Theoretical Population Biology 79 (2011) 82–96

Contents lists available at ScienceDirect

Theoretical Population Biology

journal homepage: www.elsevier.com/locate/tpb

Effects of genetic architecture on the evolution of assortative mating under frequency-dependent disruptive selection

Agnes Rettelbach [a,b,*], Joachim Hermisson [b], Ulf Dieckmann [a], Michael Kopp [b]

[a] Evolution and Ecology Program, International Institute for Applied Systems and Analysis, Schlossplatz 1, A-2361 Laxenburg, Austria

[b] Mathematics and Biosciences Group, Max F. Perutz Laboratories, University of Vienna, Dr. Bohr-Gasse 3, A-1030 Vienna, Austria

ARTICLE INFO

Article history:
Received 26 May 2010
Available online 28 December 2010

Keywords:
Competitive sympatric speciation
Assortative mating
Frequency-dependent selection
Mutational effect size
Invasion fitness
Costs of choosiness

ABSTRACT

We consider a model of sympatric speciation due to frequency-dependent competition, in which it was previously assumed that the evolving traits have a very simple genetic architecture. In the present study, we numerically analyze the consequences of relaxing this assumption. First, previous models assumed that assortative mating evolves in infinitesimal steps. Here, we show that the range of parameters for which speciation is possible increases when mutational steps are large. Second, it was assumed that the trait under frequency-dependent selection is determined by a single locus with two alleles and additive effects. As a consequence, the resultant intermediate phenotype is always heterozygous and can never breed true. To relax this assumption, here we add a second locus influencing the trait. We find three new possible evolutionary outcomes: evolution of three reproductively isolated species, a monomorphic equilibrium with only the intermediate phenotype, and a randomly mating population with a steep unimodal distribution of phenotypes. Both extensions of the original model thus increase the likelihood of competitive speciation.

1. Introduction

Sympatric speciation has long been a hotly debated topic. The reason is that speciation processes of this kind are both difficult to analyze theoretically and hard to demonstrate empirically. On the empirical side, the few clear cases rely on fortuitous circumstances that rule out alternative scenarios (for reviews, see Coyne and Orr, 2004; Bolnick and Fitzpatrick, 2007). One of the most widely cited examples is the speciation of cichlids in small and isolated crater lake environments (Schliewen et al., 1994). On the theoretical side, models to explain sympatric speciation necessarily contain a multitude of environmental and genetic factors (for reviews, see Via, 2001; Kirkpatrick and Ravigné, 2002; Gavrilets, 2004; Bolnick and Fitzpatrick, 2007). This complexity makes it difficult to understand the observed behavior or to analyze more than a tiny part of the large parameter space.

One widely studied scenario for sympatric speciation is competitive speciation (Rosenzweig, 1978), where intraspecific competition for resources induces frequency-dependent disruptive selection, which in turn favors the evolution of assortative mating. The basic idea, which already goes back to Darwin (1859), is that extreme phenotypes gain increased fitness by specializing on underutilized resources. In such a setting, mating between different extreme types produces intermediate offspring with reduced fitness. Hence, females can increase their offspring's fitness by mating with males of like phenotype. If assortative mating is sufficiently strong, it will result in isolated phenotypic clusters and sympatric speciation.

Dieckmann and Doebeli (1999) used numerical analyses to conclude that sympatric speciation is theoretically plausible and even quite easy. Their model has been criticized for making unrealistic choices for some of the biological parameters, such as a high mutation rate, the absence of costs for being choosy (Matessi et al., 2001) or unnaturally polymorphic initial conditions (Gavrilets, 2005). Responses to these points were provided by Doebeli and Dieckmann (2005). It has also been pointed out that speciation is not the only possible evolutionary response to disruptive selection (Waxman and Gavrilets, 2005; Rueffler et al., 2006). Other possibilities include the evolution of dominance (van Dooren, 1999; Peischl and Bürger, 2008; Peischl and Schneider, 2010) or sexual dimorphism (Bolnick and Doebeli, 2003; van Dooren et al., 2004). During the past decade, many models have been published that come to partly contradictory conclusions (reviewed by Bolnick and Fitzpatrick, 2007). The problem is that most of these conclusions are based on limited numerical analyses and that it is often not clear how the results may be affected by specific assumptions.

* Corresponding author at: Evolution and Ecology Program, International Institute for Applied Systems and Analysis, Schlossplatz 1, A-2361 Laxenburg, Austria.

E-mail address: agnes.rettelbach@univie.ac.at (A. Rettelbach).

0040-5809/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.tpb.2010.12.001

To resolve this problem, several studies have recently developed simplified models, which are analytically tractable and have helped explain some of the controversial results from previous numerical studies (Matessi et al., 2001; de Cara et al., 2008; Pennings et al., 2008; Kopp and Hermisson, 2008; Otto et al., 2008; Ripa, 2009). To achieve analytical tractability, however, the new models needed to rely on simplifying genetic assumptions: Typically, it is assumed that assortative mating evolves in infinitesimal (i.e., infinitely small) steps (making it possible to use fitness-gradient techniques) and that the ecological trait (i.e., the trait under frequency-dependent disruptive selection) is determined by a single locus with two alleles. These assumptions are unlikely to be met in natural situations, but their effects on the model results remain unclear.

In this study, therefore, we take a combined approach. We use a discrete-time version of the analytically well-understood model by Pennings et al. (2008) as a basis, but extend the analysis by targeted numerical simulations to remove the most severe restrictions of the analytical approach. Our focus is on how the genetic architecture of the ecological trait and of female choosiness influence the evolution of reproductive isolation. In particular, we study the evolution of assortative mating in large steps, and we extend the genetic basis of the ecological trait from one to two loci. Below we provide further background on these two extensions.

The analytical models by Pennings et al. (2008) and others (see above) explicitly or implicitly assumed that female choosiness evolves in very small steps. In many cases, this assumption will be innocuous. In particular, if assortative mating is cost-free, the conditions for invasion of large and small choosiness modifiers are identical, as long as they point into the same direction (Pennings et al., 2008; Otto et al., 2008; Durinx and Van Dooren, 2009). But for modifiers with large effect, invasion does not ensure fixation, and the further course of evolution cannot be predicted from invasion analysis alone (Geritz et al., 2002). Peischl (2010) finds that the size of the effect of modifiers for dominance and assortative mating can indeed influence the model outcome. Matessi et al. (2001) showed that assortative mating, if it evolves from random mating in small steps, can "get stuck" at intermediate levels of choosiness resulting in only partial isolation, even if complete isolation is locally stable and could be reached if choosiness is initially high. In other words, the evolution of choosiness in this model may have two stable equilibria: an intermediate equilibrium with partial isolation and a high equilibrium with complete isolation. These results were later confirmed by Pennings et al. (2008), Otto et al. (2008), and Ripa (2009), who showed that the bistability is a consequence of positively frequency-dependent sexual selection. For small mutational step sizes, the intermediate equilibrium forms a barrier against speciation. But could this barrier be overcome by means of a large mutation, which would bring the population into the domain of attraction of the alternative, complete-isolation equilibrium? Limited simulations by Pennings et al. (2008) suggest that such a jump is indeed possible, but the underlying conditions are not well understood. Here, we confirm these earlier results and provide a detailed mechanistic explanation of why and when mutations leading to a large increase in choosiness will go to fixation. More generally, we show that the evolution of assortative mating depends on the genetic architecture of the mating modifier locus, and that complete isolation evolves most easily if it can be reached in a single large mutational step.

The models by Pennings et al. (2008), de Cara et al. (2008), Otto et al. (2008), and Ripa (2009) are based on the simplifying assumption that the ecological trait is determined by a single locus with two alleles and additive effects. Such simplifications, which are made to achieve mathematical tractability, can be problematic if they introduce a bias or if they rule out important outcomes. A potential problem with the one-locus two-allele model is that the intermediate phenotype is always heterozygous. Even if it has the highest fitness, it can reach at most a frequency of 50%, simply because heterozygotes will always have 50% homozygous offspring. Here, we analyze the simplest model that allows the intermediate phenotype to be homozygous. In this model, the ecological trait is controlled by two diallelic loci with equal and additive effects. As we shall show, this extension allows additional outcomes, in particular the evolution of three species and the maintenance of a single monomorphic species with intermediate phenotype.

2. Model and methods

Our model builds on the approach by Pennings et al. (2008) and Kopp and Hermisson (2008), which in turn is based on the so-called Roughgarden model of intraspecific competition (May and MacArthur, 2001; Roughgarden, 1972). As such, it corresponds to the one-allele version of the model in Dieckmann and Doebeli (1999), but with a simplified genetic architecture.

2.1. Ecological assumptions

We consider a sexually reproducing population, whose individuals are diploid and hermaphroditic. For simplicity, we refer to hermaphrodites in female or male roles as females or males. The individuals have two traits of interest: an ecological trait X and a mating trait M, with the latter determining the degree of female choosiness.

The ecological trait X determines specialization on a certain type of resource (for example, prey of different size) and can take values between -1 and 1. It is subject to two sources of natural selection: (1) resource availability, which is measured by a phenotype-specific carrying capacity function, and (2) density- and frequency-dependent competition among individuals with similar phenotypes, which is measured by a phenotype-specific competition function.

For the carrying-capacity function $K(X)$, we assume a generalized Gaussian shape,

$$K(X) = K_0 \exp\left(-\frac{1}{2}X^2 \Big/ \sigma_K^2\right) =: K_0(1-k)^{X^2}, \tag{1}$$

where K_0 is the carrying capacity of the phenotype $X = 0$. We use the parameter $k := 1 - \frac{K(1)}{K_0}$ (where $:=$ is the definition sign) to measure the strength of the selection component resulting from the shape of K. For positive $k < 1$, selection is stabilizing and $K(X)$ is a standard Gaussian with variance $\sigma_K^2 = -1/(2\ln(1-k))$. However, we will also consider negative k. In this case, the intermediate phenotype has the lowest carrying capacity and selection is disruptive.

Individuals experience competition with other individuals. The amount of competition experienced by phenotype X can be expressed by an ecologically effective population size, given by

$$C(X) = \sum_Y \gamma(|X-Y|)N(Y). \tag{2}$$

Here, $N(Y)$ is the number of individuals with phenotype Y, and $\gamma(|X-Y|)$ measures the strength of competition between phenotypes X and Y. In accordance with previous work (e.g., Roughgarden, 1972; Dieckmann and Doebeli, 1999), we assume that γ is a Gaussian function of the phenotypic distance, with variance σ_γ^2,

$$\gamma(|X-Y|) = \exp\left(-\frac{1}{2}(X-Y)^2 \Big/ \sigma_\gamma^2\right) =: (1-c)^{(X-Y)^2}. \tag{3}$$

This competition induces frequency-dependent disruptive selection, and we will use the parameter $c := 1 - \gamma(1)$ to measure the

strength of this frequency dependence. Note that c is also an inverse measure of an individual's "range of competition or "niche width" (Bolnick et al., 2003): large c means that competition is short-ranged and operates only between very similar phenotypes. In consequence, different phenotypes can coexist by occupying different niches, and the number of available niches increases with c. The competition and carrying capacity functions are combined in the phenotype-specific death rate,

$$d(X) = \frac{C(X)}{K(X)}, \tag{4}$$

which summarizes the action of natural selection in our model (Roughgarden, 1972). In addition, individuals are also affected by sexual selection due to assortative mating by females, as described below.

Females may prefer to mate with males whose ecological phenotype is similar to their own. We assume a mating probability $\mu(|X - Y|)$ that depends on the phenotypic distance, and, once again, has a Gaussian shape,

$$\mu(|X - Y|) = \exp\left(-\frac{1}{2}(X - Y)^2 \Big/ \sigma_\mu^2\right) =: (1 - m)^{(X-Y)^2}. \tag{5}$$

In particular, the parameter $m := 1 - \mu(1)$ is the probability that a female rejects a male with a phenotypic distance of 1. Similarly, $m' := 1 - (1 - m)^4$ is the rejection probability for phenotypes with a distance of 2. Both parameters take values between 0 and 1. The value $m = 0$ means that the female is not choosy at all, corresponding to random mating at the population level. Conversely, $m = 1$ means that females mate exclusively with males of their own ecological phenotype. At the population level, this corresponds to complete isolation. Finally, intermediate values of m correspond to partial isolation.

In contrast to k and c, m (and m') are not fixed parameters of the model, but instead serve as the phenotypic values of the mating trait M. Only females are choosy, and the mating genotype is not expressed in males. Based on the mating probabilities, we can assign mating rates $\phi(X)$ to different ecological phenotypes. In particular, the mating rate of phenotype X is the average of the female and male mating rates, $\phi(X) = (\phi_{\text{female}}(X) + \phi_{\text{male}}(X))/2$, where

$$\phi_{\text{female}}(X) = \sum_Y N(Y)\mu(|X - Y|)Q(X)$$

$$\phi_{\text{male}}(X) = \sum_Y N(Y)\mu(|X - Y|)Q(Y). \tag{6}$$

Here, $Q(X)$ is a standardization factor, which can be interpreted as the mating activity of females with phenotype X (Pennings et al., 2008). If females pay no cost for being choosy, $Q(X)$ is defined such that $\phi_{\text{female}}(X) = 1$ (Dieckmann and Doebeli, 1999),

$$Q(X) = \frac{1}{\sum_Y N(Y)\mu(X, Y)}. \tag{7}$$

Even though all females have equal mating rate, this is not true for males ($\phi_{\text{male}}(X) \neq 1$ if females are choosy). More precisely, female choosiness induces sexual selection against males with rare ecological phenotypes (Dieckmann and Doebeli, 1999). In Appendix C, we also consider models in which females experience a cost of choosiness (Kopp and Hermisson, 2008; Doebeli and Dieckmann, 2003, 2005).

From the death rate in Eq. (4) and the mating rates in Eq. (6), we now construct the total invasion fitness $W(X)$ of a phenotype X. As a discrete-time version of the continuous-time fitness function in Pennings et al. (2008), we use a Ricker model (Ricker, 1954)

$$W(X) = \phi(X)\exp(\rho(1 - d(X))), \tag{8}$$

where $\exp(\rho)$ is the intrinsic growth factor. After viability selection but before reproduction, the number of individuals with phenotype X is thus given by

$$\tilde{N}(X) = N(X)\exp(\rho(1 - d(X))), \tag{9}$$

and the number of newborns with phenotype X (i.e., $N(X)$ in the next generation) is

$$B(X) = \sum_{Y,Z} \tilde{N}(Y)\tilde{N}(Z)\mu(Y, Z)\tilde{Q}(Z)R_{YZ\to X}, \tag{10}$$

where $R_{YZ\to X}$ is the probability that a mating between phenotypes Y and Z results in phenotype X. All mated individuals are assumed to produce the same average number of offspring. We will use $\rho = 2\log 2$, for which the resultant domain boundaries for the discrete-time model correspond well to those for the continuous-time model by Pennings et al. (2008) and Kopp and Hermisson (2008) (see Appendix A).

2.2. Genetic assumptions

The discrete-time version of the model by Pennings et al. (2008) will be used as our baseline model and thus as a reference for comparisons. This model assumes a minimal genetic architecture. In particular, the ecological trait X is determined by a single diploid locus with two alleles, and the mating trait M is modeled by sequential rare invasions of mutant alleles with very small effect into an otherwise monomorphic resident population. The main aim of this study is to compare the baseline model to an extended model in which two of its key assumptions are relaxed in several ways.

For the ecologic trait, we compare genetic architectures with one and two diploid loci and with two or more alleles per locus. Our main focus will be on the case of two identical diallelic loci. All alleles are additive and their effects equally spaced (see below). The total phenotype range is always the interval $[-1, 1]$. As a consequence, the phenotype range covered by a single locus scales inversely with the number of loci. In particular, for a single locus with two alleles, $+$ and $-$, there are three ecological genotypic values: -1 (genotype $--$), 0 (genotype $-+$), and 1 (genotype $++$). With two diallelic loci, there are five possible ecological genotypic values: -1, -0.5, 0, 0.5, and 1. We ignore environmental contributions to the phenotype and equate genotypic values with phenotypic values (with the latter usually being referred to simply as phenotypes).

The mating trait M is determined by a single locus. This reflects the assumption that evolution of female choosiness occurs by invasion and potential fixation of rare modifier alleles with additive effect and, therefore, is effectively a single-locus problem. In the extended model, we relax the assumption of infinitesimal mutational steps that is made in the baseline model. Instead, we assume that the M locus has a given finite number of equally spaced alleles chosen so that the resulting phenotypes span the range [0, 1]. For example, three alleles would take the values 0, 0.25, and 0.5. With additive genetics (i.e., no dominance), this leads to five diploid phenotypes with values 0, 0.25,0.5, 0.75, and 1. We will refer to these phenotypes as m-values, which determine the degree of choosiness, and to the alleles as m-alleles. Finally, we assume free recombination between the mating locus and the ecological loci, and also among the ecological loci.

2.3. Additional assumptions for numerical analysis

To address questions about these models that cannot be answered analytically, we numerically iterate the dynamics of

genotype frequencies, using modified versions of Eqs. (1)–(10), in which phenotypes have been replaced by genotypes. This works in two modes. In the deterministic mode, all effects of genetic drift are ignored. If recurrent mutation is included, a fixed proportion of the total population is designated as new mutants in every generation. In the stochastic mode, there is an additional sampling step in each generation, so as to include genetic drift through multinomial sampling from the expected genotype distribution. Since we want to link our results to analytical predictions, which all ignore genetic drift, most of the analysis will use the deterministic mode.

Each generation starts with natural selection. Death rates for all genotypes are calculated according to Eq. (4) and selection occurs according to Eq. (9). The genotype frequencies in the next generation are calculated using Eq. (10) (the function $R_{YZ\to X}$ captures the effects of recombination). Mutation at the mating locus follows a stepwise mutation model with steps in m of equal size (see above). In most simulations with recurrent mutations, we use a step size of 0.25 (four steps, or five alleles). In each generation, a fixed proportion (given by the mutation rate u) of each mating-allele class is shifted one step up or down, for example, from $m = 0$ to $m = 0.25$; if the mutant occurs in a homozygote for the $m = 0$ allele, the (diploid) phenotype then changes to $m = 0.125$.

For the highest level of choosiness, corresponding to complete isolation, we usually use a value slightly less than $m = 1$, such as $m = 0.99$ or $m = 0.999999$, which we indicate by the symbol $m = 1^-$. The reason is that the case $m = 1$ differs from the limit $m \to 1$, leading to an anomaly: due to sexual selection, (male) invaders into an empty phenotype class (e.g., heterozygotes) usually suffer a cost of rarity. For $m = 1$, however, they will always find a mate, since all females of the same class are forced to mate within this class, even if there are infinitely more mating partners outside it (Pennings et al., 2008). Note, however, that too low a value for the maximum m can prevent speciation, if reproductive isolation between neighboring phenotypes is too weak (see Results).

Our criterion for concluding that a simulation has reached complete isolation is the virtual absence of heterozygotes, defined in terms of heterozygote frequencies that drop below 0.01. In the two-locus case, we call an outcome "two species" if the frequencies of the three intermediate phenotypes are all <0.01 and "three species" if only the frequencies of the two heterozygous phenotypes are <0.01. We call an outcome "partial isolation" if the mean $m > 0.1$ and the frequencies of heterozygotes are >0.01, and we call an outcome "random mating" if the mean $m < 0.1$. Finally, we call an outcome "monomorphic" if all phenotypes but one have a frequency <0.01.

Since we are mainly interested in whether assortative mating can evolve from random mating, we usually start our simulations with a population at $m = 0$. To analyze local stability of complete isolation, we also perform simulations with initial choosiness set to $m = 1^-$. The initial allele frequencies at the ecological loci are chosen to be either almost monomorphic or nearly symmetric (with frequencies 0.51 and 0.49, to avoid artifacts caused by exact symmetry in the deterministic simulations). If not stated otherwise, the simulations are run until complete isolation is reached (see above) or for a maximum of 10,000 or 50,000 generations for the one-and two-locus model, respectively. Each plot like Fig. 2 is based on about 2000 simulations to cover the whole parameter range of k and c.

3. Results

We consider the evolution of assortative mating (or female choosiness), determined by the value m of the mating trait, for given ecological conditions and trait architectures. The ecological conditions are specified by the (fixed) parameters k and c for the strength of natural selection and for the frequency dependence of competition, respectively.

To discuss and compare the results for the various genetic architectures, we first introduce some terminology. As described above, the model populations can reach several qualitatively different equilibria or outcomes (e.g., random mating, partial isolation, or complete isolation). We dissect parameter space into regimes: Here, a regime refers to the set of all parameters that allow for the same set of stable equilibria (or evolutionary outcomes). Some regimes are monostable, that is, there is only one possible outcome. Others are bistable or multistable and the outcome depends on the initial conditions. We are particularly interested in the outcomes reached from our standard initial conditions, that is, random mating and a (slightly asymmetric) polymorphism at the ecological loci. We will label these outcomes by an asterisk and refer to the corresponding parameter regions as their domains (e.g., the C* domain is the set of parameters for which evolution from standard initial conditions leads to complete isolation). Note that monostable regimes are, by definition, subsets of the corresponding domains. For example, the P* domain consists of the P* regime and the P*/C regime (see below).

3.1. Baseline model

For later reference, we first describe results obtained for our baseline model, in which the ecological trait is determined by a single locus and choosiness evolves in infinitesimal steps. As in the analogous continuous-time model (Pennings et al., 2008), the equilibrium structure can be determined analytically. The key result is that (in the absence of costs of choosiness) evolution of the mating trait M is determined by the fitness difference between heterozygotes and homozygotes,

$$\begin{aligned}\Delta w &= W_{\text{hom}} - W_{\text{het}} \\ &= \exp(\rho(1-d_{\text{hom}}))\phi_{\text{hom}} - \exp(\rho(1-d_{\text{het}}))\phi_{\text{het}}. \end{aligned} \tag{11}$$

A mutant m allele for increased choosiness can invade a monomorphic resident population if and only if homozygotes are favored, that is if $\Delta w > 0$. Analogously, a mutant that decreases choosiness can invade if and only if $\Delta w < 0$. As pointed out in the Introduction, this criterion does not depend on the absolute effect size of the mutant allele, but only on the direction of change it induces; in our model, the invasion conditions for small and large mutations pointing into the same direction are identical. However, only in the limit of small modifiers does invasion imply fixation of the mutant allele.

The proof of criterion (11) is somewhat technical, but entirely analogous to the continuous-time case analyzed in Pennings et al. (2008), and thus is not repeated here. Intuitively, the criterion arises because females with higher m will more often mate with their own type and produce more homozygous offspring than females with lower m. If homozygotes have higher fitness than heterozygotes, the offspring of a mutant female with increased m will, therefore, be fitter than the offspring of resident females, and the mutation will spread. In consequence, we can use the sign of Δw to determine the direction of selection on m. In particular, random mating is a locally stable equilibrium if $\Delta w < 0$ at $m = m' = 0$, and complete isolation is locally stable if $\Delta w > 0$ at $m = m' = 1^-$ (see also Appendix A). Locally stable equilibria with (monomorphic) intermediate m are characterized by the condition $\Delta w = 0$.

For the baseline model, we obtain six different evolutionary regimes depending on the parameters c and k (see Fig. 1):

Complete isolation (C*). For small k and intermediate c, evolution of assortative mating in small steps always leads to the formation of two reproductively isolated species. From any initial condition, the

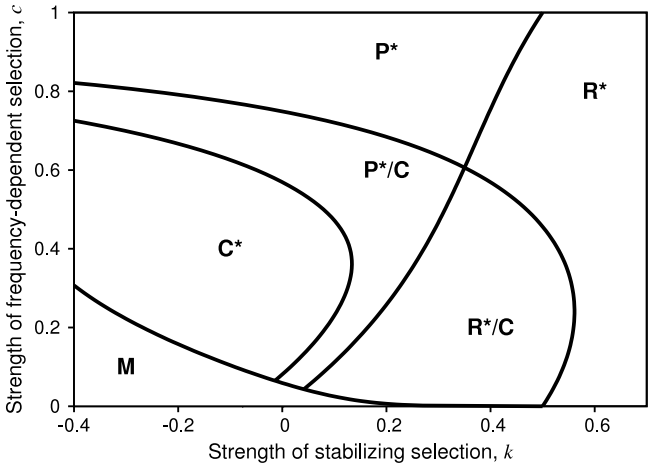

Fig. 1. Evolutionary regimes in the baseline model. The ecological trait is determined by a single locus, and choosiness evolves in infinitesimal steps. Note that negative k implies that selection due to the carrying capacity function is disruptive. C*: complete isolation; R*: random mating; P*: partial isolation; R*/C: random mating or complete isolation; P*/C: partial or complete isolation; M: monomorphic outcome with only one extreme phenotype. The asterisk indicates that the outcome is reached from random mating. The boundary for the local stability of random mating and complete isolation is calculated according to Appendix A. The boundary of the C* regime is calculated numerically using Eq. (11), and the boundary of the M regime is determined by numerical stability analysis as in Pennings et al. (2008).

population evolves towards $m = 1$ and a heterozygote frequency of $p_{\text{het}} = 0$.

Random mating (R*). For sufficiently large k, no amount of assortative mating can evolve. Instead, stabilizing selection is so strong that random mating (with $p_{\text{het}} = 0.5$) evolves from arbitrary initial values of m. Note that we do not allow for disassortative mating.

Partial isolation (P*). If frequency-dependent selection is very strong (large c) and stabilizing selection is at most moderate, a third niche opens up for intermediate phenotypes, which is filled by heterozygotes. As a consequence, the preferred phenotype distribution has an intermediate frequency of heterozygotes, $0 < p_{\text{het}} < 0.5$, which leads to the evolution of an intermediate level of choosiness, $0 < m < 1$, and hence, to partial reproductive isolation.

Random mating or complete isolation (R*/C). For low c and intermediate k, the outcome depends on the initial condition. If female choosiness is initially low, natural selection leads to the evolution of random mating ($m = 0$). However, if initial choosiness is high, the dominating force is sexual selection against rare male heterozygotes, which drives the population towards complete isolation ($m = 1$).

Partial isolation or complete isolation (P*/C). For intermediate c and k, there is another bistable regime, in which the population reaches either partial or complete isolation (Matessi et al., 2001). Starting at random mating, natural selection is disruptive and favors increased female choosiness (because choosy females have more homozygous offspring). Choosiness, however, induces sexual selection, which favors heterozygotes as long as they are common. If choosiness evolves in small steps, a partial-isolation equilibrium is reached where natural and sexual selection are balanced (Pennings et al., 2008). In contrast, if the initial m is already high, heterozygotes are rare, and sexual selection (and, potentially but not necessarily, also natural selection) favors homozygotes, resulting in evolution towards complete isolation. The existence of the P*/C regime shows that the instability of random mating and the stability of complete isolation are not sufficient conditions for complete isolation to evolve in small steps.

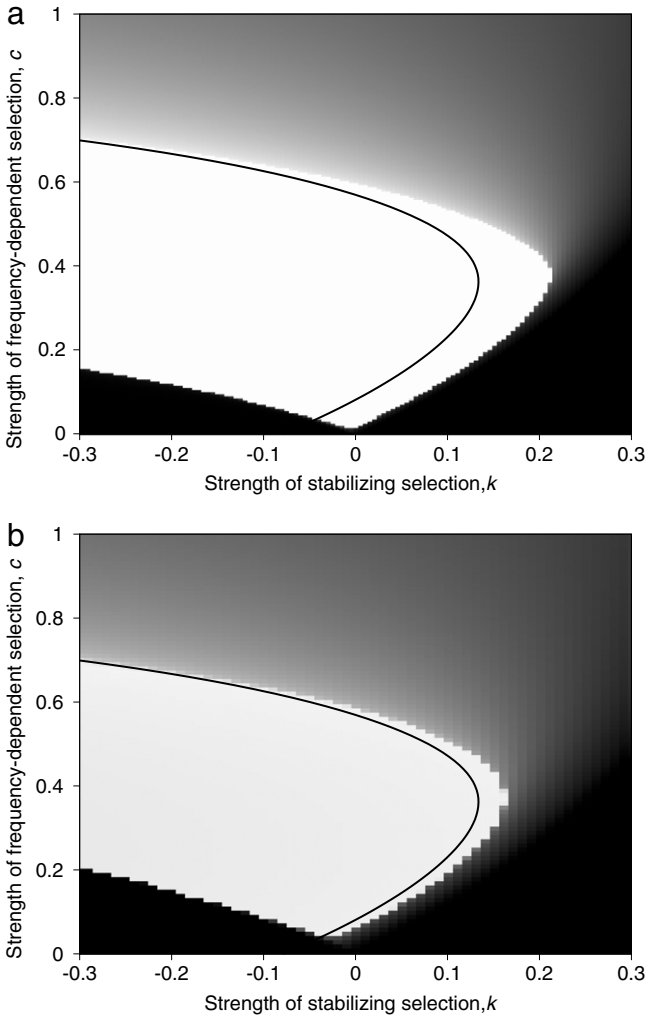

Fig. 2. Evolution of assortative mating in the single-locus model with large mutational steps. (a) Single-step model: An isolation mutant with $m = 1^- = 0.99$ is introduced at frequency $p = 10^{-4}$ into a resident population with $m_r = 0$. (b) Stepwise mutation model with five mating alleles ($m = 0, 0.25, 0.5, 0.75, 0.99$) and a mutation rate of $u = 10^{-6}$ (see text). Shades of gray indicate the mean choosiness $\bar{m}$ in the population at the end of the simulation (white: $\bar{m} = 1^-$, black: $\bar{m} = 0$). The continuous line is the boundary of the C* regime for infinitesimal mutational steps, as in Fig. 1. The white area outside this line shows the additional region where complete isolation is possible via "jumping". Note that the range of k values shown along the horizontal axis is smaller than in Fig. 1.

Monomorphic equilibrium (M). For small c and k, the ecological polymorphism is lost, and the population reaches a stable monomorphic equilibrium with only a single allele. Once such an equilibrium is reached, all selection for assortative mating ceases. The domain of attraction of the monomorphic equilibrium (with respect to the initial frequencies of the ecological alleles) depends on m. For some m, it may be globally stable, meaning that the polymorphic equilibrium is unstable. This is the case if sexual selection and disruptive selection resulting for negative k are stronger than negative frequency-dependent selection due to competition. The M area in Fig. 1 shows the range of ecological parameters for which the ecological polymorphism is lost for at least some values of $m \in [0, 1]$ (see Pennings et al., 2008, for more details). If evolution of choosiness leads to such an m-value, the monomorphic equilibrium may be the stable outcome from random mating. This is the case for sufficiently small k. However, for larger k, the population may also stay at random mating or partial isolation without losing the polymorphism. (For this reason, we do not write M*; note also that, as we do not resolve the various possibilities, the M area does not fully fit our definition of a regime.)

3.2. Genetic architecture of the mating trait

We now ask to what extent the results of the baseline model are robust with respect to the genetic architecture of the mating trait. In particular, we are interested in whether large mutations can help the population "jump" over the intermediate equilibrium in the P*/C regime.

We first study when a randomly mating resident population can reach complete isolation in a single mutational step. To this end, we introduce a mutant "isolation allele" (with $m = 1^-$) at frequency $p = 10^{-4}$ into a resident population with mating genotype $m_r = 0$. All further evolution is deterministic, and there is no recurrent mutation. As in the baseline model, the ecological trait is determined by a single diallelic locus.

The results of these simulations are shown in Fig. 2(a). In accordance with predictions from invasion-fitness analysis, the isolation allele invades in the complete isolation (C*), partial isolation (P*), and bistable partial/complete isolation (P*/C) regimes. However, it rises to fixation in only part of this parameter range. In the P* regime and in parts of the P*/C regime, the outcome is a stable polymorphism between the two m-alleles, leading to an intermediate m in the population average. In contrast, in the C* regime and in part of the P*/C regime close to the C* regime, the isolation allele reaches fixation. We thus see that "jumping" across the stable intermediate equilibrium is indeed possible in a part of the P*/C regime. As a consequence, the total parameter range in which complete isolation is reached from random mating (C* domain) is somewhat extended relative to the case with infinitesimal mutational steps.

Second, we consider the fate of an initially rare isolation allele ($m = 1^-$) in a resident population with a non-zero initial level of choosiness ($m_r > 0$). In a series of simulations (conducted as described for $m_r = 0$ above), we find that both invasion and fixation of the isolation allele becomes more difficult when the resident population has already evolved an intermediate level of choosiness. For example, when starting at $m_r = 0.1$ instead of $m_r = 0$, there are some parameter combinations for which the $m = 1^-$ allele no longer goes to fixation. We consistently find that larger jumps to complete isolation are easier. That is, if the jump $m_1 \to 1^-$ is possible then this implies that $m_2 \to 1^-$ is possible for $m_2 < m_1$, but not vice versa. In particular, we find that fixation of the isolation allele is no longer possible in the P*/C regime if the resident population is already at (or sufficiently near) the stable equilibrium point with intermediate m_r. Thus, evolution of assortative mating can indeed get stuck in this regime.

In a third series of simulations, we consider the evolution of assortative mating in multiple steps of finite size. In contrast to the single-step case, we now assume recurrent mutation at the mating locus and a stepwise mutation model (see model description). An example with five mating alleles (step size 0.25) is shown in Fig. 2(b). We see that isolation can still evolve in parts of the P*/C regime, but also that the total parameter range is smaller than for a single jump from random mating to complete isolation. We obtain similar results with different variations of the genetic architecture (not shown). Evolution of complete isolation is consistently easier with larger steps.

So far, no costs have been assigned to the mating alleles. As discussed in Kopp and Hermisson (2008), however, female choosiness can easily lead to different types of costs. In a fourth step of our analysis, we therefore consider the evolution of reproductive isolation under two types of costs (Kopp and Hermisson, 2008; see also Doebeli and Dieckmann, 2005): (i) mating costs, resulting from females having only a finite number of mating trials (i.e., choosy females run a risk of remaining unmated); and (ii) viability costs, resulting from a negative direct physiological effect of increased m-values. Results from simulations with both types of costs are reported in Appendix C. They generally confirm our basic conclusion that speciation is facilitated by large mutations at the mating locus. Two additional findings are noteworthy. First, mating costs can promote speciation via "jumping", provided they are primarily paid by heterozygous females (which is the case for large c, see Discussion). In some cases, this has the paradoxical effect that speciation is possible with costs but not without. Second, for some types of viability costs, speciation is impossible in the limit of infinitesimal steps (see Kopp and Hermisson, 2008), but occurs in a sizeable parameter range if mutational steps are large.

Finally, we repeat our previous analysis with a sampling step after each generation to simulate the effects of genetic drift. We find that drift introduces some stochasticity, but does not lead to significant shifts in the regime boundaries (not shown). For small population sizes, newly introduced mutants are often lost, but the qualitative conclusions from the large-population limit still hold. A population size of 5000 nearly recovers the deterministic case.

Summarizing these observations, we consistently find that evolution of complete isolation from random mating is most difficult for very small step sizes and easiest for a single jump from $m_r = 0$ to $m = 1^-$. "Realistic" genetic architectures with various intermediate step sizes show an intermediate behavior.

3.3. Genetic architecture of the ecological trait

We now turn to the genetic architecture of the ecological trait and its influence on the evolution of reproductive isolation. Our main focus is on the case of two diallelic loci with equal effect. In addition, we consider a model with a single ecological locus and multiple alleles. In all cases, evolution at the mating locus is modeled using a stepwise mutation model with five equidistant alleles (implying a step size of 0.25 in homozygotes) and a mutation rate of $u = 10^{-5}$. The ecological loci are initiated in a fully polymorphic state (but with slightly asymmetric allele frequencies of 0.49 and 0.51, respectively). Note that, for $m = 0$, full polymorphism is always the only stable equilibrium if $c > 0$ and $k > 0$. While we cannot exclude the existence of equilibria that cannot be reached from these initial conditions in allele frequency, our extensive numerical explorations have not revealed any such equilibria. Furthermore, potential additional equilibria do not play a role for our main question, that is, under which conditions complete isolation can evolve from random mating.

For two identical diallelic loci, the deterministic simulations exhibit a total of six qualitatively different evolutionary equilibria (as opposed to merely four for the single-locus model). The new outcomes are an intermediate monomorphic equilibrium (in addition to the extreme monomorphic equilibrium already present in the one-locus model) and a three-species equilibrium (in addition to the two-species equilibrium). As both the two- and the three-species equilibrium are locally stable over a large parameter range, the system is often bi- or even tristable. Indeed, Fig. 3 reveals no less than thirteen evolutionary regimes with qualitatively different equilibrium structures (i.e., sets of locally stable equilibria), as opposed to merely six for the single-locus model. Instead of enumerating all of these different possibilities, we will discuss them in terms of four broad categories.

Regimes with stable monomorphic equilibria. Unlike the one-locus model, the two-locus model has two different monomorphic equilibria. In addition to the familiar equilibrium with one extreme phenotype (ME, $x = \pm 1$), there is a second equilibrium in which only the intermediate phenotype is present (MI, $x = 0$) and which does not exist in the one-locus case. Fig. 3 shows that the two monomorphic equilibria are reached from random mating in different parameter regions separated by the line $k = 0$. That is, an extreme monomorphic equilibrium is only reached if selection originating from the resource distribution is disruptive ($k < 0$),

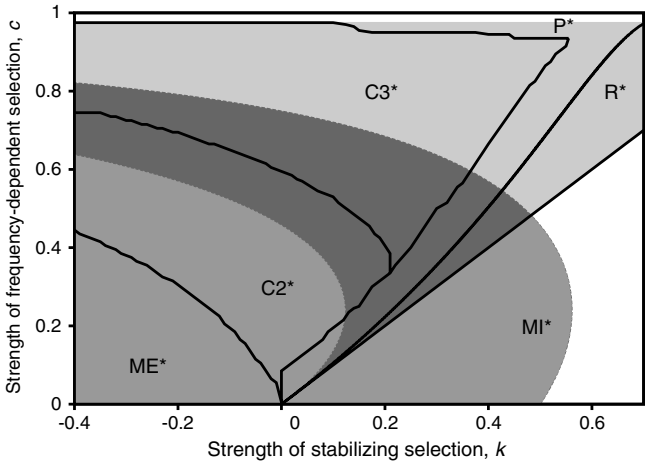

Fig. 3. Evolutionary regimes for the model with two diallelic ecological loci of equal effect. Continuous lines and labels with asterisks indicate domains in which different evolutionary equilibria are reached when simulations are started from random mating ($m = 0$). C2*: two species; C3*: three species; P*: partial isolation; R*: random mating; ME*: monomorphic equilibrium with extreme phenotype; MI*: monomorphic equilibrium with intermediate phenotype. Shades of gray indicate the local stability of complete-isolation equilibria with either two or three species, as determined from simulations started from fully assortative mating ($m = 1^- = 0.999999$). In the light gray area, only the three-species equilibrium is locally stable. In the medium gray area, only the two-species equilibrium is locally stable. In the dark gray area, both the two-and the three-species equilibria are locally stable. Simulations started with three species in the area in which only two species are locally stable reach the intermediate monomorphic equilibrium in the part overlapping with the MI* domain and the two-species equilibrium otherwise. The boundaries of the medium gray and dark gray areas are derived analytically as described in Appendix B. In particular, the upper boundary (local stability of two species) is identical to the corresponding line in Fig. 1 and is, indeed, independent of the genetic architecture of the ecological trait.

while the internal monomorphic equilibrium is only reached if this selection component is stabilizing.

In Appendix B, we show that the intermediate monomorphic equilibrium is locally stable if

$$\exp\left(\rho\left(1-\sqrt[4]{\frac{1-c}{1-k}}\right)\right) < \frac{2}{1+\sqrt[4]{1-m}}. \tag{12}$$

Here, local stability refers to the invasion of additional ecological alleles at a given value of m. For $m = 0$, this condition reduces to $k > c$. Indeed, for this parameter range, the intermediate equilibrium is always reached in our simulations, independent of initial conditions, except in the area in which complete isolation (with two species) is locally stable due to sexual selection (Fig. 3). Eq. (12) shows that, for $m > 0$, the intermediate monomorphic equilibrium can be locally stable also for $c > k$ (e.g., for $m = 1$ and $\rho = 2\log 2$, one obtains $c < (15 + k)/16$). However, over this parameter range, this equilibrium is never reached in simulations when starting at random mating.

Appendix B also gives the conditions for local stability of the extreme monomorphic equilibrium. In contrast to the intermediate equilibrium, the extreme equilibrium is not always reached over the whole parameter range in which it is locally stable for $m = 0$. This is because the outcome also depends on the initial conditions at the ecological loci. In Fig. 3, the upper boundary of the MI* domain is shown for fully polymorphic initial conditions (ecological allele frequencies close to 0.5).

Regimes with stable random-mating equilibrium. In the two-locus model, random mating ($m \approx 0$) is locally stable in a narrow band close to the MI* domain, where selection still favors an excess of intermediate phenotypes (R* domain in Fig. 3). An analytical approximation for the upper boundary of this band is derived in Appendix B. It is based on comparing the fitness of extreme and intermediate phenotypes ($x = \pm 1$ vs. $x = 0$) and follows the same logic as Eq. (11). Simulations show that this approximation is very accurate (not shown). Note that most of the parameter range with a stable random-mating equilibrium belongs to bi-or even tristable regimes, for which also complete isolation with either two or three species (or both) is locally stable.

The random-mating equilibrium in the two-locus model differs from that in the single-locus model in an important way: to maintain symmetry, the allele frequencies at a single ecological locus must always be equal to 0.5. With two ecological loci, however, they are generally shifted away from 0.5 in opposite directions at the two loci, that is, $0.5 \pm \alpha$ with $0 < \alpha < 0.5$. This leads to a symmetric phenotype distribution with a sharp peak at $x = 0$ and an excess of intermediate phenotypes relative to the case with $\alpha = 0$. For $\alpha \to 0.5$, the random-mating equilibrium turns into the MI equilibrium. Indeed, we find that, in the R* domain of Fig. 3, α increases gradually from 0 at the P*–R* boundary to 0.5 at the R*–MI* boundary (Fig. 5).

Regimes with stable partial-isolation equilibrium. There are two connected parameter ranges for which partial-isolation ($0 < m < 1$) is a stable equilibrium of the two-locus model (P* domain in Fig. 3). First, for very high c, competition is sufficiently short-ranged to create (at least) five distinct ecological niches. As a consequence, all five possible phenotypes coexist, with their relative frequencies being determined by the degree of choosiness. In part of this range, complete isolation with three species is also locally stable, owing to sexual selection (see below). Second, in a narrow band between the R* and the C* domains, partial isolation is stabilized by a balance between natural and sexual selection (as in the P*/C regime of the single-locus model). In this area, natural selection favors complete isolation with either two or three species, but these equilibria can only be reached from a high initial value of m.

Regimes with stable complete-isolation equilibrium. In the two-locus model, evolution of complete isolation ($m \approx 1$) can lead to the formation of either two or three species. As shown in Fig. 3, the ranges of local stability of these equilibria are overlapping, and they extend into the R* and the MI* domains. However, our main interest here is in the parameter range over which complete isolation can evolve from random mating in small steps. We denote the domain where this leads to two species by C2* and the domain where it leads to three species by C3* (Fig. 3). The C2* domain corresponds to the C* regime of the single-locus model (Fig. 4). Here, the two extreme phenotypes $x = -1$ and $x = 1$ form two reproductively isolated species, and the three intermediate phenotypes go extinct. The speciation process takes somewhat longer than for a single ecological locus (with the difference being on the order of a few hundred generations, relative to a total time to speciation of about 1000–3000 generations). In contrast, the C3* domain has no direct correspondence in the single-locus case. With two loci, three species can evolve due to symmetry breaking at the individual loci. One of the two $+-$ haplotypes goes extinct, such that the $x = 0$ phenotype consists only of one haplotype and does not produce heterozygous offspring. As a consequence, the allele frequencies at the ecological loci necessarily deviate from 0.5 (the rarer allele at each locus has the same frequency as one of the extreme phenotypes). Since the deviation at both loci occurs in opposite directions, symmetry at the phenotypic level is maintained.

The formation of three species requires a very high degree of female choosiness (usually $m > 0.999$, as opposed to $m > 0.95$ for the two-species regime). The reason is that the three species are phenotypically closer than the two species, so that female choosiness must be highly effective to prevent hybridization. If the maximal possible m is too small, symmetry breaking will not occur and evolution will stop at a state of partial isolation, for which the frequency of heterozygotes is still relatively high (between 0.02

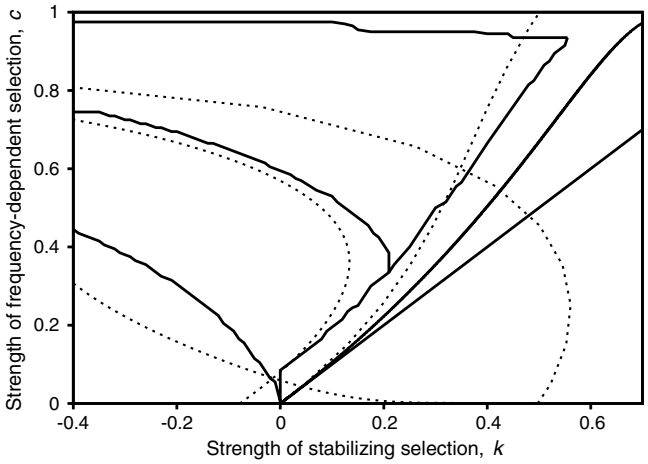

Fig. 4. Comparison of the one-and two-locus models. The continuous lines are the domain boundaries for the two-locus model (delineating the outcomes from random mating), as shown in Fig. 3. The dashed lines are the domain boundaries for the one-locus model, as shown in Fig. 1.

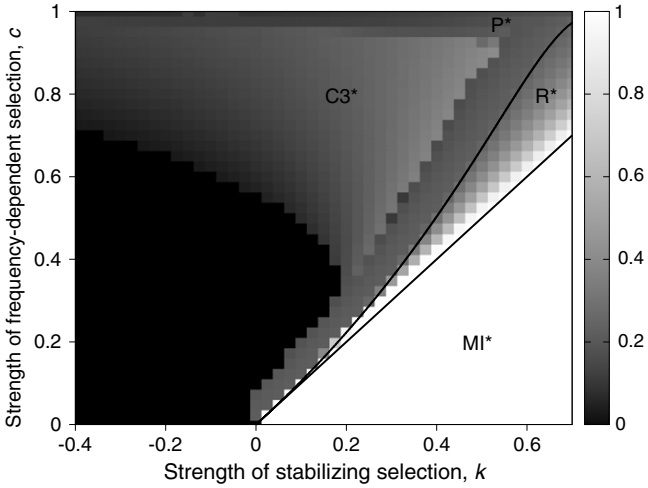

Fig. 5. Frequency of the intermediate phenotype in the two-locus model, when simulations are started from random mating. Shades of gray indicate the frequency of the phenotype $x = 0$, ranging from 0 (black) to 1 (white). The upper continuous line shows the border between the R* and P* domains, while the lower continuous line shows the border between the R* and the MI* domains. Across the R* domain, the frequency of the intermediate phenotype increases from 0.5 (at the boundary to the P* domain) to 1 (at the boundary to the MI* domain). This is achieved by an increasing asymmetry of the allele frequencies at the ecological loci, $0.5 \pm \alpha$, with α increasing from 0 to 0.5.

and 0.05). This is true even in the presence of genetic drift (not shown).

Dynamics underlying the formation of three species. The formation of three species is illustrated in Fig. 6. Notably, this process involves a phase transition, i.e., a fast change after a period of apparent stasis. The typical sequence of evolutionary events is as follows. First, the mean choosiness rapidly evolves to a relatively high value (Fig. 6(a)), for which the population attains a state of partial isolation, typically with a low frequency of the intermediate phenotype (Fig. 6(b)). In all cases observed, there is a polymorphism involving the highest and the lowest available mating alleles (for example, in Fig. 6(a), the high-frequency phenotypes are the $m = 1$ homozygote and the $m = 0/1$ heterozygote). At this stage, the allele frequencies at the ecological loci are still symmetric (Fig. 6(c)). In the deterministic simulations, this transient state can last for many (up to about 4000) generations without conspicuous changes. Then, the transition to speciation is initiated by symmetry breaking at

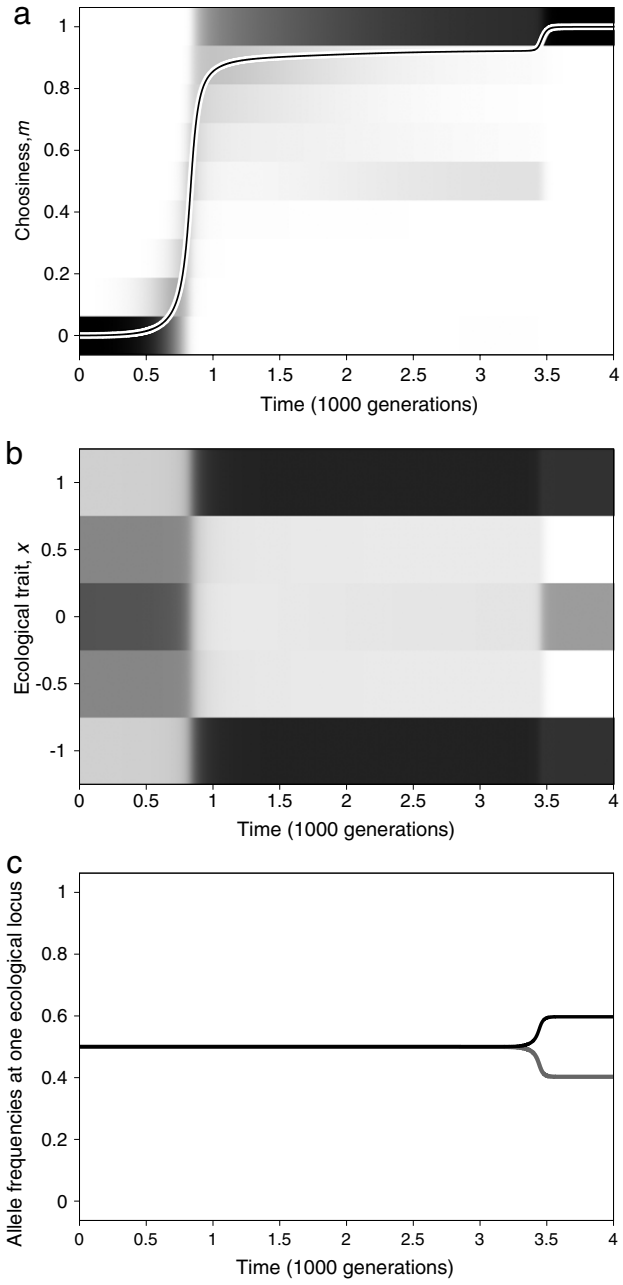

Fig. 6. Evolution of three species, in a stepwise-mutation model with five mating alleles. (a) Evolution of mating phenotypes and mean choosiness. Shades of gray depict the frequencies of the nine mating phenotypes (with white indicating a frequency of 0 and with black indicating a frequency of 1). The continuous line shows the mean choosiness m. (b) Evolution of ecological phenotypes. Shades of gray depict the frequencies of the five ecological phenotypes (with white indicating a frequency of 0 and with black indicating a frequency of 0.5). (c) Evolution of the frequencies of the two alleles at one ecological locus. Parameters: $k = 0.05$, $c = 0.6$, and $u = 10^{-5}$.

the ecological loci (compare Fig. 6(c) and (b)). Only after one of the mixed haplotypes ($+-$ or $-+$) has gained dominance, does the mean m quickly increase up to (almost) 1, the allele frequencies at the ecological loci reach their final values, and speciation occurs. These two-stage dynamics, with an extended transient phase, do not depend on choosing symmetric initial conditions. Even if the initial allele frequencies at the ecological loci deviate strongly from 0.5, symmetry is rapidly restored within the first few generations, before it is broken again at the transition point. When genetic drift is included, the transient phase tends to be shortened, but, for pa-

rameters close to the P* domain, the population may also get stuck at partial isolation (Appendix D).

Comparison with the single-locus model. Fig. 4 shows how the introduction of a second ecological locus changes the regime boundaries relative to the single-locus case. Four main observations are worth highlighting:

1. Stability of random mating ($m = 0$). The boundary for the local stability of random mating (upper boundary of the R* domain) is only slightly tilted by the introduction of a second ecological locus. The main difference is that, in the two-locus model, a large part of the former R* domain is taken up by the MI* domain.
2. Stability of partial isolation (intermediate m). The parameter region featuring a stable partial-isolation equilibrium (P* domain) is strongly restricted in the two-locus model, mainly due to the new three-species equilibrium.
3. Stability of complete isolation ($m = 1$). Similarly, the parameter region implying local stability of complete isolation (gray areas in Fig. 3) extends to much higher values of c in the two-locus model than in the single-locus model. This is entirely due to the new three-species equilibrium. Indeed, local stability of the two-species equilibrium is independent of the genetic architecture (Appendix B).
4. Two-species equilibrium. Finally, it is remarkable that the upper boundary of the parameter regions where two species can evolve from random mating are almost identical in the two models (C* domain with one locus, C2* domain with two loci; compare the middle continuous and dotted lines in the left-hand part of Fig. 4). This is because the three-species equilibrium in the two-locus model exists only in the parameter range of the former P* domain, but not of the former C* domain.

One locus with multiple alleles. To complement the diallelic two-locus model, we consider a single-locus model with multiple alleles. For the cases we tested, no new phenomena occurred. With three or five alleles, the intermediate alleles go extinct in the C* domain and in part of the P* domain, recovering outcomes already observed for the diallelic single-locus model. Thus, evolution of three species does not occur. In the MI* domain of the two-locus model, only the intermediate allele remains in the population.

Genetic drift. Finally, including genetic drift does not significantly change the results of deterministic simulations (Appendix D).

4. Discussion

We have investigated how the evolution of reproductive isolation in a model of competitive speciation depends on the genetic architectures of the ecological trait and of the mating trait that determines female choosiness. Our main results are that speciation is easiest if choosiness can evolve through a single large mutational step and that additional ecological loci enlarge the number of evolutionary outcomes. Compared to models with one ecological locus and infinitesimal mutations for the mating trait (Pennings et al., 2008; Kopp and Hermisson, 2008; Otto et al., 2008; de Cara et al., 2008; Ripa, 2009), both effects increase the range of parameters for which speciation is possible.

4.1. Genetic architecture of the mating trait

For the single-locus model, we have shown that, in part of the bistable P*/C regime, complete reproductive isolation can evolve through a single large mutational step, but not through a series of small steps. In this regime, there are two stable equilibria, partial and complete isolation, and evolution of assortative mating from random mating in small steps always leads to the partial-isolation equilibrium. In contrast, a large mutation can lead to the complete-isolation equilibrium, by "jumping" over the partial-isolation equilibrium, and thus to speciation.

Role of sexual selection. To understand this result, it helps to first consider the case of mating evolution through small steps, assuming that all individuals carry the same m-allele and therefore express the same degree of choosiness. Assume that the population is at the intermediate equilibrium. As detailed in the Results section, the intermediate equilibrium is maintained by a balance between natural and sexual selection (Pennings et al., 2008; Otto et al., 2008; Ripa, 2009). In particular, sexual selection favors heterozygotes (and thus a decrease in choosiness), because they constitute the largest phenotypic group.

Let us now imagine a population with the same genotype distribution at the ecological locus and the same mean m, but consisting of individuals with $m = 0$ alleles and $m = 1^-$ alleles. We assume that the ecological locus and the mating locus initially are at linkage equilibrium. The important point is that this situation cannot be stable. Since offspring with $m = 1^-$ are more likely to be homozygous than offspring with $m = 0$, the $m = 1^-$ mutants will accumulate disproportionally in the homozygotes. The resulting linkage disequilibrium between the polymorphic m-allele and the diploid ecological genotype increases sexual selection for homozygotes (and decreases sexual selection for heterozygotes) relative to the scenario with monomorphic m. Put simply, if the linkage disequilibrium is strong enough, sexual selection favors homozygote males, because only part of the female population is choosy, but most of the choosy females are homozygotes. Furthermore, the level of linkage disequilibrium is proportional to the allelic step size at the mating locus. This explains why a single step from $m = 0$ to $m = 1^-$ is most conducive to speciation.

In most of the P*/C regime, the $m = 1^-$ allele goes to fixation when introduced into a population with $m = 0$. In a small parameter range with large c or k, however, fixation does not occur, and the two alleles are maintained at a stable polymorphism. In this parameter range, as the number of homozygotes increases, natural selection against them becomes strong enough to offset the effect of sexual selection in their favor. This is also the range in which, when mutational steps are small, the basin of attraction of the complete-isolation equilibrium is very small (Kopp and Hermisson, 2008).

Under some conditions, speciation via "jumping" may be further facilitated if females have only a limited number of mating trials (i.e., choosiness incurs mating costs). The reason is that the risk of remaining unmated is largest for choosy females with a rare ecological phenotype. If frequency-dependent competition is strong (large c), most individuals are homozygotes, and mating costs are primarily paid by heterozygous females. This additional source of sexual selection can combine with sexual selection against heterozygous males, promoting fixation of the high-m allele.

Large mutations also enable speciation if choosiness incurs absolute viability costs (Appendix C; Fig. A.2). With this type of cost, speciation is impossible in small steps, because, as heterozygotes become very rare, the costs outweigh any potential benefit of being even choosier. However, speciation is still possible if a modifier inducing complete isolation invades a population in which heterozygotes are still common.

Limiting scenarios. As a consequence of the considerations above, evolution of assortative mating through infinitely small steps can be considered the "worst-case scenario" for speciation, and evolution in a single large step the "best-case scenario". Any realistic genetic architecture will lie in between these two extremes. We have demonstrated this for a single locus with intermediate step sizes (Fig. 2(b)), but the same conclusion should also apply more

generally to polygenic mating traits. In agreement with our results, previous studies have consistently found that speciation is easier (and faster) if the mating trait is determined by a small number of loci with large individual effects than if it depends on a large number of loci with small effects (e.g., Dieckmann and Doebeli, 1999; Gavrilets et al., 2007; Gavrilets and Vose, 2007). This is true even in the monostable C* regime (and thus holds independent of "jumping"), simply because large-effect alleles are under stronger selection. Our results reveal a second independent advantage of large step sizes at the mating locus, as such sizes facilitate the establishment of the linkage disequilibrium necessary for "jumping". On the other hand, if mutation rates are high such linkage disequilibrium may readily build up between polymorphic loci, so that multiple small mutations may combine to yield a larger step. At the far end of this scale, Doebeli et al. (2007) showed that the evolution of complete isolation occurs rather easily if the choosiness trait shows continuous quantitative genetic variation and population sizes are sufficiently large (under these assumptions, the so-called infinitesimal model applies, which becomes exact when the evolving trait is based on an infinite number of loci with infinitely small effects). Whether quantitative variation facilitates "jumping" when population sizes are finite is an open question that deserves further study.

For modeling purposes, detailed knowledge about real trait architectures will usually not be available. As an alternative strategy, it therefore seems advisable to cover the range of possibilities by considering the extreme cases, both of which are relatively easy to treat. For our present model of intraspecific competition, the differences between these limiting scenarios turn out to be relatively modest.

4.2. Genetic architecture of the ecological trait

Arguably the strongest assumption of the analytical models by Pennings et al. (2008), de Cara et al. (2008), Otto et al. (2008) and Ripa (2009) is that the ecological trait that underlies competition and mate choice is determined by a single diallelic locus. Here, we have studied the effect of adding a second locus for the ecological trait in the competitive-speciation model of Pennings et al. (2008). This additional locus creates enhanced flexibility for shaping the distribution of phenotypes. In addition to the evolutionary equilibria that are already known from the single-locus model, we therefore find three new outcomes: (1) a monomorphic equilibrium featuring only the intermediate phenotype, (2) a strongly peaked unimodal phenotype distribution with random mating, and (3) a three-species equilibrium.

Intermediate monomorphic equilibrium. Our analytical and numerical results show that a monomorphic equilibrium at the intermediate phenotype is a stable evolutionary outcome if and only if net selection under random mating ($m = 0$) is stabilizing, that is, if the stabilizing component of natural selection is stronger than frequency-dependent disruptive selection due to competition ($k \geq c$). In this parameter range, there is only a single ecological niche, and evolutionary branching (which requires net disruptive selection) is not possible. In part of this range, complete isolation can be maintained by sexual selection, but can never evolve from random mating, irrespective of the details of the genetic architecture. In contrast, for $k < c$, evolutionary branching can occur, and the intermediate monomorphic equilibrium does not play a role, at least in the deterministic case. With genetic drift, the monomorphic equilibrium is sometimes reached close to the low-c border of the C2* domain.

High frequency of the intermediate phenotype. For $k < c$, the intermediate monomorphic equilibrium is no longer stable. However, close to this parameter region, polymorphic equilibria still have a phenotype distribution with a high frequency of intermediate phenotypes. In the two-locus model, such a distribution can be achieved by an asymmetric shift of the allele frequencies at the ecological loci. Since this shift is exactly opposite at the two ecological loci, the resulting phenotype distribution is symmetric. This highlights the increased flexibility of the two-locus model relative to the single-locus model, for which the frequency of the intermediate phenotype cannot exceed 0.5 in the absence of dis-assortative mating.

Three species. In the two-locus model, three species evolve over nearly the whole parameter range for which the one-locus model predicts partial isolation. This is because short-range competition creates three ecological niches. In the one-locus model, the intermediate niche can be filled only by heterozygotes (whose frequency is determined by the degree of assortative mating). In the two-locus model, in contrast, the only way to achieve a symmetric solution with three phenotypic clusters (instead of five) is via complete isolation. Partial isolation evolves only over a small parameter range with very strong frequency dependence (c near 1), in which five phenotype clusters are favored over three, or when partial isolation is stabilized by sexual selection). The formation of three species also fails if the maximal possible m is too low to ensure reproductive isolation between neighboring species. In this sense, speciation is easier in the two-than in the three-species domains.

In contrast to the two-species equilibrium, evolution of three species requires symmetry breaking: while one intermediate haplotype (i.e., $+-$ or $-+$) goes extinct, the other one is preserved and constitutes the third species. As described in more detail in the Results section, this can lead to the delayed evolution of complete isolation, with an extended "pre-speciation" phase during which all haplotypes are still present in the population. Similar threshold phenomena have also been described in other models of speciation (e.g., Bolnick, 2006; Heinz et al., 2009), although it is not clear whether the underlying mechanism is the same in each case. In our model, the long stagnation phase can be explained by the fact that the population passes close to a saddle point. Near such a point, selection is very weak, and it only becomes stronger again once the population has moved on into a new dimension of state space. This escape requires symmetry breaking, which takes a long time in a deterministic system, but can easily occur due to stochastic fluctuations. Therefore, genetic drift enables the system to leave the neighborhood of the saddle point faster. Bolnick (2006) speculates that, in his model, the stagnation phase is used for a process of genotype sorting (into groups that breed true). In our model, however, this sorting is not a lengthy process, but rather a sudden event that is observable only directly before the symmetry breaking. Once one of the mixed haplotypes has gained dominance, it quickly spreads through the population by a positive feedback (because the more frequent haplotype has fewer heterozygous offspring and this advantage increases as it becomes more frequent).

While all new possibilities for shaping the phenotype distribution are realized in our two-locus model, none of these changes the evolutionary outcomes over the parameter range in which two species evolve from random mating. Indeed, we find that the C2* domain remains remarkably stable with respect to the genetic architecture of the ecological trait. Given appropriate ecological parameters, two clusters at the edges of the phenotype range are favored. In our model, a sexual and recombining population can reach such a phenotype distribution only by evolving complete isolation. This result is complemented by the finding that local stability of the two-species equilibrium does not depend on the genetic architecture of the ecological trait as long as the total range of possible phenotypes remains constant (Appendix B).

4.3. Other genetic architectures

Competitive speciation results from a combination of factors. Most importantly, the ecological conditions must induce frequency-dependent disruptive selection. With a restricted phenotype range (as assumed here), such selection favors the extreme phenotypes. All loci that contribute to spanning this range maintain polymorphism. However, under random mating, recombination and segregation produce intermediate phenotypes. Finally, selection against these intermediates entails selection for assortative mating (unless frequency dependence is strong enough to create a large number of niches, in which intermediates are protected from competition).

Since more polymorphic loci contributing to the ecological trait lead to more intermediates, we then expect selection for assortative mating to occur over a larger parameter range. Indeed, this is what we see if we compare the models with one and two diallelic loci. In the two-locus model, the two-species domain is complemented by a three-species domain, in which evolution of complete isolation is driven by selection against additional heterozygotes (which do not exist in the one-locus model). With even more polymorphic loci and sufficiently strong frequency dependence, we can expect additional equilibria with four or more species, as observed by Bolnick (2006).

On the other hand, increasing the number of ecological loci may also entail effects that oppose speciation. In particular, with more loci, neighboring phenotypes become more similar to each other, and female choosiness needs to evolve to higher levels before gene flow is effectively stopped. Indeed, the level of choosiness required to maintain three species is very high (see above). Furthermore, also in the two-species domain, speciation takes longer in the two-locus model than in the one-locus model.

Of course, there are many more ways to increase the complexity of the trait architecture than just increasing the number of loci. An obvious extension is to vary the number of alleles per locus. To cover this case, we performed numerical tests with one ecological locus and three to five equally spaced alleles. The three-allele case, in particular, interestingly complements the diallelic two-locus model, since it produces the same set of diploid phenotypes. Somewhat counter-intuitively, however, the three-allele model does not have a three-species regime. Instead, in the relevant parameter range, the intermediate allele goes extinct, and the population evolves to a partial-isolation equilibrium, as in the P* regime of the one-locus two-allele model. The reason is that, at random mating, the intermediate allele is selected against, because the intermediate niche is already occupied by heterozygotes carrying the two extreme alleles. Note also that, unlike the two-locus two-allele model, the one-locus three-allele model has two different ways of producing three coexisting phenotypes: either with three alleles and complete isolation or with two alleles and partial isolation. In simulations starting at random mating, the initial low fitness of the intermediate allele steers the population towards the second solution. The intermediate allele also goes extinct in the C2* regime, after which the population evolves to complete isolation between two species. This is in line with results by Schneider (2007), who showed analytically that, in a general multilocus model, all intermediate alleles go extinct if the fitness function is quadratic (which, assuming that phenotypes are bounded from above and below, may be regarded as a suitable approximation for situations with exactly two niches). Indeed, in our model, three alleles were maintained only in simulations with very high values of c, in which short-ranged competition creates more than three niches.

Our model does not consider epistasis and dominance, but these can potentially have important effects. In particular, complete dominance can prevent the production of intermediate phenotypes in heterozygotes. Therefore, evolution of dominance has been suggested as an alternative to the evolution of assortative mating in diploid sexual populations (Durinx and Van Dooren, 2009). A recent study by Peischl and Schneider (2010) shows that this is sometimes possible, but the detailed analysis is complex. Similarly, epistasis generally leads to an evolving trait architecture. Two studies (Kopp and Hermisson, 2006; van Doorn and Dieckmann, 2006) show that disruptive selection will usually favor a trait architecture with only few polymorphic loci of large effect, such that few intermediate heterozygote phenotypes are produced. In the light of these studies, assuming a small number of polymorphic loci may be more realistic than extensive polymorphism at many loci with small individual effects. Indeed, for situations with two or three niches, the models by van Doorn and Dieckmann (2006) and Kopp and Hermisson (2006) predict evolution towards only a single diploid polymorphic locus. However, this result rests on the assumption that there are no constraints on the effect a single locus can have. Therefore, if a one-locus architecture is prevented by an upper limit on individual locus effects, a two-locus architecture, as studied here, is the logical next alternative. More generally, evolution of genetic architecture and evolution of assortative mating are alternative responses to the challenge presented by frequency-dependent disruptive selection against intermediate phenotypes (Rueffler et al., 2006). Often, evolution of one of these responses will weaken the selection pressure for the other (Durinx and Van Dooren, 2009; but see Peischl and Schneider, 2010). In the three-species regime, for example, evolution of a one-locus architecture (if possible) relaxes selection on the mating trait to a degree that full isolation no longer evolves. If, in addition, the range of allelic effects at this locus is unconstrained, selection pressure on the mating locus may be reduced even further (Ripa, 2009). In this sense one might conclude that evolution of assortative mating is driven by constraints on the genetic architecture of the ecological trait. Interestingly, with regard to the "problem" of intermediate phenotypes, a two-locus model turns out to be more constrained than a one-locus model with the same total phenotypic range because, with random mating, two polymorphic additive loci inevitably produce three intermediate phenotypes.

4.4. Conclusions

We have investigated how competitive speciation depends on the genetic architecture of both an ecological under frequency-dependent disruptive selection and of a mating trait regulating female choosiness. Our analyses have revealed a pattern of opposite effects: in general, speciation is easiest if allelic effects are small for the ecological trait and large for the mating trait. This contrast arises because these traits are under different forms of selection: frequency-dependent disruptive (and eventually potentially stabilizing) selection on the ecological trait as opposed to directional (and eventually potentially stabilizing) selection on the mating trait. More specifically, frequency-dependent disruptive selection tends to maintain polymorphism at multiple loci, which in turn creates low-fitness intermediate phenotypes (due to segregation and recombination). Eliminating these phenotypes by means other than assortative mating becomes more difficult if the genetic architecture is complex. In contrast, no widespread polymorphism is maintained at the mating trait. Here, large mutations are under stronger (directional) selection, and in addition, they allow the population to reach complete isolation by "jumping" over an alternative partial-isolation equilibrium. Finally, the original motivation for our study was to test the robustness of previous analytical models, which assumed a single ecological locus and evolution of choosiness in infinitesimal steps. It turns out that both of these assumptions are conservative with respect to the possibility of competitive speciation.

Acknowledgments

We thank three anonymous reviewers for helpful comments on the manuscript. AR, JH, and MK were supported by a grant by the Vienna Science and Technology Fund (WWTF) to JH. AR was also supported by the Evolution, Ecology and Systematics program at the University of Munich. UD gratefully acknowledges financial support by the European Commission, through the Marie Curie Research Training Network FishACE and the Specific Targeted Research Project FinE, funded under the European Community's Sixth Framework Program. UD received additional support by the European Science Foundation, the Austrian Science Fund, the Austrian Ministry of Science and Research, and the Vienna Science and Technology Fund.

Appendix A. Derivations for single-locus model

In this Appendix, we derive the conditions for local stability of random mating and complete isolation in the discrete-time single-locus model.

Stability of random mating

According to condition (11), random mating is stable if $W_{\text{het}} > W_{\text{hom}}$ at $m = 0$. Since there is no sexual selection with random mating, all mating rates are $\phi(X) = 1$ in this case. All differences in the fitness values (8) are therefore due to differences in the death rates (4). In particular, random mating is stable if $d_{\text{hom}} > d_{\text{het}}$, and thus

$$\frac{(1+(1-c)^4)N_{\text{hom}} + (1-c)N_{\text{het}}}{K_0(1-k)} > \frac{2(1-c)N_{\text{hom}} + N_{\text{het}}}{K_0}.$$

With $\frac{N_{\text{het}}}{N_{\text{hom}}} = 2$ for random mating, this leads to

$$k > \frac{1-(1-c)^4}{4-2c},$$

which is identical to the condition in the continuous-time model (Pennings et al., 2008) with a Gaussian shape of the competition function.

Stability of complete isolation

Complete isolation is stable if $W_{\text{het}} < W_{\text{hom}}$ at $m = 1$. We can assume $N_{\text{het}} \to 0$. As μ for complete isolation is 0 (5), Eq. (6) gives us $\phi_{\text{male}} = 0$. From Eq. (8), we thus get

$$W_{\text{het}} = \exp(\rho(1-d_{\text{het}}))\frac{1}{2} < \exp(\rho(1-d_{\text{hom}}))\frac{1+N_{\text{hom}}Q_{\text{hom}}}{2} = W_{\text{hom}}.$$

As d_{hom} at equilibrium is 1 and $Q_{\text{hom}} = \frac{1}{N_{\text{hom}}}$ according to Eq. (7), we can simply write $\exp(\rho(1-d_{\text{het}})) < 2$ or $d_{\text{het}} > 1 - \frac{\log(2)}{\rho}$. Using Eq. (4), $d_{\text{het}} = \frac{2(1-c)N_{\text{hom}}}{K_0}$, where $N_{\text{hom}} = \frac{K_0*(1-k)}{1+(1-c)^4}$, that is

$$\frac{2(1-c)(1-k)}{1+(1-c)^4} > 1 - \frac{\log(2)}{\rho}.$$

This leads to the condition

$$k < 1 - \frac{1-\frac{\log(2)}{\rho}(1+(1-c)^4)}{2(1-c)}.$$

With the choice $\rho = 2\log(2)$, this again matches the condition in the continuous-time model (Pennings et al., 2008),

$$k < \frac{3-4c+(1-c)^4}{4-4c}.$$

Appendix B. Derivations for multi-locus model

In this Appendix, we summarize several analytical results for models in which the ecological trait is influenced by either two loci or an arbitrary number of loci.

Stability of monomorphic equilibria

The stability of the monomorphic equilibria can be calculated for an arbitrary genetic basis of the ecological trait. Only the monomorphic states of the intermediate phenotype $x = 0$ (MI) and the extreme phenotypes $x = \pm 1$ (ME) are of interest. The equilibrium is stable if and only if no mutant with a different phenotype can invade. Since the fitness of rare mutants is necessarily dominated by the heterozygotes, this is equivalent to the condition that the fitness of all heterozygote single mutant invaders into the monomorphic resident population is less than one. Let us assume that the phenotypic effect of a given mutant is $1/n$. In particular, this is the effect of a single mutant in a model with n identical diallelic loci spanning the phenotype range $[-1; 1]$. For $n = 1$ and $n = 2$, this covers the models considered in the bulk of the paper.

For the internal monomorphic equilibrium MI with $N(0) = K_0$ and $N(x) = 0$ for $x \neq 0$, the mating rate of a rare mutant with phenotype $1/n$ is given by $\phi(1/n) = (1+(1-m)^{1/n^2})/2$, and its fitness follows from (8) as

$$W_{\frac{1}{n}} = \exp\left[\rho\left(1-\left(\frac{1-c}{1-k}\right)^{1/n^2}\right)\right]\frac{1+(1-m)^{1/n^2}}{2},$$

which leads to the stability condition

$$\exp\left[\rho\left(1-\left(\frac{1-c}{1-k}\right)^{\frac{1}{n^2}}\right)\right] < \frac{2}{1+(1-m)^{\frac{1}{n^2}}}.$$

For $m = 0$, the MI equilibrium is stable for $k > c$, independent of the genetic architecture of the mating trait.

For the extreme monomorphic equilibrium ME at $x = 1$, we have $N(1) = K_0(1-k)$ and $N(x) = 0, x \neq 1$. For a mutant with phenotype $1 - 1/n$, the condition $W_{1-1/n} < 1$ leads to the stability condition

$$\exp\left[\rho\left(1-(1-c)^{(1/n^2)}(1-k)^{(2n-1)/n^2}\right)\right] < \frac{2}{1+(1-m)^{1/n^2}}.$$

For $m = 0$, the right-hand side is 1, and the extreme monomorphic equilibrium is stable for

$$c < 1 - \frac{1}{(1-k)^{2n-1}}.$$

In particular, the equilibrium is never stable at $m = 0$ if $k \geq 0$. For negative k, the domain of stability increases with increasing n, that is, with decreasing mutational effect. This shows that the stability of the ME equilibrium depends on the largest mutation that is possible for the ecological trait. It is, therefore, not necessarily true that the equilibrium is more stable for a polygenic trait, unless the genetic architecture excludes (even occasional) mutations with large effect.

Stability of random mating

Unlike in the single-locus model, we were not able to derive an analytical condition for the stability of the random mating equilibrium. The reason is that the simple criterion (11), which compares the fitness values of two phenotypic classes (homozygotes and heterozygotes), can no longer be applied when the number of classes increases. Nevertheless, the following

heuristic leads to a quite accurate analytical approximation for the two-locus case:

Assume that mutants with non-zero choosiness segregate at low frequency in linkage equilibrium with the ecological loci. As in the single-locus case, these mutants will, on average, produce offspring with more extreme phenotypes (in particular, more offspring with phenotypes ± 1 and fewer with phenotype 0). Suppose that among residents $W(\pm 1) > W(\pm 0.5) > W(0)$. Then the offspring of rare mutants will be fitter than the residents, and the mutation will spread. In the opposite case, $W(\pm 1) < W(\pm 0.5) < W(0)$, mutant offspring are less fit, and the mutation will decrease in frequency. Unfortunately, this heuristic is inconclusive if the fitness values are not monotonic. However, for a resident population at random mating ($m_r = 0$), it turns out that the fitness values are monotonic in almost the entire parameter space spanned by c and k. In fact, the conditions $W(\pm 1) = W(0)$ and $W(\pm 0.5) = W(0)$ lead to boundaries in the c–k plane that are almost indistinguishable by eye. We can use the fact that, at $m = 0$, differences in fitness are only due to differences in the death rates. From $d(\pm 1) = d(0)$, we derive

$$k(c) = \frac{1 - 4(1-c)^{(1/4)} + 6(1-c) + 4(1-c)^{(9/4)} + (1-c)^4 + 1}{2(4-c) + 8(1-c)^{(1/4)}},$$

which is the boundary line used in Fig. 3.

Stability of complete isolation

Also the limit $m \to 1$ allows for further analytical derivations. Note, first, that the condition for local stability of the two-species equilibrium with respect to invasion of rare heterozygotes with phenotype $x = 0$ is the same as in the single-locus case. An analogous calculation shows that invasion of mutants with phenotype $x > 0$ is always more difficult and can thus be ignored. We can conclude that the boundary for local stability of two species with complete isolation is independent of the genetic architecture of the ecological trait, as long as the total phenotypic range is kept constant.

For the three-species equilibrium, the equilibrium frequencies can be obtained from the condition that all three phenotypes must have equal fitness. Since the mating rates are necessarily equal at $m = 1$, this condition reduces to equal death rates, $d(x = \pm 1) = d(x = 0)$. Using Eq. (4), we obtain $d(0) = 1 - c + p_0 c$ and $d(\pm 1) = ((1-p_0)(1+(1-c)^4) + 2p_0(1-c))/(2-2k)$, where p_0 is the frequency of the third species at $x = 0$. The above system of equations evaluates to

$$p_0 = 1 - \frac{2(c-k)}{2c(c-k) + c^2(2-c)^2}. \tag{B.1}$$

We find that $p_0 < 1$ for $c > k$, which is consistent with the finding that the intermediate monomorphic equilibrium (i.e., $p_0 = 1$) is always stable for $c < k$. The condition $p_0 > 0$ leads to

$$k > c\left(1 - \frac{c(2-c)^2}{2(1-c)}\right). \tag{B.2}$$

For smaller k, only a two-species equilibrium is (locally or globally) stable. Finally, a condition for invasion of phenotypes at $x = (\pm 0.5)$ into the three-species equilibrium can be formulated, but leads to higher order polynomials that can only be analyzed numerically.

Note that the stability results depend on the scaling of the locus effects chosen in our model, which leaves the total phenotype range invariant when going from one to two loci. An alternative scaling, which keeps the single-locus effects constant, would result in a doubling of the phenotype range with two loci. Compared to the single locus case, the emerging species at the boundaries of this range have much more extreme phenotypes. We can obtain the corresponding regime picture by a simple rescaling $k \to 1 - (1-k)^4$ and $c \to 1 - (1-c)^4$ of the model (not shown). Since an enlarged phenotype range increases the effective strength of both stabilizing and frequency-dependent selection, the area where two species are stable shifts to smaller values of k and c.

Appendix C. Costs of choosiness

Until now, we have assumed that assortative mating does not affect female mating success or viability, that is, there are no costs of choosiness. This may be not realistic if, for example, rare choosy females have problems of finding an acceptable mate or if choosiness increases the death rate. In this Appendix, we investigate how costs of choosiness affect our conclusions from the main text, focusing on the effect of mutational step size at the mating locus in the model with a single ecological locus.

We model costs according to Kopp and Hermisson (2008), distinguishing between *mating costs* and *viability costs*. With mating costs, a female has only a finite number (T) of mating trials in a breeding season. This affects the female mating rate according to

$$\phi_{\text{female}}(X) = 1 - \left[1 - \left(\frac{1}{N}\sum_Y N(Y)\mu(X,Y)\right)\right]^T. \tag{C.1}$$

The sum on the right-hand side is the probability that an encounter of a female with a random male leads to mating. Mating costs are frequency-dependent and will most strongly affect females with a rare phenotype, which need many trials to find a matching male. They are, therefore, an example of relative costs (Otto et al., 2008; Kopp and Hermisson, 2008).

Other types of costs are absolute and do not depend on the number of mating trials. These can occur, for example, if choosiness comes with a physiological cost and, thus, reduces female viability. Absolute viability costs can be included as an extra term in the death rate,

$$d(X) = \frac{C(X)}{K(X)} + f_\delta(m - \bar{m}), \tag{C.2}$$

where f_δ is a costs function that depends only on the difference between the mutant choosiness m and the mean resident choosiness $\bar{m}$. The latter assumption implies soft selection, which can arise if costs do not affect the carrying capacity. In this case, costs are only paid by mutants, and their magnitude depends on the mutational step size. For example, we can assume a linear costs function

$$f_\delta = \delta(m - \bar{m}) + \delta'(m' - \bar{m}'), \tag{C.3}$$

where δ and δ' are the costs linked to m and m', respectively.

Evolution of complete isolation

We investigate how (relative) mating costs and (absolute) viability costs affect the conditions under which complete isolation can evolve from random mating in large steps.

Mating costs. To test the effect of mating costs, we performed simulations with costs of choosiness modeled according to Eq. (C.1). Other than restricting the number of mating trials to $T = 10$ and $T = 5$, respectively, we used the same assumptions as in the simulations without costs. The results are shown in Fig. A.1. The parameter range where complete isolation can evolve (C* regime) is shifted towards lower values of k compared to the simulations without costs (compare Fig. A.1 to Fig. 2), but it is still larger with

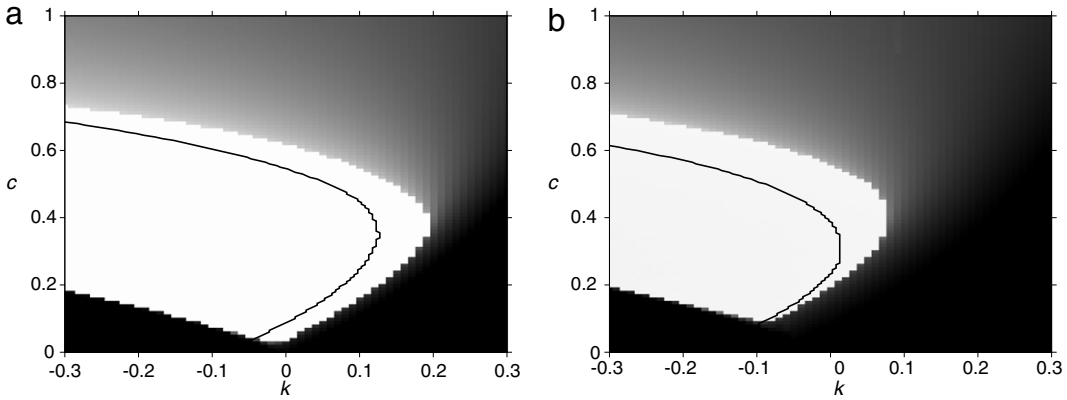

Fig. A.1. Evolution of assortative mating in the single-locus model with mating costs, assuming a single large mutational step. The number of mating trials per female is $T = 10$ in (a) and $T = 5$ in (b). Shades of gray indicate the equilibrium frequency (black: 0, white: 1) of an "isolation allele" with $m = 1^- = 0.99$ introduced at the low frequency $p = 10^{-4}$ into a randomly mating resident population with $m_r = 0$. The continuous lines show the corresponding boundaries of the complete-isolation regime for infinitesimally small steps (calculated as in Kopp and Hermisson, 2008).

one step and costs than with infinitesimal steps and costs. Furthermore, mating costs appear to facilitate "jumping" near the high c-boundary of the C* regime (where it is almost absent without costs, see Fig. 2), to the point that there are some parameter combinations for which complete isolation can only evolve in the presence of (moderate) costs. As detailed in the Discussion, this result arises because mating costs induce sexual selection against rare females (Kopp and Hermisson, 2008). If most males (and females) are homozygous due to strong frequency-dependent competition, this effect combines with sexual selection against heterozygous males to facilitate fixation of the high-m allele.

Viability costs. With absolute viability costs and soft selection (Eq. (C.2)), evolution of complete isolation in infinitesimal steps is impossible for costs functions $f_\delta(m - \bar{m})$ with a positive first derivative at $m - \bar{m} = 0$ (as is the case for the linear costs function Eq. (C.3)). The reason is that, for $m \to 1$, heterozygotes disappear from the population, and the invasion fitness gradient (which measures the selection pressure for increasing m) vanishes even in the absence of costs. Unless the derivative of the costs function vanishes, too, the invasion fitness gradient becomes negative, which precludes evolution of complete isolation (Kopp and Hermisson, 2008).

However, the above reasoning applies only to infinitesimal steps. With large steps, an $m = 1^-$ isolation mutant can invade while heterozygotes are still frequent, such that a fitness advantage for homozygote offspring can potentially offset the negative effect of costs. Whether this is easier for (finitely) small or large mutations depends on the details of the costs function (C.2). Recall that this function depends only on the difference between the mutant and resident m. If costs are low for small differences but high for large differences, small steps will be preferred. But if costs for small differences are sizeable and costs for large differences are only moderate, evolution of complete isolation might be possible only in large steps. In summary, it is possible to construct costs functions that favor either finitely small or large steps, but evolution of complete isolation via infinitesimal small costly steps is always impossible.

Fig. A.2 demonstrates evolution of reproductive isolation in a single large step, using as an example the linear cost function (C.3) with $\delta = 0.01$ and $\delta' = 0$. The parameter range where complete isolation evolves is quite large. This is true even though the boundary for the stability of random mating is shifted slightly towards lower k as compared to the boundary for infinitesimal steps. Note, however, that since the invasion fitness gradient at $m \to 1$ is negative (see above), the whole C* regime is susceptible

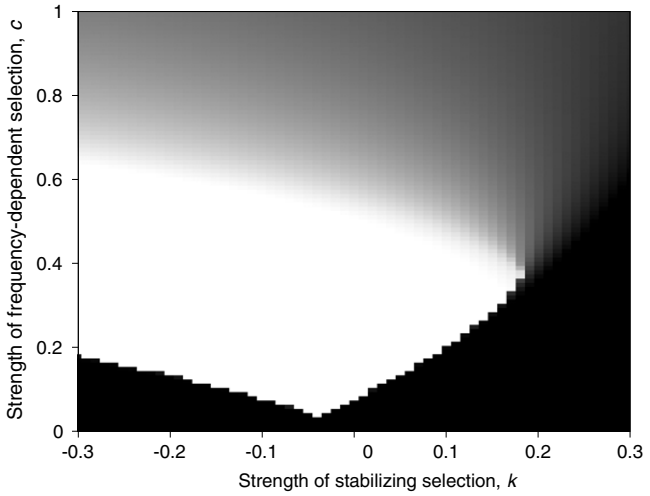

Fig. A.2. Evolution of assortative mating in the single-locus model with absolute viability costs and a single large mutational step (see Fig. A.1 for more details). Shades of gray indicate the equilibrium frequency (black: 0, white: 1) of an "isolation allele" with $m = 1^- = 0.99$ introduced at the low frequency $p = 10^{-4}$ into a randomly mating resident population with $m_r = 0$. The figure shows simulation results for the linear cost function (C.3) with $\delta = 0.01$ and $\delta' = 0$. Note that, with this costs function, evolution of complete isolation through infinitesimally small mutational steps is always impossible.

to invasion by small (but not large) modifiers for lower m. It is, therefore, possible that nearly complete reproductive isolation is reached in a large step, but then decreases again slightly due to invasion and fixation of alleles with small negative effect.

Appendix D. Genetic drift in the two-locus model

Including genetic drift in the two-locus model yields roughly the same results as the deterministic model. In the three-species domain C3*, the stagnation phase described in the main text is shorter and, during this phase, the polymorphism is not between the $m = 1$ and the $m = 0$ allele but between the $m = 1$ allele and the lowest m allele that is still present in the population (usually, the $m = 0$ has died out at this point, something which is not possible in the infinite population limit). Near the border to the P* domain, the formation of three species sometimes fails altogether, and the population gets stuck at partial isolation. Note that this may be a transient state, however, and the phase transition to three species might still occur eventually. Near the boundary of the C2*

domain, evolution from random mating with drift sometimes leads to the formation of two species instead of three. Furthermore, in the presence of drift, the intermediate monomorphic equilibrium MI sometimes evolves also with parameters pertaining to the R^* or $C2^*$ domain, and the extreme monomorphic equilibrium ME is reached in a larger parameter range, too.

References

Bolnick, D.I., 2006. Multi-species outcomes in a common model of sympatric speciation. J. Theoret. Biol. 241, 734–744.

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

Bolnick, D.I., Fitzpatrick, B.M., 2007. Sympatric speciation: models and empirical evidence. Annu. Rev. Ecol. Evol. Syst. 38, 459–487.

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

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

Darwin, C., 1859. On the Origin of Species by Means of Natural Selection or the Preservation of Favored Races in the Struggle for Life, 1st ed. J. Murray, London.

de Cara, M.A.R., Barton, N.H., Kirkpatrick, M., 2008. A model for the evolution of assortative mating. Am. Nat. 171, 580–596.

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

Doebeli, M., Blok, H.J., Leimar, O., Dieckmann, U., 2007. Multimodal pattern formation in phenotype distribution of sexual populations. Proc. R. Soc. Lond. Ser. B 274, 347–357.

Doebeli, M., Dieckmann, U., 2003. Speciation along environmental gradients. Nature 421, 259–264.

Doebeli, M., Dieckmann, U., 2005. Adaptive dynamics as a mathematical tool for studying the ecology of speciation processes. J. Evol. Biol. 18, 1194–1200.

Durinx, M., Van Dooren, T., 2009. Assortative mate choice and dominance modification: alternative ways of removing heterozygote disadvantage. Evolution 63, 334–352.

Gavrilets, S., 2004. Fitness Landscapes and the Origin of Species. Princeton University Press, Princeton, NJ.

Gavrilets, S., 2005. "Adaptive speciation"—it is not that easy: a reply to Doebeli et al. Evolution 59, 696–699.

Gavrilets, S., Vose, A., 2007. Case studies and mathematical models of ecological speciation. 2. Palms on an oceanic island. Mol. Ecol. 16, 2910–2921.

Gavrilets, S., Vose, A., Barluenga, M., Salzburger, W., Meyer, A., 2007. Case studies and mathematical models of ecological speciation. 1. Cichlids in a crater lake. Mol. Ecol. 16, 2893–2909.

Geritz, S.A.H., Gyllenberg, M., Jacobs, F.J.A., Parvinen, K., 2002. Invasion dynamics and attractor inheritance. J. Math. Biol. 44, 548–560.

Heinz, S., Mazzucco, R., Dieckmann, U., 2009. Speciation and the evolution of dispersal along environmental gradients. Evol. Ecol. 23, 53–70.

Kirkpatrick, M., Ravigné, V., 2002. Speciation by natural and sexual selection: models and experiments. Am. Nat. 159, S22–S35.

Kopp, M., Hermisson, J., 2006. The evolution of genetic architecture under frequency-dependent disruptive selection. Evolution 60, 1537–1550.

Kopp, M., Hermisson, J., 2008. Competitive speciation and costs of choosiness. J. Evol. Biol. 21, 1005–1023.

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

May, R., MacArthur, R., 2001. Niche overlap as a function of environmental variability. Proc. Natl. Acad. Sci. USA 69, 1109–1113.

Otto, S.P., Servedio, M., Nuismer, S., 2008. Frequency-dependent selection and the evolution of assortative mating. Genetics 179, 2091–2112.

Peischl, S., 2010. Mathematical models of freqency-dependent selection with dominance. Ph.D. Thesis. University of Vienna.

Peischl, S., Bürger, R., 2008. Evolution of dominance under frequency-dependent intraspecific competition. J. Theoret. Biol. 251, 210–226.

Peischl, S., Schneider, K., 2010. Evolution of dominance under frequency-dependent intraspecific competition in an assortatively mating population. Evolution 64, 561–582.

Pennings, P.S., Kopp, M., Meszéna, G., Dieckmann, U., Hermisson, J., 2008. An analytically tractable model of competitive speciation. Am. Nat. 171, E44–E71.

Ricker, W., 1954. Stock and recruitment. J. Fish. Res. Board Can. 11, 559–623.

Ripa, J., 2009. When is sympatric speciation truly adaptive? An analysis of the joint evolution of resource utilization and assortative mating. Evol. Ecol. 23, 31–52.

Rosenzweig, M.L., 1978. Competitive speciation. Biol. J. Linn. Soc. 10, 275–289.

Roughgarden, J., 1972. Evolution of niche width. Am. Nat. 106, 683–718.

Rueffler, C., van Dooren, T.J.M., Leimar, O., Abrams, P.A., 2006. Disruptive selection and then what? Trends Ecol. Evol. 21, 238–245.

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

Schneider, K., 2007. Long-term evolution of polygenic traits under frequency-dependent intraspecific competition. Theor. Popul. Biol. 71, 342–366.

van Dooren, T.J.M., 1999. The evolutionary ecology of dominancerecessivity. J. Theoret. Biol. 198, 519–532.

van Doorn, G.S., Dieckmann, U., 2006. The long-term evolution of multilocus traits under frequency-dependent disruptive selection. Evolution 60, 2226–2238.

van Dooren, T.J.M., Durinx, M., Demon, I., 2004. Sexual dimorphism or evolutionary branching? Evol. Ecol. Res. 6, 857–871.

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

Waxman, D., Gavrilets, S., 2005. 20 questions on adaptive dynamics: a target review. J. Evol. Biol. 18, 1139–1154.