EVOLUTION

INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

PUBLISHED BY

THE SOCIETY FOR THE STUDY OF EVOLUTION

Vol. 60 **November 2006** **No. 11**

Evolution, 60(11), 2006, pp. 2185–2206

THE CONDITIONS FOR SPECIATION THROUGH INTRASPECIFIC COMPETITION

REINHARD BÜRGER,[1,2,3] KRISTAN A. SCHNEIDER,[1,4] AND MARTIN WILLENSDORFER[3,5,6]

[1]*Department of Mathematics, University of Vienna, Nordbergstrasse 15, 1090 Vienna, Austria*

[2]*E-mail: reinhard.buerger@univie.ac.at*

[3]*Program for Evolutionary Dynamics, Department of Mathematics and Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138*

[4]*E-mail: kristan.schneider@univie.ac.at*

[5]*Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138*

[6]*E-mail: willensd@fas.harvard.edu*

Abstract.—It has been shown theoretically that sympatric speciation can occur if intraspecific competition is strong enough to induce disruptive selection. However, the plausibility of the involved processes is under debate, and many questions on the conditions for speciation remain unresolved. For instance, is strong disruptive selection sufficient for speciation? Which roles do genetic architecture and initial composition of the population play? How strong must assortative mating be before a population can split in two? These are some of the issues we address here. We investigate a diploid multilocus model of a quantitative trait that is under frequency-dependent selection caused by a balance of intraspecific competition and frequency-independent stabilizing selection. This trait also acts as mating character for assortment. It has been established previously that speciation can occur only if competition is strong enough to induce disruptive selection. We find that speciation becomes more difficult for very strong competition, because then extremely strong assortment is required. Thus, speciation is most likely for intermediate strengths of competition, where it requires strong, but not extremely strong, assortment. For this range of parameters, however, it is not obvious how assortment can evolve from low to high levels, because with moderately strong assortment less genetic variation is maintained than under weak or strong assortment—sometimes none at all. In addition to the strength of frequency-dependent competition and assortative mating, the roles of the number of loci, the distribution of allelic effects, the initial conditions, costs to being choosy, the strength of stabilizing selection, and the particular choice of the fitness function are explored. A multitude of possible evolutionary outcomes is observed, including loss of all genetic variation, splitting in two to five species, as well as very short and extremely long stable limit cycles. On the methodological side, we propose quantitative measures for deciding whether a given distribution reflects two (or more) reproductively isolated clusters.

Key words.—Assortative mating, frequency-dependent selection, genetic variation, intraspecific competition, linkage disequilibrium, multilocus genetics, sympatric speciation.

Received June 1, 2006. Accepted July 27, 2006.

Sympatric speciation has been and continues to be a controversial topic. The disagreement concerns both empirical evidence and theoretical plausibility. The diverse views on the empirical issues were expressed in several recent publications. In accordance with the classical view advanced by Mayr (1963), Coyne and Orr (2004) took allopatric speciation as the null hypothesis in their systematic analysis of the evidence on speciation and considered only three cases to be promising for sympatric speciation. By contrast, Via (2001) and Dieckmann et al. (2004) argued strongly for the importance and frequent occurrence of sympatric speciation. We will be concerned here only with aspects of theoretical plausibility.

Maynard Smith (1966) was apparently the first to show that sympatric speciation is theoretically possible. Later work substantiated this view but also showed that rather specific, though highly diverse, conditions are required (for reviews, see Turelli et al. 2001; Gavrilets 2003, 2004). The most complex and appealing model proposed so far is that of Dieckmann and Doebeli (1999). It involves a polygenic trait subject to intraspecific competition and an evolving, polygenic mating character. They demonstrated numerically, guided by analytical reasoning, that if competition is sufficiently strong to induce disruptive selection, assortative mating can evolve to the extent that reproductively isolated clusters emerge, which may be interpreted as speciation. They concluded that their theoretical evidence generally suggests a prominent role for ecologically driven speciation in sympatry. They further strengthened and extended their plea for ''adaptive speciation'' in subsequent publications (e.g., Doebeli and Dieck-

mann 2000; Dieckmann et al. 2004). The Dieckmann-Doebeli model, as well as the so-called adaptive-dynamics approach in general, has received considerable attention, inspired researchers to evaluate various aspects of ecologically driven speciation (e.g., Drossel and McKane 2000; Geritz and Kisdi 2000; Matessi et al. 2001; Bolnick 2004b; Gourbiere 2004; Kirkpatrick and Nuismer 2004; Polechová and Barton 2005; Bürger and Schneider 2006), and caused vivid, sometimes fierce debate (e.g., Waxman and Gavrilets 2005; the many commentaries following in the same issue of the *Journal of Evolutionary Biology*, Doebeli et al. 2005; Gavrilets 2005).

An important aspect of the disagreement between the adherents of the adaptive-dynamics approach and workers rooted more firmly in the population-genetics tradition revolves around the practice of adaptive dynamics to predict the evolution of diversity and speciation by identifying so-called branching points from purely ecological considerations. Their identification clearly is essential in determining whether disruptive selection can be expected. But to what extent are ecological considerations sufficient to predict evolutionary diversification? In particular, what is the role of the genetic architecture and the initial composition of the population? How strong must assortment be to induce fission of a population? These are some of the issues we address here in an attempt to shed more light on the genetic and ecological conditions under which speciation driven by intraspecific competition can and does occur. Despite a plethora of theoretical literature on this topic, these conditions are still not well understood.

We confine our attention to models of disruptive selection caused by intraspecific competition as the driving force. The ecological part of the Dieckmann-Doebeli model coincides with the model of intraspecific competition for a unimodally distributed resource devised by Roughgarden (1972) and further analyzed by Slatkin (1979). However, as discussed below, this model has some questionable features. We instead employ a version of a model introduced by Bulmer (1974, 1980) that is very similar to the Roughgarden model if selection is weak (e.g., Bürger and Gimelfarb 2004; Bürger 2005), but behaves differently for strong selection. Whereas the full Dieckmann-Doebeli model assumes a genetically variable mating character, we consider the strength of assortative mating as a fixed parameter and explore how strong it must be so that reproductively isolated clusters can evolve and, if so, at which frequency compared with other evolutionary outcomes. A further simplification is our assumption that assortative mating is based on the trait under natural selection. Gavrilets (2004) called this a ''magic'' trait. This makes speciation easier than if the mating trait was genetically unrelated to the ecological trait (Dieckmann and Doebeli 1999). Examples of such magic traits include body size in sticklebacks (McKinnon et al. 2004) and in sea horses (Jones et al. 2003). Our trait is determined by a finite number of loci that, in contrast to most previous studies, may have arbitrary but additive effects. For assortative mating, we adopt a model of Matessi et al. (2001) that includes a parameter for costs to being choosy.

This study extends and complements Bürger and Schneider (2006) in several ways. Here, we consider a diploid population instead of a haploid; we include strong natural selection, that is, strong competition and strong stabilizing selection; we pay attention to edge effects that may arise if the phenotypic range is limited, as it is the case if the trait is determined by a given, finite set of genes (Polechová and Barton 2005); and we study the roles of initial conditions and the distribution of locus effects in more detail. A major part of this study is numerical but quite systematic and comprehensive in determining the parameter regions where population splitting occurs. In particular, for every combination of ecological and mating parameters, evolution of a large number of genetic architectures (distribution of allelic effects) is simulated, each from 10 different initial conditions. From a theoretical point of view, it is not always simple to decide when a distribution of phenotypes can be interpreted as speciation. We propose a single measure of linkage disequilibrium for this purpose that does a surprisingly good job, unless competition and assortment are so strong that more than two species can coexist. We complement numerical work by analytical work on the stability of monomorphic equilibria. Convergence to a monomorphic equilibrium is, of course, quite the opposite of speciation.

The Model

We consider a sexually reproducing population of diploid organisms with discrete generations in which both sexes have the same genotype distribution among zygotes. Its size, N, is density regulated but sufficiently large so that random genetic drift can be ignored. Natural selection acts through differential viabilities on an additive polygenic trait such that individual fitness is determined by two components: frequency-independent stabilizing selection on this trait and frequency- and density-dependent competition among individuals of similar phenotype. Assortative mating may induce sexual selection.

Ecological Assumptions

The first fitness component is frequency independent and reflects some sort of stabilizing selection on the trait, for example, by differential supply of a resource whose utilization efficiency is phenotype dependent. As most previous studies, we ignore environmental variation and deal directly with the fitnesses of genotypic values, g. Therefore, we use the terms genotypic value and phenotype synonymously.

We model stabilizing selection by the Gaussian function

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

where s measures its strength and θ is the position of the optimum. We model competition between phenotypes g and h by

$$\alpha(g, h) = \exp[-c(g - h)^2]. \tag{2}$$

This implies that competition between individuals of similar phenotype is stronger than between individuals of very different phenotype, as it will be the case if different phenotypes preferentially utilize different food resources. Large c implies a strong frequency-dependent effect of competition, whereas frequency dependence vanishes in the limit $c \to 0$. 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 in a population with distribution $\mathcal{P}$, is given by

$$\bar{\alpha}(g) = \sum_h \alpha(g, h)\mathcal{P}(h). \quad (3)$$

There is no simple expression for $\bar{\alpha}(g)$. To leading order in c, it can be approximated by

$$\bar{\alpha}(g) = 1 - c[(g - \bar{g})^2 + V_A], \quad (4)$$

where $\bar{g}$ and V_A denote the mean and additive genetic variance, respectively, of the distribution $\mathcal{P}$ of genotypic values.

We include density-dependent population growth which, in the absence of genetic variation, follows the logistic equation

$$N' = \begin{cases} N(\rho - N/\kappa) & 0 \le N < \rho\kappa \\ 0 & N \ge \rho\kappa. \end{cases} \quad (5)$$

The carrying capacity 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 $2 < \rho < 3$. Other forms of population regulation may be used as well (Appendix 2; Bürger 2005).

Following Bulmer (1974), we assume that the absolute fitness of an individual with genotypic value (phenotype) g is

$$W(g) = [\rho - \bar{\alpha}(g)N/\kappa)]S(g), \quad (6)$$

where the dependence of $W(g)$ on N and $\mathcal{P}$ is omitted. Closely related ecological models, with quadratic instead of Gaussian functions in (1) and (2), were studied under different assumptions and with an other focus by Bürger and Gimelfarb (2004) and Schneider and Bürger (2006). The Gaussian choice has the advantage that weak and strong selection can be modeled, but is prohibitive to a general mathematical analysis. The fitness function (6) generates disruptive selection if c is large enough and the variance of the phenotypic distribution is not too large (see Theoretical Background and Analytical Results).

If stabilizing selection and competition are both sufficiently weak, so that terms of order s^2, c^2, cs, and smaller can be ignored, then all known functional forms of fitness that have been used in modeling intraspecific competition for a unimodally distributed resource converge on the same simple quadratic fitness function,

$$W_{app}(g) = \left(\rho - \frac{N}{\kappa}\right)\left\{1 - s(g - \theta)^2 + c\left(\frac{\rho\kappa}{N} - 1\right)^{-1}[(g - \bar{g})^2 + V_A]\right\}. \quad (7)$$

Here, frequency dependence enters fitness only through the mean and the variance of the distribution, and fitness is either convex or concave but never multimodal. At least under random mating, the equilibrium structure resulting from selection according to equation (7) is very similar to that resulting from equation (6) if s and c are not too large (e.g., $s \le 0.4$ and $[\rho - 1]c/s \le 2$; Bürger 2005; see also Kopp and Hermisson 2006; Bürger and Schneider 2006).

Assortative Mating

We assume that individuals mate assortatively with respect to the trait 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 with that of their prospective mating partner. We choose the preference function

$$\pi(g - h) = \exp[-a(g - h)^2]. \quad (8)$$

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

Females mate only once, whereas males may participate in multiple matings. If an encounter was not successful, in which case a female remains unmated, she may try again unless the total number of encounters has reached a number M. This reflects the idea that choosiness has costs, for instance, because the mating period is limited. If $M = \infty$, there are no costs to assortative mating because every female is sure to find a mating partner; if $M = 1$, these costs are very high. The probability that an encounter of a female of type g with a random male results in mating is $\bar{\pi}(g) = \Sigma_h \pi(g - h)\mathcal{P}(h)$, and the probability that she eventually mates with a male of type h is given by $Q(g, h)\mathcal{P}(h)$, where

$$Q(g, h) = \sum_{m=0}^{M-1} [1 - \bar{\pi}(g)]^m \pi(g - h). \quad (9)$$

In general Q is not symmetric in g and h, and the first argument refers to the female.

If $M = 1$, then $Q(g, h) = \pi(g - h)$. This leads to strong sexual selection in both sexes. It admits a number of different interpretations, for instance, that both sexes are choosy, and has been used in a variety of studies. If the encounter rate is very high, M may be chosen to be infinity, and we obtain $Q(g, h) = \pi(g - h)/\bar{\pi}(g)$. Then, $\Sigma_h Q(g, h)\mathcal{P}(h) = 1$ for all g, and assortative mating does not induce sexual selection among females. It does, however, induce sexual selection among males. For a more detailed discussion of this model and its relation to other work, see Bürger and Schneider (2006).

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 $\frac{1}{2}\gamma_i$ and $-\frac{1}{2}\gamma_i$ ($\gamma_i > 0$), 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 homozygous effects (the effect of a substitution) is $2\gamma_i$ because constants that determine the mean phenotype can be absorbed by θ. We assume that θ is within the range of possible genotypic values, that is, $-\Gamma < \theta < \Gamma$, where $\Gamma = \Sigma_{i=1}^{n} \gamma_i$ and the scale is chosen such that $\Gamma = \frac{1}{2}$. This excludes frequency-independent directional selection.

The multilocus dynamics has to be described in terms of diploid genotype frequencies because zygotes (offspring) are generally not in Hardy-Weinberg proportions because of non-

random mating. Genotypes are unordered. Let t represent an offspring genotype and u and v parental genotypes. The corresponding phenotypes are denoted by g_t, g_u, and g_v. The frequency of genotype t (among zygotes) in consecutive generations is denoted by p_t and p_t'. The frequency of t after natural selection is $p_t^* = p_t W_t/\bar{W}$, where $Wt = W(g_t)$ and $\bar{W} = \Sigma_t W_t p_t$ is the mean viability. After selection, mating and recombination occur. Let $R(uv \to t)$ designate the probability that parents with genotypes u and v produce a zygote with genotype t. $R(uv \to t)$ is determined by the pattern of recombination between loci.

The genetic dynamics is given by a system of $2^{2n-1} + 2^{n-1}$ recursion relations that can be written as

$$p_t' = \frac{\bar{W}^2}{\tilde{W}} \sum_{u,v} p_u^* p_v^* Q_{uv}^* R(uv \to t), \tag{10}$$

where $Q_{uv}^* = Q^*(g_u, g_v)$ and the asterisk indicates that Q is calculated from the genotypic frequencies after selection and $\tilde{W} = \bar{W}^2 \Sigma_{t,u,v} p_u^* p_v^* Q_{uv}^* R(uv \to t)$. The demographic dynamics follows the standard recursion

$$N' = N(\tilde{W}/\bar{W}). \tag{11}$$

Thus, for a genetically 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 (10) and (11). We set $N' = 0$ (population extinction) if $\tilde{W}/\bar{W} \le 0$.

Theoretical Background and Analytical Results

Equilibrium Structure and the Maintenance of Genetic Variability

The complexity of this model prohibits a detailed and comprehensive analytic treatment. However, for special or limiting cases some useful results can be derived. Also the stability of monomorphic equilibria can be treated in full generality. Therefore, conditions for the maintenance of genetic variation are available: The results in this section are valid for a more general model than introduced above: the number of alleles per locus and their effects may be arbitrary, and population regulation can be more general (see Appendix 2).

Random mating and weak selection.—For a randomly mating population ($a = 0$) and if the population size is assumed to be at demographic equilibrium, the equilibrium and stability structure can be determined completely provided selection is sufficiently weak, so that the fitness function (6) can be approximated by the quadratic function (7), and linkage disequilibrium can be ignored (Bürger 2005; Schneider 2006a). Because these results are important in guiding our intuition, we briefly summarize them for the case of logistic population growth.

Result 1: If, approximately, $c(\rho - 1) < s$, then at most one locus can be polymorphic at a stable equilibrium and, typically, multiple stable equilibria coexist. At a polymorphic locus two alleles with neighboring effects are segregating (i.e., there is no allele with effect in between).

Result 2: If, approximately, $c(\rho - 1) > s$, then there exists a unique asymptotically stable equilibrium that is globally stable. At least one locus is polymorphic at this equilibrium. The polymorphic loci can be determined (those with large effects are polymorphic) and the allele frequencies can be calculated. At a polymorphic locus the two alleles with the largest and smallest effect are segregating. If the optimum is symmetric ($\theta = 0$), then all loci are polymorphic.

As discussed below, $c(\rho - 1) > s$ implies that a population with small variance and mean at the optimum θ is under disruptive selection. Otherwise, it experiences stabilizing selection. Thus, roughly, results 1 and 2 show that high genetic variability is maintained in a randomly mating population if the combined strength of frequency- and density-dependent selection is greater than that of stabilizing selection so that overall disruptive selection is induced. Otherwise, little or no variation is maintained. Comparison with exact results from numerical iteration of the recursion relations for the gamete frequencies, so that linkage disequilibrium and population regulation are admitted, shows that the linkage-equilibrium approximation is very accurate if linkage is not too tight (Bürger 2005; Schneider 2006).

Weak assortative mating and weak selection.—If assortative mating and selection are both weak, such that only terms of order a, s, and c need to be retained, but all others, including interaction terms such as as, can be ignored, simple and intuitive conditions for the stability of monomorphic equilibria can be derived. It can be shown (Appendix 2) that monomorphic equilibria sufficiently close to θ are stable if (approximately)

$$s + \frac{a}{2} > c(\rho - 1) \quad \text{if } M \ge 2 \qquad \text{or} \tag{12}$$

$$s + a > c(\rho - 1) \quad \text{if } M = 1. \tag{13}$$

Otherwise, all monomorphic equilibria are unstable. Further, stable monomorphic equilibria exist in any genetic system (i.e., even if no genotype is close to θ) if

$$\begin{aligned} &\frac{a}{2} > c(\rho - 1) \text{ and } M \ge 2 \qquad \text{or if} \\ &a > c(\rho - 1) \text{ and } M = 1. \end{aligned} \tag{14}$$

The conditions for $M \ge 2$ are more accurate if M is large.

Moderately strong assortment promotes stability of monomorphic equilibria because, compared with random mating, it counteracts competition by inducing local stabilizing selection around monomorphic states. If a monomorphic genotype is frequent, it becomes more difficult for deviating types to find a mating partner. This effect is more pronounced for strongly selective mating (M small).

For $M = \infty$, condition (12) is equivalent to equation (3.3) of Kirkpatrick and Nuismer (2004) with $k = \frac{1}{2}$ (their animal model, which corresponds to our $M = \infty$); condition (13) is equivalent to their (3.3) with $k = 1$ (their plant model). Their ecological model coincides with ours if we ignore population regulation and assume $N = K = \kappa(\rho - 1)$. Their a, c_1, c_2, and m correspond to our s, $1 - (1/\rho)$, c, and a, respectively. Kirkpatrick and Nuismer's (3.3) quantifies when natural and sexual selection together (their lifetime fitness) become disruptive near the optimum of stabilizing selection. As our (12–14), it assumes weak natural selection and weak assortment. Conditions (12–14) are also valid if individuals are haploid (Bürger and Schneider 2006) or if there is a single locus with multiple alleles (Schneider and Bürger 2006).

Strong selection.—With strong selection, we have to resort to the fitness function (6), and in general we cannot ignore linkage disequilibrium. Fitness now depends on s and c separately and not only on c/s. The above results are no longer valid if stabilizing selection becomes very strong, not even for random mating. On the one hand, if Gaussian stabilizing selection is strong, then already in the absence of competition (i.e., if $c = 0$), stable fully polymorphic equilibria can be maintained (Willensdorfer and Bürger 2003). This does not require linkage, but is facilitated by it. On the other hand, for some genetic systems monomorphic equilibria can be stable for arbitrarily strong competition. Indeed, from the eigenvalues of the monomorphic equilibria, the following sufficient condition for the stability of the monomorphic equilibrium with genotype $\boldsymbol{uu}$ fixed can be derived

$$\frac{\ln \rho}{(g_{\boldsymbol{ut}} - \theta)^2} < s < \frac{\ln \rho}{(g_{\boldsymbol{uu}} - \theta)^2} \quad \text{for all } \boldsymbol{t} \neq \boldsymbol{u} \tag{15}$$

(see Appendix 2; here and in Appendix 2 we use $\boldsymbol{u}$, $\boldsymbol{t}$, etcetera to label gametes and $\boldsymbol{uu}$, $\boldsymbol{ut}$, etcetera to label genotypes). Notably, this condition is independent of c and implies that for strong enough stabilizing selection monomorphic equilibria whose phenotype is close to θ become stable. If $s \to \infty$, they lose their stability (except when $g_{\boldsymbol{uu}} = \theta$). Because in our model, we have $(g_{\boldsymbol{ut}} - \theta)^2 \leq 1$, condition (15) can be fulfilled only if $s > \ln \rho$. If $\theta = 0$, $s > 4 \ln \rho$ is required. For random mating, it can also be shown that if $2 \leq \rho < 3$, $\theta = 0$, and $s = 0.4$, then no monomorphic equilibrium can be stable if $c(\rho - 1) > s$.

The general case.—The following simple conclusions can be drawn. For proofs and more details, see Appendix 2. First, increasing strength of assortment (large a) and high costs to being choosy (small M) promote the stability of the monomorphic equilibria, whereas increasing frequency dependence (large c) tends to reduce their stability. Second, unless costs are absent, that is, $M < \infty$, all monomorphic equilibria at which a positive population size can be maintained become stable for sufficiently large a. Third, high recombination rates favor the stability of monomorphic equilibria. The reason is that recombination generates intermediate phenotypes, which promotes stability of monomorphic equilibria with an intermediate phenotype relative to equilibria representing strong divergence. Finally, the stability properties of the monomorphic equilibria depend on the genetic architecture, that is, on the distribution of allelic effects and on the recombination rates. For instance, if two genetic architectures produce the same genotypic values, for example, one with two diallelic loci and one with a single locus with four alleles, the stability properties of the monomorphic equilibria can be different for the same model parameters.

Shape of the Fitness Function

Because selection is frequency dependent, the shape of the fitness function, (6), depends strongly on the distribution of phenotypes. If there is a monomorphic distribution located at $g = \theta$, then $W(g)$ displays disruptive selection in a neighborhood of θ if and only if $c(\rho - 1) > s$. By equation (4) and continuity, this continues to be true for distributions with mean θ and sufficiently small variance. With increasing variance, the shape of the fitness function may change dramatically. This is illustrated in Figures 1a,c for dimorphic distributions with type frequencies of ½ at each of the positions $-\delta$ and δ, where $0 \leq \delta \leq \Gamma = ½$ (from here on we assume $\theta = 0$). If δ is sufficiently small, there is disruptive selection near zero: fitness is U-shaped if stabilizing selection is weak ($s = 0.4$, Fig. 1a) and bimodal if it is strong ($s = 2$, Fig. 1c). As δ increases beyond a critical value (0.45 in Fig. 1a and 0.20 in Fig. 1c), the fitness function develops a local maximum at $g = 0$ and becomes trimodal. It is very flat then, so that this feature is invisible in the figures. As δ increases further (beyond 0.49 in Fig. 1a and beyond 0.25 in Fig. 1c) the fitness function becomes unimodal with mode at $g = 0$, and there is pure stabilizing selection. If $s = 0.4$ and $\rho = 2$, as in Figure 1a, and if $0.4 \leq c \leq 1.46$, then the fitness function is U-shaped on the phenotypic range for every possible dimorphic distribution.

The fitness function (6) has the (natural) property that for sufficiently extreme types fitness decays to zero. (Since, as a matter of convenience, we chose a fixed phenotypic range, namely $[-½, ½]$, this property can be achieved by making s large.) By contrast, Roughgarden's functional form of fitness (Roughgarden 1972; Slatkin 1979),

$$W_R(g) = \rho - \frac{(\rho - 1)N}{S(g)}\bar{\alpha}(g), \tag{16}$$

which has been used in many studies of intraspecific competition, may behave differently if selection becomes strong (Fig. 1). For small δ, it can have either only one local maximum, then selection is stabilizing, or one local minimum, then selection is disruptive. The latter occurs if and only if $c > s$. (Note that s in eq. 16 corresponds to our $s/[\rho - 1]$, provided the population size is close to carrying capacity; Bürger 2005.) Similar to our fitness function, the Roughgarden fitness has the property that disruptive selection becomes weaker as the variance of the phenotypic distribution increases; however, compared with ours, a much higher variance is needed to achieve this (Figs. 1b,d). For a dimorphic distribution, as above, the critical value for δ can be calculated explicitly and is $\sqrt{(c - s)/(2cs)}$. If $\rho = 2$, $s = 0.4$ and $c = 2$ (Fig. 1b), this gives 1; if $\rho = 2$, $s = 2$, and $c = 10$ (Fig. 1d), it gives 0.45.

The basic features of the fitness assignments (6) and (16) discussed above and displayed in Figure 1 remain qualitatively the same if discrete logistic growth is replaced by the Beverton-Holt (or some similar) functional form of population regulation (results not shown). Of course, this is true only if ρ is small enough that no cycling can occur.

Methods

Numerical Approach

We use the approach of Bürger and Gimelfarb (2004) with the obvious modifications required by modeling assortative mating. Its basic idea is to evaluate the quantities of interest for many randomly chosen genetic parameter sets and initial conditions and then calculate various statistics. In this sense, we obtain statistical results, although each single result is obtained by numerical iteration of the deterministic system of recursion relations (10) and (11).

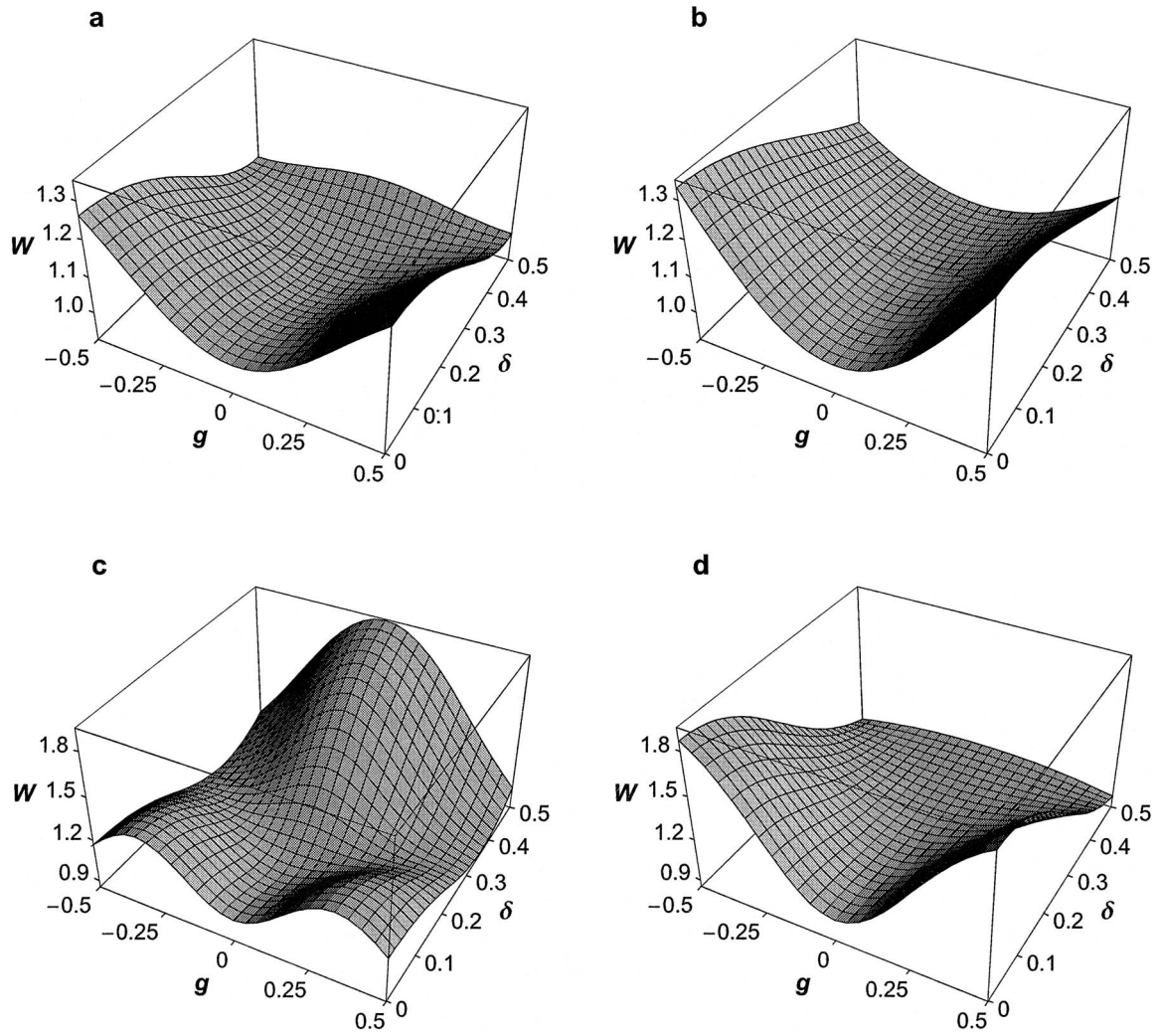

FIG. 1. The Bulmer and Roughgarden fitness landscapes. Panels (a) and (c) display Bulmer's fitness (eq. 6), as function of the phenotypic value g and the distance, δ, from the optimum $\theta = 0$ of each of the two phenotypes constituting a dimorphic distribution. Panels (b) and (d) are analogous for the Roughgarden fitness (eq. 16). Panels (a) and (b) use $s = 0.4$ and $c = 2$, (c) and (d) use $s = 2$ and $c = 10$. The other parameters are $\theta = 0$, $\rho = 2$, $\kappa = 10{,}000$. The values of δ for which fitness near $g = 0$ becomes stabilizing (at least locally) are 0.45, 2.00, 0.20, and 0.45 for panels (a–d), respectively. Clearly, $\delta = 2$ can never be realized. The fitness functions are normalized such that $\bar{W} = 1$, that is, demographic equilibrium is assumed.

For a given number n of loci we constructed more than $\psi(n)$, what we call ''genetic parameter sets'' (allelic effects of loci and recombination rates between adjacent loci from the given range). We have $\psi(n) = 1000$ if $n = 2$, 3, and $\psi(4) = 300$ because it turned out that $\psi(4) = 1000$ was much too time consuming. (The total CPU time on PCs with 3 GHZ was more than 2.5 years.) For each genetic parameter set, allelic effects were obtained by generating values β_i ($i = 1, 2, \ldots, n$) as independent random variables, uniformly distributed between 0.1 and 1.0 and transforming them into the actual allelic effects via $\gamma_i = \frac{1}{2}\beta_i/\Sigma_k \beta_k$. By restricting the maximum ratio of locus effects to 10, the speed of the computations was increased substantially because the time to equilibration may be extremely long if one or several loci have a very small effect. The consequences of admitting more variation in allelic effects are discussed below. Genotypic values were obtained by assuming additivity of allelic effects (no epistasis or dominance). Our scaling implies that the range of phenotypic values is always $[-\Gamma, \Gamma] = [-\frac{1}{2}, \frac{1}{2}]$, and the extreme values are assumed. Recombination between loci is free. We always have a symmetric optimum, $\theta = 0$, as well as $\rho = 2$ and $\kappa = 10{,}000$.

For a given ecological parameter combination ($\rho = 2$, $\kappa = 10{,}000$, $\theta = 0$, s, c), given mating parameters (a, M), and a given number of loci (n), the recursion relations (10) and (11) were numerically iterated starting from 10 different, randomly chosen initial genotype distributions in Hardy-Weinberg proportions for each of more than $\psi(n)$ genetic parameter sets. To avoid clustering of the initial distributions, the constituting gamete distributions were chosen such that the Euclidean distance between any two of them was no less than a predetermined value (0.25, 0.30, and 0.35 for two, three, and four loci, respectively). Each genetic system had different initial distributions. An iteration was stopped after generation t when either an equilibrium was reached (in the sense that the distance between genotype distributions in two consec-

utive generations was less than 10^{-8}) or t exceeded 10^6 generations. Equilibria were classified as different if their Euclidean distance exceeded 2×10^{-3}. If equilibrium was not reached, the parameter combination was excluded from the analysis. The proportion of excluded runs was exceedingly small (less than 1%, most 0%) and did not induce a bias. If convergence within the specified maximum number of generations did not occur, it was mainly because of extremely slow convergence. Complicated dynamic behavior was detected in a few runs under extremely strong competition and assortment (see below).

For each combination of ecological and mating parameters and given number of loci, all statistics are based on $\psi(n)$ genetic parameter sets, each with 10 initial conditions, that led to equilibration. For each parameter set we recorded the number of different equilibria that were reached, the genotype frequencies and the population size at equilibrium, and the number of trajectories converging to each equilibrium. Using this database, we also calculated the following quantities for each equilibrium: the proportion of trajectories converging to an equilibrium with a given number of polymorphic loci; the (additive) genetic variance V_A; the genic variance $V_{gen} = 2 \Sigma_i \gamma_i^2 P_i(1 - P_i)$ (the variance that would be observed under Hardy-Weinberg and linkage equilibrium); the relative variance $V_R = V_A/V_{max}$, where $V_{max} = \frac{1}{2} \Sigma_i \gamma_i^2$ is the maximum possible variance in the given genetic system under the assumption of Hardy-Weinberg and linkage equilibrium (this normalization of the additive genetic variance enables proper comparison of genetic systems with different locus effects or number of loci); two criteria for speciation (see below); and the following compound measure of disequilibrium (linkage and Hardy-Weinberg),

$$D = \frac{V_D}{V_{Dmax}}, \tag{17}$$

where $V_D = V_A - V_{gen}$, and

$$V_{Dmax} = \Gamma^2 - \frac{1}{2} \sum_{i=1}^{n} \gamma_i^2 \tag{18}$$

is the value of V_D if only the two extreme genotypes, with values $-\Gamma$ and Γ, are present, each at frequency ½. Disequilibrium D was calculated only if at least one locus was polymorphic.

Important properties of D are summarized in Appendix 1; see also Figure A1. As described below, we use D in one criterion for determining when speciation occurs.

Quantitative Criteria for Speciation

In our model, speciation can occur only by prezygotic isolation. For speciation we require that individuals of the two putative species mate with probability less than $p = 0.01$. For given strength a of assortative mating, equation (8) implies that the distance between individuals constituting the two prospective species must be at least

$$d_{crit} = \sqrt{\frac{-\ln p}{a}}. \tag{19}$$

With $p = 0.01$ and because $d_{crit} \leq 2\Gamma = 1$, speciation cannot occur if $a \leq 4.6$. Of course, a small mating probability between two types of individuals is not sufficient to guarantee speciation because hybrids could increase in frequency. Therefore, we need a more elaborate criterion. We define two homozygous genotypes, G_1 and G_2, to be the crystallization points of different species if they satisfy the following three conditions: (i) their phenotypic distance is larger than d_{crit}; (ii) their frequency is higher than 0.05 each; and (iii) intermediate genotypes are at sufficiently low frequency. Specifically, for G_1 (G_2) we calculate the next closest phenotypic value toward G_2 (G_1) and call it G_1' (G_2'). The phenotypic distance between G_1' and G_2' is required to be $\geq 0.9 d_{crit}$, and the cumulative frequency of individuals between G_1' and G_2' must be less than 10% of the sum of the frequencies of G_1, G_1', G_2 and G_2'. This criterion can be applied repeatedly so that several species can be identified. It turned out that it leads to nearly perfect coincidence with decision by inspection of the distributions. Therefore, we used it as our formal criterion.

The following simpler criterion is based on calculating the disequilibrium measure D. Sufficiently large D, especially $D = 1$, corresponds to speciation, provided there is sufficiently strong assortative mating. But what is a sufficiently large D? If a is very high and competition sufficiently strong, then species can be rather close together, as quantified by d_{crit}, for example, and, in principle, several species can coexist. Therefore, we calculated a critical value D_{crit} by assuming that only two types are present, each with frequency ½, and their distance being d_{crit} (Appendix 1). Thus, D_{crit} is defined as the value of D obtained from (A5) by assuming $p_1 = ½$ and $d = d_{crit}$. In a two-locus system, the resulting D_{crit} is uniquely determined and given by

$$D_{crit}(2) = \frac{d_{crit}^2}{1 + 2d_{crit} - 2d_{crit}^2}; \tag{20}$$

see (A7). For three or more loci, D_{crit} depends on the distribution of locus effects, but good approximations can be found (A8), namely

$$D_{crit}(3) \approx d_{crit}^{2.45} \quad \text{and} \tag{21a}$$

$$D_{crit}(4) \approx d_{crit}^{2.28}. \tag{21b}$$

Consequently, a simple criterion for determining whether a distribution represents speciation is

$$D > D_{crit}. \tag{22}$$

TABLE 1. Critical distance for two species to exist, d_{crit} (eq. 19), and critical disequilibrium, D_{crit} (2) (eq. 20), and D_{crit} (3), D_{crit} (4) (eq. 21), for two, three, and four locus systems, respectively, as a function of the strength of assortative mating, a.

a	d_{crit}	D_{crit} (2)	D_{crit} (3)	D_{crit} (4)
5	0.96	0.86	0.90	0.91
6	0.88	0.63	0.72	0.74
8	0.76	0.42	0.51	0.53
10	0.68	0.32	0.39	0.41
12	0.62	0.26	0.31	0.34
14	0.57	0.22	0.26	0.28
16	0.54	0.19	0.22	0.24
32	0.38	0.10	0.09	0.11
100	0.22	0.03	0.02	0.03

TABLE 2. Equilibrium structure and frequency of speciation as a function of the strength of assortative mating. Presented are the proportion of trajectories converging to an equilibrium with the indicated number of polymorphic loci and the averages (over 300 genetic systems, each with 10 initial conditions) of the following quantities: number of equilibria, #(E); relative genetic variance, V_R; measure D of disequilibrium (calculated only among equilibria with at least one polymorphic locus); population size, N; proportion of trajectories leading to speciation, sp; proportion of trajectories for which the two criteria for speciation disagree, dis. The following parameters are fixed: $\theta = 0$, $\rho = 2$, $\kappa = 10{,}000$, $n = 4$, $s = 0.4$, $c = 2$, and $M = \infty$. A number 0 (1) means that this event was never (always) observed; 0.00 (1.00) means that the frequency of this event was less (more) than 0.005 (0.995).

	Polymorphism										
a	0	1	2	3	4	#(E)	$\bar{V}_R$	$\bar{D}$	$\bar{N}$	sp	dis
0.0	0	0	0	0	1	1	1.05	0.01	11,275	0	0
1.0	0	0	0	0	1	1	1.11	0.02	11,312	0	0
2.0	0	0	0	0	1	1	1.18	0.03	11,366	0	0
3.0	0	0.48	0.52	0	0	4.5	0.60	0.01	10,734	0	0
4.0	0.54	0.46	0.00	0	0	5.0	0.11	0.00	10,140	0	0
6.0	0.89	0.07	0.02	0.00	0.02	4.8	0.16	0.15	10,206	0	0
7.0	0.63	0.01	0	0	0.36	4.4	2.09	0.90	12,007	0.24	0.12
8.0	0.30	0.00	0	0	0.70	3.3	4.60	0.99	14,015	0.70	0.00
10.0	0.07	0	0	0.00	0.93	1.7	6.33	1.00	15,374	0.93	0.00
12.0	0.02	0	0	0.01	0.97	1.3	6.67	1.00	15,648	0.98	0
14.0	0.00	0	0	0.03	0.97	1.3	6.75	0.99	15,719	1.00	0
16.0	0.00	0	0	0.05	0.95	1.5	6.77	0.99	15,726	1.00	0
32.0	0	0	0.00	0.18	0.82	2.3	6.59	0.96	15,662	1	0
100.0	0	0	0.02	0.23	0.76	7.1	6.14	0.88	15,571	1	0

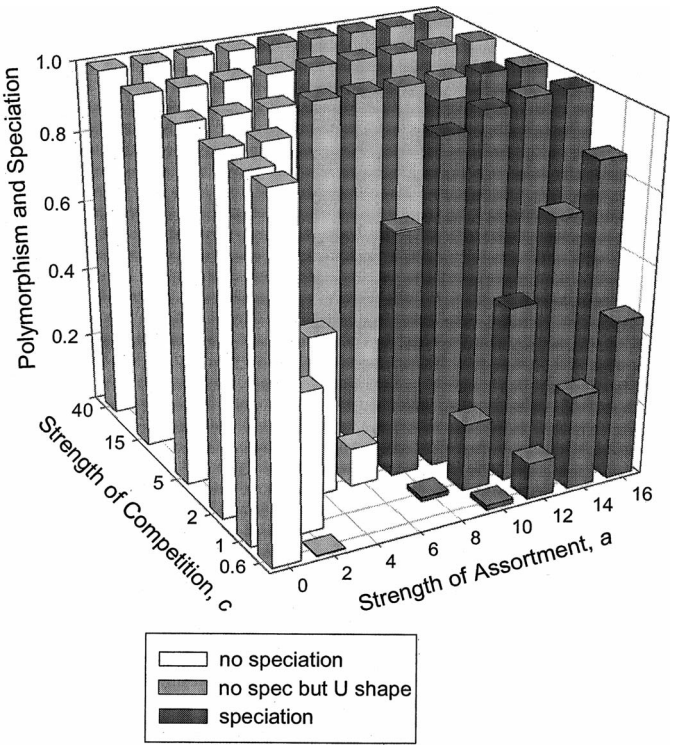

FIG. 2. Polymorphism and speciation as a function of c and a in four-locus systems. The height of each column gives the proportion of trajectories that converged to an equilibrium with at least one polymorphic locus. The white columns represent the parameter combinations for which no speciation occurred and linkage disequilibrium is low ($D < 0.3$), so that the distribution may be spread out but is not U-shaped. The gray columns indicate that no speciation occurred but $D > 0.3$, so that most distributions are U-shaped. The black columns show the proportion of trajectories converging to an equilibrium distribution representing speciation. A gray top indicates the very small proportion of trajectories that did not give rise to speciation but converged to a U-shaped distribution. There are no columns where a change in shading is invisible. The fixed parameters are $\theta = 0$, $\rho = 2$, $\kappa = 10{,}000$, $s = 0.4$, $M = \infty$.

This is not expected to work if more than two species can coexist. Table 1 lists the values of d_{crit} and D_{crit} that are used in our numerical study. In general, both criteria agree well as long as no more than two species can exist (see below).

NUMERICAL RESULTS

First, we explore the possibility of speciation under conditions considered to be favorable, that is, we assume absence of costs to female preferences ($M = \infty$) and weak stabilizing selection ($s = 0.4$). Therefore, the phenotypic range is narrower than the width of the stabilizing selection function $S(g)$, the fitness of the extreme phenotypes is reduced by only 10%, and the fitness function (7) is a very accurate approximation to the actual fitness, equation (6), if $c \leq 1$. We investigate the roles of the strength of competition, c, and assortment, a, of the number of loci, n, of the distribution of allelic effects, γ_i, and of the initial conditions. Then we study the role of costs ($M < \infty$) and strong stabilizing selection ($s = 2$) or, equivalently, a wide phenotypic range. Whenever we use the term ''equilibrium,'' we mean an equilibrium that was approached by at least one trajectory in our numerical study. Generically, such equilibria are locally asymptotically stable.

Strength of Competition and Assortment

The equilibrium structure, the properties of the equilibrium distributions, and the probability of speciation all depend in a highly nonlinear way on c and a. Here, we assume that four loci contribute to the trait. Table 2 documents how the most important quantities depend on the strength of assortment if $c = 2$. It shows that a complex reorganization of the equilibrium structure occurs as the strength of assortment increases. In particular, for intermediate values of a ($a = 4, 6$) very little genetic variation is maintained.

Figure 2 visualizes how polymorphism and speciation depend on both the strength of assortment and competition. According to this figure, $a \geq 8$ seems necessary for speci-

TABLE 3. The role of the number of loci for speciation. Shown is the proportion of trajectories that converge to an equilibrium classified as representing speciation. The fixed parameters are: $\theta = 0$, $\rho = 2$, $\kappa = 10{,}000$, $s = 0.4$, $M = \infty$. For $a \leq 5$, no speciation was observed. For $a = 32$ and $a = 100$, often more than two species arise. If $a = 32$, two or three species coexist if $c = 2$, 5, or 15. If $a = 100$, two or three species coexist if $c = 0.6$ or 2 (but not if $c = 1$), two to four species if $c = 5$, and up to five species if $c = 15$ or 40. For $c = 40$, populations went extinct in some cases. This is likely due to the property of discrete logistic growth that overcrowding can lead to extinction. The frequency of extinction events are: (1) 9%; (2) 60%; (3) 6%.

		a							
c	*n*	6	8	10	12	14	16	32	100
0.6	2	0.02	0.41	0.62	0.75	0.81	0.86	1	1
	3	0	0.03	0.21	0.44	0.61	0.73	1.00	1
	4	0	0	0.01	0.10	0.26	0.45	0.99	1
1.0	2	0.38	0.66	0.79	0.88	0.93	0.97	1	1
	3	0	0.28	0.62	0.79	0.89	0.95	1	1
	4	0	0.01	0.19	0.50	0.73	0.86	1	1
2.0	2	0.74	0.93	0.95	0.97	0.98	0.99	1	1
	3	0	0.92	0.97	0.99	1.00	1.00	1	1
	4	0	0.70	0.93	0.98	1.00	1.00	1	1
5.0	2	0	0.14	0.34	0.52	0.67	0.81	1	1
	3	0	0.00	0.02	0.89	0.94	0.93	1	1
	4	0	0	0	0.94	1	1	1	1
15.0	2	0	0	0	0	0	0	0.63	0.89
	3	0	0	0	0	0	0	0.83	0.99
	4	0	0	0	0	0	0	0.80	1
40.0	2	0	0	0	0	0	0	0 (1)	0.60 (2)
	3	0	0	0	0	0	0	0	0.91 (3)
	4	0	0	0	0	0	0	0	0.94

ation, but for $c = 2$ (and, among the values of c we used, only for $c = 2$), speciation occurs already if $a = 7$ (Table 2). Interestingly, if $c = 5$, stronger assortment is required to induce population splitting than for $c = 0.6$, 1, and 2 (but for $c = 0.6$ or 1, speciation is rare if $a = 8$, 10). No speciation at all was observed for $c = 15$ and $c = 40$ if $a \leq 16$. As shown in Table 3, speciation may occur for $c = 15$ and $c = 40$ if assortment is extremely strong, that is, if $a = 32$ or 100. Then, more than two species may emerge, if $a = 100$, up to five were observed. Of course, if $c \leq s = 0.4$, speciation is infeasible because then overall selection is stabilizing, not disruptive.

The reason why very strong competition (e.g., $c = 15$, 40) requires extremely strong assortment to induce population splitting is that with $a \leq 16$ and two clusters at a distance $\geq d_{\text{crit}}$, individuals in the middle do not suffer a fitness reduction from competition, that is, large c opens up one (or even more) new niche(s) in the middle. If assortment is too weak to prevent mating between individuals in the (or a) middle niche with the extreme types, three species cannot be maintained and a complicated U-shaped or multimodal distribution may result (e.g., see Fig. 3c).

The question arises whether the strong assortment needed for speciation in this model is biologically realistic. Already $a = 16$ represents very strong assortative mating: the probability that two individuals with extreme and opposite phenotype mate is about 10^{-7}, the probability for individuals that are half the phenotypic range apart is 0.018. With $a = 100$, individuals that are only 20% of the total phenotypic range apart mate with probability 0.018. Individuals differing by $\sqrt{1/12} \approx 0.29$, one standard deviation of the uniform distribution on $[-\frac{1}{2}, \frac{1}{2}]$, mate with probability 2.4×10^{-4}. (Note that any unimodal distribution has a smaller variance than the uniform distribution.) If the strength of assortment is compared with the strength of competition, then $a \geq 32$ can also be regarded as extremely strong compared with $c \leq 5$ because an individual's niche width ($1/\sqrt{2c}$) is then more than 2.5 times as large as its mating width ($1/\sqrt{2a}$).

In summary, speciation clearly occurs, but strong assortment is necessary and an intermediate strength of competition is most favorable. The purely ecological condition $c(\rho - 1) > s$, or one of its equivalents for other forms of population regulation (Bürger 2005), is by no means sufficient.

Another interesting feature, already noted earlier in simpler models (Bürger and Schneider 2006; Schneider and Bürger 2006), is the depletion of genetic variation by moderately strong assortment, as shown by Figure 2 if $c \leq 2$ (see also Table 2; Figs. 4, 5a). This is predicted on theoretical grounds (see Theoretical Background and Analytical Results) and relatively easy to understand intuitively. For random mating ($a = 0$), polymorphism is maintained whenever $c(\rho - 1) > s$ and stabilizing selection is not extremely strong. In comparison with random mating, weak assortment counteracts competition because it induces local stabilizing selection around monomorphic states because for deviating types it is more difficult to find a mating partner. This is more pronounced if mating is strongly selective (M is small), and it is reflected by equations (12), (13), and (14). Thus, to maintain polymorphism with certainty, competition must be stronger than stabilizing selection and assortment together. If assortment becomes very strong, then two clusters that are sufficiently far apart (so that individuals of different type do not mate) can coexist because no intermediate types are produced (cf. Bürger and Schneider 2006). Without showing all the results, we note that equations (12–14) predict the onset of the loss of variation very well if $c \leq 2$ and $s = 0.4$ (e.g., Table 2).

For $c \geq 5$, equations (12–14) are no longer applicable. Then, independently of a, all equilibria have at least two polymorphic loci and the relative genetic variance is increas-

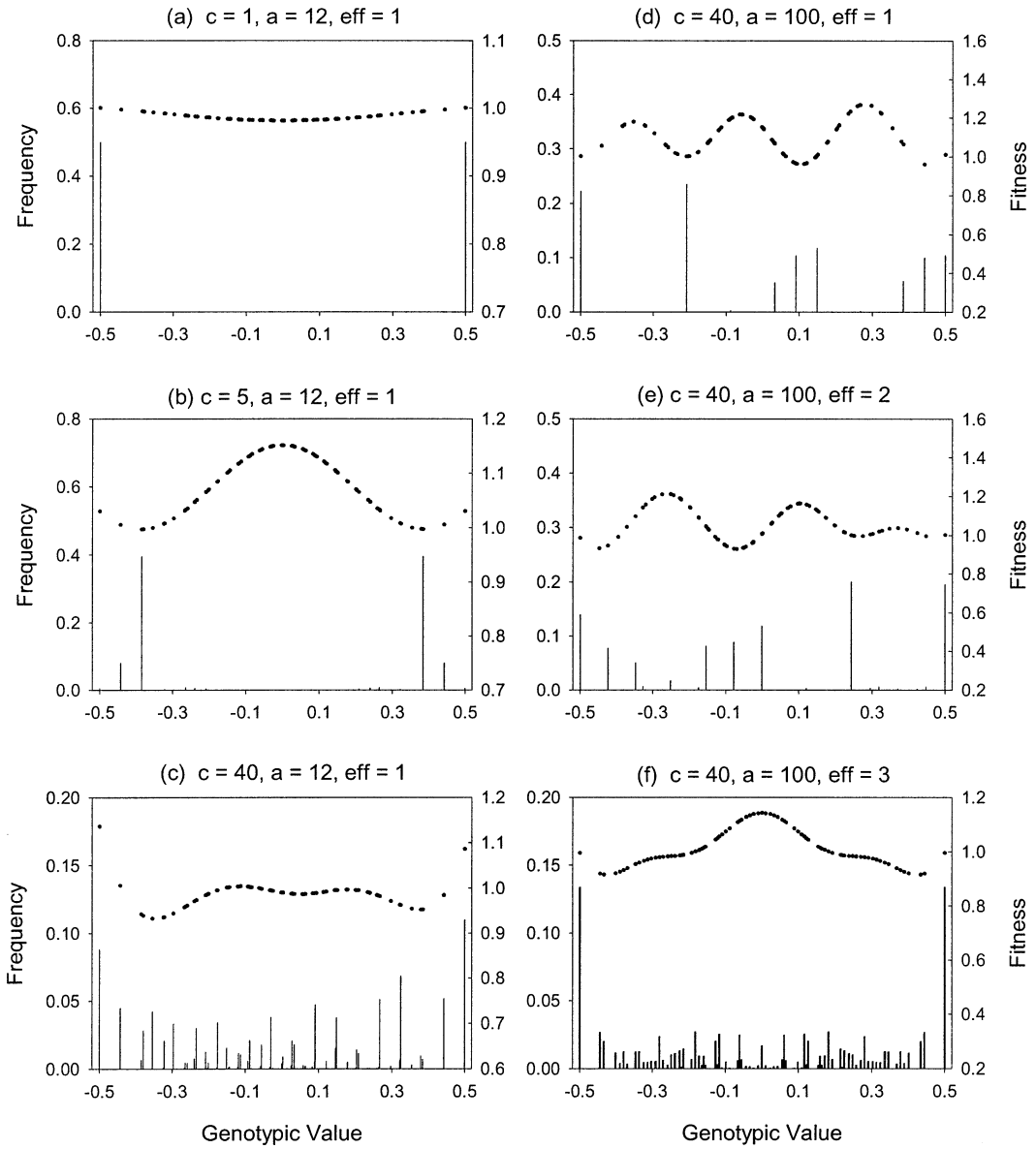

FIG. 3. Equilibrium distributions and fitness landscapes in four-locus systems. Each panel displays an equilibrium distribution and the corresponding fitness landscape for the varied parameters as indicated. In all cases we have $\theta = 0$, $\rho = 2$, $\kappa = 10{,}000$, $s = 0.4$, $M = \infty$. For the panels with eff = 1, the locus effects are (sorted by size, as is admissible for free recombination) 0.058, 0.121, 0.146, 0.176; if eff = 2, the effects are 0.052, 0.076, 0.173, 0.198; if eff = 3, they are 0.055, 0.066, 0.098, 0.280. In panels (a) and (b) there are two species, in (c) and (f) there is one, in (d) there are four species, and in (e) there are three. In (a–f), the measure D of disequilibrium assumes the values 1.00, 0.55, 0.35, 0.44, 0.41, 0.35, respectively.

ing as a function of a (though this is due mainly to an increase in linkage disequilibrium). The proportion of fully polymorphic equilibria is >0.98 if $c = 5$ and $0 \le a \le 8$, if $c = 15$ and $0 \le a \le 32$, or if $c = 40$ and $0 \le a \le 100$. Still, the proportion of highly polymorphic equilibria is reduced if a is sufficiently large. (Note that the height of the bars in Fig. 2 gives the proportion of trajectories that do not converge to a monomorphism.) For instance, if $c = 5$ and $10 \le a \le 16$, the proportion of fully polymorphic equilibria is down to about 0.35; for larger a, it increases again. If $c = 15$ and $a = 100$, this proportion is still very high, namely 0.96. Thus, although very strong competition is not beneficial for speciation, it is a potent force in maintaining high levels of genetic variation.

If $a \le 16$ and speciation occurs, the shape of the corresponding distribution is in general very simple. It typically consists of two spikes, representing homozygous genotypes that are situated at or close to the boundary of the phenotypic range. In the vast majority of cases, each of them has a frequency of ≥ 0.4, often close to 0.5. Examples are displayed in Figure 3. The average distance between these spikes is often much larger than predicted by d_{crit} (Table 1). For instance, if $a = 16$, their average distance is 0.85, 0.95, 0.99, and 0.81 for $c = 0.6$, 1.0, 2.0, and 5.0, respectively (recall

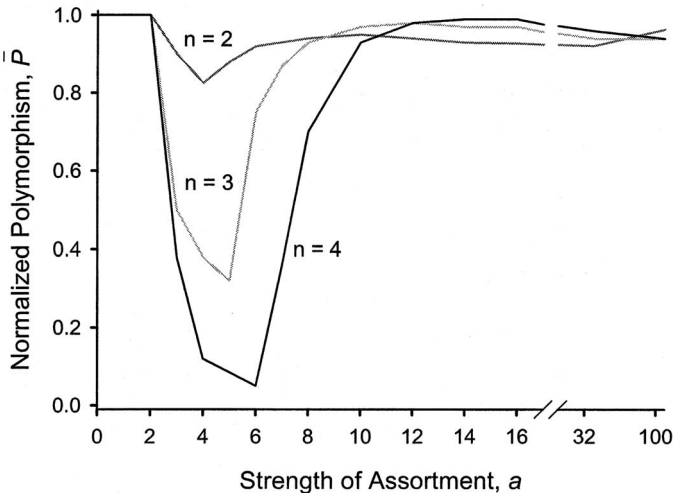

FIG. 4. Influence of the number of loci on polymorphism as a function of a. The average normalized polymorphism is shown, that is, the number of polymorphic loci at equilibrium as fraction of the number of loci. The fixed parameters are $\theta = 0$, $\rho = 2$, $\kappa = 10{,}000$, $s = 0.4$, $M = \infty$, and $c = 2$. Note the different scale for $a > 16$.

that speciation does not occur if $c = 15$ or 40, and $a \leq 16$). If $a > 16$, the shape of the distributions, whether speciation occurs or not, is usually complex and often more than two species coexist (Table 3; Fig. 3). The distinctive fitness maxima occurring in Figures 3b,d, and e indicate open niches. They are not filled up because assortment is too weak to prevent mating between individuals from such a niche with those from neighboring niches, and hybrid individuals would lead to increased competition.

Number of Loci

The number of loci contributing to the trait has a significant effect on the amount and pattern of variation if $c \leq 5$. Figure 4 is representative for $c \leq 2$, but for smaller c the attenuating effect of moderately strong assortment on genetic variation is stronger. If $c = 5$, all stable equilibria have at least one polymorphic locus. Nevertheless, the degree of polymorphism is reduced if assortment exceeds a certain strength, and more so if the number of loci is higher. If $n = 4$, the proportion of fully polymorphic equilibria is much lower at intermediate a than if $n = 2$ or 3 (cf. Bürger and Schneider 2006).

The number of loci also affects the likelihood of speciation (Table 3). Independently of c, in two-locus systems the first speciation events occur at smaller values of a than in three- or four-locus systems. If $c = 0.6$ or $c = 1$, the likelihood of speciation decreases with increasing number of loci. If $c = 2$ or 5, this is so only below a certain value of a; above this value, the likelihood of speciation is independent of n or even increases with n. An intuitive explanation is given in the Discussion.

Allelic Effects and Initial Conditions

Empirical data suggest a highly skewed distribution of locus effects (e.g., Bürger 2000; Barton and Keightley 2002). By contrast, most previous studies of sympatric speciation that admit multiple loci assumed equal effects among loci (e.g., Dieckmann and Doebeli 1999; Bolnick 2004b; Gavrilets 2004, ch. 10.3; Gourbiere 2004; Kirkpatrick and Nuismer 2004). Therefore, we explored the role of allelic effects and initial conditions by comparing the following seven scenarios which all assume $n = 4$. The first is our standard scenario in which 300 sets of allelic effects, each with 10 random initial conditions, are chosen randomly as described in Numerical Approach. In scenarios 2, 3, and 4, the allelic effects at the four loci are in proportion 1:2:4:8, and either 10 initial conditions are chosen randomly; or one initial condition is chosen in linkage equilibrium with allele frequencies 0.491,

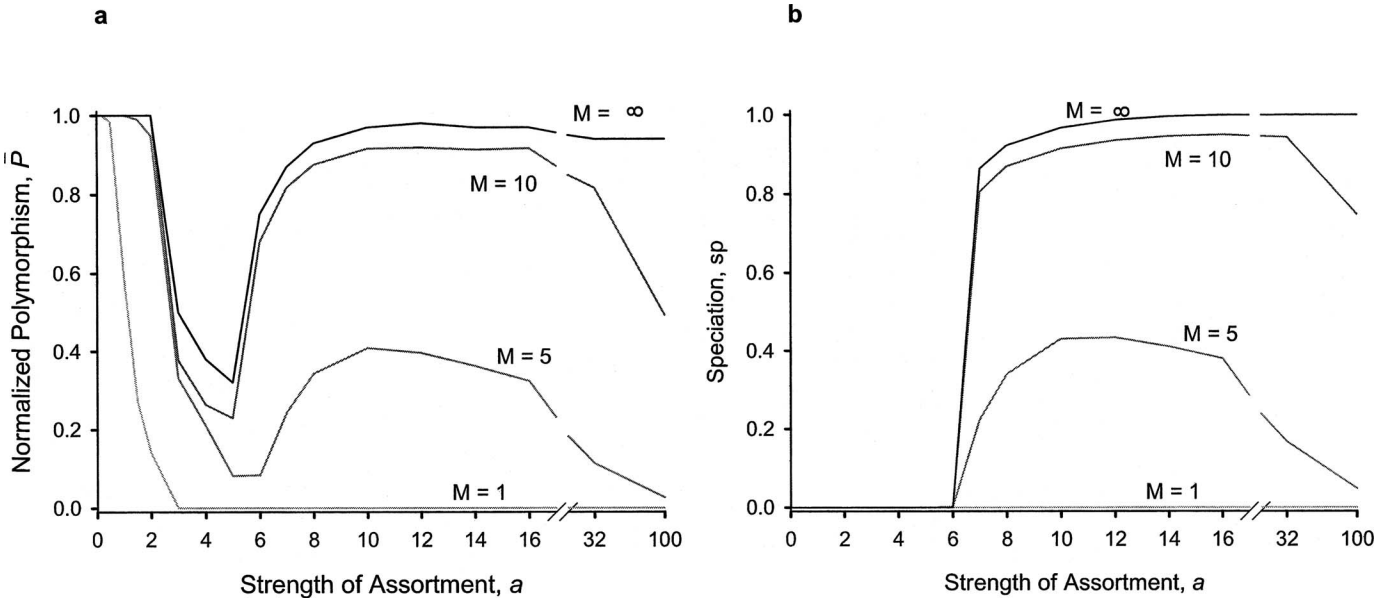

FIG. 5. Influence of the costs on polymorphism (a) and speciation (b) in three-locus systems as a function of a. Other parameters are the same as in Figure 4. Note the different scale for $a > 16$.

TABLE 4. Influence of allelic effects and initial conditions. The parameters are $n = 4$, $s = 0.4$, $a = 12$, $M = \infty$. Allelic effects (AE) can be random (RD), in proportion 1:2:4:8 (1248), or equal (EQ). Initial conditions (IC) can be random (RD), all alleles at nearly the same frequency and close to linkage equilibrium (LE), or have variation only due to rare alleles (RA). MP indicates the proportion of trajectories converging to a multilocus polymorphism (i.e., two or more loci are polymorphic); $\bar{P}$ is the average normalized polymorphism, that is, the number of polymorphic loci at equilibrium divided by the number of loci of the system. Other quantities are as in Table 2.

c	AE	IC	MP	$\bar{P}$	$\bar{V}_R$	$\bar{D}$	sp
1.0	RD	RD	0.50	0.48	3.08	0.93	0.73
	1248	RD	0.90	0.88	4.50	0.94	0.80
	1248	LE	1	1	5.29	1.00	1
	1248	RA	0.25	0.19	1.15	0.84	0.25
	EQ	RD	0.20	0.20	1.60	1.00	0.20
	EQ	LE	1	1	8.00	1.00	1
	EQ	RA	0	0	0	—	0
2.0	RD	RD	0.98	0.98	6.67	1.00	0.98
	1248	RD	1	1	5.29	1.00	1
	1248	LE	1	1	5.29	1.00	1
	1248	RA	0.50	0.38	2.14	0.77	0.50
	EQ	RD	1	1	8.00	1.00	1
	EQ	LE	1	1	8.00	1.00	1
	EQ	RA	0.13	0.13	1.00	1.00	0.13
5.0	RD	RD	1	0.76	4.33	0.58	0.94
	1248	RD	1	0.50	3.36	0.56	1
	1248	LE	1	0.50	3.36	0.56	1
	1248	RA	1	0.50	3.36	0.56	1
	EQ	RD	1	1	5.66	0.67	0
	EQ	LE	1	1	5.66	0.67	0
	EQ	RA	0.63	0.50	2.86	0.54	0.50
40.0	RD	RD	1	1	2.90	0.34	0
	1248	RD	1	1	2.17	0.28	0
	EQ	RD	1	1	3.39	0.36	0

0.495, 0.504, 0.508; or eight initial conditions are chosen such that at each locus there is one allele at frequency 0.05, the other at 0.95, and there is linkage equilibrium. Scenario 3 is a small perturbation of the completely symmetric initial condition in order to avoid extreme symmetry and, possibly, nongeneric behavior. In scenarios 5 to 7, the initial conditions are analogous, but all loci have equal effects. It is interesting to note that, for a constrained phenotypic range and a freely evolving genetic architecture, Kopp and Hermisson (2006) showed numerically that locus effects evolve (approximately) to the proportions 1:2:4: . . . if competition is strong enough to provide sufficiently many niches (to be filled with homozygous genotypes and their relatives).

The most relevant results are summarized in Table 4. It is evident that the choice of allelic effects and initial conditions may have a marked influence on the evolutionary outcome. If initially one allele at each locus is rare and the other common, then much less polymorphism and genetic variance is maintained at equilibrium than if initially both alleles are at intermediate frequency. The reason is that the initial frequencies lie with higher probability in the region of attraction of a monomorphic equilibrium. Interestingly, the same is mostly, but not always, true for the probability of speciation (see the exceptional case of equal effects if $c = 5$). Without showing the results, we note that the influence of initial conditions becomes weaker as a declines. This is expected because if $a = 0$, $c(\rho - 1) > s$, and loci are only loosely linked so that linkage equilibrium can be assumed, then there exists a uniquely determined, globally stable equilibrium (Bürger 2005; Schneider 2006); thus initial conditions are irrelevant. This remains true for sufficiently small a, but what is sufficiently small depends on c: the larger c, the larger is sufficiently small. For instance, if $c = 1$, initial conditions do not influence the evolutionary outcome if $a \leq 0.8$ and very weakly so if $a \leq 6.0$. If $c = 40$, with the exception of a single trajectory, no influence at all was observed for $a \leq 16$.

There is no clear tendency by which the equilibrium structure for loci with equal effects departs from the average. For small to moderate a (approximately $a \leq 6$), they are usually very similar. For intermediate values of a, often less variation is maintained with loci of equal effect (results not shown). For large values of a ($a \geq 12$) the probability of speciation can be higher or lower than on average, depending on c and a (Table 4).

We studied the role of locus effects also in a different way by choosing the allelic effects γ_i from the interval [0.01, 1] instead of [0.1, 1]. For numerical reasons (runs with one or more loci of very small effect converge extremely slowly), this was done only for two- and three-locus systems. For two loci, admission of loci of very small effect slightly increases the proportion of polymorphic equilibria and the frequency of speciation. For three loci, and the same must be true for four, inclusion of loci of very small effect changes the data by shifting them slightly in the direction of the two-locus case. This is so because if in a four-locus system one or two loci have very small effects compared to the others, then it behaves similar to a three- or two-locus system, respectively. From a practical point of view, the exclusion of loci of very small effect does not seem to be a substantial restriction because QTLs with very small effects are hardly detectable by current methods, and their influence on genetic variation and the evolutionary dynamics is minor.

Complex Dynamics

For very strong competition ($c \geq 15$) and assortment ($a \geq 16$) complex evolutionary dynamics occurred for some parameter combinations and in some genetic systems. For instance, if $n = 3$, $s = 0.4$, $c = 15$, $a = 16$, in six of 1000 genetic systems stable limit cycles of very long period occurred (several thousands or tens of thousands of generations). Stable limit cycles continue to exist if the parameters c, a, and the locus effects are slightly varied. Their existence is very sensitive to variation in the locus effects, however. If $n = 4$, $s = 0.4$, $c = 40$, and $a = 100$, in two of 300 runs stable limit cycles of period two were observed. They continue to exist (again with period two) if parameters are slightly varied. We do not know if these cycles are caused by the genetic dynamics, the demographic dynamics, or a combination of both. Because the period of the very long cycles seems to change in a quasi-continuous way, it appears unlikely that they are generated by the demographic dynamics.

Costs to Being Choosy

The role of costs to female preferences were investigated mainly for three-locus systems. Figure 5 shows that, not sur-

TABLE 5. Equilibrium structure and frequency of speciation for strong selection. Presented are the same quantities as in Table 2, but for different values of s, c, and a. The following parameters are fixed: $\theta = 0$, $\rho = 2$, $\kappa = 10{,}000$, $n = 4$, $s = 2$, and $M = \infty$. The numbers 0, 1, 0.00, and 1.00 have the same meaning as in Table 2.

		Polymorphism										
c	a	0	1	2	3	4	#(E)	$\bar{V}_R$	$\bar{D}$	$\bar{N}$	sp	dis
5	12.0	0.75	0.25	0	0	0	4.8	0.10	0.02	10,221	0	0
	16.0	0.77	0.21	0	0	0	5.1	0.16	0.05	10,347	0.01	0.01
	32.0	0.18	0.42	0.36	0.04	0.00	7.7	1.02	0.13	12,014	0.64	0.23
	100.0	0	0.20	0.46	0.21	0.04	8.4	1.39	0.15	12,809	1	0.01
10	12.0	0	0.67	0.33	0	0	5.2	0.66	0.03	13,377	0	0
	16.0	0.00	0.77	0.23	0	0	5.5	0.78	0.05	13,807	0.02	0.02
	32.0	0	0.30	0.68	0.03	0.00	6.4	1.45	0.15	15,546	0.72	0.35
	100.0	0	0.04	0.34	0.48	0.14	8.3	1.73	0.19	16,415	1	0.00
20	12.0	0	0.13	0.67	0.19	0.01	5.5	1.14	0.09	19,268	0	0
	16.0	0	0.19	0.67	0.14	0.01	6.0	1.19	0.10	19,384	0	0
	32.0	0	0.14	0.60	0.24	0.02	6.4	1.44	0.15	19,629	0.32	0.60
	100.0	0	0	0.16	0.45	0.40	8.0	1.96	0.21	21,553	1.00	0.00
200	16.0	0	0	0	0.04	0.96	4.3	1.80	0.18	62,503	0	0
	32.0	0	0	0.00	0.11	0.89	4.5	2.03	0.22	63,509	0	1.00
	100.0	0	0	0	0.06	0.94	4.3	2.12	0.23	62,524	0.15	0.86

prisingly, highly selective mating ($M = 1$) prohibits speciation. However, for $M = 5$, which still represents relatively high costs, speciation occurs at appreciable frequency, and with $M = 10$ it is only slightly reduced, unless $a = 100$. As expected, costs become more important for stronger assortment. With four loci, $c = 2$ and $a = 16$, the probabilities of speciation in our model are 1.00, 0.95, and 0.20 for $M = \infty$, 10, and 5, respectively. Comparison with Figure 5a indicates that costs may become more important as the number of loci increases.

Quite unexpectedly, we found that for very strong competition ($c = 15$) and extremely strong assortment ($a = 32$), intermediate values of M ($M = 5$, 10) can lead to a considerably higher frequency of speciation events than $M = \infty$. The average number of polymorphic loci at equilibrium, however, is always substantially reduced with higher costs. The reason appears to be that with smaller (but not too small) M, some intermediate types are lost because females are less likely to mate with males of nonsimilar phenotype. Because for such strong competition several niches coexist, however, a highly polymorphic population can disaggregate into two or more subpopulations through loss of intermediate types.

Strong Stabilizing Selection

By making stabilizing selection sufficiently strong or, equivalently, the range of phenotypic effects sufficiently large, we can test if the evolution of reproductively isolated clusters occurs only if extreme phenotypes are selectively favored. To this aim we chose $s = 2$ (instead of $s = 0.4$ as above). For computational reasons, numerical calculations were performed primarily for strong assortment. The most relevant results are summarized in Table 5, where absence of costs and four loci are assumed. It shows that speciation can still occur if stabilizing selection is so strong that the extreme phenotypes are not maintained in the population or only at very low frequency. The conditions for speciation, however, are very restrictive. For instance, with $c(\rho - 1)/s$ as in Table 3, no speciation events were observed for $a \leq 12$. Without showing the results, we mention that equation (12) still predicts the loss of variation by increasingly strong assortment accurately if $s = 2$ and $c \leq 5$. It is not surprising that extremely strong assortment is needed if $s = 2$ to induce population splitting because the two species cannot be established near the boundaries of the phenotypic range. As it was the case with smaller s, very large c/s is again prohibitive of speciation, that is, extremely large a is needed. Additional numerical results for two specific genetic systems, four loci of equal effects and four loci with effects in proportion 1:2:4:8, show that for weak assortment, high amounts of polymorphism are maintained if $c \geq 3$ and that for intermediate values of a, polymorphism is reduced. Thus, the overall pattern of polygenic variation as a function of a is qualitatively similar to the case of weak stabilizing selection. For random or weak assortative mating and for the same values of $c(\rho - 1)/s$, the amount of variation maintained tends to be somewhat lower with large s than with small s (results not shown).

Roughgarden's Versus Bulmer's Fitness Function

As shown above, our functional form of fitness (6), introduced by Bulmer (1974), is very similar to that of Roughgarden if stabilizing selection is weak but differs if it is strong. In that case, fitness of extreme phenotypes declines in our model, whereas it levels off to its maximum value in the Roughgarden model. The latter seems unrealistic and might favor speciation unduly (Polechová and Barton 2005).

We compared evolution under our fitness assignment (6) with evolution under Roughgarden's fitness (16). As expected from the considerations above, as well as from general theoretical results (Bürger 2005), the evolutionary outcome for the two choices, for example, structure, frequency of speciation, or population size, is very similar if stabilizing selection is weak ($s = 0.4$), even if c is large (results not shown). There is a tendency that with Roughgarden's fitness slightly more variation is maintained and, in some cases, speciation is a little bit more frequent.

With strong stabilizing selection ($s = 2$), there is always considerably more variation maintained under Roughgarden's fitness (Table 6). Especially, equilibria with a high level

TABLE 6. Comparison of equilibrium structure and frequency of speciation under Roughgarden's (eq. 16) and our fitness function (eq. 6), for strong selection. Presented are the same quantities as in Table 2, but for different values of n, s, c, and a. The following parameters are fixed: $\theta = 0$, $\rho = 2$, $\kappa = 10{,}000$, $n = 3$, $s = 2$, and $M = \infty$. The numbers 0, 1, 0.00, and 1.00 have the same meaning as in Table 2. The results in this table are based on only 300 runs for each parameter combination.

		Polymorphism									
c	a	0	1	2	3	#(E)	$\bar{P}$	$\bar{V}_R$	$\bar{D}$	$\bar{N}$	sp
Roughgarden's fitness											
3	0	0	0.00	0.00	0.99	1.0	1.00	1.01	0.00	11,574	0
	12	0.66	0.32	0.01	0.01	3.2	0.12	0.29	0.10	10,408	0.12
	16	0.44	0.41	0.15	0.01	4.4	0.24	0.68	0.16	10,884	0.42
	100	0	0.06	0.40	0.54	5.9	0.83	1.64	0.23	12,180	1
10	0	0	0	0	1	1	1	1.13	0.03	16,674	0
	16	0	0.16	0.40	0.44	3.6	0.76	1.91	0.26	18,472	0.13
	32	0	0.05	0.33	0.62	4.2	0.86	2.24	0.33	18,993	0.98
20	32	0	0	0.17	0.83	3.5	0.94	2.22	0.32	24,764	0.47
	100	0	0	0.07	0.93	4.5	0.98	2.30	0.34	25,325	0.99
Bulmer's (our) fitness											
3	0	0	0.46	0.54	0.01	3.1	0.52	0.55	0.00	10,847	0
	12	0.88	0.11	0.01	0.01	2.8	0.05	0.08	0.05	10,008	0.00
	16	0.82	0.12	0.05	0.01	3.2	0.08	0.15	0.08	10,116	0.07
	100	0.07	0.69	0.17	0.06	5.1	0.41	0.58	0.05	10,832	0.90
10	0	0	0	0.07	0.93	1.1	0.98	1.03	0.01	15,432	0
	16	0.00	0.80	0.16	0.04	3.9	0.41	0.85	0.09	14,405	0.09
	32	0	0.47	0.39	0.14	4.0	0.56	1.21	0.14	15,486	0.93
20	32	0	0.12	0.52	0.37	3.5	0.84	1.46	0.18	20,234	0.55
	100	0	0.01	0.26	0.73	4.5	0.91	1.60	0.19	21,249	0.98

of polymorphism are more frequent. The reason is that, in contrast to Roughgarden's fitness, in our model fitness decreases rapidly near the boundaries of the phenotypic range (Fig. 1). Also linkage disequilibrium is considerably higher with Roughgarden's fitness, and the fitness landscape tends to be much flatter than for the fitness function in our model. Interestingly, however, the frequency of speciation is not greatly changed in Roughgarden's model. For small c it is higher, because niches can be established near the boundary of the phenotypic range, but for large c it may even be lower. Why speciation is not always enhanced under Roughgarden's fitness is not clear.

Beverton-Holt Population Regulation

We also performed some numerical work for Beverton-Holt population regulation (eq. A10). The results obtained are qualitatively similar. If the growth rate λ equals two, then much more variation is lost for intermediate a than for logistic growth and $\rho = 2$. Also the range of such values a is larger because in equation (12) $c(\rho - 1)$ has to be replaced by $c(1 - \lambda^{-1})$ (eq. A39). If $\lambda = 100$, then much less variation than for logistic growth with $\rho = 2$ is lost. For both values of λ, speciation occurs approximately in the same range of parameters as for logistic growth with $\rho = 2$. If $\lambda = 2$, the likelihood of speciation is reduced, if $\lambda = 100$ it is slightly enhanced.

Comparison of the Formal Criteria for Speciation

In general, the two criteria we used coincide very well (e.g., Table 2). Differences occurred for the lowest values of a at which speciation occurs and for $c \geq 15$. In both cases, the criterion based on disequilibrium, equation (22), indicates speciation when the more elaborate one does not. For low values of a, the discrepancy occurs because both methods yield values close to the critical values used for identifying speciation, one lower, the other higher. For $c \geq 15$, the criterion based on D yields speciation for values of a that are too low because it was designed to predict speciation into two species properly, but not when the utilization widths of phenotypes ($1/\sqrt{2c}$) is so small that more than two niches are provided.

DISCUSSION

The purpose of this study was to elucidate the conditions for population splitting, or sympatric speciation, under disruptive selection caused by intraspecific competition for resources. With the possible exception of habitat choice, this is arguably the most promising mechanism for sympatric speciation (Turelli et al. 2001; Via 2001; Dieckmann et al. 2004; Gavriltes 2004; Kawecki 2004; Kirkpatrick and Nuismer 2004). However, there are numerous possible determinants of this process. These include the nature of the resources (discrete or continuous, unimodal or multimodal), the precise choice of the functional form for fitness, the strength of frequency dependence caused by competition (inversely related to the individual's niche width), the strength of stabilizing selection if the resource is unimodal, the kind of population regulation and the associated parameters, assumptions on assortative mating and its strength and costs, whether natural selection acts directly on the assortment trait or on another, asexual or sexual reproduction, all kinds of assumptions on the underlying genetics (haploid or diploid, number of loci, number of alleles per locus, distribution of allelic effects, additivity of effects versus epistasis or dominance, recombination rates, a fixed genetic architecture or one that can evolve, e.g., because new mutants are admitted), and assumptions on the initial population distribution. Dieckmann

and Doebeli (1999) argued that with a continuous unimodal resource, Roughgarden's version of the Lotka-Voltera model of competition, and direct natural selection on the mating trait, strong assortative mating will evolve, thus leading to sympatric speciation, if the frequency-dependent effect of competition is stronger than stabilizing selection ($c > s$ in our terminology). Thus, roughly, sympatric speciation is considered to be predictable primarily from ecological conditions, with genetic or other factors playing a minor role (see also Doebeli and Dieckmann 2000).

The bottom line of our results is that population splitting and the evolution of reproductively isolated clusters may occur under disruptive selection, but even strong disruptive selection and strong assortative mating are not sufficient to guarantee it. In particular, whether speciation is a likely event or occurs only for particular genetic architectures and special initial conditions depends on a relatively delicate interplay of our main parameters, c, ρ, s, and a. If the optimum θ of stabilizing selection is near the center of the phenotypic range, then $c(\rho - 1) > s$ is a sufficient condition for the emergence of disruptive selection near that center. In ecological terms, this condition can be interpreted as the width of the individual's utilization function being narrower than that of the resource distribution. Speciation, however, requires more. If disruptive selection is weak, that is, $c(\rho - 1)$ is not much larger than s, then increasing c promotes speciation because with increasingly large c, the strength of assortment needed for population splitting decreases. Interestingly, very large c hinders speciation; that is, beyond a certain (not very large) value of c, any further increase in c leads to a much larger minimal strength of assortment necessary for speciation (to the extent that only individuals can mate that are a small fraction of a phenotypic standard deviation apart). For typical examples where no speciation occurs despite large c, see Figures 3c,f. Thus, all other parameters fixed, there is an intermediate strength of frequency-dependent competition that maximizes the likelihood for speciation (Fig. 2; Tables 3–5).

Similar observations were made by Gourbiere (2004), Kirkpatrick and Nuismer (2004), and Schneider and Bürger (2006) for somewhat different models. The reason for this nonlinearity is that with very large c, individuals use only a small range of the resource spectrum. Hence, large c opens up intermediate niches. Unless assortment is sufficiently strong to prevent mating between niches, intermediate heterozygotes counteract separation of clusters. More than two species can coexist if assortment is correspondingly high (see also Bolnick 2006; Kopp and Hermisson 2006). For an asexual population with a continuum of possible types, the niche structure can be determined numerically in a straightforward way (Appendix 3; Kopp and Hermisson 2006). If $s = 0.4$, then two niches close to or at the boundary of the phenotypic range exist if $0.4 < c < 1.6$. If $c \approx 1.6$, a third niche in the middle (at $g = 0$) opens up. This splits into two if $c \approx 6.9$. A fifth niche (again at $g = 0$) emerges if $c \approx 14.1$. If $33.3 \leq c \leq 45.0$, there are seven niches (M. Kopp and J. Hermisson, pers. comm.). To fill all these niches with reproductively isolated populations, extremely strong assortment is needed. For instance, if $a = 100$, the strongest assortment we considered, at most five clusters were observed if $c = 40$.

It has been noted previously (Drossel and McKane 2000; Gavrilets 2004; Gourbiere 2004; Kirkpatrick and Nuismer 2004; Doebeli 2005) that costs to assortative mating will reduce the likelihood of speciation. Unless competition is very strong, this is confirmed by our study. Nevertheless, if assortment is not extremely strong ($a \leq 16$), a substantial reduction was found only if $M \leq 5$, in which case mating is already highly selective. If $M = 10$, the probability of speciation is reduced only by several percent relative to $M = \infty$. For much stronger assortment (e.g., $a = 100$), as used in some previous studies (e.g., Gourbiere 2004), a substantial reduction of the likelihood of speciation occurs already if $M = 10$. These findings are in accordance with Schneider and Bürger (2006), who studied a single-locus model with many alleles. By contrast, Bolnick (2004b) reported that the waiting time to speciation is much more sensitive to costs. His model, however, differs not only in several genetic aspects from ours but, in particular, allows the strength of assortment to evolve. Somewhat surprisingly, we found that intermediate costs (e.g., $M = 5,10$) can enhance the frequency of speciation if competition is very strong ($c = 15$) and assortative mating is extremely strong ($a = 32$).

Most studies have ignored the role of the choice of allelic effects and of initial conditions on the evolutionary outcome. They assumed that all loci have the same effects and that initially both alleles at each locus occur at (nearly) equal frequency. Notable exceptions, in part for different models of speciation, are Geritz and Kisdi (2000), Kirkpatrick and Ravigné (2002), Gavrilets (2004, pp. 380–382), Kirkpatrick and Nuismer (2004), and Bürger and Schneider (2006). Our results show that, unless assortment is absent or weak, for each genetic system typically several stable equilibria, often with different number of polymorphic loci, coexist (their average number is given by the entries #(E) in Tables 2, 4, 5). Thus, the initial population distribution may be of paramount importance for the eventual outcome: for a given genetic system and given ecological and mating parameters, evolution may result in loss of all variation (because we ignore mutation) as well as in speciation with varying degree of differentiation. Convergence to a monomorphic state is particularly likely if the initial distribution has low variance and is located close to the optimum of stabilizing selection.

Similarly, the genetic architecture (number of loci, distribution of locus effects, recombination) may greatly effect the evolutionary outcome. As shown for weak stabilizing selection ($s = 0.4$), the likelihood of speciation depends in a nonstraightforward way on the number of loci. If disruptive selection is relatively weak ($c = 0.6, 1$), the likelihood of speciation decreases with increasing number of loci. A similar result was found by Gourbiere (2004). For stronger disruptive selection ($c \geq 2$), this is so only for values of a below a certain threshold. Above it, and unless speciation is already certain for two loci, speciation is more frequent if more loci contribute to the trait (Table 3). We offer the following intuitive explanation. If c (and also a) is relatively small, fission of the distribution requires that the clusters are located near the boundary of the phenotypic range because d_{crit} is large (Table 1). For two loci, this can be achieved with weaker assortative mating than for three of four loci, because the average distance among phenotypes is larger, thus mating

probabilities are lower. If c is large (and also a), then clusters do not occur near the boundary of the phenotypic range, because individuals situated in the middle would experience little competition. Therefore, there appears to be an optimal distance between the two prospective species that depends on c. To position the clusters (species) in the right distance, better fine tuning is an advantage, and this occurs more readily with more loci. The same is true if competition and assortment are so strong that more than two species can coexist. Mostly, speciation involves differentiation at two or more loci. If, however, more than two species coexist or if stabilizing selection is very strong, neighboring species may differ only at a single locus.

Especially for strong assortment, the evolutionary outcome may depend heavily on the particular choice of allelic effects, as exemplified in Table 4. Unless assortment is weak, there are always genetic systems (sets of allelic effects) that depart substantially from the average, but no clear pattern could be detected. In addition, the pattern of recombination may have a marked influence (for numerical results in the case of haploid genetics and weak natural selection, see Bürger and Schneider 2006; see Appendix 2 for analytical results).

Compared with the haploid case investigated in Bürger and Schneider (2006), speciation in diploids requires stronger assortment, in part much stronger. The reason is simply that with diploid genetics many more different phenotypes exist, thus discrimination must be more efficient to allow splitting.

The widely used Roughgarden fitness function (16) has some unrealistic features. It has the nongeneric property that, if reproduction is asexual, a continuous equilibrium distribution exists (Gyllenberg and Meszéna 2005). This is not the case for our fitness function (6) (Appendix 3; see also Kopp and Hermisson 2006). For a Gaussian, and even for a dimorphic distribution, the Roughgarden fitness is always U-shaped and achieves its maximum at the boundaries of the phenotypic range provided competition is strong enough to induce disruptive selection. Otherwise, fitness is stabilizing with a unique mode (Fig. 1). Polechová and Barton (2005) criticized the conclusion of Dieckmann and Doebeli (1999) that sympatric speciation is driven by disruptive selection on the following grounds. The character range in the Dieckmann-Doebeli model is limited and, with their parameters, no population can evolve the large variance required to realize the Gaussian equilibrium distribution predicted for the asexual Roughgarden model (Slatkin 1979). Polechová and Barton argued that because of this limitation, the lack of competitors outside the phenotypic range gives an advantage to extreme phenotypes. Hence, strong disruptive selection is perpetuated, and strong assortment can evolve. Thus, effectively, speciation is caused by edge effects.

We showed that this argument is not necessarily valid. With our fitness function, extreme phenotypes always have reduced fitness if stabilizing selection is strong enough (fitness is bimodal). In this case, speciation can still occur, with up to four species coexisting for our parameters, but it becomes more difficult to achieve, that is, extremely strong assortment is needed (Table 5). This is not surprising because the two species cannot be established near the boundaries of the phenotypic range but have to be closer together. As for weaker stabilizing selection, extremely strong competition is not favorable for speciation. For the Roughgarden fitness, the evolutionary outcome is very similar if stabilizing selection is weak. For strong stabilizing selection, much more variation is maintained by the Roughgarden fitness, but the frequency of speciation is still quite similar to that for our choice of fitness, though mostly somewhat higher (Table 6). The fitness function at equilibrium is usually multimodal then and rather complex (results not shown).

As already noted previously for related models (Bürger and Schneider 2006; Schneider and Bürger 2006), an obstacle for the evolution of assortative mating in small steps is posed by the fact that moderately strong assortment depletes (almost) all genetic variation, unless competition is very strong ($c(\rho - 1) > 10s$). Thus, as assortment evolves from very weak to moderate, the population will lose genetic variation (if initially it was variable). Populations with little genetic variation, however, have an elevated likelihood to remain in a state of depauperate variation under strong assortment instead of evolving toward speciation. For a single locus under frequency-dependent disruptive selection and a modifier locus that increases the strength of assortment by a small amount, Matessi et al. (2001) showed that at an intermediate level of assortment a polymorphic evolutionarily stable strategy ceases to exist. Taken together, these observations suggest that the evolution of strong assortment in small steps may face difficulties in this parameter region.

Developing formal criteria to decide whether a given distribution reflects two (or more) species is a difficult task, especially because biologists will often disagree on the extent of the required reproductive isolation. Nevertheless, we used a formal procedure to decide objectively if speciation has occurred; and if yes, into how many species (see Methods). Ideally, one would like to have a simple measure describing the degree of speciation. Traditionally, in two-locus models linkage disequilibrium has been used as such a measure (e.g., Kirkpatrick and Ravigné 2002; Gavrilets 2004). In multilocus models, linkage disequilibrium cannot be quantified by a single measure. Nevertheless, we proposed a relatively simple compound measure of linkage disequilibrium (eq. 17) that provides a useful guide to deciding whether speciation has occurred (eq. 22). It is based on comparing the excess of the total genetic variance over the variance expected under Hardy-Weinberg and linkage equilibrium with the corresponding quantity calculated under the assumption of maximum possible divergence (two spikes at the boundaries of the phenotypic range). It is very reliable unless c is so large that more than two species can be maintained for sufficiently strong assortment. Whether a U-shaped distribution without a gap in the middle (thus not termed as speciation by our criterion) can be interpreted as incipient speciation may be a matter of taste and depends on the potential of assortment to evolve to higher levels.

There is one simple and rather general conclusion from this and related investigations (Bürger and Gimelfarb 2004; Bürger 2005; Bürger and Schneider 2006; Schneider 2006). Frequency-dependent disruptive selection caused by intraspecific competition is a potent force in maintaining genetic variation at multiple loci. This occurs almost whenever $c(\rho - 1) > s$, with the possible exception of moderately strong assortment or very asymmetric selection (the optimum of

stabilizing selection close to one of the boundaries of the phenotypic range). For other fitness functions and assumptions on population regulation, analogous conditions exist (Bürger 2005).

Our results, namely that typically several stable equilibria coexist, most of which are asymmetric, show that the use of the hypergeometric model (e.g., Doebeli 1996; Kondrashov and Kondrashov 1999; Gourbiere 2004), which forces all allele frequencies to be the same, can be misleading and should be avoided in this context (see also Barton and Shpak 2000).

We ignored mutation in this study. Including sufficiently weak forward and backward mutation ($\mu/s < 0.01$ is most likely to be sufficient from what is known about mutation-selection balance) will not change our results qualitatively but only perturb the equilibrium distributions slightly. Different results might be obtained if mutation to new alleles is allowed, so that the genetic architecture can evolve. Such work is in progress.

Populations may respond to frequency-dependent disruptive selection in several ways. One is evolution of assortment leading to speciation, others include the evolution of high genetic variation, sexual dimorphism (Bolnick and Doebeli 2003), or genetic architecture. For instance, dominance relations can rapidly evolve if heterozygotes are at a selective disadvantage (Bürger 1983; Wilson and Turelli 1989; Otto and Bourguet 1999; Van Dooren 1999), or the distribution of allelic effects can evolve (van Doorn and Dieckmann 2004; Kopp and Hermisson 2006; K. A. Schneider, unpubl. ms.). Results in Bürger and Schneider (2006) suggest that another possibility to achieve optimal resource usage might be the evolution of tight linkage. It would be of interest to study competition between such mechanisms, such as simultaneous evolution of genetic architecture and assortment, to find out which is more likely.

In summary, speciation is a feasible evolutionary outcome, and in certain regions of the parameter space it is also very likely. The strength of assortment needed to induce population fission depends in a nonlinear way on the strength of frequency-dependent competition, population growth rate, and the strength of stabilizing selection. The conditions for the evolution of assortment definitely need to be studied in more detail. Even if all these parameters are in the appropriate range, genetic architecture and initial conditions can prevent a particular population from speciating. Among the main achievements of adaptive dynamics are that it has drawn the interest of evolutionary biologists to frequency-dependent selection and provided a relatively simple theoretical framework for determining branching points. For sexual populations, this does not necessarily imply population splitting. However, evolution to such a branching point apparently always leads to some kind of genetic or ecological diversification.

The present study raises questions about the strength of assortative mating in natural populations and about the mechanisms that could achieve the strength required for speciation. If assortative mating occurs through preferences, as assumed here, then individuals must be able to evaluate their potential mates with high accuracy, and they must have (precise) information about their own phenotype. Under very strong competition, speciation can evolve only if mating is restricted to occur between individuals that differ by considerably less than a phenotypic standard deviation. It may depend on the specific mechanism by which assortment is mediated if sufficiently intense assortment is possible at all. Studies of the strength and costs of assortment in natural populations are rare (Jones et al. 2003). Jones et al. obtained (in our notation) the estimates $M = 36$ and $M = 50$ in sea horses. Their estimates of the strength of assortment are difficult to translate precisely to our model but correspond to $a > 2$.

Somewhat more, but still not enough, is known about the ratio c/s. A close analog of $\sqrt{s/c}$ is WIC/TNW, the ratio of within-individual niche width to the total niche width. The values that have been measured are larger than 1/3, corresponding roughly to $c/s \leq 9$ (Bolnick et al. 2003; Bolnick 2006). This is sufficient to induce strong disruptive selection. However, only a few studies have shown that intraspecific competition indeed induces disruptive selection (Swanson et al. 2003; Bolnick 2004a). Even frequency-dependent selection has been established rarely in the context of resource competition (Schluter 2003). Although disruptive selection may occur more frequently in nature than previously thought (Endler 1986; Kingsolver et al. 2001), the mechanisms by which it is generated have not been explored. In addition, only the curvatures have been reported, which is not sufficient to conclude disruptive selection, and many of the measurements are statistically not significantly different from zero. The same, however, applies to the reported estimates of stabilizing selection (Johnson and Barton 2005). Thus, much more empirical information on these issues is needed before the biological relevance of frequency-dependent disruptive selection and its possible evolutionary consequences can be taken for granted.

Acknowledgments

We thank M. Doebeli, J. Hermisson, and S. Otto for many constructive comments on a previous version. This work was supported by grant P16474-N04 of the Austrian Science Foundation (FWF) and by a grant of the Vienna Science and Technology Fund (WWTF). The Program for Evolutionary Dynamics is supported by Jeffrey Epstein.

Literature Cited

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

Barton, N. H., and M. Shpak. 2000. The stability of symmetric solutions to polygenic models. Theor. Popul. Biol. 57:249–263.

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

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

———. 2006. Multi-species outcomes in a common model of sympatric speciation. J. Theor. Biol. 241:734–744.

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

Bolnick, D. I., R. Svanbäck, J. A. Fordyce, L. H. Yang, J. M. Davis, C. D. Hulsey, and M. L. Forrister. 2003. The ecology of individuals: incidence and implications of individual specialization. Am. Nat. 161:1–28.

Bulmer, M. G. 1974. Density-dependent selection and character displacement. Am. Nat. 108:45–58.
———. 1980. The mathematical theory of quantitative genetics. Clarendon Press, Oxford, U. K.
Bürger, R. 1983. Nonlinear analysis of some models for the evolution of dominance. J. Math. Biol. 16:269–280.
———. 2000. The mathematical theory of selection, recombination, and mutation. John Wiley and Sons, Chichester, U. K.
———. 2005. A multilocus analysis of intraspecific competition and stabilizing selection on a quantitative trait. J. Math. Biol. 50:355–396.
Bürger, R., and A. Gimelfarb. 2004. The effects of intraspecific competition and stabilizing selection on a polygenic. Genetics 167:1425–1443.
Bürger, R., and K. A. Schneider. 2006. Intraspecific competitive divergence and convergence under assortative mating. Am. Nat. 167:190–205.
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 Univ. Press, Cambridge, U. K.
Doebeli, M. 1996. A quantitative genetic competition model for sympatric speciation. J. Evol. Biol. 9:893–909.
———. 2005. Adaptive speciation when assortative mating is based on female preference for male marker traits. J. Evol. Biol. 18: 1587–1600.
Doebeli, M., and U. Dieckmann. 2000. Evolutionary branching and sympatric speciation caused by different types of ecological interaction. Am. Nat. 156:S77–S101.
Doebeli, M., U. Dieckmann, J. A. J. Metz, and D. Tautz. 2005. What we have also learned: adaptive speciation is theoretically plausible. Evolution 59:691–695.
Drossel, B., and A. McKane. 2000. Competitive speciation in quantitative genetics models. J. Theor. Biol. 204:467–478.
Endler, J. A. 1986. Natural selection in the wild. Princeton Univ. Press, Princeton, NJ.
Gavrilets, S. 2003. Models of speciation: what have we learned in 40 years? Evolution 57:2197–2215.
———. 2004. Fitness landscapes and the origin of species. Princeton Univ. Press, Princeton, NJ.
———. 2005. ''Adaptive speciation'': it is not that easy. A reply to Doebeli et al. Evolution 59:696–699.
Gavrilets, S., and C. R. B. Boake. 1998. On the evolution of premating isolation after a founder event. Am. Nat. 152:706–716.
Geritz, S. A. H., and E. Kisdi. 2000. Adaptive dynamics in diploid, sexual populations and the evolution of reproductive isolation. Proc. R. Soc. London B 267:1671–1678.
Gourbiere, S. 2004. How do natural and sexual selection contribute to sympatric speciation? J. Evol. Biol. 17:1297–1309.
Gyllenberg, M., and G. Meszéna. 2005. On the impossibility of coexistence of infinitely many strategies. J. Math. Biol. 50: 133–160.
Hassell, M. P. 1975. Density-dependence in single-species populations. J. Animal Ecol. 44:283–296.
Johnson, T., and N. Barton. 2005. Theoretical models of selection and mutation on quantitative traits. Phil. Trans. R. Soc. B, 360: 1411–1425.
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. Proc. Natl. Acad. Sci., USA, 100: 6598–6603.
Kawecki, T. 1997. Sympatric speciation by habitat specialization driven by deleterious mutations. Evolution 51:1751–1763.
———. 2004. Genetic theories of sympatric speciation. Pp. 36–53 *in* U. Dieckmann, M. Doebeli, J. A. J. Metz, and D. Tautz, eds. Adaptive speciation. Cambridge Univ. Press, Cambridge, U.K.
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. Am. Nat. 157:245–261.
Kirkpatrick, M., and S. L. Nuismer. 2004. Sexual selection can constrain sympatric speciation. Proc. R. Soc. Lond. B 271: 687–693.
Kirkpatrick, M., and V. Ravigné. 2002. Speciation by natural and sexual selection: models and experiments. Am. Nat. 159: S22–S35.
Kondrashov, A. S., and F. A. Kondrashov. 1999. Interactions among quantitative traits in the course of sympatric speciation. Nature 400:351–354.
Kopp, M., and J. Hermisson. 2006. The evolution of genetic architecture under frequency-dependent disruptive selection. Evolution 60:1537–1550.
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.
Maynard Smith, J. 1966. Sympatric speciation. Am. Nat. 100: 637–650.
———. 1974. Models in ecology. Cambridge Univ. Press, Cambridge, U. K.
Mayr, E. 1963. Animal species and evolution. Belknap Press, Cambridge, MA.
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.
Meszéna, G., M. Gyllenberg, L. Pásztor, and J. A. J. Metz. 2006. Competitive exclusion and limiting similarity: a unified theory. Theor. Popul. Biol. 69:68–87.
Nagylaki, T., J. Hofbauer, and P. Brunovský. 1999. Convergence of multilocus systems under weak epistasis or weak selection. J. Math. Biol. 38:103–133.
Otto, S. P., and D. Bourguet. 1999. Balanced polymorphism and the evolution of dominance. Am. Nat. 153:561–574.
Polechová, J., and N. H. Barton. 2005. Speciation through competition: a critical review. Evolution 59:1194–1210.
Ricker, W. E. 1954. Stock and recruitment. J. Fisheries. Res. Board Can. 11:559–623.
Roughgarden, J. 1972. Evolution of niche width. Am. Nat. 106: 683–718.
Schluter, D. 2003. Frequency-dependent natural selection during character displacement in sticklebacks. Evolution 57:1142–1150.
Schneider, K. A. 2006. A multilocus-multiallele analysis of frequency-dependent selection induced by intraspecific competition. J. Math. Biol. 52:483–523.
Schneider, K. A., and R. Bürger. 2006. Does competitive divergence occur if assortative mating is costly? J. Evol. Biol. 19:570–588.
Slatkin, M. 1979. Frequency- and density-dependent selection on a quantitative character. Genetics 93:755–771.
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.
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 Ecol. Evol. 16:330–343.
Van Dooren, T. J. M. 1999. The evolutionary ecology of dominance-recessivity. J. Theor. Biol. 198:519–532.
van Doorn, G. S., and U. Dieckmann. 2004. The long-term evolution of multi-locus traits under frequency-dependent disruptive selection. *In* G. S. van Doorn. Sexual selection and sympatric speciation. Ph. D. diss., University of Groningen.
Via, S. 2001. Sympatric speciation in animals: the ugly duckling grows up. Trends Ecol. Evol. 16:381–390.
Waxman, D., and S. Gavrilets. 2005. Twenty questions on adaptive dynamics. J. Evol. Biol. 18:1139–1154.
Willensdorfer, M., and R. Bürger. 2003. The two-locus model of Gaussian stabilizing selection. Theor. Popul. Biol. 64:101–117.
Wilson, D. S., and M. Turelli. 1986. Stable underdominance and the evolutionary invasion of empty niches. Am. Nat. 127: 835–850.

Corresponding Editor: S. Otto

Appendix 1. Properties of the Measure D of Disequilibrium

First, in the case of two loci we relate our measure D of disequilibrium to well known measures of linkage and Hardy-Weinberg disequilibrium. Deviating from the notation in the main text, let the alleles at the first locus be denoted by A and a and at the second by B and b. Let the allelic effects be $\frac{1}{2}\gamma_i$ and $-\frac{1}{2}\gamma_i$ ($i = 1, 2$), and let p_{Ab} denote the frequency of gamete Ab, p_{Aa} the frequency of the ordered one-locus genotype Aa (thus, $p_{AA} + 2p_{Aa} + p_{aa} = 1$), and $p_{AB/ab}$ of the ordered two-locus genotype AB/ab. Further, let $D_g = p_{AB}p_{ab} - p_{Ab}p_{aB}$ denote the ordinary measure of linkage (gametic) disequilibrium, let $H_1 = p_{AA}p_{aa} - p_{Aa}^2$ and $H_2 = p_{BB}p_{bb} - p_{Bb}^2$ be the measures of Hardy-Weinberg disequilibrium at loci 1 and 2, and define the Hardy-Weinberg disequilibrium between gametes AB and ab by $H_{AB,ab} = p_{AB/ab} - p_{AB}p_{ab}$ and, similarly, $H_{Ab,aB} = p_{Ab/aB} - p_{Ab}p_{aB}$. Then, it is straightforward to check with Mathematica (after homogenizing the resulting polynomial) that

$$D = \frac{8}{\gamma_1^2 + 4\gamma_1\gamma_2 + \gamma_2^2} \times \left[2\gamma_1\gamma_2 D_g + \gamma_1\gamma_2(p_{Ab/aB} - p_{AB/ab}) + \frac{1}{2}(\gamma_1^2 H_1 + \gamma_2^2 H_2)\right] \quad \text{and} \tag{A1}$$

$$= \frac{8}{\gamma_1^2 + \gamma_1\gamma_2 + \gamma_2^2} \times \left[\gamma_1\gamma_2 D_g + \gamma_1\gamma_2(H_{Ab,aB} - H_{AB,ab}) + \frac{1}{2}(\gamma_1^2 H_1 + \gamma_2^2 H_2)\right]. \tag{A2}$$

In Hardy-Weinberg equilibrium and for equivalent loci, this simplifies to $D = (4/3)D_g$. If only locus 1 is polymorphic, we obtain

$$D = \frac{4\gamma_1^2}{\gamma_1^2 + 4\gamma_1\gamma_2 + \gamma_2^2}(p_{AA}p_{aa} - p_{Aa}^2). \tag{A3}$$

Next, we derive a simple expression for D if there is an arbitrary number of loci ($n \geq 2$), but only two genotypes are present in the population. This will be useful in quantifying when we can speak of speciation. Let g_1 and g_2 denote the two types, and p_1 and $p_2 = 1 - p_1$ their frequencies. Let L denote the set of all loci and let J be the subset of L in which these two types differ. Because we are interested in deriving a measure for speciation, we can assume that if these two types differ at a locus, then one is homozygous for the $-$ allele and the other homozygous for the $+$ allele. Therefore, defining their distance by $d = 2\,\Sigma_{i\in J}\,\gamma_i$, we can write

$$g_1 = -d/2 + C, \quad \text{and} \tag{A4a}$$

$$g_2 = d/2 + C, \tag{A4b}$$

where C is a constant depending on the effects of the fixed loci. Then we obtain for the mean, the additive genetic variance, and the genic variance $\bar{g} = d(\frac{1}{2} - p_1) + C$, $V_A = d^2 p_1(1 - p_1)$, and $V_{gen} = 2\,\Sigma_{i\in J}\,\gamma_i^2 p_1(1 - p_1)$, respectively. Combining this with equation (18), we obtain

$$D = 4p_1(1 - p_1)\frac{d^2 - 2\sum_{i\in J}\gamma_i^2}{1 - 2\sum_{i\in L}\gamma_i^2}, \tag{A5}$$

where we have used $\Gamma = \frac{1}{2}$. If all loci have equal effects, that is, $\gamma_i = 1/(2n)$, and the two types differ in j loci, then (A5) simplifies to

$$D = 4p_1(1 - p_1)\frac{2j^2 - j}{2n^2 - n} = \frac{d^2 - d/(2n)}{1 - 1/(2n)}. \tag{A6}$$

If $n = 4$ and $j = 1, 2, 3$, we obtain $D = 0.036, 0.143, 0.536$, respectively, provided $p_1 = \frac{1}{2}$.

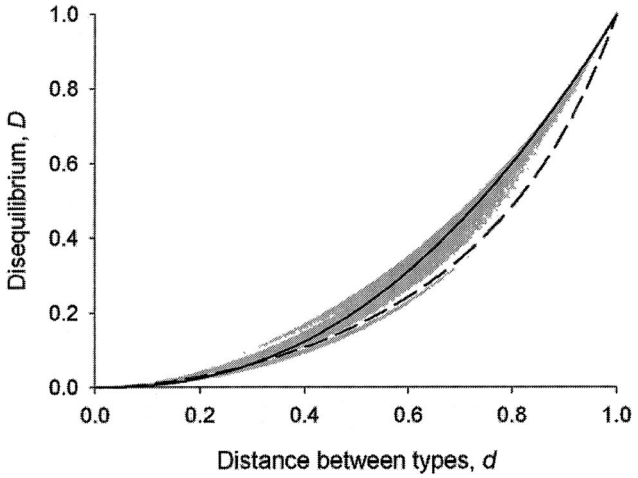

Fig. A1. Relation between d, the distance of two types in a population, each with frequency $\frac{1}{2}$, and the measure D of disequilibrium defined in equation (17). The datapoints (in gray) represent pairs (d, D) from 10,000 randomly chosen genetic systems with four loci (the γ_i are chosen as described in The Numerical Approach). The solid line is given by $d^{2.28}$ and is the best fit of the form d^α to the data (several other functions of different structure yielded a poorer fit). The dashed line gives the exact relation between d and D for the two-locus case (see Appendix 1). If the locus effects are drawn from [0, 1] instead of [0.1, 1], the gray area extends to slightly below the dashed line (such as for low values of d).

If, in a two-locus system, only locus 1 is polymorphic, hence only the two homozygous types AB/AB and aB/aB are present (or, equivalently, Ab/Ab and ab/ab), and if we assume that each occurs with frequency $\frac{1}{2}$, then D becomes

$$D = \frac{d^2}{1 + 2d - 2d^2}. \tag{A7}$$

Thus, in the two-locus case a simple relationship between D and the distance of two types exists. For three or more loci, this is more complicated because D is no longer a function of d alone. It turns out that by fitting a power function d^α to D, excellent approximations can be obtained, with R^2 values above 0.995. For three and four loci we obtain

$$D \approx \begin{cases} d^{2.45} & \text{if } n = 3 \\ d^{2.28} & \text{if } n = 4. \end{cases} \tag{A8}$$

For an illustration, see Figure A1.

Appendix 2. Stability of Monomorphic Equilibria

We assume an assortatively mating diploid population under frequency-dependent selection as described in The Model. However, we include two generalizations: at each locus there may be an arbitrary number of alleles, and population regulation may be more general. Here, density-dependent population growth occurs, in the absence of genetic variation, according to

$$N' = NF(N), \tag{A9}$$

where N and N' are the population sizes in consecutive generations and $F: [0, \infty) \to [0, \infty)$ is a strictly decreasing, twice differentiable function in N (on the interval of admissible values, i.e., all N such that $F(N) = 0$) 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 simple demographic dynamics, that is, convergence to K occurs for all (admissible) initial conditions (for general conditions on F see Thieme 2003, ch.9). Typical examples include (discrete) logistic growth (5), a model generalizing those of Hassell (1975) and Maynard Smith (1974), given by

$$F(N) = \frac{\lambda}{(1 + bN^{\xi})^{\vartheta}}, \tag{A10}$$

and the Ricker (1954) model

$$F(N) = \exp(r - N/K). \tag{A11}$$

Here, λ, ξ, ϑ, r, and K are positive constants. In the latter two models convergence to the carrying capacity is guaranteed if $\vartheta\xi \leq 2$ or $r \leq 2$, respectively. The model (A10) reduces to the one of Hassell (1975) if $\xi = 1$, to one suggested by Maynard Smith (1974) if $\vartheta = 1$, and to the Beverton-Holt model if $\vartheta = \xi = 1$ (see Bürger 2005).

Following Bürger (2005), the absolute fitness of an individual with genotypic value (phenotype) g is assumed to be

$$W(g) = F(N\bar{\alpha}(g))S(g). \tag{A12}$$

We shall use the following notation (Nagylaki et al. 1999). On locus k ($1 \leq k \leq n$) the m_k alleles $A_{i_k}^{(k)}$ can occur ($i_k = 1, \ldots, m_k$). The multi-index $\boldsymbol{i} = (i_1, \ldots, i_n)$ is used as an abbreviation for the gamete $A_{i_1}^{(1)} A_{i_2}^{(2)} \cdots A_{i_n}^{(n)}$. We order the multi-indices $\boldsymbol{i} = (i_1, \ldots, i_n)$ and $\boldsymbol{j} = (j_1, \ldots, j_n)$ such that $\boldsymbol{i} < \boldsymbol{j}$ if and only if $i_k < j_k$ for the largest integer k with $i_k \neq j_k$. We have $\boldsymbol{i} = \boldsymbol{j}$ if and only if $i_k = j_k$ for all k.

We assume no position effect. The frequency of the unordered genotype that carries allele $A_{i_k}^{(k)}$ at locus k on the first chromosome and allele $A_{j_k}^{(k)}$ at locus k on the second is denoted by $p_{\boldsymbol{ij}}$. We posit the convention $\boldsymbol{i} \leqslant \boldsymbol{j}$. During an individual's lifetime, viability selection acts. We denote the fitness of genotype $\boldsymbol{ij}$ by $W_{\boldsymbol{ij}}$, its genotypic value by $g_{\boldsymbol{ij}}$. The genotype frequencies after viability selection are given by

$$p_{\boldsymbol{ij}}^{*} = \frac{p_{\boldsymbol{ij}} W_{\boldsymbol{ij}}}{\bar{W}}, \tag{A13}$$

where

$$\bar{W} = \sum_{\boldsymbol{i} \leqslant \boldsymbol{j}} p_{\boldsymbol{ij}} W_{\boldsymbol{ij}} \tag{A14}$$

is the mean fitness and $\Sigma_{\boldsymbol{i} \leqslant \boldsymbol{j}}$ indicates summation over all pairs $(\boldsymbol{i}, \boldsymbol{j})$ such that $\boldsymbol{i} \leqslant \boldsymbol{j}$. Moreover, $R(\boldsymbol{ij}, \boldsymbol{kl} \to \boldsymbol{uv})$ denotes the probability that an $\boldsymbol{ij}$ individual and a $\boldsymbol{kl}$ individual produce a $\boldsymbol{uv}$ individual. Mating occurs as described in The Model. Therefore, the frequency of $\boldsymbol{st}$ individuals in the next generation is

$$p'_{\boldsymbol{st}} = \frac{\bar{W}^2}{\tilde{W}} \sum_{\boldsymbol{i} \leqslant \boldsymbol{j}} \sum_{\boldsymbol{k} \leqslant \boldsymbol{l}} p_{\boldsymbol{ij}}^{*} p_{\boldsymbol{kl}}^{*} Q_{\boldsymbol{ij},\boldsymbol{kl}}^{*} R(\boldsymbol{ij}, \boldsymbol{kl} \to \boldsymbol{st}) =: \frac{W_{\boldsymbol{st}}^{*}}{\tilde{W}}, \tag{A15}$$

where

$$\tilde{W} = \bar{W}^2 \sum_{\boldsymbol{s} \leqslant \boldsymbol{t}} \sum_{\boldsymbol{i} \leqslant \boldsymbol{j}} \sum_{\boldsymbol{k} \leqslant \boldsymbol{l}} p_{\boldsymbol{ij}}^{*} p_{\boldsymbol{kl}}^{*} Q_{\boldsymbol{ij},\boldsymbol{kl}}^{*} R(\boldsymbol{ij}, \boldsymbol{kl} \to \boldsymbol{st}). \tag{A16}$$

Here,

$$Q_{\boldsymbol{ij},\boldsymbol{kl}}^{*} = \frac{1 - (1 - \bar{\pi}_{\boldsymbol{ij}}^{*})^{M}}{\bar{\pi}_{\boldsymbol{ij}}^{*}} \pi_{\boldsymbol{ij},\boldsymbol{kl}} \tag{A17}$$

with $\pi_{\boldsymbol{ij},\boldsymbol{kl}} = \pi(g_{\boldsymbol{ij}} - g_{\boldsymbol{kl}})$ as in equation (8) and

$$\bar{\pi}_{\boldsymbol{ij}}^{*} = \sum_{\boldsymbol{k} \leqslant \boldsymbol{l}} \pi_{\boldsymbol{ij},\boldsymbol{kl}} p_{\boldsymbol{kl}}^{*}. \tag{A18}$$

The population size evolves according to

$$N' = N\tilde{W}/\bar{W}. \tag{A19}$$

The population size at (demographic) equilibrium is denoted by $\hat{N}$.

Derivation of the Eigenvalues

We want to derive the stability properties of the monomorphic equilibria. Choosing one of them arbitrarily and relabeling the alleles accordingly, we can assume without loss of generality that the corresponding genotype is $\mathbf{11}$, where $\mathbf{1} = (1, \ldots, 1)$. Clearly, we have $\tilde{W}\big|_{p_{\mathbf{11}}=1} = \bar{W}\big|_{p_{\mathbf{11}}=1} = W_{11}\big|_{p_{\mathbf{11}}=1} = 1$ and $W_{\boldsymbol{st}}^{*}\big|_{p_{\mathbf{11}}=1} = 0$ if $\boldsymbol{st} \neq \mathbf{11}$ (we assume $N = \hat{N}$ when evaluating at $p_{\mathbf{11}} = 1$). We express the dynamics by the genotype frequencies $p_{\boldsymbol{ij}}$ for $\boldsymbol{ij} \neq \mathbf{11}$, that is, we eliminate the (redundant) variable $p_{\mathbf{11}} = 1 - \Sigma_{\boldsymbol{i} \leqslant \boldsymbol{j}, \boldsymbol{ij} \neq \mathbf{11}} p_{\boldsymbol{ij}}$. In the following we always assume $\boldsymbol{uv} \neq \mathbf{11}$ and, unless otherwise mentioned, $\boldsymbol{st} \neq \mathbf{11}$. Then, the entries of the Jacobian are

$$\left.\frac{\partial p'_{\boldsymbol{st}}}{\partial p_{\boldsymbol{uv}}}\right|_{p_{\mathbf{11}}=1} = \left.\frac{\tilde{W}\dfrac{\partial W_{\boldsymbol{st}}^{*}}{\partial p_{\boldsymbol{uv}}} - W_{\boldsymbol{st}}^{*}\dfrac{\partial \tilde{W}}{\partial p_{\boldsymbol{uv}}}}{\tilde{W}^2}\right|_{p_{\mathbf{11}}=1} = \left.\frac{\partial W_{\boldsymbol{st}}^{*}}{\partial p_{\boldsymbol{uv}}}\right|_{p_{\mathbf{11}}=1}, \tag{A20a}$$

$$\left.\frac{\partial p'_{\boldsymbol{st}}}{\partial N}\right|_{p_{\mathbf{11}}=1} = \left.\frac{\tilde{W}\dfrac{\partial W_{\boldsymbol{st}}^{*}}{\partial N} - W_{\boldsymbol{st}}^{*}\dfrac{\partial \tilde{W}}{\partial N}}{\tilde{W}^2}\right|_{p_{\mathbf{11}}=1} = \left.\frac{\partial W_{\boldsymbol{st}}^{*}}{\partial N}\right|_{p_{\mathbf{11}}=1}, \tag{A20b}$$

$$\left.\frac{\partial N'}{\partial p_{\boldsymbol{uv}}}\right|_{p_{\mathbf{11}}=1} = \left.\frac{N}{\bar{W}^2}\left(\bar{W}\frac{\partial \tilde{W}}{\partial p_{\boldsymbol{uv}}} - \tilde{W}\frac{\partial \bar{W}}{\partial p_{\boldsymbol{uv}}}\right)\right|_{p_{\mathbf{11}}=1} = N\left(\frac{\partial \tilde{W}}{\partial p_{\boldsymbol{uv}}} - \frac{\partial \bar{W}}{\partial p_{\boldsymbol{uv}}}\right)\Bigg|_{p_{\mathbf{11}}=1}, \quad \text{and} \tag{A20c}$$

$$\left.\frac{\partial N'}{\partial N}\right|_{p_{\mathbf{11}}=1} = \frac{\tilde{W}}{\bar{W}} + \left.\frac{N}{\bar{W}^2}\left(\bar{W}\frac{\partial \tilde{W}}{\partial N} - \tilde{W}\frac{\partial \bar{W}}{\partial N}\right)\right|_{p_{\mathbf{11}}=1} = 1 + N\left(\frac{\partial \tilde{W}}{\partial N} - \frac{\partial \bar{W}}{\partial N}\right)\Bigg|_{p_{\mathbf{11}}=1}. \tag{A20d}$$

First we derive an explicit formula for (A20a). For this purpose we need the following derivatives:

$$\left.\frac{\partial p_{\boldsymbol{ij}}^{*}}{\partial p_{\boldsymbol{uv}}}\right|_{p_{\mathbf{11}}=1} = \left.\frac{\left(\dfrac{\partial p_{\boldsymbol{ij}}}{\partial p_{\boldsymbol{uv}}} W_{\boldsymbol{ij}} - p_{\boldsymbol{ij}}\dfrac{\partial W_{\boldsymbol{ij}}}{\partial p_{\boldsymbol{uv}}}\right)\bar{W} - p_{\boldsymbol{ij}} W_{\boldsymbol{ij}}\dfrac{\partial \bar{W}}{\partial p_{\boldsymbol{uv}}}}{\bar{W}^2}\right|_{p_{\mathbf{11}}=1}. \tag{A21}$$

Thus, for $\boldsymbol{ij} \neq \mathbf{11}$ we have

$$\left.\frac{\partial p_{\boldsymbol{ij}}^{*}}{\partial p_{\boldsymbol{uv}}}\right|_{p_{\mathbf{11}}=1} = \delta_{\boldsymbol{ij},\boldsymbol{uv}} W_{\boldsymbol{ij}}\big|_{p_{\mathbf{11}}=1}, \tag{A22}$$

which implies $(\partial p_{\boldsymbol{ij}}^{*}/\partial p_{\boldsymbol{uv}})\big|_{p_{\mathbf{11}}=1} = 0$ if $\boldsymbol{ij} \neq \boldsymbol{uv}$. Moreover, because $p_{\mathbf{11}}^{*} = 1 - \Sigma_{\boldsymbol{i} \leqslant \boldsymbol{j}, \boldsymbol{ij} \neq \mathbf{11}} p_{\boldsymbol{ij}}^{*}$, we have

$$\left.\frac{\partial p_{\mathbf{11}}^{*}}{\partial p_{\boldsymbol{uv}}}\right|_{p_{\mathbf{11}}=1} = -\left.\frac{\partial p_{\boldsymbol{uv}}^{*}}{\partial p_{\boldsymbol{uv}}}\right|_{p_{\mathbf{11}}=1} = -W_{\boldsymbol{uv}}\big|_{p_{\mathbf{11}}=1}. \tag{A23}$$

We obtain from (A15) by using (A22), (A23), and the fact that $R(\mathbf{11}, \mathbf{11} \to \boldsymbol{st}) = 0$,

$$\begin{aligned}
\left.\frac{\partial W_{\boldsymbol{st}}^{*}}{\partial p_{\boldsymbol{uv}}}\right|_{p_{\mathbf{11}}=1} &= \bar{W}^2 \sum_{\boldsymbol{i} \leqslant \boldsymbol{j}} \sum_{\boldsymbol{k} \leqslant \boldsymbol{l}} \left(\frac{\partial p_{\boldsymbol{ij}}^{*}}{\partial p_{\boldsymbol{uv}}} p_{\boldsymbol{kl}}^{*} + \frac{\partial p_{\boldsymbol{kl}}^{*}}{\partial p_{\boldsymbol{uv}}} p_{\boldsymbol{ij}}^{*}\right) Q_{\boldsymbol{ij},\boldsymbol{kl}}^{*} R(\boldsymbol{ij}, \boldsymbol{kl} \to \boldsymbol{st})\Bigg|_{p_{\mathbf{11}}=1} \\
&\quad + \bar{W}^2 \sum_{\boldsymbol{i} \leqslant \boldsymbol{j}} \sum_{\boldsymbol{k} \leqslant \boldsymbol{l}} p_{\boldsymbol{ij}}^{*} p_{\boldsymbol{kl}}^{*} \frac{\partial Q_{\boldsymbol{ij},\boldsymbol{kl}}^{*}}{\partial p_{\boldsymbol{uv}}} R(\boldsymbol{ij}, \boldsymbol{kl} \to \boldsymbol{st}) + \frac{W_{\boldsymbol{st}}^{*}}{\bar{W}^2}\left.\frac{\partial \bar{W}^2}{\partial p_{\boldsymbol{uv}}}\right|_{p_{\mathbf{11}}=1} \\
&= \sum_{\substack{\boldsymbol{k} \leqslant \boldsymbol{l} \\ \boldsymbol{kl} \neq \mathbf{11}}} \delta_{\boldsymbol{kl},\boldsymbol{uv} W_{\boldsymbol{kl}}} Q_{\mathbf{11},\boldsymbol{kl}}^{*} R(\mathbf{11}, \boldsymbol{kl} \to \boldsymbol{st})\big|_{p_{\mathbf{11}}=1} \\
&\quad + \sum_{\substack{\boldsymbol{i} \leqslant \boldsymbol{j} \\ \boldsymbol{ij} \neq \mathbf{11}}} \delta_{\boldsymbol{ij},\boldsymbol{uv}} W_{\boldsymbol{ij}} Q_{\boldsymbol{ij},\mathbf{11}}^{*} R(\boldsymbol{ij}, \mathbf{11} \to \boldsymbol{st})\big|_{p_{\mathbf{11}}=1}.
\end{aligned} \tag{A24}$$

Therefore, (A20a) can be written as

$$\left.\frac{\partial p'_{st}}{\partial p_{uv}}\right|_{p_{11}=1} = \left.\frac{\partial W^*_{st}}{\partial p_{uv}}\right|_{p_{11}=1}$$

$$= W_{uv}(Q^*_{11,uv} + Q^*_{uv,11})R(\mathbf{11}, uv \to st)\big|_{p_{11}=1}. \quad \text{(A25)}$$

To calculate the right side of (A20b), we first observe

$$\left.\frac{\partial p^*_{st}}{\partial N}\right|_{p_{11}=1} = \left.\frac{p_{st}}{\bar{W}^2}\left(\frac{\partial W_{st}}{\partial N}\bar{W} - \frac{\partial \bar{W}}{\partial N}W_{st}\right)\right|_{p_{11}=1} = 0. \quad \text{(A26)}$$

In the following we have to include the case $st = \mathbf{11}$. Because genotype frequencies after selection sum to one, we also have $\partial p^*_{ij}/\partial N|_{p_{11}=1} = 0$. We obtain, by using $R(\mathbf{11}, \mathbf{11} \to st) = \delta_{\mathbf{11},st}$,

$$\left.\frac{\partial W^*_{st}}{\partial N}\right|_{p_{11}=1}$$

$$= \bar{W}^2 \sum_{i\leqslant j}\sum_{k\leqslant l}\left(\frac{\partial p^*_{ij}}{\partial N}p^*_{kl} + \frac{\partial p^*_{kl}}{\partial N}p^*_{ij}\right)Q^*_{ij,kl}R(ij, kl \to st)\Big|_{p_{11}=1}$$

$$+ \bar{W}^2 \sum_{i\leqslant j}\sum_{k\leqslant l} p^*_{ij}p^*_{kl}\frac{\partial Q^*_{ij,kl}}{\partial N}R(ij, kl \to st)\Big|_{p_{11}=1} + \left.\frac{W^*_{st}}{\bar{W}^2}\frac{\partial \bar{W}^2}{\partial N}\right|_{p_{11}=1}$$

$$= \sum_{i\leqslant j}\frac{\partial p^*_{ij}}{\partial N}(Q^*_{ij,\mathbf{11}} + Q^*_{\mathbf{11},ij})R(ij, \mathbf{11} \to st)\Big|_{p_{11}=1}$$

$$+ \frac{\partial Q^*_{\mathbf{11},\mathbf{11}}}{\partial N}R(\mathbf{11}, \mathbf{11} \to st)\big|_{p_{11}=1} + \left.\frac{W^*_{st}}{\bar{W}^2}\frac{\partial \bar{W}^2}{\partial N}\right|_{p_{11}=1}$$

$$= \delta_{\mathbf{11},st}\left(\left.\frac{\partial Q^*_{\mathbf{11},\mathbf{11}}}{\partial N}\right|_{p_{11}=1} + 2\left.\frac{\partial \bar{W}}{\partial N}\right|_{p_{11}=1}\right). \quad \text{(A27)}$$

Thus, if $st \neq \mathbf{11}$, equation (A20b) simplifies to

$$\left.\frac{\partial p'_{st}}{\partial N}\right|_{p_{11}=1} = 0. \quad \text{(A28)}$$

Finally, we investigate (A20d). (Because of (A28), we will not need (A20c).) We have

$$\left.\frac{\partial \bar{W}}{\partial N}\right|_{p_{11}=1} = \sum_{i\leqslant j} p_{ij}\left.\frac{\partial W_{ij}}{\partial N}\right|_{p_{11}=1} = \left.\frac{\partial W_{\mathbf{11}}}{\partial N}\right|_{p_{11}=1}$$

$$= \left.\frac{\partial F(N\bar{\alpha}(g_{\mathbf{11}}))S(g_{\mathbf{11}})}{\partial N}\right|_{p_{11}=1}$$

$$= S(g_{\mathbf{11}})F'(N\bar{\alpha}(g_{\mathbf{11}}))\left[\bar{\alpha}(g_{\mathbf{11}}) + N\frac{\partial \bar{\alpha}(g_{\mathbf{11}})}{\partial N}\right]\Big|_{p_{11}=1}$$

$$= S(g_{\mathbf{11}})F'(\hat{N}), \quad \text{(A29)}$$

where $\hat{N}$ is the population size at this equilibrium, and

$$\left.\frac{\partial \bar{W}}{\partial N}\right|_{p_{11}=1}$$

$$= \sum_{i\leqslant j}\left.\frac{\partial W^*_{ij}}{\partial N}\right|_{p_{11}=1} = \left.\frac{\partial Q^*_{\mathbf{11},\mathbf{11}}}{\partial N}\right|_{p_{11}=1} + 2\left.\frac{\partial \bar{W}}{\partial N}\right|_{p_{11}=1}$$

$$= \frac{\partial}{\partial N}\left[\frac{1-(1-\bar{\pi}^*_{\mathbf{11}})^M}{\bar{\pi}^*_{\mathbf{11}}}\right]\Big|_{p_{11}=1} + 2\left.\frac{\partial \bar{W}}{\partial N}\right|_{p_{11}=1}$$

$$= \left[\frac{\bar{\pi}^*_{\mathbf{11}}M(1-\bar{\pi}^*_{\mathbf{11}})^{M-1} - 1 + (1-\bar{\pi}^*_{\mathbf{11}})^M}{\bar{\pi}^{*2}_{\mathbf{11}}}\right]\left.\frac{\partial \bar{\pi}^*_{\mathbf{11}}}{\partial N}\right|_{p_{11}=1}$$

$$+ 2\left.\frac{\partial \bar{W}}{\partial N}\right|_{p_{11}=1}$$

$$= -\left.\frac{\partial \bar{\pi}^*_{\mathbf{11}}}{\partial N}\right|_{p_{11}=1} + 2\left.\frac{\partial \bar{W}}{\partial N}\right|_{p_{11}=1} = 2\left.\frac{\partial \bar{W}}{\partial N}\right|_{p_{11}=1}. \quad \text{(A30)}$$

Therefore, equation (A20d) is equivalent to

$$\left.\frac{\partial N'}{\partial N}\right|_{p_{11}=1} = 1 + \hat{N}F'(\hat{N})S(g_{\mathbf{11}}). \quad \text{(A31)}$$

Now we have enough information to argue that the eigenvalues of the Jacobian are its diagonal elements. This is best seen in the following way. We order the genotypes ij and kl so that $ij \leqslant' kl$ if and only if $i < k$, or $j < l$ and $i = k$. We have $ij = kl$ if and only if $i = k$ and $j = l$. We order the Jacobian in this way, and place the partial derivatives with respect to N in the last column, and the derivatives of the population size in the last row. Clearly, (A28) informs us that all entries of the last column of the Jacobian are zero except the last one, $\partial N'/\partial N$, which consequently must be an eigenvalue. By developing the Jacobian with respect to the last column, it is obvious that the remaining matrix contains only the derivatives $\partial p'_{st}/\partial p_{uv}$ (A25). This matrix is indeed an upper triangular matrix because for $uv <' st$ we have $R(\mathbf{11}, uv \to st) = 0$, because the genotypes $\mathbf{11}$ and uv cannot produce an st-individual. Therefore, the eigenvalues are the diagonal elements of the Jacobian, that is, $\lambda N = 1 + \hat{N}F'(\hat{N})S(g_{\mathbf{11}})$, and

$$\lambda_{st} = \left.\frac{\partial p'_{st}}{\partial p_{st}}\right|_{p_{11}=1}$$

$$= W_{st}(Q^*_{\mathbf{11},st} + Q^*_{st,\mathbf{11}})R(\mathbf{11}, st \to st)\big|_{p_{11}=1} \quad \text{if } s \leqslant t. \quad \text{(A32)}$$

Clearly all eigenvalues are nonnegative. Furthermore, if $\mathbf{1} < s$, then $\lambda_{st} = 0$, whereas if $\mathbf{1} = s$, then $\lambda_{st} > 0$.

To obtain the eigenvalues at an equilibrium $p_{uv} = 1$, we just have to replace $\mathbf{1}$ by u in the above equations. The equilibrium population size is then given by

$$\hat{N} = F^{-1}(S^{-1}(g_{uu})) = F^{-1}(e^{s(g_{uu}-\theta)^2}). \quad \text{(A33)}$$

Admissibility of this monomorphic equilibrium requires $\hat{N} > 0$, a condition that is independent of the strength of frequency dependence and of the mating pattern.

Hence, we obtain

$$\hat{N} = \begin{cases} \kappa[\rho - e^{s(g_{uu}-\theta)^2}] & \text{for discrete logistic growth,} \\ \left\{\dfrac{[\lambda e^{-s(g_{uu}-\theta)^2}]^{1/\vartheta} - 1}{b}\right\}^{1/\xi} & \text{for the Hassell and Maynard Smith model,} \\ K\left[1 - \dfrac{s}{r}(g_{uu} - \theta)^2\right] & \text{for the Ricker model.} \end{cases} \quad \text{(A34)}$$

Therefore, the equilibrium $p_{uu} = 1$ is admissible if and only if

$$(g_{uu} - \theta)^2 < \frac{\ln \Lambda}{s}, \quad \text{(A35)}$$

where $\Lambda = \rho$, λ, or e^r for the discrete logistic, the generalized Hassell and Maynard Smith, and the Ricker model, respectively. For our choice of Q, W, S, α, and π, the nonzero eigenvalues are given by (A31), with $\hat{N}$ as in (A34), and by

$$\lambda_{ut} = F(\hat{N}\alpha_{uu,ut})e^{-s(g_{ut}-\theta)^2}$$

$$\times \varphi(a, M, g_u - g_t)R(uu, ut \to ut), \qquad t \neq u, \quad \text{(A36)}$$

which is nonnegative and where $\varphi(a, M, x) = 1 + e^{-ax^2} - (1 - e^{-ax^2})^M$. For the three cases in (A34), we have

$$F(\hat{N}\alpha_{uu,ut})$$

$$= \begin{cases} \rho - e^{-c(g_u-g_t)^2}[\rho - e^{s(g_{uu}-\theta)^2}], \\ \{\lambda^{-1/\vartheta} + e^{-\xi c(g_u-g_t)^2}[e^{-s(g_{uu}-\theta)^{2/\vartheta}} - \lambda^{-1/\vartheta}]\}^{-\vartheta}, \\ \exp\{r + [s(g_{uu} - \theta)^2 - r]e^{-c(g_u-g_t)^2}\}, \end{cases} \quad \text{(A37)}$$

and (A31) becomes

$$\lambda_N = \begin{cases} 2 - \rho e^{-s(g_{uu}-\theta)^2}, \\ 1 - \vartheta\xi\{1 - [\lambda e^{-s(g_{uu}-\theta)^2}]^{-1/\vartheta}\}, \\ 1 - r + s(g_{uu} - \theta)^2. \end{cases} \quad \text{(A38)}$$

We note that $-1 < \lambda_N < 1$ if $1 < \rho < 3$, $\xi\vartheta \le 2$, or $r \le 2$, for the respective choices of population growth models, provided $\hat{N} > 0$. Hence, demographic stability is guaranteed if the equilibrium is admissible, that is, if (A35) holds, and if convergence to the carrying capacity occurs. Because the eigenvalues λ_{ut} are always nonnegative, local asymptotic stability can be inferred if $\lambda_{ut} < 1$ for all $\boldsymbol{t} \neq \boldsymbol{u}$.

Conclusions

We can now draw a few simple but important conclusions from (A36) and (A38).

(i) The eigenvalues are decreasing in modulus as a function of a, because φ is decreasing in a. Thus, large a promotes stability of monomorphic equilibria, and sufficient conditions for their stability can be obtained by setting $a = 0$, so that $\varphi(0, M, x) = 2$.

(ii) If $M < \infty$, then $\lim_{a\to\infty} \varphi(a, M, x) = 0$. Hence, all admissible monomorphic equilibria become stable if a is sufficiently large. This is in general not the case if $M = \infty$ because then $\lim_{a\to\infty} \varphi(a, \infty, x) = 1$.

(iii) Because φ is increasing in M, high costs (small M) favor stability of monomorphic equilibria.

(iv) We always have $R(\boldsymbol{uu}, \boldsymbol{ut} \to \boldsymbol{ut}) \le \frac{1}{2}$ and $R(\boldsymbol{uu}, \boldsymbol{ut} \to \boldsymbol{ut}) = \frac{1}{2}$ if there is no recombination or if $\boldsymbol{u}$ and $\boldsymbol{t}$ differ in only one locus. Therefore, simpler (sufficient) conditions for stability of monomorphic equilibria can be obtained by setting $R(\boldsymbol{uu}, \boldsymbol{ut} \to \boldsymbol{ut}) = \frac{1}{2}$. High recombination promotes the stability of monomorphic equilibria because $R(\boldsymbol{uu}, \boldsymbol{ut} \to \boldsymbol{ut})$ decreases with increasing recombination if $\boldsymbol{u}$ and $\boldsymbol{t}$ differ by more than one locus.

(v) By developing the eigenvalues (A36) into a Taylor series and retaining only terms of order s, c, and a (and no interaction terms), it follows as in Bürger and Schneider (2006, online appendix, p. 4) that for $M \ge 2$ monomorphic equilibria sufficiently close to θ become stable if (approximately)

$$s + \frac{a}{2} > c\Phi, \tag{A39}$$

where $\Phi = \rho - 1$ for discrete logistic growth, $\Phi = \vartheta\xi(1 - \lambda^{-1/\vartheta})$ for the Hassell and Maynard Smith model, and $\Phi = r$ for the Ricker model. In the first case, this gives equation (12). The estimates in (13) and (14) can be generalized analogously. We further note that the (approximate) condition $c(\rho - 1) > s$ for disruptive selection can be generalized to $c\Phi > s$ (Bürger 2005, appendix C).

(vi) The monomorphic state $p_{\boldsymbol{uu}} = 1$ is an admissible equilibrium if $\hat{N} > 0$ and (A35) holds. Therefore, the eigenvalues at an (admissible) monomorphic equilibrium are increasing as a function of c. By letting $c \to \infty$ in (A36), we infer

$$\lambda_{\boldsymbol{ut}} \le \Lambda e^{-s(g_{\boldsymbol{ut}}-\theta)^2}\varphi(a, M, g_{\boldsymbol{u}} - g_{\boldsymbol{t}})R(\boldsymbol{uu}, \boldsymbol{ut} \to \boldsymbol{ut}). \tag{A40}$$

Hence,

$$s > \frac{\ln[\Lambda\varphi(a, M, g_{\boldsymbol{u}} - g_{\boldsymbol{t}})R(\boldsymbol{uu}, \boldsymbol{ut} \to \boldsymbol{ut})]}{(g_{\boldsymbol{ut}} - \theta)^2} \quad \text{for all } \boldsymbol{t} \neq \boldsymbol{u} \tag{A41}$$

is a sufficient condition for asymptotic stability of $p_{\boldsymbol{uu}} = 1$. We note that large a, in particular if $M < \infty$, and high recombination rates (small R) facilitate validity of (A41). By (iv) and because $\varphi(a, M, g_{\boldsymbol{u}} - g_{\boldsymbol{t}}) \le 2$, we obtain (15) as a simple sufficient condition for stability of this monomorphic equilibrium, valid for every c; the left estimate in (15) comes from (A35).

(vii) Simple rearrangement of $\lambda_{\boldsymbol{ut}} \le 1$ (A36) shows that $p_{\boldsymbol{uu}} = 1$ is locally stable if

$$c < \frac{\Omega}{(g_{\boldsymbol{u}} - g_{\boldsymbol{t}})^2} \tag{A42}$$

for all $\boldsymbol{t} \neq \boldsymbol{u}$, where

$$\Omega = \begin{cases} \ln[\rho - e^{s(g_{\boldsymbol{uu}}-\theta)^2}] - \ln\left[\rho - \dfrac{e^{s(g_{\boldsymbol{ut}}-\theta)^2}}{\varphi(a, M, g_{\boldsymbol{u}} - g_{\boldsymbol{t}})R(\boldsymbol{uu}, \boldsymbol{ut} \to \boldsymbol{ut})}\right], \\ \dfrac{1}{\xi}(\ln\{[\lambda e^{-s(g_{\boldsymbol{uu}}-\theta)^2}]^{1/\vartheta} - 1\} \\ \quad - \ln\{[\lambda e^{-s(g_{\boldsymbol{ut}}-\theta)^2}\varphi(a, M, g_{\boldsymbol{u}} - g_{\boldsymbol{t}})R(\boldsymbol{uu}, \boldsymbol{ut} \to \boldsymbol{ut})]^{1/\vartheta} - 1\}), \\ \ln[r - s(g_{\boldsymbol{uu}} - \theta)^2] \\ \quad - \ln\{r - s(g_{\boldsymbol{ut}} - \theta)^2 + \ln[\varphi(a, M, g_{\boldsymbol{u}} - g_{\boldsymbol{t}})R(\boldsymbol{uu}, \boldsymbol{ut} \to \boldsymbol{ut})]\}. \end{cases} \tag{A43}$$

For an admissible equilibrium, the argument of the first logarithm is always positive. If that of the second is ≤ 0, which is the case if and only if (A41) is satisfied, the equilibrium is stable, as may be seen directly from (A36). In this case, we set the right side to infinity. For large a, φ may be very small if $M < \infty$; hence, the right side of (A42) will often be infinity, and stable monomorphic equilibria will exist for arbitrarily large c.

(viii) In the absence of stabilizing selection ($s = 0$) no monomorphic equilibrium is stable for sufficiently large M and c provided $\Lambda \ge 2$.

Appendix 3. Niche Structure for Our Fitness Function

For an asexually reproducing population that evolves subject to Roughgarden's fitness function (16), a continuous equilibrium density exists which is Gaussian if, as usual, $S(g)$ and $\alpha(g)$ are assumed Gaussian (Slatkin 1979). As shown by Gyllenberg and Meszéna (2005), this property is structurally unstable; if either S or α are perturbed by an arbitrarily small amount, no equilibrium density exists. This property also contrasts the common principle of competitive exclusion and limiting similarity (Meszéna et al. 2006). We show that for our fitness function (6), no continuous equilibrium density exists. Thus, it exhibits generic behavior.

Let $\mathcal{P}(g)$ denote the population density of an asexual population. Here, we do not assume that α or S are Gaussian. It is sufficient that they are nonnegative, bounded, and vanish at infinity (and, of course, are measurable). Equilibrium requires that $W(g) = 1$ for all $g \in (-\infty, \infty)$. Observing that $\bar{\alpha}(g) = \alpha * \mathcal{P}(g)$, where $*$ denotes convolution, the equilibrium condition can be written as

$$\alpha * \mathcal{P}(g) = \frac{\kappa}{\hat{N}}[\rho - S(g)^{-1}], \tag{A44}$$

where $\hat{N}$ is determined from the condition $\bar{W} = 1$. The left side is a convolution of two nonnegative functions, one is a density ($\mathcal{P}$), the other is bounded and vanishes at infinity. Therefore, it is also nonnegative, bounded, and vanishes at infinity. The right side, however, tends to $-\infty$ if $|g| \to \infty$ because $S(g) \to 0$ in that case. Hence, an equilibrium density cannot exist. Kopp and Hermisson (2006) described how discrete equilibrium distributions can be calculated and how their stability can be checked. This is based on analyzing a discrete version of (A44), that is, a system of linear equations.