mating between a $g$ female and an $h$ male is $\pi(g - h)$. This corresponds to the "plant model" of Kirkpatrick and Nuismer (2004). Also, Drossel and McKane (2000) and Gourbiere (2004) consider models of selective assortative mating that are equivalent to the case $M = 1$.

If the encounter rate is very high, $M$ may be chosen to be infinity, and we obtain

$$Q(g, h) = \frac{\pi(g - h)}{\bar{\pi}(g)}. \tag{16}$$

Then, $\sum_h Q(g, h)\mathcal{P}(h) = 1$ for all $g$, and assortative mating does not induce sexual selection among females. It does, however, induce sexual selection among males.

### *Genetic Assumptions and Evolutionary Dynamics*

The trait value $g$ of an individual is determined additively by $n$ diallelic loci. We denote the alleles at locus $i$ by $A_i$ and $a_i$, their effects by $\gamma_i$ and $-\gamma_i$ ($\gamma_i > 0$), and their frequencies by $P_i$ and $1 - P_i$. As noted by Turelli and Barton (2004), this choice of effects is general if the difference of effects (the effect of a substitution) is $2\gamma_i$ because constants that determine the mean phenotype can be absorbed by $\theta$. We assume that $\theta$ is within the range of possible genotypic values (i.e., $|\theta| < \Gamma = \sum_{i=1}^{n} \gamma_i$) and call the optimum symmetric if $\theta = 0$.

The multilocus dynamics is described in terms of gamete frequencies. In accordance with some population genetic models of haploid populations (Feldman 1971; Rutschman 1994; but see Kirzhner and Lyubich 1997), the frequencies of gametes are measured among adults after selection and before mating. Gametes are denoted by $r$, $u$, and $v$ and their phenotypic values by $g_r$, $g_u$, and $g_v$, respectively. The frequencies of gamete $r$ in consecutive generations are denoted by $p_r$ and $p_r'$. We designate by $R(uv \to r)$ the probability that haploid parents with genotypes $u$ and $v$ produce an offspring with genotype $r$. The function $R$ is determined by the pattern of recombination between loci, for which we shall consider several scenarios.

Therefore, the genetic dynamics is given by the system of $2^n$ recursion relations:

$$p_r' = \frac{W_r^*}{\overline{W}} \sum_{u,v} p_u p_v Q_{uv} R(uv \to r), \tag{17}$$

where $W_r^* = W^*(g_r)$ and the asterisk indicates that fitness is calculated from gamete frequencies after recombination. Moreover, $Q_{uv} = Q(g_u, g_v)$ and $\overline{W} = \sum_r W_r^* \sum_{u,v} p_u p_v Q_{uv} R(uv \to r)$. The demographic dynamics follows the standard recursion relation

$$N' = N\overline{W}. \tag{18}$$

Thus, for a monomorphic population ($V_A = 0$) with $g = \bar{g} = \theta$, population growth follows equation (5). The complete evolutionary dynamics is given by the coupled system in equations (17) and (18).

### Analytical Results

The complexity of this model prohibits a detailed and comprehensive analytic treatment. However, for a randomly mating population, the equilibrium and stability structure can be determined if linkage disequilibrium is ignored and the population size is assumed to be constant and at demographic equilibrium. For logistic population growth (eq. [6]), the following results are obtained (see appendix in the online edition of the *American Naturalist* for detailed and more general versions, valid for any population regulation obeying eq. [5]). (1) If, approximately, $f(\rho - 1) < 1$, then at most one locus can be polymorphic at a stable equilibrium, and, typically, multiple stable equilibria coexist. (2) If, approximately, $f(\rho - 1) > 1$, then there exists a unique asymptotically stable equilibrium that, therefore, is globally stable. At least one locus is polymorphic at this equilibrium. If the optimum is symmetric ($\theta = 0$), then all loci are polymorphic. For an arbitrary optimum, the polymorphic loci can be determined (those with large effects are polymorphic) and the allele frequencies can be calculated.

The condition $f(\rho - 1) > 1$ is (approximately) equivalent to the exact and general condition $\eta(\hat{N}) > 1$, where $\eta(\hat{N})$ is defined in equation (9) and $\hat{N}$ is the population size at the stable (polymorphic) equilibrium. As shown by inequalities (11), selection is disruptive if $\bar{g} = \theta$ and $\eta(N) > 1$. Thus, roughly, results 1 and 2 show that high genetic variability is maintained in a randomly mating population if frequency- and density-dependent selection together are stronger than stabilizing selection and overall disruptive selection is induced. Otherwise, little or no variation is maintained. Comparison with exact results from numerical iteration of the recursion relations of the full model show that the linkage equilibrium approximation is very accurate if linkage is not too tight (see table A1 in the online appendix).

For the general model with population regulation, linkage disequilibrium, and assortative mating, the stability conditions for the monomorphic equilibria can be derived. For logistic population growth and $f(\rho - 1) > 1$, the most interesting parameter range, the following conclusions can be drawn (proofs are given in "Stability of Monomorphic Equilibria under Assortative Mating" in the online appendix). (*a*) If $a = 0$ (random mating), no genetic system can have a stable monomorphic equilibrium. This complements result 2, which assumes linkage equilibrium and constant size. (*b*) For given costs of choosiness, $M$, increasingly strong assortment promotes stability of monomorphic equilibria. For given $a > 0$, increasing costs (decreasing $M$) promote stability of monomorphic equilibria. If $M < \infty$, then all monomorphic equilibria with $\hat{N} > 0$ (see eq. [A13] in the online appendix) become stable for sufficiently strong assortment. (*c*) If $M \geq 2$ and if, in addition to $s$ (and $c$), $a$ is assumed to be small enough that terms of order $a^2$ and $as$ can be neglected, then monomorphic equilibria close to $\theta$ become stable if (approximately)

$$s + \frac{a}{2} > c(\rho - 1). \tag{19}$$

Stable monomorphic equilibria exist in any genetic system (i.e., even if no genotype is close to $\theta$) if

$$\frac{a}{2} > c(\rho - 1). \tag{20}$$

These approximate conditions are more accurate for large $M$ than for small $M$. If $M = 1$, then in the above inequalities $a/2$ has to be replaced by $a$.

The proportion of trajectories converging to a monomorphic equilibrium, which can be used as a measure for the size of their basin of attraction, will be determined numerically. Convergence to such an equilibrium clearly prevents divergence and speciation.

### Statistical and Numerical Approach

We use the approach of Bürger and Gimelfarb (2004), with the obvious modifications required by modeling assortative mating and assuming a haploid population. Its basic idea is to evaluate the quantities of interest for many randomly chosen genetic parameter sets and initial conditions and then calculate averages, standard deviations, and other statistics. In this sense, we obtain statistical results, although each single result is obtained by iterating numerically the deterministic system of recursion relations (17) and (18). All numerical results are based on logistic population growth (model [6]). All results presented use $\rho = 2$, $\kappa = 10{,}000$, and $s = 0.4$. The choice of $\kappa$ has no effect on the dynamics because it enters fitness only by the ratio $N/\kappa$; see equations (8) and (10). Therefore, $\kappa$ affects the population size multiplicatively. The computer program was developed by A. Gimelfarb.

For a given number of loci $n$ and a given range of recombination rates, we constructed more than 1,000 of what we call "genetic" parameter sets (allelic effects of loci and recombination rates between adjacent loci from the given range). For each genetic parameter set, allelic effects were obtained by generating values $\beta_\ell$ ($\ell = 1, 2, \ldots, n$) as independent random variables, uniformly distributed between

0 and 1, and transforming them into the actual allelic effects, $\gamma_\ell = (1/2)\beta_\ell / \sum_k \beta_k - (1/2)$. The additivity assumption yields the genotypic values. Therefore, the range of phenotypic values is always $[-\Gamma, \Gamma] = [-1/2, 1/2]$, with the extreme values actually assumed. Recombination rates between adjacent loci, $r_{\ell,\ell+1}$ ($\ell = 1, \ldots, n-1$), are assumed to be either all 1/2 or obtained as independent random variables, uniformly distributed between 0.01 and 0.1 or between 0 and 0.01. We assumed absence of interference and refer to these three scenarios as free recombination, tight linkage, and very tight linkage, respectively.

For a given ecological parameter combination ($\rho$, $\kappa$, $\theta$, $s$, $f$), given mating parameters ($a$, $M$), number of loci ($n$), and recombination scenario, the recursion relations (17) and (18) were numerically iterated starting from 10 different, randomly chosen initial gamete distributions for each of more than 1,000 genetic parameter sets. To make the initial distributions more evenly distributed in the gamete state 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 0.35 for 2, 3, 4, and 5 loci, respectively). Each genetic system has different initial distributions. An iteration was stopped after generation $t$ when either an equilibrium was reached (in the sense that the euclidean distance between gamete distributions in two consecutive generations was $<10^{-10}$), or $t$ exceeded $10^6$ generations (for some parameter sets, even $10^7$). Equilibria were classified as different if their euclidean distance exceeded $5 \times 10^{-4}$. If equilibrium was not reached, the parameter combination was excluded from the analysis. The proportion of excluded runs was small enough not to introduce a bias. If convergence within the specified maximum number of generations did not occur, it was because of extremely slow convergence of allele frequencies. The main reason for this is the presence of alleles of extremely small effect. No instance of complicated dynamic behavior (such as limit cycles or chaos) was detected.

For each combination of ecological and mating parameters, number of loci, and recombination scenario, all statistics are based on 1,000 genetic parameter sets, each with 10 initial conditions, that led to equilibration. For each parameter set, we recorded the number of different equilibria, the gamete frequencies, and the population size at each equilibrium, as well as the number of trajectories (initial distributions) converging to each equilibrium. Using this database, we calculated the following quantities for each equilibrium: the normalized polymorphism $P$ (the number of polymorphic loci divided by the number of loci of the genetic system); the (additive) genetic variance $V_A$; the genic variance $V_{LE}$ (the variance that would be observed under linkage equilibrium); the relative variance $V_R = V_A/V_{max}$, where $V_{max} = \sum_i \gamma_i^2$ is the maximum possible variance in the given genetic system under the assumption of linkage equilibrium (this normalization of the additive genetic variance enables proper comparison of genetic systems with different locus effects or number of loci); two measures of (overall) linkage disequilibrium, $V_D = (V_A - V_{LE})/V_{LE}$ and

$$D = \frac{V_D}{V_{Dmax}}, \tag{21}$$

where

$$V_{Dmax} = \frac{(\sum_i \gamma_i)^2}{\sum_i \gamma_i^2} - 1; \tag{22}$$

and the population size $N$. Whereas $V_D$ is an obvious measure for linkage disequilibrium, the definition of $D$ needs explanation. It is the ratio of $V_D$ and the maximum possible value of $V_D$, $V_{Dmax}$. Therefore, it is a normalized and dimension-free measure with values between $-1$ and 1. For a given genetic system, $V_D$ and $D$ are both maximized if and only if the two gametes that generate the two extreme phenotypes, $-\Gamma$ and $\Gamma$, are each at frequency 1/2. A simple calculation involving equation (A1) in the online appendix yields equation (22). The relation between distributions of genotypic values and $D$ is illustrated in "The Measure $D$ of Linkage Disequilibrium" in the online appendix. A high value of $D$ ($D > 0.8$ or 0.9) indicates that the population consists of two clusters around the extreme phenotypes.

The values $P$, $V_A$, $V_R$, and $N$ were averaged over all equilibria resulting from the 10 initial conditions of all of the 1,000 genetic parameter sets that led to equilibration, and standard deviations were calculated. The linkage disequilibrium measure $D$ was averaged only over the trajectories that reached an equilibrium with at least two polymorphic loci. This yielded our quantities of interest for each set of ecological and mating parameters, number of loci, and recombination scenario. The data presented in the figures and tables are such averages. We denote the averages of $V_A$, $V_R$, $P$, $N$, and $D$ by $\overline{V}_A$, $\overline{V}_R$, $\overline{P}$, $\overline{N}$, and $\overline{D}$, respectively. Similarly, the average proportion of equilibria with $k$ ($k = 0, \ldots, n$) polymorphic loci was calculated.

We note that for a given number of loci, the relative genetic variance $\overline{V}_R$ and the (average) genetic variance $\overline{V}_A$ behave very similarly. Multiplying $\overline{V}_R$ by $E[V_{max}]$, which can be computed explicitly (Bürger and Gimelfarb 2004), yields an estimate of $\overline{V}_A$ that typically is within about 10% of the true value (results not shown).

For a given number of loci and a given recombination scenario, the genetic parameter sets as well as the initial conditions are the same for all ecological parameter sets.

**Table 1:** Equilibrium structure for increasingly strong assortative mating

| | Polymorphism | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $a$ | 0 | 1 | 2 | 3 | 4 | #($E$) | $\overline{P}$ | $\overline{V}_R$ | $\overline{D}$ | $\overline{N}$ |
| 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1.06 | .03 | 10,849 |
| .25 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1.07 | .04 | 10,861 |
| .50 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1.09 | .04 | 10,872 |
| .60 | .01 | .02 | 0 | 0 | .98 | 1.2 | .98 | 1.07 | .05 | 10,861 |
| .75 | .33 | .53 | 0 | 0 | .14 | 4.1 | .28 | .22 | .07 | 10,218 |
| .90 | .62 | .36 | 0 | 0 | .03 | 3.7 | .12 | .07 | .08 | 10,064 |
| 1.00 | .75 | .24 | 0 | 0 | .01 | 3.6 | .07 | .04 | .07 | 10,027 |
| 1.25 | .98 | .01 | 0 | 0 | .00 | 3.3 | .01 | .01 | .06 | 9,991 |
| 1.50 | 1.00 | .00 | 0 | 0 | 0 | 3.5 | .00 | .00 | … | 9,984 |
| 2.00 | 1.00 | .00 | 0 | 0 | 0 | 4.0 | .00 | .00 | … | 9,980 |
| 3.00 | .98 | .00 | .01 | 0 | .01 | 4.7 | .01 | .02 | .41 | 10,004 |
| 4.00 | .90 | .00 | .00 | .00 | .09 | 5.4 | .10 | .19 | .65 | 10,221 |
| 5.00 | .70 | .00 | .00 | .00 | .29 | 5.7 | .29 | .76 | .92 | 10,948 |
| 6.00 | .53 | .00 | .01 | .00 | .46 | 5.2 | .47 | 1.33 | .97 | 11,615 |
| 8.00 | .31 | .00 | .01 | .04 | .63 | 4.3 | .67 | 2.02 | .96 | 12,356 |
| 12.00 | .10 | .00 | .02 | .14 | .74 | 3.4 | .86 | 2.57 | .92 | 12,918 |
| 16.00 | .02 | .00 | .04 | .17 | .76 | 3.2 | .91 | 2.76 | .90 | 13,087 |

Note: Presented are the proportion of trajectories converging to an equilibrium with the indicated number of polymorphic loci, the average number of equilibria, #($E$), and the (average) normalized polymorphism, relative variance, linkage disequilibrium (averaged only over multilocus polymorphisms), and population size. The parameters are $\rho = 2$, $\kappa = 10{,}000$, $n = 4$, $s = 0.4$, $f = 1.5625$ ($c = 0.625$), and $M = \infty$, and there is free recombination. A 0 (1) means that this event never (always) occurred, and a .00 means that it occurred but with frequency <0.005.

Therefore, variation among quantities of interest comes almost exclusively from variation in the ecological and mating parameters. Only the exclusion of slow runs leads to some variation among the genetic parameter sets used for different ecological parameter combinations. Whenever we use the term equilibrium without qualification, we mean a locally asymptotically stable equilibrium.

## Numerical Results

We explore the effects of the interplay of frequency-dependent selection and assortative mating on the population genetic structure, as well as the role of the other parameters (e.g., number of loci, recombination rates, and position of the optimum). We are especially interested in the conditions under which two reproductively isolated clusters, corresponding to evolutionary splitting, emerge. For this purpose, we study the equilibrium structure, the genetic variance maintained, and the amount of linkage disequilibrium. Equilibria with a high amount of linkage disequilibrium $D$ (approximately $D > 0.8$) represent strong competitive divergence because the distribution of phenotypes is concentrated near the ends of the phenotypic range (see "The Measure $D$ of Linkage Disequilibrium" in the online appendix). If, in addition, $a$ is large ($a > 4.6$), we speak of two reproductively isolated clusters because then the probability of mating between the extreme phenotypes is low (<1%).

Our analytical results show that in a randomly mating population at linkage equilibrium, multilocus polymorphisms, that is, equilibria with at least two polymorphic loci, are maintained only if (approximately) $f(\rho - 1) > 1$. Only in this case, frequency dependence is sufficiently strong to induce disruptive selection near equilibrium. Because our numerical computations use logistic population growth with $\rho = 2$, we focus on parameter values such that $f > 1$.

### *No Costs*

We start by presenting the results for the case when there are no costs associated with assortative mating ($M = \infty$). This is the scenario that has been most frequently studied in the context of sympatric speciation (see introductory text).

*Free Recombination and a Symmetric Optimum.* Let us first assume free recombination between all loci and a symmetric optimum ($\theta = 0$). Table 1 shows the effect of increasingly strong assortment on the equilibrium distribution for strong, but not very strong, frequency dependence ($f = 1.56$). If $a \geq 0.5$, increasingly strong assortment decreases genetic variation. Linkage disequilib-

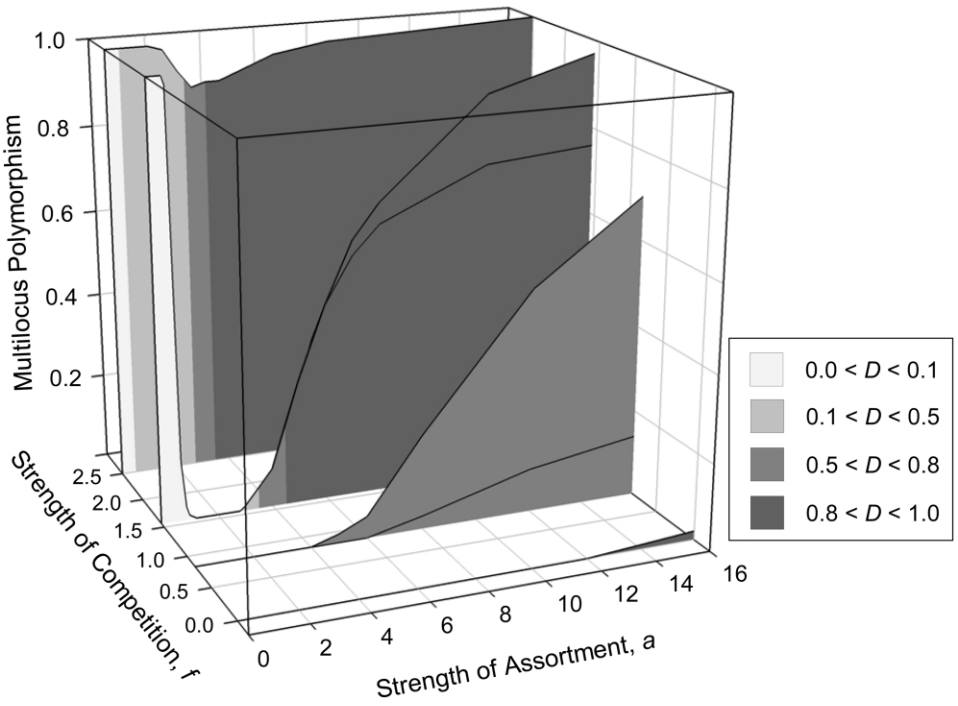


**Figure 1:** Multilocus polymorphism and linkage disequilibrium in four-locus genetic systems as a function of the strength of assortment, $a$, and the strength of frequency dependence, $f$. The four values of $f$ are 0, 0.833, 1.5625, and 2.5 (resulting from $s = 0.4$ and $c = 0$, 1/3, 5/8 = 0.625, and 1). In addition, we have $\rho = 2, \kappa = 10{,}000, \theta = 0, M = \infty$, and free recombination. The "waterfall" plots show the fraction of trajectories converging to an equilibrium with two or more polymorphic loci. The shades in gray signify the amount of linkage disequilibrium, $\overline{D}$. The solid lines in the plains, visible for $f = 0.833$ and $f = 1.5625$, indicate the (absolute) fraction of fully polymorphic equilibria. In the two other cases, they are invisible because >99% of multilocus polymorphisms are fully polymorphic.

rium is positive but very small. For intermediate values of $a$ ($a = 1.5, 2$), almost no genetic variation is maintained, and 99.7% of the trajectories converge to a monomorphic equilibrium. Equations (19) and (20) predict the values $a$ at which the first and last genetic systems develop stable monomorphisms well (according to our theory, they are $a = 0.45$ and $a = 1.25$). If the strength of assortment is increased further, genetic variation is built up again. For strong assortment, in most genetic systems several types of stable equilibria coexist, among them monomorphic equilibria, and high linkage disequilibrium prevails at equilibria with multilocus polymorphism. If $a > 4.6$, these are the equilibria representing two reproductively isolated clusters.

For very large values of $a$, only a small fraction of trajectories converge to a monomorphic equilibrium, but a larger proportion of genetic systems has a stable monomorphic state (e.g., if $a = 16$, convergence to a monomorphic equilibrium occurs in about 22% of genetic systems). This shows that for very strong assortment, the basin of attraction of monomorphic equilibria becomes small. The number of simultaneously stable equilibria is high, higher than the reported #($E$), which is an underestimate because it is calculated from only 10 initial conditions. We further mention (data not shown) that if $a \geq 8$, all genetic systems have a stable fully polymorphic equilibrium (with all gene frequencies $= 1/2$ and large $D$). Thus, depending on the initial conditions, the evolution of reproductively isolated clusters can occur in every genetic system. Interestingly, if $a = 12, 16$, in some genetic systems two stable, fully polymorphic equilibria coexist, one with $D > 0$ and the other with $D < 0$ (about 0.1% if $a = 12$ and 1.5% if $a = 16$). Therefore, even if competition and assortment are both strong, the presence of much initial variation is not sufficient to guarantee divergence because $D < 0$ means that intermediate phenotypes are overrepresented. For $a = 16$, a genetic system was found in which stable equilibria of all kinds of polymorphism (i.e., 0–4 polymorphic loci) coexist.

Figure 1 demonstrates the combined influence of the

strength of competition and of assortment on the equilibrium structure. For the two smaller values of $f$ ($f = 0, 5/6 = 0.833$), multilocus polymorphism is maintained only if assortment is sufficiently strong. Then, however, $\overline{D} < 0.8$ except for $f = 0$ and $a = 12, 14$, when fewer than 2% of the trajectories converge to fully polymorphic equilibria, which have $\overline{D} > 0.9$. If $f = 0$, then no genetic variation at all is maintained if $a < 12$. In the absence of any natural selection ($s = c = 0$), no variation is maintained if $a \leq 10$, whereas the fraction of multilocus polymorphism reaches 0.40 if $a = 16$. If $f = 0.833$, some variation is maintained by single-locus polymorphisms if $a \leq 0.5$, but none if $1 \leq a \leq 4$. The plot for $f = 1.56$ corresponds to table 1 (but is based on more data points). Also, if competition is much stronger than stabilizing selection ($f = 2.5$), the amount of polymorphism is reduced for intermediate values of $a$. If $a \geq 3.5$, then $\overline{D} \geq 0.9$ and strong competitive divergence occurs with a high probability. If $a > 4.6$, according to our definition, the clusters become reproductively isolated. However, even under such favorable conditions for divergence, trajectories starting close enough to a monomorphic equilibrium may converge to it. For strong assortment, the amount of linkage disequilibrium is higher at fully polymorphic equilibria than at other multilocus polymorphisms because they lack at least one of the extreme phenotypes.

We note that $a \geq 8$ implies extremely strong assortment. With $a = 8$ and 16, the probability that an encounter of the two extreme phenotypes ($-1/2$ and $1/2$) leads to mating is $3.35 \times 10^{-4}$ and $1.13 \times 10^{-7}$, respectively. The probability that an extreme phenotype mates with a phenotype in the middle of the phenotypic range is 0.135 if $a = 8$ and 0.018 if $a = 16$.

Although figures 1–4 and tables 1 and 2 report only the properties of an "average genetic system," the particular evolutionary path depends both on the allelic effects and on the initial conditions, and substantial deviations from the average behavior occur in some genetic systems. The standard deviations of the average amount of polymorphism and of the (relative) genetic variance are large, on the order of the mean. Hence, the proportion of initial conditions that lead to strong divergence or to the evolution of reproductively isolated clusters varies substantially among genetic systems. Conditional on convergence to a multilocus polymorphism, strong divergence occurs for most such trajectories in most genetic systems if $a$ is large because the standard deviation of $D$ is quite small (<30% of the mean, often on the order of 10%).

An intuitive explanation for the reduction of variation by intermediately strong assortment is the following. In comparison to random mating, weak assortment counteracts competition because it induces local stabilizing selection around monomorphic states because for deviating types it is harder to find a mating partner. This is more pronounced if mating is strongly selective ($M$ small), and it is reflected by equations (19) and (A17) in the online appendix. Thus, competition must be stronger than stabilizing selection and assortment together to maintain polymorphism with certainty. If assortment is extremely strong, then two monomorphisms that are sufficiently far apart, so that individuals of different type do not mate, can coexist because no intermediate types (which have lower fitness) are produced.

*Number of Loci.* Figure 2 displays the influence of the number of loci on multilocus polymorphism and linkage disequilibrium. The pattern that intermediately strong assortment always decreases genetic variation, and hence impedes divergence, is clearly visible. With two loci, the effect is much weaker than with three or more loci; four- and five-locus systems behave similarly. However, the range of values $a$ for which no multilocus polymorphism (and in fact, no variation at all) is maintained is wider with five loci than with four. For strong assortment ($a > 7$) and four or five loci, high positive linkage disequilibrium ($\overline{D} > 0.9$) is maintained among multilocus polymorphisms, and reproductively isolated clusters evolve frequently.

The data suggest that for very strong assortment, a larger number of loci determining the trait is favorable for the evolution of two reproductively isolated clusters. By contrast, weak or moderately strong assortment is more efficient in attenuating genetic variation of traits determined by many loci. This remains true for very strong competition. If $f = 2.5$, then the (average) fraction of multilocus polymorphisms in a five-locus system is minimized near $a = 3$, its value being 0.58, whereas in two-, three-, and four-locus systems, the minima are assumed near $a = 4$, 4, and 3 with values 0.93, 0.92, and 0.90, respectively. (For these values of $a$, all polymorphic equilibria found numerically are fully polymorphic.)

With increasing number of polymorphic loci, the number of phenotypes in the population increases rapidly. Therefore, it is easier to find a mating partner than when only few types exist in the population (most of them quite different). Thus, monomorphic states can attract a larger proportion of the state space, and assortment is more effective in counteracting competition. Therefore, the gap becomes deeper with more loci. It also becomes slightly wider because with more loci it is more likely to have a phenotype near the optimum; hence condition (19) applies to more genetic systems.

*Linkage.* Figure 3 displays the effects of linkage on the equilibrium properties of a population under very strong competition ($f = 2.5$). Relative to random mating, mod-

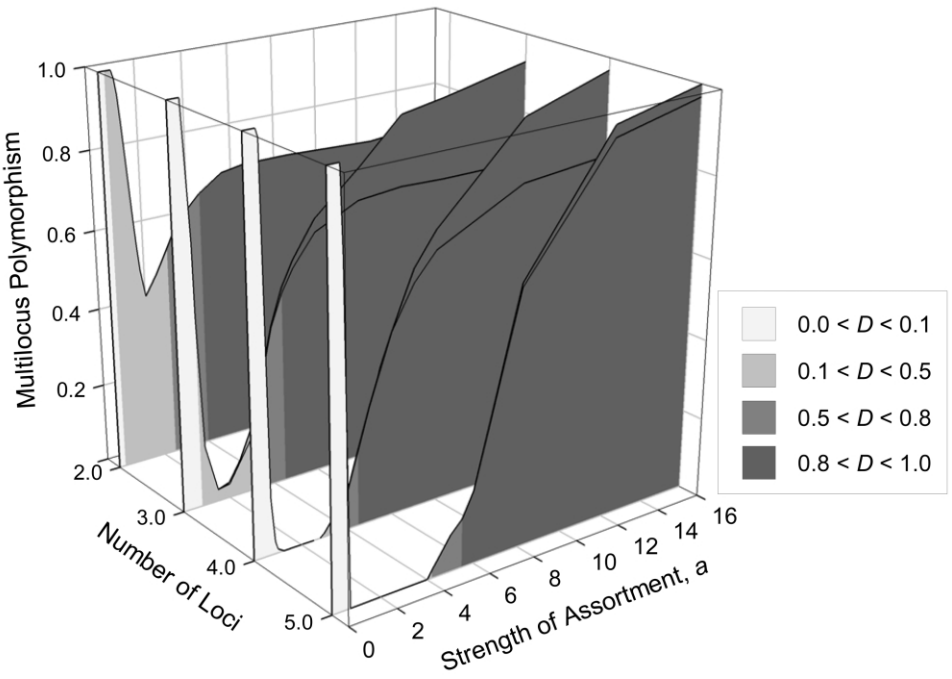


**Figure 2:** Multilocus polymorphism and linkage disequilibrium for strong, but not very strong, frequency dependence ($f = 1.5625$) as a function of $a$ and the number of loci $n$. We have $\rho = 2$, $\kappa = 10{,}000$, $s = 0.4$, $\theta = 0$, $M = \infty$, and free recombination. The "waterfall" plots give the fraction of trajectories converging to an equilibrium with two or more polymorphic loci. The shades in gray signify the amount of linkage disequilibrium $\overline{D}$. The solid lines in the plains signify the (absolute) fraction of fully polymorphic equilibria.

erately strong assortment leads to a much higher loss of multilocus polymorphism if loci are linked than if they are freely recombining. Also, with linked loci, the genetic variance is minimized for intermediate values of $a$, whereas for freely recombining loci, it is increasing (with weaker competition, $f = 1.56$, it has a square-root-sign-like shape; results not shown). Linkage disequilibrium is extremely high for linked loci if assortment is absent or moderate (hence the high variance) and, interestingly, may be lower if assortment is strong or very strong. The reason linkage disequilibrium decreases as assortment becomes strong is that an increasing fraction of the stable multilocus equilibria is not fully polymorphic. Thus, for moderate or strong assortment, increasingly strong linkage reduces the likelihood of convergence to a multilocus polymorphism, and it reduces the amount of linkage disequilibrium and hence of bimodality and strong competitive divergence (recall that $\overline{D}$ is calculated only from trajectories that reached a multilocus polymorphic equilibrium). For weaker competition ($f = 1.56$), the effects of linkage are even more pronounced (results not shown). Thus, with random mating, very tight linkage induces the evolutionary emergence of two isolated clusters with higher probability than with strong assortative mating. Still, there is a difference: with tight linkage and random mating, there is no reproductive isolation; only the offspring are (largely) of one or the other type.

*Free Recombination and an Asymmetric Optimum.* We also investigated a model with asymmetrical selection, where the optimum $\theta$ of stabilizing selection was set to $\theta = 0.25$. The analytical results in the online appendix demonstrate that with random mating, this asymmetry reduces the proportion of genetic systems for which a fully polymorphic equilibrium exists relative to $\theta = 0$. Consequently, our numerical results (fig. 4) show that for weak to moderately strong assortment, the proportion of fully polymorphic equilibria is substantially reduced relative to the symmetric case. The genetic variance and the linkage disequilibrium, however, are only slightly reduced. The only slight reduction of the variance is in accordance with the analytical results for random mating because with an asymmetric optimum, loci of small effect tend to be fixed, whereas those of large effect are polymorphic. With strong assortment, the position of the optimum has very little influence on the equilibrium structure and hence on the evolution of divergence. Qualitatively similar results were obtained for weaker frequency dependence (not shown).

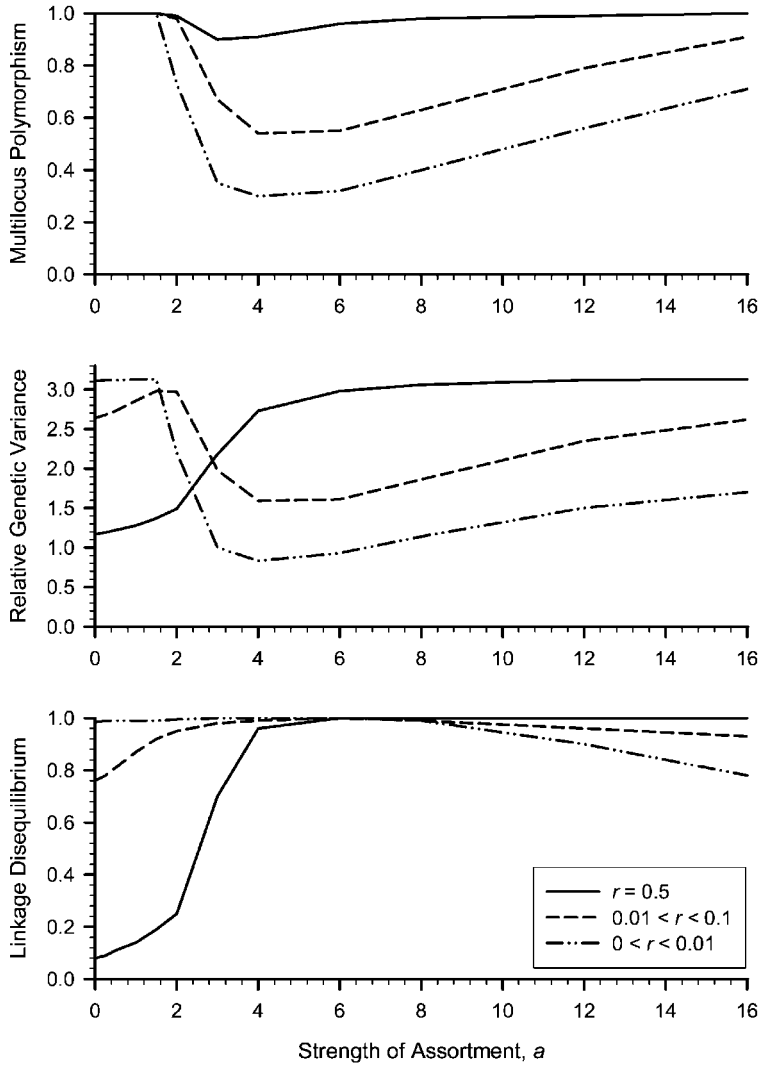


**Figure 3:** Influence of linkage on the equilibrium structure for $n = 4$ loci and very strong frequency dependence ($f = 2.5$). The upper panel displays the fraction of trajectories converging to an equilibrium with two or more polymorphic loci as a function of $a$, the middle panel the relative genetic variance $\overline{V}_R$ (averaged over all trajectories), and the lower panel the measure $\overline{D}$ of linkage disequilibrium. The three lines in each panel are for the three recombination scenarios indicated. We have $\rho = 2$, $\kappa = 10{,}000$, $s = 0.4$, $n = 4$, $\theta = 0$, and $M = \infty$.

### *Equal versus Unequal Locus Effects*

Because the assumption of equal locus effects is widespread in this field, we investigated it in some detail for $f = 1.56$ and $f = 2.5$. We used 20 initial conditions to obtain a more precise characterization of the equilibrium and stability structure. If the optimum is symmetric ($\theta = 0$), then, as a function of $a$, the structure is quite similar; the range of values $a$ that reduce variability most is somewhat broader with equal effects. For an asymmetric optimum ($\theta = 0.25$), much more variation is maintained with equal effects than on average (with randomly assigned effects); in particular, and in contrast to the general case (fig. 4), the symmetric, fully polymorphic equilibrium is globally stable if $a \leq 1.5$. However, little linkage disequilibrium is maintained; thus no divergence occurs. If $2 \leq a \leq 10$, the proportion of multilocus polymorphisms is reduced but always in excess of 0.5. Hence, it is higher than for the average genetic system, especially if $f = 1.56$. If $a \geq 4$, then linkage disequilibrium is high ($\overline{D} > 0.8$); that is, multilocus polymorphisms represent reproductively isolated clusters. If $a \geq 12$, all trajectories converge to fully polymorphic equilibria; hence splitting occurs with certainty. For other genetic systems, this is not necessarily the case (fig. 4 and results for $f = 1.56$ that are not shown).

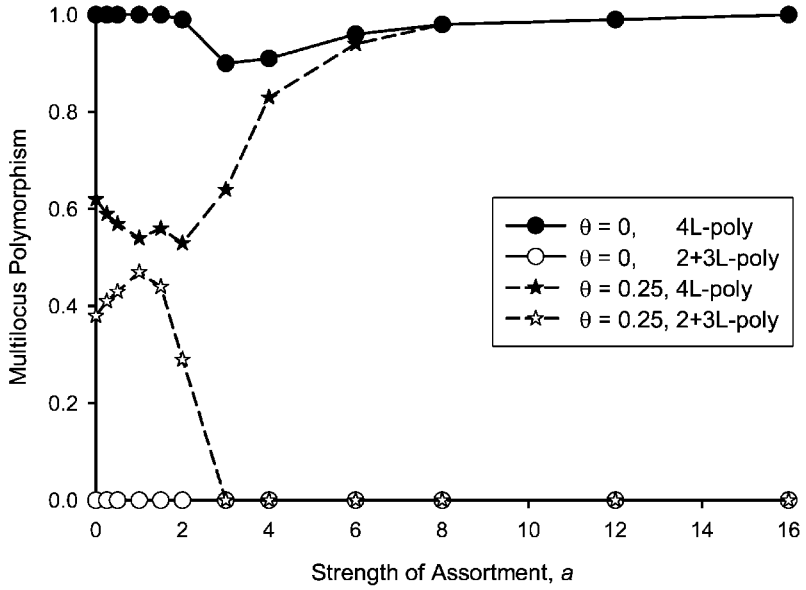


**Figure 4:** Influence of the position $\theta$ of the optimum on multilocus polymorphism (i.e., two or more loci are polymorphic) for the different scenarios indicated. We have $\rho = 2, \kappa = 10{,}000, s = 0.4, n = 4, M = \infty, f = 2.5$, and free recombination. 4L-poly indicates the fraction of trajectories converging to an equilibrium with four polymorphic loci; 2+3L-poly indicates the fraction converging to an equilibrium with two or three polymorphic loci.

### *Population Size*

The equilibrium population size depends strongly on the combination of the selection and mating parameters. If two (reproductively) isolated clusters evolve, it may be twice as high as for a monomorphic population. The reason is that a genetically variable population can exploit the resources more efficiently. As shown in "Population Size at Equilibrium" in the online appendix, total resource utilization depends strongly, and in a nonlinear way, on the parameters *f*, *a*, and the pattern of recombination.

### *Costs*

Radically different results are obtained for moderate and strong assortment if choosiness is very costly. We studied the case $M = 1$ numerically. It leads to sexual selection on both sexes and has several alternative interpretations, for instance, as both sexes being choosy (see above). The results for $n = 4$ and strong frequency dependence ($f = 2.5$) are summarized in table 2. The message is clear: high costs associated with assortative mating deplete any genetic variation and lead to genetic uniformity. Equations (A17) and (A18) in the online appendix predict the values *a* at which the first and last genetic systems develop stable monomorphisms very well (according to our theory, they are $a = 0.6$ and $a = 1.0$). In contrast to the case $M = \infty$, here apparently only monomorphic equilibria are stable if *a* is large. We note that costly assortative mating counteracts competition also indirectly because the population size after mating is reduced, which weakens the frequency-dependent effect of competition (i.e., $\eta(N)$ is reduced).

Schneider and Bürger (2005) performed a detailed analysis of the role of costs for a single-locus haploid model with multiple alleles that allows for arbitrarily strong selection. They showed that the equilibrium structure in models with $M \geq 10$ is very similar to that for $M = \infty$.

**Table 2:** Equilibrium structure for assortative mating with costs ($M = 1$)

| | Polymorphism | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *a* | 0 | 1 | 2 | 3 | 4 | #(*E*) | $\overline{P}$ | $\overline{V}_R$ | $\overline{D}$ | $\overline{N}$ |
| 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1.17 | .08 | 11,933 |
| .25 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1.15 | .07 | 11,297 |
| .50 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1.13 | .06 | 10,695 |
| .60 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1.12 | .06 | 10,457 |
| .70 | .34 | .65 | .00 | 0 | 0 | 4.4 | .17 | .11 | .01 | 10,019 |
| .80 | .63 | .37 | 0 | 0 | 0 | 3.6 | .09 | .04 | … | 9,987 |
| .90 | .87 | .13 | 0 | 0 | 0 | 3.3 | .03 | .01 | … | 9,980 |
| 1.00 | 1 | 0 | 0 | 0 | 0 | 3.3 | 0 | 0 | … | 9,978 |
| 2.00 | 1 | 0 | 0 | 0 | 0 | 5.4 | 0 | 0 | … | 9,920 |
| 4.00 | 1 | 0 | 0 | 0 | 0 | 7.0 | 0 | 0 | … | 9,822 |
| 8.00 | 1 | 0 | 0 | 0 | 0 | 8.2 | 0 | 0 | … | 9,735 |
| 16.00 | 1 | 0 | 0 | 0 | 0 | 9.0 | 0 | 0 | … | 9,686 |

Note: The parameters are $\rho = 2$, $\kappa = 10{,}000$, $n = 4$, $s = 0.4$, and $f = 2.5$ ($c = 1$), and there is free recombination (see table 1 for details).

Unless assortment is extremely strong, the case $M = 5$ is more similar to $M = \infty$ than to $M = 1$, but the likelihood of converging to the equilibrium at which only the two extreme phenotypes are present is lower compared with $M = \infty$. The case $M = 2$ is, if averaged over many genetic systems, similar to $M = 1$. Although divergence can occur if $M = 2$, it is extremely unlikely. Therefore, it seems that weak or moderate costs for being choosy have only a slight negative effect on the evolution of competitive divergence.

## Discussion

This work is concerned with frequency-dependent selection and its ability to induce sympatric speciation in the presence of assortative mating. In our model, frequency-dependent selection is induced because individuals of similar phenotype compete for a continuous resource. The same trait causes assortative mating because the probability of mating is based on phenotypic similarity with respect to this trait. The strength of assortment is tuned by a parameter, as are the costs for being choosy (Gavrilets and Boake 1998; Matessi et al. 2001). Thus, we are dealing with a "magic" trait, an assumption that is considered to be most conducive for sympatric speciation (Gavrilets 2004, chap. 10). The main motivation for this work results from the need to characterize the conditions that enable competitively driven divergence within a population.

Our focus here is on the identification of the conditions that lead to the evolution of two reproductively isolated clusters of phenotypes, which may be interpreted as an important step toward sympatric speciation. Because the strength of assortment is controlled by a parameter, we can determine the strength that is necessary for divergence as a function of the other parameters. A simplifying assumption is haploidy of individuals, which has, however, been made by other authors too (e.g., in some of the models of Gavrilets 2003, 2004; Gourbiere 2004). The main reason for this is limitations of computing time because with diploid individuals, genotype frequencies have to be followed. In a diploid population, stronger competition and assortment will be required to evolve strong divergence and reproductively isolated clusters because a much higher proportion of intermediate types is produced. Preliminary numerical results confirm this expectation (R. Bürger, K. A. Schneider, and M. Willensdorfer, unpublished data).

In contrast to previous studies (Dieckmann and Doebeli 1999; Bolnick 2004*b*; Gavrilets 2004, chap. 10.3; Gourbiere 2004; Kirkpatrick and Nuismer 2004), we do not assume that loci have equal effects. Rather, we explore randomly sampled genetic systems with two alleles per locus. Of course, we do not use the hypergeometric model (e.g., Doebeli 1996; Kondrashov and Kondrashov 1999), which even posits identical allele frequencies across loci. This extreme symmetry assumption often forces evolutionary trajectories to converge to equilibria representing strong divergence and reproductive isolation, whereas the correct evolutionary dynamics may converge to a different equilibrium, often with very little variation (see also Gavrilets 2004, pp. 377–380).

A further simplifying assumption we have made is that of weak selection. This has the advantage that the resulting fitness function, equation (8), is the second-order approximation (i.e., to order $s^2$) to all models of competition for a continuous and unimodally distributed resource that the authors are aware of (see Bürger 2005 for details). Hence, our results are representative of all these models unless natural selection is too strong, that is, if the fitness function can be approximated by a quadratic on the whole range of phenotypes. This model is also easily amenable to mathematical analysis. In an earlier work (Schneider and Bürger 2005), we studied a single-locus model with multiple alleles in which stabilizing selection and competition are modeled by Gaussian functions and thus can be arbitrarily strong, and assortment is modeled as here. Interestingly, the evolution of maximum divergence and complete reproductive isolation becomes less likely as competition gets extremely strong.

Summarizing, it seems that the model assumptions underlying this work are very favorable for the evolution of competitively driven divergence. Nevertheless, our results demonstrate that the evolution of divergence and eventual splitting is less likely than suggested by Dieckmann and Doebeli (1999) and Kondrashov and Kondrashov (1999). It is constrained by the genetic assumptions and the initial genetic composition of the population. Assortative mating, even if it has no costs, has a tendency to reduce polymorphism relative to random mating (table 1; figs. 1–3). Hence, it impedes or may even prohibit divergence. This effect is most pronounced for moderately strong assortment, which can lead to complete depletion of genetic variation. The attenuating effect of moderate assortment increases with the number of loci contributing to the trait. It provides an explanation for the finding of Matessi et al. (2001) that under disruptive selection and if costs for assortment are absent or low, there may be a barrier to a transition in small steps from weak to very strong assortment, although a modifier enhancing assortment can invade a randomly mating population, and complete reproductive isolation is evolutionarily stable. This is in contradistinction to the numerical results of Dieckmann and Doebeli (1999), who showed that strong assortment does evolve even if there are several "mating" loci. However, their figure 5*b* shows that the waiting time to branching increases rapidly with the number of mating loci. Apparently, in Dieckmann and Doebeli's model, several (or

all) mating loci have to contribute together to induce a more or less sudden sweep from random mating to very strong assortment. This is greatly facilitated by their assumption of initial allele frequencies of 1/2 at all loci, including the mating loci, and their high mutation rates of $10^{-3}$.

Another important aspect of our results is that the evolutionary outcome may depend heavily on the initial conditions, especially if assortment is strong. In fact, we proved analytically that with moderate or strong assortment, there always exist stable monomorphic equilibria. Therefore, a fraction of trajectories converge to a monomorphic equilibrium. Also, typically, equilibria with fewer than $n$ polymorphic loci, and thus less than maximum divergence (and reduced reproductive isolation), are stable. Populations with little genetic variance are much less likely to split than populations with high initial variability; rather, assortment will tend to further reduce their genetic variation because it induces stabilizing selection locally. Indeed, as shown by some of our numerical results, even under very strong competition and assortment, a high amount of initial genetic variation does not necessarily lead to the evolution of reproductively isolated clusters but may lead to fully polymorphic equilibria with negative linkage disequilibrium. In general, our choice of initial conditions favors evolutionary divergence because they are sampled uniformly from all frequency distributions. Hence, initially, the majority of alleles is at intermediate frequency. Empirical evidence lends stronger support to the maintenance of genetic variation by rare alleles of large effects (Barton and Turelli 1987; Mackay 2001; Barton and Keightley 2002). If mutation were included, a monomorphic equilibrium in this model would correspond to a distribution with relatively little variation maintained by mutation. Still, heritabilities up to 20% can be explained by mutation-selection balance (Bürger 2000). Interestingly, frequency dependence together with assortment can give rise to such equilibria, to highly polymorphic ones where alleles are at intermediate frequency, and to evolutionary divergence. Thus, it has the potential to generate a large variety of evolutionary outcomes.

The dependence on the initial conditions was ignored by most earlier studies and would deserve closer examination. For some such results, see the works by Geritz and Kisdi (2000), Kirkpatrick and Ravigné (2002), Gavrilets (2004, pp. 380–382), Kirkpatrick and Nuismer (2004), and Schneider and Bürger (2005). In addition to the initial conditions, the evolutionary fate of a population depends on the genetic basis of the trait, that is, on the locus effects and the recombination rates. Among genetic systems differing in the number of loci and locus effects, there is considerable variation in the equilibrium structure and hence in the likelihood of divergence and splitting.

Linkage reduces the fraction of trajectories converging to equilibria representing strong divergence (i.e., those with high linkage disequilibrium) unless assortment is weak (fig. 3). In fact, linkage disequilibrium and population size are maximized in a random mating population with tightly linked loci. The two emerging clusters are, of course, not reproductively isolated, but the phenomenon shows that strong assortment and speciation are not the only means to use resources most efficiently. Frequency-dependent disruptive selection can also lead to the evolution of dominance (Van Dooren 1999) or sexual dimorphism (Bolnick and Doebeli 2003). A well-documented example from the empirical literature that competition for resources does not lead to the evolution of assortative mating and speciation is provided by the African finch *Pyrenestes ostrinus* (Smith 1990, 1993). Here, a single-locus polymorphism with complete dominance underlies the two bill morphs specialized on the different resources, and mating with respect to bill traits appears to be at random.

We also investigated the effects of an asymmetric resource distribution. For weak assortment, asymmetry leads to loss of genetic variation. For strong assortment, asymmetry affects the likelihood of the evolution of divergence and the establishment of reproductively isolated clusters only slightly.

Finally, we explored the role of costs for being choosy. We showed analytically that increasing costs enhance stability of monomorphic equilibria and thus reduce the probability of divergence. We showed numerically that with high costs ($M = 1$) or, equivalently, if both sexes are choosy, moderate or strong assortment depletes all genetic variation; that is (in the absence of mutation), it leads to convergence to one of the monomorphic states. Similar results have been found by Drossel and McKane (2000), Bolnick (2004*b*), Gavrilets (2004, chap. 9.2), Gourbiere (2004), and Kirkpatrick and Nuismer (2004). Schneider and Bürger (2005) performed a detailed study of the role of costs for a single-locus model with multiple alleles. It appears that only high costs, or strong selective mating in both sexes, inhibit divergence.

Our results as well as those from previous studies show that the condition $\sigma_C < \sigma_K$ proposed by Dieckmann and Doebeli (1999) to imply evolutionary branching is insufficient in various ways (their condition is equivalent to our $f(\rho - 1) > 1$). It is necessary, however, because otherwise frequency dependence is too weak to generate disruptive selection and maintain (much) variation (Bulmer 1974, 1980; Slatkin 1979; Christiansen and Loeschcke 1980; Bürger and Gimelfarb 2004; Bürger 2005). In addition to sufficiently strong frequency-dependent competition, appropriate initial conditions and very strong assortment are necessary prerequisites for evolutionary

splitting or speciation. How this can evolve in view of the reduction of variation by moderately strong assortment remains unclear and deserves further investigation. Despite the various limitations identified by this and previous work, sympatric speciation driven by competition or some other form of frequency-dependent or balancing selection that induces disruptive selection continues to be not only a fascinating but also a viable hypothesis. More work, theoretical and empirical, will be necessary to reach a definite conclusion. For instance, it would be important to have empirical estimates of the strength of assortment from populations that are candidates for strong frequency-dependent selection or incipient sympatric speciation. On the theoretical side, the importance of sympatric speciation relative to other evolutionary strategies as a response to disruptive selection awaits further evaluation.

In contrast to the popularity among theoreticians of models of sympatric speciation driven by competition, there is only sparse empirical support for them. Often-cited examples, for which more or less convincing evidence for sympatric speciation has been provided, include crater lake cichlids (Schliewen et al. 1994) and host races in the apple maggot fly *Rhagoletis pomonella* (Bush 1994; see also Via 2001). In the first case, the driving force of speciation has not been established, whereas in the second case there is disruptive selection caused by the discreteness of resources. Only few studies (Swanson et al. 2003; Bolnick 2004*a*) have shown that intraspecific competition indeed induces disruptive selection, and most empirical evidence on the role of resource competition in evolutionary diversification concerns interspecific competition (Schluter 2000). Even then, frequency-dependent selection has been established only rarely (Schluter 2003). Although disruptive selection may occur more frequently in nature than previously thought (Kingsolver et al. 2001), the mechanisms that generate it have not been explored.

Recently, it has been demonstrated that "the parallel build-up of mating incompatibilities between stickleback populations can be largely accounted for by assortative mating based on one trait, body size, which evolves predictably according to environment" (McKinnon et al. 2004, pp. 294–295). In other words, body size in sticklebacks is likely to be both ecologically important and a mediator of assortative mating. This lends some support to magic-trait models. Another case in point may be body size in seahorses (Jones et al. 2003). Empirical estimates for the costs of choosiness are rare, but existing data suggest that $M$ is not small (Jennions and Petrie 1997; Jones et al. 2003). There are good arguments, however, that $M$ can range from very small to very large values (Bolnick 2004*b*; Gavrilets 2004). More empirical studies of these issues are needed before theoreticians can be confident of the biological relevance of their models of competitively driven sympatric speciation.

## Acknowledgments

This work would not have been feasible without the contribution of S. Gimelfarb, who provided the computer program and worked to improve it until shortly before his death. It is his will not to be listed as a coauthor. We are grateful to J. Hermisson and T. Kawecki for many useful comments. This work was supported by grant P16474-N04 of the Austrian Science Foundation (FWF).

## Literature Cited

Barton, N. H., and P. Keightley. 2002. Understanding quantitative genetic variation. Nature Reviews Genetics 3:11–21.

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

Bodmer, W. F. 1965. Differential fertility in population genetic models. Genetics 51:411–424.

Bolnick, D. I. 2004*a*. Can intraspecific competition drive disruptive selection? an experimental test in natural populations of sticklebacks. Evolution 58:608–613.

———. 2004*b*. Waiting for sympatric speciation. Evolution 58:895–899.

Bolnick, D. I., and M. Doebeli. 2003. Sexual dimorphism and adaptive speciation: two sides of the same ecological coin. Evolution 57:2433–2449.

Bulmer, M. G. 1974. Density-dependent selection and character displacement. American Naturalist 108:45–58

———. 1980. The mathematical theory of quantitative genetics. Clarendon, Oxford.

Bürger, R. 2000. The mathematical theory of selection, recombination, and mutation. Wiley, Chichester.

———. 2002. On a genetic model of intraspecific competition and stabilizing selection. American Naturalist 160:661–682.

———. 2005. A multilocus analysis of intraspecific competition and stabilizing selection on a quantitative trait. Journal of Mathematical Biology 50:355–396.

Bürger, R., and A. Gimelfarb. 2004. The effects of intraspecific competition and stabilizing selection on a polygenic. Genetics 167: 1425–1443.

Bush, G. L. 1994. Sympatric speciation in animals: new wine in old bottles. Trends in Ecology & Evolution 9:285–288.

Christiansen, F. B., and V. Loeschcke. 1980. Evolution and intraspecific exploitative competition. II. One-locus theory for small additive gene effects. Theoretical Population Biology 18:297–313.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer, Sunderland, MA.

Dieckmann, U., and M. Doebeli. 1999. On the origin of species by sympatric speciation. Nature 400:354–357.

Dieckmann, U., M. Doebeli, J. A. J. Metz, and D. Tautz. 2004. Adaptive speciation. Cambridge University Press, Cambridge.

Doebeli, M. 1996. A quantitative genetic competition model for sympatric speciation. Journal of Evolutionary Biology 9:893–909.

Doebeli, M., and U. Dieckmann. 2000. Evolutionary branching and

sympatric speciation caused by different types of ecological interaction. American Naturalist 156(suppl.):S77–S101.

Drossel, B., and A. McKane. 2000. Competitive speciation in quantitative genetics models. Journal of Theoretical Biology 204:467–478.

Feldman, M. W. 1971. Equilibrium studies of two locus haploid populations with recombination. Theoretical Population Biology 2:299–318.

Gavrilets, S. 1998. One-locus two-allele models with maternal (paternal) selection. Genetics 149:1147–1152.

———. 2003. Models of speciation: what have we learned in 40 years? Evolution 57:2197–2215.

———. 2004. Fitness landscapes and the origin of species. Princeton University Press, Princeton, NJ.

Gavrilets, S., and C. R. B. Boake. 1998. On the evolution of premating isolation after a founder event. American Naturalist 152:706–716.

Gavrilets, S., and D. Waxman. 2002. Sympatric speciation by sexual conflict. Proceedings of the National Academy of Sciences of the USA 99:10533–10538.

Geritz, S. A. H., and E. Kisdi. 2000. Adapative dynamics in diploid, sexual populations and the evolution of reproductive isolation. Proceedings of the Royal Society of London B 267:1671–1678.

Gourbiere, S. 2004. How do natural and sexual selection contribute to sympatric speciation? Journal of Evolutionary Biology 17:1297–1309.

Hadeler K. P., and U. Liberman 1975. Selection models with fertility differences. Journal of Mathematical Biology 2:19–32.

Higashi, M., G. Takimoto, and N. Yamamura. 1999. Sympatric speciation by sexual selection. Nature 402:523–526.

Jennions, M. D., and M. Petrie. 1997. Variation in mate choice and mating preferences: a review of causes and consequences. Biological Reviews of the Cambridge Philosophical Society 72:283–327.

Jones, A. G., G. I. Moore, C. Kvarnemo, D. Walker, and J. C. Avise. 2003. Sympatric speciation as a consequence of male pregnancy in seahorses. Proceedings of the National Academy of Sciences of the USA 100:6598–6603.

Kawecki, T. 1997. Sympatric speciation by habitat specialization driven by deleterious mutations. Evolution 51:1751–1763.

Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gibert, and P. Beerli. 2001. The strength of phenotypic selection in natural populations. American Naturalist 157:245–261.

Kirkpatrick, M., and S. L. Nuismer. 2004. Sexual selection can constrain sympatric speciation. Proceedings of the Royal Society of London B 271:687–693.

Kirkpatrick, M., and V. Ravigné. 2002. Speciation by natural and sexual selection: models and experiments. American Naturalist 159(suppl.):S22–S35.

Kirzhner, V., and Y. Lyubich. 1997. Multilocus dynamics under haploid selection. Journal of Mathematical Biology 35:391–408.

Kondrashov, A. S., and F. A. Kondrashov. 1999. Interactions among quantitative traits in the course of sympatric speciation. Nature 400:351–354.

Loeschcke, V., and F. B. Christiansen. 1984. Evolution and intraspecific exploitative competition. II. A two-locus model for additive gene effects. Theoretical Population Biology 26:228–264.

Mackay, T. F. C. 2001. The genetic architecture of quantitative traits. Annual Review of Genetics 35:303–339.

Matessi, C., A. Gimelfarb, and S. Gavrilets. 2001. Long-term buildup of reproductive isolation promoted by disruptive selection: how far does it go? Selection 2:41–64.

McKinnon, J. S., S. Mort, B. K. Blackman, L. David, D. M. Kingsley, L. Jamieson, J. Chou, and D. Schluter. 2004. Evidence for ecology's role in speciation. Nature 429:294–298.

Parker, G. A., and L. Partridge. 1998. Sexual conflict and speciation. Philosophical Transactions of the Royal Society of London B 353: 261–274.

Rutschman, D. H. 1994. Dynamics of the two-locus haploid model. Theoretical Population Biology 45:167–176.

Schliewen, U. K., D. Tautz, and S. Pääbo. 1994. Sympatric speciation suggested by monophyly of crater lake cichlids. Nature 368:629–632.

Schluter, D. 2000. The ecology of adaptive radiation. Oxford University Press, Oxford.

———. 2003. Frequency dependent natural selection during character displacement in sticklebacks. Evolution 57:1142–1150.

Schneider, K. A. 2005. Competitive divergence in non-random mating populations. Theoretical Population Biology 68:105–118.

Schneider, K. A., and R. Bürger. 2005. Does competitive divergence occur if assortative mating is costly? Journal of Evolutionary Biology, doi:10.1111/j1420-9101.2005.01001.x.

Slatkin, M. 1979. Frequency- and density-dependent selection on a quantitative character. Genetics 93:755–771.

Smith, T. B. 1990. Natural selection on bill characters in two bill morphs of the African finch *Pyrenestes ostrinus*. Evolution 44:832–842.

———. 1993. Disruptive selection and the genetic basis of bill size polymorphism in the African finch *Pyrenestes*. Nature 363:618–620.

Swanson, B. O., A. C. Gibb, J. C. Marks, and D. A. Hendrickson. 2003. Trophic polymorphism and behavioral differences decrease intraspecific competition in a cichlid, *Herichthys minckleyi*. Ecology 84:1441–1446.

Thieme, H. 2003. Mathematics in population biology. Princeton University Press, Princeton, NJ.

Turelli, M., and N. H. Barton. 2004. Polygenic variation maintained by balancing selection: pleiotropy, sex-dependent allelic effects and G × E interactions. Genetics 166:1053–1079.

Turelli, M., N. H. Barton, and J. A. Coyne. 2001. Theory and speciation. Trends in Ecology & Evolution 16:330–343.

Van Dooren, T. J. M. 1999. The evolutionary ecology of dominance-recessivity. Journal of Theoretical Biology 198:519–532.

van Doorn, G. S., A. J. Noest, and P. Hogeweg. 1998. Sympatric speciation and extinction driven by environment dependent sexual selection. Proceedings of the Royal Society of London B 265:1915–1919.

Via, S. 2001. Sympatric speciation in animals: the ugly duckling grows up. Trends in Ecology & Evolution 16:381–390.

Associate Editor: Sergey Gavrilets
Editor: Michael C. Whitlock

# Appendix from R. Bürger and K. A. Schneider, "Intraspecific Competitive Divergence and Convergence under Assortative Mating" (Am. Nat., vol. 167, no. 2, p. 190)

## Analytical Results and Derivations

### The Linkage Equilibrium Approximation for a Randomly Mating Population

We assume a randomly mating population of constant size, for instance, at demographic equilibrium. The proofs are analogous to those for the diploid case treated in Bürger (2005), actually slightly simpler. Therefore, we summarize only the relevant results and refer to this article for more details. In linkage equilibrium (LE), the evolutionary dynamics can be described by the allele frequencies $P_i$. The functional forms of $W(g)$ and $\overline{W}$ are affected by the LE assumption only insofar as the genetic variance $V_A$ has to be replaced by the linkage equilibrium variance,

$$V_{LE} = 4\sum_i \gamma_i^2 P_i(1 - P_i). \tag{A1}$$

The population mean is not affected by this assumption and can be written as

$$\bar{g} = \sum_i \gamma_i(2P_i - 1). \tag{A2}$$

We denote the values of the functions $F$ and $\eta$ at demographic equilibrium ($\hat{N}$) by $\hat{F}$ and $\hat{\eta}$, respectively. Then, the LE approximation, or the weak-selection limit, of the full genetic dynamics (eq. [17]) is of the much simpler form

$$\dot{P}_i = 4\hat{F}s\gamma_i^2 P_i(1 - P_i)\left[(2P_i - 1)(1 - \hat{\eta}) - \frac{\Delta}{\gamma_i}\right], \tag{A3}$$

where $\Delta = \bar{g} - \theta$, which is a function of the $P_i$ ($i = 1, \ldots, n$). Thus, frequency dependence enters the dynamics and equilibrium structure only through the mean phenotype $\bar{g}$. The system of equation (A3) is closely related to models of Zhivotovsky and Gavrilets (1992) and Turelli and Barton (2004). The system (A3) is a generalized gradient system with potential $V(g) = \hat{F}[1 - s\Delta^2 - (1 - \hat{\eta})V_{LE}]$. Therefore, all solutions converge to the set of equilibria and, if there is a single locally stable equilibrium, it must be globally stable (Bürger 2005). The main results are the following.

*Result 1*. Let $\hat{\eta} < 1$. Then at most one locus can be polymorphic at a stable equilibrium, and, typically, multiple stable equilibria coexist. In addition, for every given set of locus effects ($\gamma_i$) such that no genotype matches the optimum $\theta$, only single-locus polymorphisms can be stable if $\hat{\eta}$ is sufficiently close to 1.

*Result 2*. Let $\hat{\eta} > 1$. Then there exists a unique (globally) asymptotically stable equilibrium. At least one locus is polymorphic. (i) This equilibrium is fully polymorphic if and only if

$$\gamma_i > \frac{|\theta|}{n + \hat{\eta} - 1} \quad \text{for all } i. \tag{A4}$$

The allele frequencies are

$$\hat{P}_i = \frac{1}{2} + \frac{1}{2\gamma_i}\frac{\theta}{n + \hat{\eta} - 1}. \tag{A5}$$

(ii) A stable, fully polymorphic and symmetric equilibrium ($\hat{P}_i = 1/2$ for all $i$) exists if and only if $\theta = 0$. The only stable equilibrium at which $\hat{\bar{g}} = \theta$ is the fully polymorphic symmetric equilibrium. (iii) If the unique stable equilibrium is not fully polymorphic (hence, $\theta \neq 0$), we order loci such that $\gamma_1 \leq \gamma_2 \leq \ldots \leq \gamma_n$. Then (*a*) there exists a unique integer $m$, where $1 \leq m \leq n-1$, such that loci $1, \ldots, m$ are fixed, and loci $m+1, \ldots, n$ are polymorphic. This $m$ is the largest integer $\leq n-1$ that satisfies

$$\gamma_m(n - m + \hat{\eta}) + \sum_{i=1}^{m-1}\gamma_i < |\theta|, \tag{A6}$$

and (*b*) the allele frequencies of the fixed loci are all 1 (i.e., the + allele is fixed) if $\theta > 0$, and all 0 if $\theta < 0$; the allele frequencies of the polymorphic loci are

$$\hat{P}_i = \frac{1}{2} + \frac{1}{2\gamma_i}\frac{\theta - \delta_\theta \sum_{j=1}^{m}\gamma_j}{n - m + \hat{\eta} - 1}, \tag{A7}$$

where $\delta_\theta = 1$ if $\theta > 0$, and $\delta_\theta = -1$ if $\theta < 0$. We have $\hat{P}_i > 1/2$ if and only if $\theta > 0$.

Simple calculations show that the deviation of the mean from the optimum is

$$\hat{\Delta} = \frac{\delta_\theta \sum_{i=1}^{m}\gamma_i - \theta}{1 + (n-m)/(\hat{\eta} - 1)}, \tag{A8}$$

and the variance is

$$\hat{V}_{\mathrm{LE}} = \sum_{i=m+1}^{n}\gamma_i^2 - (n-m)\frac{(\sum_{i=1}^{m}\gamma_i - |\theta|)^2}{(n - m + \hat{\eta} - 1)^2}. \tag{A9}$$

The above results demonstrate that the equilibrium structure depends crucially on whether $\hat{\eta} > 1$ or not. For the model with population regulation, this condition is deceptively simple because $\hat{\eta}$ depends on the population size, which is not (exactly) known without solving the full model. It is readily shown that in demographic equilibrium

$$\hat{N} = K + sK\left[2fV_{\mathrm{A}} + \frac{V_{\mathrm{A}} + \Delta^2}{KF'(K)}\right] + O(s^2). \tag{A10}$$

Therefore, simple approximations for $\hat{\eta}$ can be derived. For the discrete logistic model, one obtains

$$\hat{\eta} = f(\rho - 1) + sf\rho\{[2f(\rho - 1) - 1]V_{\mathrm{A}} - \Delta^2\} + O(s^2). \tag{A11}$$

An explicit expression for the first-order approximation of $\hat{\eta}$ in $s$ is obtained from equation (A11) by substituting $\hat{\Delta}$ and $\hat{V}_{\mathrm{LE}}$, calculated from equations (A8) and (A9) by using the zeroth-order approximation $\hat{\eta} = f(\rho - 1)$, for $\Delta$ and $V_{\mathrm{A}}$, respectively. By numerical solution of a cubic, even the exact value of $\hat{\eta}$ can be calculated (Bürger 2005).

As a consequence, simple approximations for the allele frequencies $\hat{P}_i$ and other derived quantities are easily obtained, even in the presence of population regulation. Comparison with exact results from numerical iteration of the recursion relations of the full model show that in the haploid case, the LE approximation is accurate as long as linkage is not too tight (see table A1 for a typical example). In the diploid case, the LE approximation is accurate for a larger range of recombination rates because linkage disequilibria are smaller than in the corresponding haploid case. The reason is that, because the phenotypic range is the same, selection on alleles is

stronger in the haploid case. If the population size is not assumed constant but is regulated according to equation (18), the equilibrium structure is identical except for a small interval of size $O(s)$ around $\hat{\eta} = 1$, in which equilibria described in results 1 and 2 may coexist (Bürger 2005). Results 1 and 2 can be generalized to include multiple alleles. It can be shown that if a locus is polymorphic at equilibrium, then exactly two alleles are segregating. If $\hat{\eta} > 1$, these are the two alleles with the most extreme effects, each on one side of the optimum (Schneider 2006).

## Stability of Monomorphic Equilibria under Assortative Mating

For the monomorphic equilibria, the eigenvalues can be calculated in the general model with population regulation, linkage disequilibrium, and assortative mating. For other equilibria this seems to be impossible because no explicit formulas are available. Let $v$ denote an arbitrary gamete. Then the eigenvalues $\lambda_r$ ($r \neq v$) at the monomorphic equilibrium $p_v = 1$ are given by

$$\lambda_r = \frac{1 - s(g_r - \theta)^2 + s\hat{\eta}(g_v - g_r)^2}{1 - s(g_v - \theta)^2}\{1 + \pi(g_v - g_r) - [1 - \pi(g_v - g_r)]^M\}R(rv \to r), \tag{A12}$$

where $\hat{\eta} = f[\rho - 1 - s\rho(g_v - \theta)^2]$ because the equilibrium population size is

$$\hat{N} = \kappa\{\rho - [1 - s(g_v - \theta)^2]^{-1}\} \tag{A13}$$

(calculated from $\overline{W} = 1$). The equilibrium $p_v = 1$ is locally asymptotically stable if and only if $|\lambda_r| < 1$ for all $r \neq v$. We note that $R(rv \to r) \leq 1/2$ and $R(rv \to r) = 1/2$ if $v$ and $r$ differ at only one locus or if there is no recombination at all. With recombination and if $v$ and $r$ differ in many loci, $R(rv \to r)$ can be very small. The demographic equilibrium $\hat{N}$ exists (i.e., $\hat{N} > 0$) and is locally stable if and only if

$$1 < \rho[1 - s(g_v - \theta)^2] < 3, \tag{A14}$$

where the right inequality holds trivially if $\rho < 3$, as is assumed throughout. The left inequality means that if $\rho$ is small or $s$ is large, equilibria with an effect $g_v$ deviating too far from the optimum cannot be realized (because $\hat{N} = 0$). The proofs are given below in "Derivation of Eigenvalues for Monomorphic Equilibria."

These results generalize those of Gavrilets (2004, pp. 380–382) for a two-locus model with loci of equal effects and $M = 1$ that becomes a special case of ours if in his fitness function, terms of order $cs$ are ignored (i.e., our fitness function [8] is the weak-selection limit of his). In addition, he showed that several polymorphic equilibria may exist and discussed the consequences for sympatric speciation.

Next, we prove the conclusions (*a*)–(*c*) in "Analytical Results." Unfortunately, the magnitude of the eigenvalues does not permit any conclusion on the size of the basin of attraction.

Proof of (*a*): Let $f(\rho - 1) > 1$. Because the term in braces in equation (A12) equals 2 if $a = 0$, and $R(vr \to r) = 1/2$ if $r$ and $v$ differ at only one locus, it is sufficient to show that for every $v$ an $r$ exists such that the fraction in equation (A12) is >1. Write $g_v = \theta + e$ with $e \geq 0$ and let $g_r = \theta + e - h$ with $h > 0$. Then this fraction is >1 if and only if $h[1 - f(\rho - 1) + e^2 f\rho s] < 2e$. If $e = 0$, this holds for all $h$. The left inequality in (A14) informs us that it is sufficient to consider equilibria satisfying $e^2 s < (\rho - 1)/\rho$. This implies that $1 - f(\rho - 1) + e^2 f\rho s$ is monotone decreasing in $f$ for all possible $e$, and it is sufficient to have $he^2 s\rho/(\rho - 1) < 2e$ (because $f > 1/(\rho - 1)$). This is fulfilled if $h < 2e$ can be chosen. It surely holds for all $h$ (which have to satisfy $h \leq \Gamma + |\theta|$) and any $e$ ($e \leq \Gamma + |\theta|$) if $s < 2(1 - 1/\rho)(\Gamma + |\theta|)^{-2}$. Unless $\rho$ is close to 1, this restriction on $s$ is only slightly stronger than our general restriction $s < (\Gamma + |\theta|)^{-2}$. An analogous proof applies if $e < 0$, when $h < 0$ has to be chosen.

Proof of (*b*): The first and second statements are valid because $\lambda_r > 0$ and $\lambda_r$ is a decreasing function in $a$ and an increasing function in $M$. For the third statement, observe that if $M < \infty$, then $1 + \pi(g_v - g_r) - [1 - \pi(g_v - g_r)]^M$ decreases to 0 as $a \to \infty$ if $g_r \neq g_v$. Therefore, all monomorphic equilibria with $\hat{N} > 0$ (eq. [A13]) are stable for sufficiently strong assortment, even if $f(\rho - 1) > 1$. Exceptions can occur only if the phenotype is not uniquely determined by the gametic type, for example, because there are loci of identical effect. If $M = \infty$, then $1 + \pi(g_v - g_r) - [1 - \pi(g_v - g_r)]^M = 1 + \pi(g_v - g_r)$ decreases to 1 as $a \to \infty$ if $g_r \neq g_v$. Therefore, only certain monomorphic equilibria become stable if $a$ is sufficiently large.

Proof of (*c*): If $M \geq 2$ and $\lambda_r$ is developed into a Taylor series, we obtain

$$\lambda_r \approx 1 + \left[sf(\rho - 1) - s - \frac{1}{2}a\right]z^2 - 2sz(g_v - \theta), \tag{A15}$$

where $z = g_r - g_v$ and terms of order $s^2$, $a^2$, and $sa$ (and higher) are ignored. If $g_v = \theta$, then $\lambda_r < 1$ for all $r \neq v$ if and only if condition (19) holds. The approximation (A15) shows that the first monomorphic states that become stable as $a$ increases are those closest to the optimum $\theta$. If $g_v \neq \theta$, a stronger condition may be necessary. In the extreme case that the genotypic value that is closest to $\theta$ is at the boundary of the phenotypic range ($g_v = \Gamma$; thus all but one locus have negligible effect), the condition for stability becomes (20) because $z = -2\Gamma$ is the largest possible deviation. Numerical computations suggest that (20) is always sufficient for the stability of at least one monomorphic state (even if $s$, $c$, and $a$ are not small). The parameter $M$ enters the eigenvalues only through terms of order $a^2$ and higher.

If $M = 1$, then we obtain

$$\lambda_r \approx 1 + [sf(\rho - 1) - s - a]z^2 - 2sz(g_v - \theta) \tag{A16}$$

instead of (A15). Therefore, (19) and (20) have to be replaced by

$$s + a > c(\rho - 1) \tag{A17}$$

and

$$a > c(\rho - 1), \tag{A18}$$

respectively. Thus, monomorphisms become stable for weaker assortment. For analogous single-locus results with multiple alleles, see Schneider (2005) and Schneider and Bürger (2005).

*Derivation of Eigenvalues for Monomorphic Equilibria*

Let us first assume an arbitrary, but constant, population size. Let

$$p_r^* = \frac{\tilde{p}_r}{\bar{p}} \tag{A19}$$

denote the frequency of gamete $r$ after assortative mating and recombination, where

$$\tilde{p}_r = \sum_{u,v} p_u p_v Q_{uv} R(uv \to r) \text{ and } \bar{p} = \sum_r \tilde{p}_r. \tag{A20}$$

After viability selection the frequency of gamete $r$ is

$$p_r' = \frac{p_r^* W_r^*}{\overline{W}}, \tag{A21}$$

where $W_r^*$ depends on the gamete frequencies $p_u^*$ and $\overline{W} = \sum_r p_r^* W_r^*$.

We derive the Jacobian at the monomorphic equilibrium $\hat{p}$ with $p_v = 1$. To this aim, we eliminate the redundant variable $p_v$ by setting $p_v = 1 - \sum_{r \neq v} p_r$. We need the following partial derivatives:

$$\begin{aligned}\frac{\partial p_r'}{\partial p_u}\bigg|_{p_v=1} &= p_r^* \frac{(\partial W_r^*/\partial p_u)\overline{W} - W_r^*(\partial \overline{W}/\partial p_u)}{\overline{W}^2}\bigg|_{p_v=1} + \frac{W_r^*}{\overline{W}}\frac{\partial p_r^*}{\partial p_u}\bigg|_{p_v=1} \\ &= \frac{W_r^*}{W_v^*}\bigg|_{p_v=1}\frac{\partial p_r^*}{\partial p_u}\bigg|_{p_v=1} = \frac{W_r}{W_v}\bigg|_{p_v=1}\frac{\partial p_r^*}{\partial p_u}\bigg|_{p_v=1}.\end{aligned} \tag{A22}$$

Hence, we also need the following derivatives:

$$\left.\frac{\partial p_r^*}{\partial p_u}\right|_{p_v=1} = \left.\frac{(\partial\tilde{p}_r/\partial p_u)\bar{p} - \tilde{p}_r(\partial\bar{p}/\partial p_u)}{\bar{p}^2}\right|_{p_v=1} = \left.\frac{\partial\tilde{p_r}}{\partial p_u}\right|_{p_v=1}, \tag{A23}$$

where

$$\begin{aligned}
\left.\frac{\partial\tilde{p}_r}{\partial p_u}\right|_{p_v=1} &= \sum_{i,j}\left[\left(\frac{\partial p_i}{\partial p_u}p_j + p_i\frac{\partial p_j}{\partial p_u}\right)Q_{ij} + p_i p_j\frac{\partial Q_{ij}}{\partial p_u}\right]\Bigg|_{p_v=1} R(ij\to r)\\
&= \sum_{\substack{i,j\\ i,j\neq v}}\left[(\delta_{iu}\delta_{jv} + \delta_{iv}\delta_{ju})Q_{ij} + \delta_{iv}\delta_{jv}\frac{\partial Q_{ij}}{\partial p_u}\right]\Bigg|_{p_v=1} R(ij\to r)\\
&\quad + \sum_{\substack{i\\ i\neq v}}\left[(\delta_{iu}-\delta_{iv})Q_{iv} + \delta_{iv}\frac{\partial Q_{iv}}{\partial p_u}\right]\Bigg|_{p_v=1} R(iv\to r)\\
&\quad + \sum_{\substack{j\\ j\neq v}}\left[(\delta_{ju}-\delta_{jv})Q_{vj} + \delta_{jv}\frac{\partial Q_{vj}}{\partial p_u}\right]\Bigg|_{p_v=1} R(vj\to r)\\
&\quad - \left(2Q_{vv} - \left.\frac{\partial Q_{vv}}{\partial p_u}\right|_{p_v=1}\right) R(vv\to r)\\
&= R(uv\to r)(Q_{uv}+Q_{vu})|_{p_v=1}.
\end{aligned} \tag{A24}$$

Here, $\delta_{iu} = 1$ if $i = u$ and $\delta_{iu} = 0$ otherwise. Therefore, we obtain

$$\left.\frac{\partial p_r'}{\partial p_u}\right|_{p_v=1} = R(uv\to r)\left.\frac{W_r}{W_v}\right|_{p_v=1}(Q_{uv}+Q_{vu})|_{p_v=1}. \tag{A25}$$

Next we show that the eigenvalues of the Jacobian are its diagonal elements; that is,

$$\begin{aligned}
\lambda_r = \left.\frac{\partial p_r'}{\partial p_r}\right|_{p_v=1} &= \frac{1-s(g_r-\theta)^2 + s\hat{\eta}(g_r-g_v)^2}{1-s(g_v-\theta)^2}\\
&\quad\times\{1+\pi(g_v-g_r) - [1-\pi(g_v-g_r)]^M\}R(rv\to r).
\end{aligned} \tag{A26}$$

This is best seen as follows. We order the multi-indices $r = (r_1, \ldots, r_n)$ and $u = (u_1, \ldots, u_n)$ such that $r \prec u$ if and only if $r_k < u_k$ for the largest integer $k$ with $r_k \neq u_k$. We have $r = u$ if and only if $r_k = u_k$ for all $k$. Furthermore, without loss of generality we label the alleles at each locus so that $v = (1, \ldots, 1)$. We arrange the elements of the Jacobian according to this ordering. This ordering implies that for a given row $r$ of the Jacobian, $r \prec u$ or $r = u$ are necessary for $R(uv\to r) \neq 0$. Thus, the Jacobian can be rearranged as an upper triangular matrix, implying that its eigenvalues are its diagonal elements. Clearly, we have $\lambda_r > 0$ for every $r$.

If the population size is not assumed to be constant, we additionally need the following derivatives:

$$\left.\frac{\partial p_r'}{\partial N}\right|_{(\hat{N},\hat{p})} = p_r^*\left.\frac{(\partial W_r^*/\partial N)\overline{W} - W_r^*(\partial\overline{W}/\partial N)}{\overline{W}^2}\right|_{(\hat{N},\hat{p})} = 0 \text{ for } r\neq v, \tag{A27}$$

and

$$\left.\frac{\partial N'}{\partial N}\right|_{(\hat{N},\hat{p})} = \left.\left[\overline{W} + N\frac{\partial \overline{W}}{\partial N}\right]\right|_{(\hat{N},\hat{p})}. \tag{A28}$$

Equation (A27) implies that $(\partial N'/\partial N)\big|_{(\hat{N},\hat{p})}$ is an eigenvalue. Because

$$\begin{aligned}\left.\frac{\partial \overline{W}}{\partial N}\right|_{(\hat{N},\hat{p})} &= \left.\left[\sum_r p_r^* \frac{\partial W_r^*}{\partial N}\right]\right|_{(\hat{N},\hat{p})} = \left.\frac{\partial W_v^*}{\partial N}\right|_{(\hat{N},\hat{p})} \\ &= (F'(N)\{1 - s(g_v - \theta)^2 + s\eta(N)[(g_v - \bar{g}^*)^2 + V_A^*]\} \\ &\quad + F(N)s\eta'(N)(g_v - \bar{g}^*)^2 + V_A^*)|_{(\hat{N},\hat{p})} \\ &= \frac{F'(\hat{N})}{F(\hat{N})},\end{aligned}$$

we have

$$\left.\frac{\partial N'}{\partial N}\right|_{(\hat{N},\hat{p})} = 1 + \frac{\hat{N}F'(\hat{N})}{F(\hat{N})}. \tag{A29}$$

For discrete logistic growth we obtain

$$\left.\frac{\partial N'}{\partial N}\right|_{(\hat{N},\hat{p})} = 1 - \frac{\hat{N}}{\rho\kappa - \hat{N}}. \tag{A30}$$

Therefore, we have $-1 < (\partial N'/\partial N)|_{(\hat{N},\hat{p})} < 1$ if and only if $2\rho\kappa > 3\hat{N}$. Hence, the demographic equilibrium (eq. [A13]) exists and is locally stable if and only if (A14) holds.

## The Measure $D$ of Linkage Disequilibrium

First, we note that for two diallelic loci, the following relation between $D$ and the classical measure of linkage disequilibrium (here denoted by $D_2 = p_1p_4 - p_2p_3$) holds:

$$D = 4D_2\frac{V_{\max}}{V_{\mathrm{LE}}}.$$

For a fully polymorphic symmetric equilibrium, $D = 4D_2$ is obtained.

Figure A1 illustrates the relation between various distributions of genotypic values and $D$. In figure A1*a*, there are four loci with allelic effects such that the 16 phenotypic values are equally spaced. In figure A1*b*, there are eight loci with equal effects, hence only nine different phenotypes. It should be noted that in the first case, a uniform distribution is in linkage equilibrium, whereas in the second case, a binomial has this property. Clearly, a high value of $D$ indicates that the population consists of two clusters around the extreme phenotypes.

## Population Size at Equilibrium

It is of interest to consider the population size at equilibrium because it provides a measure of how efficiently the resources are utilized and, in some sense, for the degree of adaptation to the environment. Figure A2 displays the average equilibrium population size as a function of the strength of assortment for four different scenarios: strong frequency dependence ($f = 2.5$) with free recombination, tight linkage, and very tight linkage, as well as moderately strong frequency dependence ($f = 1.56$) and free recombination. For a monomorphic population at

the phenotypic optimum, the equilibrium population size would be at the carrying capacity of $K = 10{,}000$. As the figure shows, the population size is always higher except for free recombination and moderately strong assortment, when only monomorphic equilibria are stable. Then $\overline{N}$ is between 9,990 and 10,000. Tables 1 and 2 and, especially, comparison of figures 3 and A2 show that the equilibrium population size is closely related to the equilibrium variance. In addition, if $\overline{D}$ is close to 1, so that the population clusters around the two extreme phenotypes, $\overline{N}$ is close to $2K = 20{,}000$.

## Literature Cited Only in the Appendix

Schneider, K. A. 2006. A multilocus-multiallele analysis of frequency-dependent selection induced by intraspecific competition. Journal of Mathematical Biology (forthcoming).

Zhivotovsky, L. A., and S. Gavrilets. 1992. Quantitative variability and multilocus polymorphism under epistatic selection. Theoretical Population Biology 42:254–283.

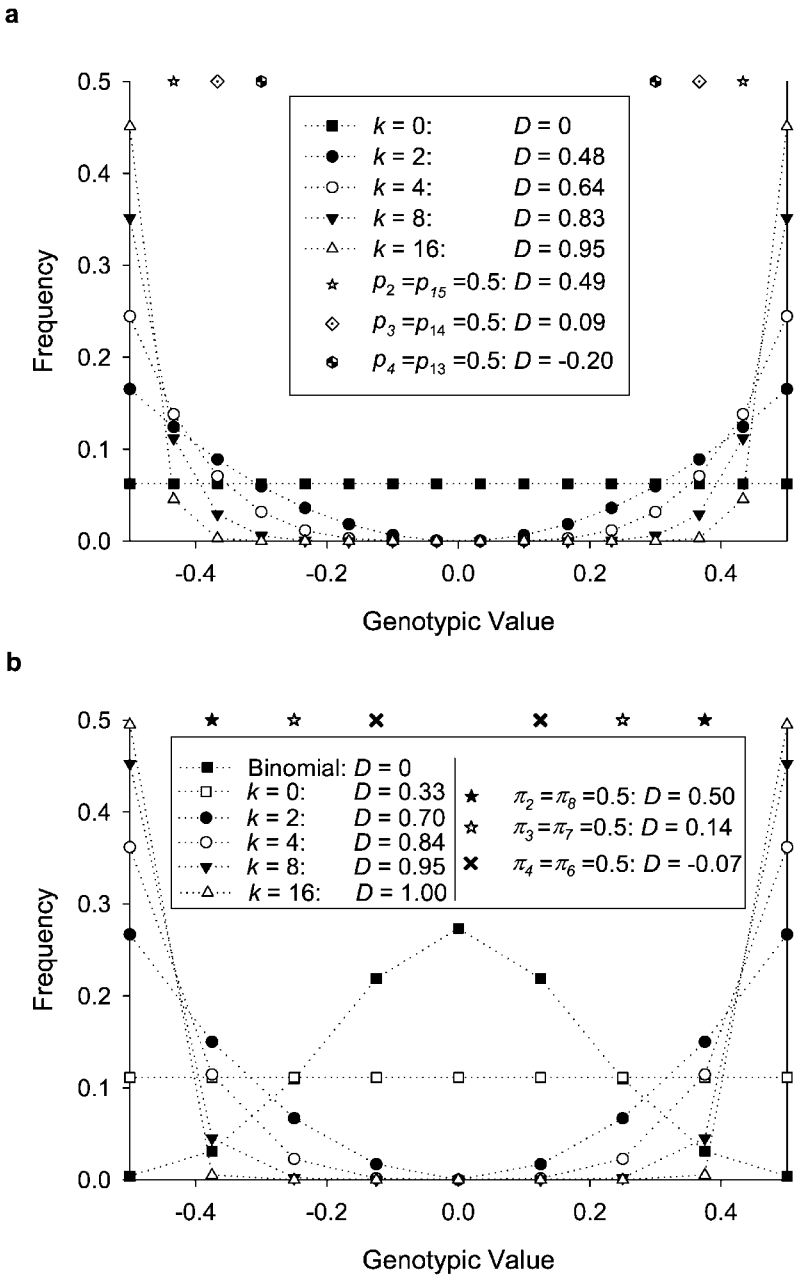


**Figure A1:** Some phenotypic distributions and the associated overall linkage disequilibrium, $D$ (eq. [21]). The distributions that are given by dotted lines and symbols are proportional to $x^k$ with $k$ as indicated. In addition, there are three distributions for which only two types are present at equilibrium (and at equal frequency). *a*, There are four loci with allelic effects $\pm 1/30$, $\pm 2/30$, $\pm 4/30$, and $\pm 8/30$. *b*, There are eight loci with equal effects.

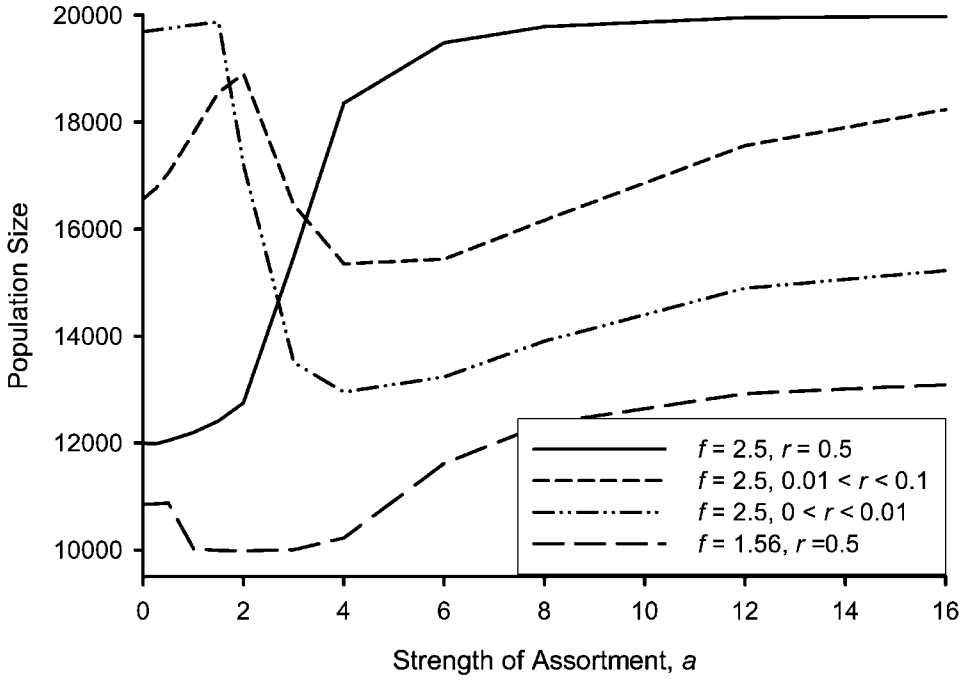


**Figure A2:** Equilibrium population size as a function of the strength of assortment for different strengths of frequency dependence and different recombination scenarios, as indicated. We have $\kappa = 10{,}000$, $\rho = 2$, $n = 4$, $s = 0.4$, $\theta = 0$, and $M = \infty$.

**Table A1**

Equilibrium structure in a typical six-locus system for strong competition ($\hat{\eta} > 1$) and an asymmetric optimum ($\theta = 0.25$)

| | $\hat{P}_1$ | $\hat{P}_2$ | $\hat{P}_3$ | $\hat{P}_4$ | $\hat{P}_5$ | $\hat{P}_6$ | $\hat{\Delta}$ | $\hat{V}_A$ | $\hat{V}_{LE}$ |
|---|---|---|---|---|---|---|---|---|---|
| Effects $\gamma_i$ | .019 | .124 | .015 | .056 | .127 | .159 | | | |
| Equation (A7): | | | | | | | | | |
| $N = K$ | 1.000 | .691 | 1.000 | .926 | .685 | .648 | −.027 | | .051 |
| $N = \hat{N}$ | 1.000 | .685 | 1.000 | .911 | .679 | .643 | −.033 | | .052 |
| $r = .5$ | 1.000 | .684 | 1.000 | .903 | .679 | .644 | −.034 | .054 | .052 |
| $r = .1$ | 1.000 | .696 | 1.000 | .841 | .665 | .646 | −.041 | .060 | .052 |
| $.005 \le r \le .05$ | .765 | .676 | .670 | .662 | .654 | .660 | −.083 | .097 | .054 |
| $r = .01$ | .622 | .608 | .600 | .605 | .606 | .607 | −.143 | .194 | .058 |
| $r = .001$ | .582 | .582 | .582 | .582 | .582 | .582 | −.168 | .239 | .059 |
| $r = 0$ | .580 | .580 | .580 | .580 | .580 | .580 | −.170 | .244 | .059 |

**Note:** $r = 0.5$ means that recombination rates between adjacent loci are 0.5 (no interference), and similarly for $r = 0.1$, etc. For $0.005 \le r \le 0.05$, recombination rates between adjacent loci were chosen randomly from a uniform distribution and are $r = 0.0122$, 0.0089, 0.0372, 0.0292, and 0.0415 in this example. The following parameters are fixed in all cases: $\rho = 2$, $s = 0.4$, $c = 0.625$, hence $f = 1.5625$. With the (true) numerically determined $\hat{N}$ (= 10,493), we have $\hat{\eta} = 1.7246$.