A Model for the Evolution of Assortative Mating

M. A. R. de Cara,[1,*] N. H. Barton,[1] and M. Kirkpatrick[2]

1. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom;
2. Section of Integrative Biology, University of Texas, Austin, Texas 78712

Submitted December 21, 2006; Accepted November 15, 2007; Electronically published March 21, 2008

Online enhancements: Mathematica notebook, appendixes.

ABSTRACT: Many animals and plants show a correlation between the traits of the individuals in the mating pair, implying assortative mating. Given the ubiquity of assortative mating in nature, why and how it has evolved remain open questions. Here we attempt to answer these questions in those cases where the trait under assortment is the same in males and females. We consider the most favorable scenario for assortment to evolve, where the same trait is under assortment and viability selection. We find conditions for assortment to evolve using a multilocus formalism in a haploid population. Our results show how epistasis in fitness between the loci that control the focal trait is crucial for assortment to evolve. We then assume specific forms of assortment in haploids and diploids and study the limiting cases of selective and nonselective mating. We find that selection for increased assortment is weak and that where increased assortment is costly, it does not invade.

Keywords: assortative mating, disruptive selection, speciation.

Assortative mating occurs when there is a correlation between phenotypes of mated individuals with respect to a trait that is expressed in both. For instance, the flowering time of plants that fertilize each other and the body size of mating individuals are frequently found to be correlated. Assortative mating is important because it alters the variance of quantitative traits (Fisher 1918) and the effective rate of recombination and can generate sexual selection. One important role that assortment can play (and one that has received a lot of recent theoretical attention) is as a source of reproductive isolation between emergent lineages within species, which is a prerequisite for sympatric speciation and for reinforcement between diverged species after secondary contact (reviewed in Kirkpatrick and Ravigné 2002; Coyne and Orr 2004; Gavrilets 2004, pt. III).

However, despite the vast literature on the consequences of assortative mating, the conditions under which assortment evolves are still not clearly understood. Most of the available theory is based on simulations with so many parameters that only a fraction of their values could be investigated and does not provide a clear picture of the general principles under which assortative mating evolves. Our goal here is to provide a simple model from which we can derive analytical expressions while making as few assumptions as possible.

We focus here on assortative mating generated by a focal trait, such as body size or flowering time in plants, controlled by a single set of loci and expressed in males and females. This contrasts with nonrandom mating generated via preferences in one sex (typically females) for a display trait expressed on the other sex, where preference and display trait are regulated by different sets of genes.

With our definition of assortative mating, we can distinguish between one- and two-trait models (Fry 2003), which correspond to magic-trait and similarity-based mating models in the classification of Gavrilets (2004, chap. 9). The difference between these two kind of models relies on whether the set of loci under viability selection and that involved in assortative mating are the same. This distinction is important because the conclusions derived from different genetic setups can be surprisingly different.[1]

Two-trait models distinguish two sets of loci, those controlling a trait under natural and sexual selection, such as body size or eye color, and a separate set of loci for a trait under assortative mating (Maynard Smith 1966; Felsenstein 1981; Diehl and Bush 1989; Dieckmann and Doebeli 1999; Kondrashov and Kondrashov 1999). The main conclusion of these models is that if linkage is tight and assortment is strong, the assortment trait can become as-

* Corresponding author; e-mail: angeles.decara@ed.ac.uk.

Am. Nat. 2008. Vol. 171, pp. 580–596.
DOI: 10.1086/587062

[1] Viability selection causes individuals with different genotypes to produce different numbers of surviving offspring. Sexual selection is caused by the differential success of individuals in finding a mate. Natural selection includes both components.

sociated with the selected trait, bringing together the components of reproductive isolation.

In contrast, we focus on one-trait models. These assume one set of genes that are simultaneously under natural and sexual selection and assortative mating (Rice 1984; Doebeli 1996; Kondrashov and Shpak 1998; Dieckmann and Doebeli 1999; Kirkpatrick 2000; Matessi et al. 2001; Kirkpatrick and Nuismer 2004). Unfortunately, the main conclusions of one-trait models do not always coincide. Rice (1984) found that disruptive natural selection or assortment separately need to be very strong, almost complete, in order for reproductive isolation to evolve. However, when assortment and disruptive selection acted jointly, intermediate strengths of each of them were required to produce an extreme reduction in gene flow. This conclusion is in agreement with later one-trait models of sympatric speciation and reinforcement (Kondrashov and Shpak 1998; Kirkpatrick 2000; Kirkpatrick and Nuismer 2004), from which it follows that with little or no disruptive natural selection, the evolution of assortative mating is difficult. These models, however, assumed fixed strength of assortment, so their conclusions on whether assortment can evolve were limited. The model of Dieckmann and Doebeli (1999) studied explicitly the evolution of strength of assortment by introducing a third set of genes that modified the strength of assortment. From their simulation results, it seemed that assortment was likely to evolve under a very broad range of parameter values. Furthermore, assortment increased indefinitely, so that sympatric speciation occurred fairly easily over such a range of parameters. To complicate things further, the study of Matessi et al. (2001) found that assortment could stabilize at intermediate levels, so that it did not always increase indefinitely.

Thus, there is no clear answer for when assortment will evolve to be complete under the one-trait scenario. A major reason for the difference in conclusions is the use of different assumptions. At the same time, most of these studies have relied mainly on numerical investigations or individual-based simulations, which makes comparison difficult. This article attempts to rectify that situation.

We study here the evolution of assortment and consequent reproductive isolation. We focus on one-trait models of assortment based on an ecological cue (such as host choice or flowering time in plants). Selection will act on one focal trait, which is under both assortative mating and natural and sexual selection.

We use analytical approximations based on a general framework for multilocus models developed by Kirkpatrick et al. (2002). The strength of this method is that it provides exact solutions via recursion equations in terms of selection coefficients that apply to arbitrary models.

In order to study the evolution of assortment, we perform a modifier analysis (Karlin and McGregor 1974). This relies on one set of loci that control the focal trait and a different set of one or more modifier loci that alter the existing interactions on the focal trait. Specifically, we study modifiers of strength of assortment, so that those individuals that carry the modifier express a stronger or weaker degree of assortment. By assuming a modifier of weak and additive effect, we can identify the strength and direction of selection (as in Barton 1995). This will allow us to study approximate conditions for the modifier to invade and fix. The models allow for any initial degree of assortment, and so the analysis allows us to see when selection favors increased or decreased assortment, starting from any initial condition, and to study the long-term evolution.

The article is divided in two parts: the first derives general equations, using the multilocus method, and the second studies specific mating schemes. In the first part, we derive general equations for the evolution of modifiers of assortative mating. We then analyze the case in which the modifier is of weak and additive effect. These general equations are illustrated with an example of one selected trait and one modifier in a haploid population. In the second part, we study eight specific schemes proposed in previous studies, using analytical expressions and numeric iterations.

To our knowledge, our article is the first to derive analytically the conditions for assortative mating to evolve, which is crucial for the origin and maintenance of species in sympatry via reproductive isolation. The implications of this model are not limited to sympatric speciation. Indeed, the processes considered here are equally relevant to reinforcement after secondary contact and divergence of host races with partial isolation.

Models and Results

Our models for the evolution of assortative mating involve two sets of genes. The first contributes to variation in one or more focal traits that are the basis of nonrandom mating, for example, the date that a flower opens or the body size of an animal. For simplicity, and to make clear the connections with previous work, we refer to a single focal trait in what follows. When the model includes more than one locus, however, these might contribute to the same or to multiple focal traits, and we will return to this point in "Discussion." Typically, we expect such traits to be under natural selection as well as any sexual selection generated by nonrandom mating. The second set of genes consists of modifiers that affect the strength or pattern of nonrandom mating based on traits controlled by the first set. Examples include genes that affect the length of time that the flower stays open or the range of body sizes that are acceptable between mates. In general, there could be

overlap between these two sets of genes. In this article, however, we follow the traditional approach of assuming that the sets are nonoverlapping. This both simplifies the math and makes the evolutionary forces acting on the modifiers clear.

Our analysis uses a general framework for multilocus models that was introduced by Barton and Turelli (1991) and elaborated by Kirkpatrick et al. (2002); those works can be consulted for further details. The motivation for using this approach here is that it allows us to derive results that apply to a wide range of genetic systems and mating behaviors without having to commit, for example, to whether the organism in question is haploid or diploid or to a particular way of choosing mates. We then explore the implications of the general results by applying them to more restrictive cases. All derivations are shown in a Mathematica notebook available through a link in the online edition of the *American Naturalist*[2] and also available at http://bartongroup.icapb.ed.ac.uk (see "Software").

Methods and Assumptions

According to this multilocus formalism, genes occupy "positions." A position is specified by the locus, the sex of the individual carrying the gene, and, in diploids, its sex of origin, because in sexual haploids the sex of origin is not relevant. For instance, a haploid locus in a dioecious population i has two positions $\{i_f, i_m\}$ ("f" for female and "m" for male). We assume that two alleles, designated 0 and 1, segregate at each position. We define $p_{\mathbf{i}}$ as the frequency of allele 1 at position **i** in the gene and $q_{\mathbf{i}} = 1 - p_{\mathbf{i}}$ as the frequency of allele 0 at position **i**. We can define indicator variables at position **i** as $X_{\mathbf{i}}$, which is 1 if an individual carries allele 1 at position **i** and 0 otherwise. Genetic associations can then be defined as $D_{\mathbf{U}} = \mathrm{E}(\prod_{\mathbf{j}\in\mathbf{U}}(X_{\mathbf{j}} - p_{\mathbf{j}}))$, where E(…) represents an expectation taken over genotype frequencies (see eq. [3] of Kirkpatrick et al. 2002). By definition, associations on single positions, such as $D_{\mathbf{i}}$, are 0, because $p_{\mathbf{i}} = \mathrm{E}(X_{\mathbf{i}})$.

We can then write down the changes in allele frequencies at one position and in genetic associations. These changes depend on selection coefficients $a_{\mathbf{A}}$, which represent the force of selection acting over the entire generation on the set of positions **A** (for a more detailed description, see Kirkpatrick et al. 2002, p. 1733).

The set of all positions that contribute to genetic variation in the focal trait is denoted **Z**, and the set of modifier positions is denoted **M**. Assortative mating evolves through changes in the frequencies of the modifiers. An exact and general expression for the change in the frequency of allele 1 at modifier position **i** from the start of the current generation to the next is derived in appendix A in the online edition of the *American Naturalist*. In general, such expressions involve a very large number of terms. So in order to derive results with simple interpretations, we make several strong assumptions.

We assume that modifiers have additive effects, which implies that all selection coefficients that include more than one modifier position are 0. This is because selection coefficients represent projections of fitness on its components. If two components, such as $\{X_i, X_j\}$, act additively on fitness, selection coefficients a_i and a_j are not 0, but a_{ij} is 0, because there is no joint contribution of $\{ij\}$ to fitness. The second assumption is that the modifiers have weak effects. Specifically, we assume that $\epsilon \ll 1$, where ϵ is the magnitude of the largest selection coefficient with a modifier position. Note that selection on the focal genes is not necessarily weak. The third assumption is that linkage between the modifier and the focal-trait loci is loose and that selection acting on the modifier is weak relative to this recombination. The fourth assumption is that allele frequencies at the focal positions are at equilibrium or are changing slowly (app. B of Barton and Turelli 1991). In this case, the system will rapidly evolve to a state of "quasi-linkage equilibrium" (QLE), at which associations involving a modifier position are of order ϵ (Barton and Turelli 1991; Kirkpatrick et al. 2002). We emphasize that we are assuming QLE for the modifier but that we are making no restrictive assumptions about the focal genes: they may be under strong selection and have large disequilibria. We denote values at equilibrium with a hat (e.g., $\hat{D}_{\mathbf{U}}$) and values at QLE with a tilde (e.g., $\tilde{D}_{\mathbf{U}}$).

To further simplify the analysis, we make the assumption of symmetry between sexes. That implies that the female and male genetic contributions and selection coefficients are the same. One example where sexes are symmetric is when individuals are hermaphrodites and there is no genomic imprinting. Hermaphroditism is widespread in both animals and plants and so is not very biologically restrictive. If there is no genomic imprinting, the gene inherited via the sperm is expressed in the same way as the gene inherited via the egg. In terms of our notation, a consequence of symmetric sexes is that at QLE, the allele frequencies at all positions for a given locus i are equal, and we can write $p_{\mathbf{i}} = p_i$. Likewise, associations with positions that have the same sex of carrier (i.e., classical linkage disequilibrium) are equal in the two sexes, and for them we can write $D_{\mathbf{V}} = D_V$. Finally, selection coefficients for different positions at a given locus are equal: $a_{\mathbf{j}} = a_j$.

The dynamics for the allele frequency at a modifier locus, as derived in appendix A, are, then,

[2] Code that appears in the *American Naturalist* has not been peer reviewed, nor does the journal provide support.

$$\Delta p_i = a_i pq_i + \sum_{\mathbf{j}:j=i} t_{\mathbf{i}\leftarrow\mathbf{j}}\left(\sum_{\mathbf{V}\subseteq\mathbf{Z}} a_{\mathbf{V}}\tilde{D}_{\mathbf{jV}} + pq_i a_{\mathbf{Vj}} D_{\mathbf{V}}\right) + O(\epsilon^2), \tag{1}$$

where $t_{\mathbf{i}\leftarrow\mathbf{j}}$ is the probability that a gene in position **i** was inherited from a gene in position **j**. The sum $\sum_{\mathbf{j}:j=i}$ is over all positions **j** such that $j = i$, that is, such that i and j are the same once the context is removed. The set **jV** is the union of sets {**j**} and **V**, and $pq_i = p_i q_i$.

Here we see that evolution of a modifier depends on three kinds of selection coefficients. The first is a_i, representing direct selection on modifier locus i. The second kind of selection coefficient appears in equation (1) as $a_{\mathbf{V}}$, which represents selection on individual positions and sets of positions in the focal set. This force of selection is transmitted to the modifier locus i if there is an association maintained between those focal positions and i (represented in eq. [1] by the association $D_{\mathbf{jV}}$). Note that **j** must be a copy of the i gene that will be transmitted to position **i** via $t_{\mathbf{i}\leftarrow\mathbf{j}}$. The third kind of selection coefficient that contributes to evolution of the modifier appears in equation (1) as $a_{\mathbf{Vj}}$. This selection coefficient represents selection acting jointly on the modifier at locus i and the focal positions in set **V**.

The dynamics of the modifier also depend on QLE values for associations, which appear as $\tilde{D}_{\mathbf{jV}}$. Their quasi-equilibrium values can be calculated using the methods of Kirkpatrick et al. (2002), given expressions that depend only on allele frequencies and selection coefficients. We give examples below.

The Evolutionary Fate of a Modifier. When will a rare modifier of assortment spread? The conditions are easily found from equation (1). One way of measuring the strength of selection on the modifier is by looking at the rate of change of the modifier frequency $\Delta p_i/pq_i$, that is, dividing equation (1) by pq_i (note that this expression is normally known as the selection coefficient in a one-locus haploid model, but in order to avoid confusion, we refer to it as the rate of change). The evolutionary fate of the modifier is determined by whether the rate of change at small frequency is positive. That is, $\Delta p_i/pq_i$ as p_i approaches 0 (and, correspondingly, as q_i approaches 1) gives the proportional change in the frequency at locus i for a vanishingly rare modifier, which is the rate of invasion of the modifier λ_i,

$$\lambda_i = \lim_{p_i\to 0}\frac{\Delta p_i}{pq_i} = a_i + \sum_{\mathbf{j}:j=i} t_{\mathbf{i}\leftarrow\mathbf{j}}\sum_{\mathbf{V}\subseteq\mathbf{Z}} a_{\mathbf{V}}\frac{\tilde{D}_{\mathbf{jV}}}{pq_i} + \sum_{\mathbf{j}:j=i} t_{\mathbf{i}\leftarrow\mathbf{j}}\sum_{\mathbf{V}\subseteq\mathbf{Z}} a_{\mathbf{Vj}}\hat{D}_{\mathbf{V}} + O(\epsilon^2), \tag{2}$$

keeping in mind that the hat denotes values at equilibrium.

If λ_i is positive, the modifier invades. The strength of selection is determined by three quantities: the strength of direct selection on the modifier (a_i), the strength of indirect selection generated by individual genes and set of genes for the focal trait (a_V), and the strength of selection on sets of focal genes in combination with the modifier (a_{Vi}). Again the result depends on the associations $\tilde{D}_{\mathbf{jV}}$ and $\hat{D}_{\mathbf{V}}$, which can be calculated for any particular cases of interest using the methods of Kirkpatrick et al. (2002).

What is the fate of a modifier that invades? Under our assumption that assortment modifiers have weak additive effects, a modifier that invades will spread to fixation under a constant selection coefficient, λ_i. This is because, as the modifier spreads to fixation, its increased frequency can have only negligible weak effects on the frequencies of all the other loci and the linkage disequilibria among them. It is these selected genotype frequencies, together with the coefficients involving the modifier ($a_{\mathbf{iU}}$, etc.), that determine the weak associations involving the modifier and hence the value of λ_i. We require that the modifier have additive effects to ensure that the coefficients that involve the modifier remain constant as the frequency of the modifier changes: plainly, any degree of dominance (represented by coefficients such as $a_{i_{\mathrm{mm}}i_{\mathrm{fm}}\mathbf{U}}$) will cause the marginal effects of the modifier to change as its frequency changes.

An Example. To illustrate how the tools developed above can be used, we now analyze a specific model consisting of the simplest possible genetic system. We find the conditions under which a modifier invades and spreads to fixation in terms of the generalized selection coefficients $a_{\mathbf{V}}$ acting on the focal positions, so that the results apply to any form of assortment. Despite the major assumptions we have made to simplify the model, a full analysis is fairly involved.

The population is hermaphroditic and haploid, with one modifier locus i and two identical loci j and k controlling the focal trait. We expect similar results if the focal trait is controlled by one diploid locus. For the sake of simplicity, we will study here only the case where the focal trait is controlled by two haploid loci.

We assume that some form of balancing selection maintains an equilibrium with equal allele frequencies at the focal-trait loci, $p_j = p_k = 1/2$. This kind of selection arises, for example, in models of competitive speciation (Doebeli 1996; Dieckmann and Doebeli 1999; Drossel and McKane 2000; Kirkpatrick and Nuismer 2004). We make this assumption because it allows us to pick out the indirect selection that acts on the modifier, disentangling this from any direct selection arising through sexual selection. By taking it as given that frequency-dependent disruptive selection maintains a stable polymorphism at intermediate

frequency, we study the most favorable situation in which increased assortment can evolve. The associations within an individual are denoted D_{jk} and D_{ijk}. If there is to be a stable equilibrium for the focal genes, then there must be no directional selection at these loci: $a_j = a_k = 0$. The modifier is assumed to be free of direct selection (i.e., all selection coefficients where the modifier and focal-trait positions have the same sex of carrier are also 0). Finally, because this is a model of assortment (as opposed to a preference-trait system), selection coefficients in which the modifier has one sex of carrier and all the focal-trait positions the other vanish. Selection coefficients like $a_{j_\mathrm{f}j_\mathrm{m}k_\mathrm{m}}$ (one focal locus has one sex of carrier and two focal loci the other) are 0 because the system is in a symmetric equilibrium ($p_j = p_k = 1/2$).

Natural selection and sexual selection acting on individuals generate selection on the focal positions. The corresponding selection coefficients are $a_{j_\mathrm{f}k_\mathrm{f}} = a_{j_\mathrm{m}k_\mathrm{m}}$, which for compactness we denote a_2. (Note that $a_{k_\mathrm{f}j_\mathrm{f}} = a_{j_\mathrm{f}k_\mathrm{f}}$ because they contain the same elements.) Nonrandom mating produces five nonzero selection coefficients. Because we have assumed that the focal loci are equivalent, four of these are equal, $a_{j_\mathrm{f}j_\mathrm{m}} = a_{k_\mathrm{f}k_\mathrm{m}} = a_{j_\mathrm{f}k_\mathrm{m}} = a_{k_\mathrm{f}j_\mathrm{m}}$, and we denote these $a_{1,1}$. The fifth selection coefficient involving only focal positions is $a_{j_\mathrm{f}k_\mathrm{f}j_\mathrm{m}k_\mathrm{m}}$, which we denote $a_{2,2}$.

In another category of selection coefficients are those that include one modifier position. These arise because the modifier changes the frequencies of mating between individuals with different genotypes at focal loci. Eight of these have two focal positions with different sexes of carrier (e.g., $a_{i_\mathrm{f}j_\mathrm{f}j_\mathrm{m}}$, $a_{i_\mathrm{f}j_\mathrm{f}k_\mathrm{f}}$, $a_{i_\mathrm{m}k_\mathrm{f}j_\mathrm{m}}$, etc.), and we denote these $a_{i1,1}$. The last nonzero selection coefficients are $a_{i_\mathrm{f}j_\mathrm{f}k_\mathrm{f}j_\mathrm{m}k_\mathrm{m}}$ and $a_{i_\mathrm{m}j_\mathrm{f}k_\mathrm{f}j_\mathrm{m}k_\mathrm{m}}$. These two are equal because of symmetry, and we denote them $a_{i2,2}$.

The dynamics of the modifier are given by equation (1), which in the notation of this example is

$$\Delta p_i = a_2\tilde{D}_{ijk} + a_{2,2}\tilde{D}_{ijk}\hat{D}_{jk} + a_{i2,2}pq_i\hat{D}_{jk}^2 + O(\epsilon^2). \quad (3)$$

This result shows that the modifier evolves as the result of three kinds of selection. The first, represented by a_2, is selection on the focal trait coming from natural and sexual selection on individuals. This occurs when there is epistasis for fitness between these genes, for example, when they contribute to a quantitative trait that is under stabilizing or disruptive selection. By epistasis for fitness, we mean that the trait may be determined additively but that fitness on the trait is not additive, that is, is not the sum of marginal fitnesses across loci. The second kind of selection, represented by $a_{2,2}$, occurs when the degree of assortment is influenced by epistatic interactions between genes expressed in both individuals of a mated pair. The third kind, represented by $a_{i2,2}$, occurs when the modifier alters the strength of that kind of epistatic selection.

The modifier's dynamics also depend on the associations, as we noted earlier. In this model there are two of them, one for the linkage disequilibrium between the focal loci ($\hat{D}_{jk}$) and one for the linkage disequilibrium between all three loci ($\tilde{D}_{ijk}$). We calculate their values in appendix B in the online edition of the *American Naturalist*. Substituting those results into equation (3) gives

$$\Delta p_i = \left[a_{i1,1}\frac{K_1}{R} + a_{i2,2}\left(\hat{D}_{jk}^2 + \frac{K_2}{R}\right)\right]pq_i, \quad (4)$$

where

$$K_1 = r_{jk}\left(\frac{1}{4} + \hat{D}_{jk}\right)^2\left(a_2 + a_{2,2}\hat{D}_{jk}\right), \quad (5)$$

$$K_2 = (1 - r_{jk})\frac{\hat{D}_{jk}}{16}\left(a_2 + a_{2,2}\hat{D}_{jk}\right), \quad (6)$$

and

$$R = 1 - (1 - r_{jk})(1 - r_{ij})\left(1 - a_2\hat{D}_{jk} - a_{2,2}\hat{D}_{jk}^2\right) - \left[r_{jk}a_{1,1}\left(\frac{1}{4} + \hat{D}_{jk}\right) + (1 - r_{jk})r_{ij}\left(a_2\hat{D}_{jk} + \frac{1}{16}a_{2,2}\right)\right] \quad (7)$$

is the chance that there will be a recombination event among loci i, j, and k, $1 - (1 - r_{ij})(1 - r_{jk})$, modulated by changes in genotype frequencies caused by selection. The recursion for associations between the selected loci j and k is given by

$$\hat{D}_{jk} = (1 - r_{jk})\left[\hat{D}_{jk} + \left(\frac{1}{16} - \hat{D}_{jk}^2\right)\left(a_2 + a_{2,2}\hat{D}_{jk}\right)\right] + r_{jk}a_{1,1}\left(\frac{1}{4} + \hat{D}_{jk}\right)^2, \quad (8)$$

which gives an implicit solution for the equilibrium $\hat{D}_{jk}$; r_{jk} and r_{ij} are recombination rates between loci j and k and between i and j, respectively. The two terms in equation (4) correspond to selection for stronger assortment between single genes and selection between pairs of genes. Both depend almost entirely on the composite coefficient of epistasis $a_2^* = a_2 + a_{2,2}\hat{D}_{jk}$, the only exception being a weak influence of $a_{2,2}$ on the effective recombination rate R. Equations (5) and (6) show that evolution of the modifier depends on how it alters the intensity of nonrandom mating between

Table 1: Relative contribution of each pair of haploid genotypes, in terms of the general coefficients $a_{1,1}$, a_2, and $a_{2,2}$

	00	01	11
00	$1 + a_{1,1} + a_2/2 + a_{2,2}/16$	$1 - a_{2,2}/16$	$1 - a_{1,1} + a_2/2 + a_{2,2}/16$
01	$1 - a_{2,2}/16$	$1 - a_2/2 + a_{2,2}/16$	$1 - a_{2,2}/16$
11	$1 - a_{1,1} + a_2/2 + a_{2,2}/16$	$1 - a_{2,2}/16$	$1 + a_{1,1} + a_2/2 + a_{2,2}/16$

single genes carried by mated pairs, $a_{i1,1}$, and between pairs of genes carried by mated pairs, $a_{i2,2}$.

Interpretation of the selection and assortment coefficients in the symmetric haploid model. The relative contributions of each of the 16 possible pairs of genotypes are given in table 1 (genotypes 10 and 01 are equivalent and so are shown together). The coefficient $a_{1,1}$ describes assortment between genotypes with similar numbers of 0 and 1 alleles. Coefficient a_2 describes disruptive selection in favor of extreme genotypes. Finally, $a_{2,2}$ describes the contribution of matings between intermediates, relative to the other genotypes. Note that with the assumption of complete symmetry, there are only three remaining degrees of freedom, and so this is the most general model possible.

Since no pair of genotypes can make a negative contribution, table 1 constrains the values of the coefficients. Necessarily, $|a_{2,2}| < 1/16$, $-4 < a_{1,1}$, and $a_2 < 4$, but the allowable region is a tetrahedral region that is considerably smaller than these ranges (fig. 1). Note that the coefficient $a_{1,1}$ is equal to the covariance in trait value between mates, divided by the square of the trait variance. This is not dimensionless, and it is not equal to the correlation coefficient.

We can gain some insight into the selection on assortment, $a_{i1,1}$, by finding the average fitness of offspring of an individual that slightly varies its strength of assortment $a_{1,1}$. The average fitness of genotypes 00, 11 is $1 + a_2[(1/4) + \hat{D}^2_{jk}]$, and that of genotypes 01, 10 is $1 - a_2[(1/4) - \hat{D}^2_{jk}] - a_{2,2}(\hat{D}^2_{jk}/4)$. Offspring of matings between 00 and 00 are entirely 00; those of matings between 00 and 01 are half 00, half 01; those of matings between 00 and 11 are 00 or 11 with probability $1 - r_{jk}$ and otherwise are 01 or 10; and so on. By multiplying the average fitness of offspring from each possible mating by the corresponding weighting of the coefficient $a_{1,1}$ in table 1 and by the genotype frequencies, we find that the increase in mean fitness of offspring gained per increase in assortment is just equal to K_1. More directly, K_1 is the increase in average relative fitness of offspring due to an additional copy of the modifier allele, per $a_{i1,1}$. We see that the selection coefficient for assortment for the trait (the first term in eq. [4]) is equal to $a_{i1,1}K_1/R$ and is just the ratio of the increase in mean offspring fitness due to the increased assortment to the rate at which the modifier is dissociated from the pair of selected loci. Regrettably, there seems to be no such simple relation with the selection on assortment $a_{2,2}$ for pairs of selected loci (or, equivalently, for variance in the selected trait).

Results for the two-locus example. Equation (4) shows that the fate of the modifier is determined by the two selection coefficients $a_{i1,1}$ and $a_{i2,2}$ and by the terms K_1/R and $K_2/R + \hat{D}^2_{jk}$ (given by eqq. [5], [6]). The latter depend, in turn, on the association between the focal loci, $\hat{D}_{jk}$, which is given by equation (8). Figure 2 shows how these three quantities depend on the selection coefficients.

Figure 2*a*–2*c* shows the maximum value at equilibrium of the association D_{jk} for free recombination ($r_{jk} = 1/2$, $r_{ij} = 1/2$). As expected, the value of this linkage disequilibrium increases with the strength of assortment ($a_{1,1}$) and with the strength of viability selection a_2 (positive values of a_2 corresponds to disruptive selection, while negative values of a_2 correspond to stabilizing selection). Using the equilibrium value $\hat{D}_{jk}$, which depends on the value of the selection coefficients acting on the focal trait ($a_{1,1}$, a_2, $a_{2,2}$), we can determine the fate of the modifier. For that we look at the signs of K_1/R and $K_2/R + \hat{D}^2_{jk}$ separately. Figure 2*d*–2*f* shows contour plots of K_1/R for free recombination, depending on the selection coefficients, and figure 2*g*–2*i* shows similar plots for $K_2/R + \hat{D}^2_{jk}$. The contour plots are such that in regions where K_1/R or $K_2/R + \hat{D}^2_{jk}$

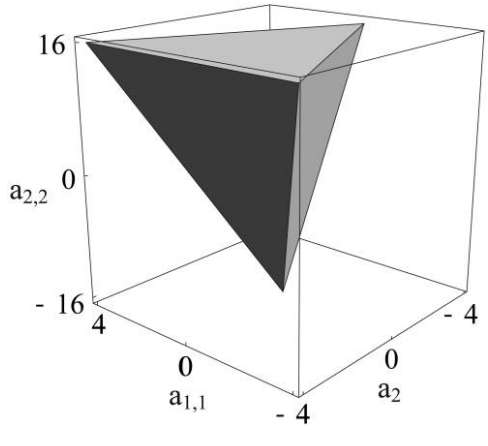

Figure 1: Allowable combinations of coefficients $a_{1,1}$, a_2, and $a_{2,2}$ in the symmetric haploid model.

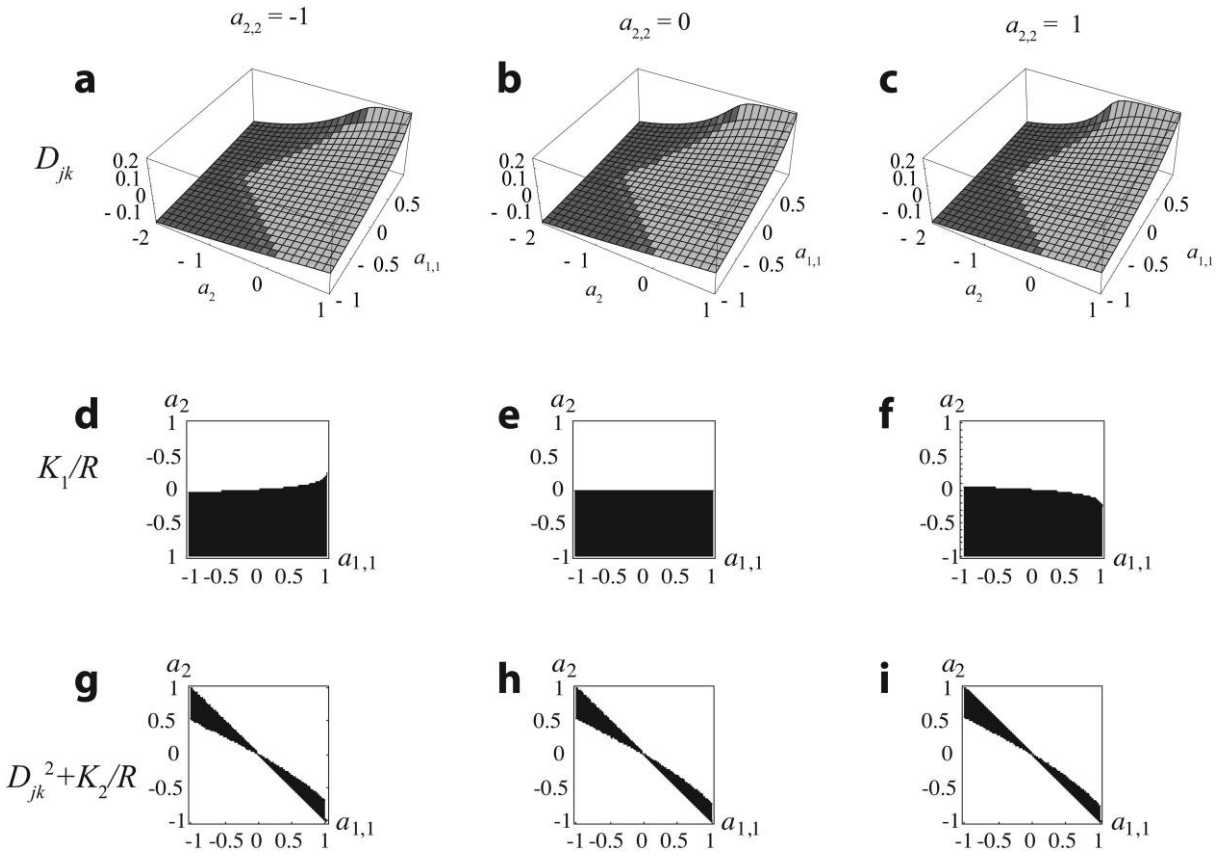

Figure 2: Modifier invasion depending on the selection coefficients $a_{1,1}$, a_2, and $a_{2,2}$. *a–c*, Maximum value of D_{jk} at equilibrium versus a_2 and $a_{1,1}$ for $a_{2,2} = -1$ (*a*), $a_{2,2} = 0$ (*b*), and $a_{2,2} = 1$ (*c*). Areas where fitness for that combination of $a_{1,1}$, a_2, $a_{2,2}$, and D_{jk} are negative (and thus biologically unrealistic) are in dark gray, and regions of positive fitness are in light gray. *d–f*, Contour plots for K_1 versus a_2 and $a_{1,1}$ for $a_{2,2} = -1$ (*d*), $a_{2,2} = 0$ (*e*), and $a_{2,2} = 1$ (*f*). Parameter combinations for which K_1 is positive (modifier invades) are in white, and combinations for which K_1 is negative (modifier does not invade) are in black. *g–i*, Contour plots for K_2 versus a_2 and $a_{1,1}$ for $a_{2,2} = -1$ (*g*), $a_{2,2} = 0$, (*h*), and $a_{2,2} = 1$ (*i*). Parameter combinations for which K_2 is positive (modifier invades) are in white, and combinations for which K_2 is negative (modifier does not invade) are in black.

are negative, the area is colored in black, while regions in which they are positive are in white.

The sign of K_1/R depends crucially on the sign of the selection coefficient a_2 (on the *Y*-axis) in figure 2*d*–2*f*. This selection coefficient represents epistasis between the loci that control the focal trait. One example of positive epistasis is disruptive selection, so positive a_2 could be due to selection against intermediate values of the traits. Negative epistasis (negative a_2) can be due to stabilizing selection. Thus, we can see from figure 2*d*–2*f* that for the modifier to invade, positive epistasis, such as disruptive selection, seems to be crucial. This seems considerably more important than the initial strength of assortment, which is measured by $a_{1,1}$, where positive values of $a_{1,1}$ are associated with assortative mating, while negative values of $a_{1,1}$ are due to disassortative mating.

The analysis of the sign of $K_2/R + \hat{D}_{jk}^2$ is less straightforward. However, although we do not present here the values of $K_2/R + \hat{D}_{jk}^2$ but only its sign, it is in general considerably much smaller than K_1/R, and thus we expect the fate of the modifier to be mainly controlled by K_1/R. We have plotted here a wide range of selection coefficients that goes from strong disassortment (negative $a_{1,1}$) to strong assortment (positive $a_{1,1}$) and from strong stabilizing selection (negative a_2) to strong disruptive selection (positive a_2).

To sum up, the fate of a modifier is determined by equation (4), which consists of two terms. Analysis shows that the first term that appears there, $a_{i1,1}K_1/R$, is typically much larger than the second, $a_{i2,2}(K_2/R + \hat{D}_{jk}^2)$. Thus, a modifier will generally spread when $a_{i1,1}K_1/R$ is positive. The sign of K_1/R, in turn, can be determined from figure 2.

This completes our analysis of arbitrary assortment

evolving in a simple genetic system. An alternative way to make progress is to commit to particular rules for how natural selection acts and how mating pairs are formed. Although restrictive, this is a useful exercise because it illuminates its relationship to our general model. We therefore turn to that approach next.

Specific Assortment Functions

In this section, we explore how assortment evolves under two simple rules for how mated pairs are formed. We do so because we want to show how the general framework developed above can be applied to specific cases. The mating rules studied in this section have been used in previous studies, which gives us an opportunity to compare results.

Simple rules for pair formation typically generate selective mating, in which some genotypes have higher reproductive success than others. Consequently, the modifiers experience direct selection. This force will often overwhelm indirect selection and dominate the evolution of the modifiers. Previous work on speciation has focused interest on how indirect selection on the focal genes can cause assortment to evolve (Doebeli 1996; Dieckmann and Doebeli 1999; Doebeli and Dieckmann 2000). These models have normalized the frequencies of mated pairs so that all genotypes have equal mating success and direct selection is eliminated. This scheme has the drawback that it is biologically quite restrictive, as it is satisfied only if the population is strictly monogamous. On the other hand, this scheme has clearly shown the role of indirect selection in the evolution of the mating system.

In this section, we consider both possibilities, that is, schemes where there is direct selection and schemes where it is eliminated by a normalization procedure. The normalization can be done only numerically, which precludes general analytic results. We are, however, able to get analytic results when selective mating is allowed.

The section begins by outlining the assumptions for the two assortment rules. We study these in the context of two genetic models, whose assumptions we next describe. The four combinations of mating rules and genetic models are then analyzed without assuming a normalization that eliminates selective mating. Last is an analysis of these four combinations with the normalization.

Assumptions. The population is such that the sexes are symmetric, as in a hermaphroditic population, and mating is governed by a quantitative trait z expressed equally in both partners. We consider two genetic systems. The haploid model is the same one studied in the previous section, with two loci (j and k) for the focal trait and one for the modifier of assortment. In the diploid model, the focal trait is governed by a single locus j. In both models, we visualize the focal genes as contributing additively to z. An individual's phenotype is the sum of its allelic values for its focal genes, so that z ranges from 0 to 2.

The life cycle consists of viability selection followed by mating. Individuals with intermediate phenotypes ($z = 1$) have viability $1 - S$ relative to those with extreme phenotypes ($z = 0$ and $z = 2$), whose viabilities are equal. A positive value of S corresponds to disruptive selection on the focal trait. As in the last section, we assume some form of selection that stabilizes frequencies for the focal alleles at 1/2. (Note that this equilibrium can be stable even when intermediates or heterozygotes have lower fitness if selection is frequency dependent, as in models of competitive speciation.)

We study two rules for how mated pairs are formed. Under Gaussian assortment, the probability that two individuals mate falls off as a Gaussian function of the difference in z between the individuals. Under quadratic assortment, that probability declines inversely with the square of the difference in z values. Under both Gaussian and quadratic selection, the parameter A measures the strength of assortment. A value of $A = 0$ corresponds to random mating, and increasing positive values produce stronger correlations between the values of z in mated pairs. Table 2 shows the probabilities that a pair of individuals with given genotypes will mate, under Gaussian assortment or quadratic assortment. The frequency of matings between individuals with genotype x and individuals with genotype y is $M_{x,y}$, which is proportional to $f_x^* f_y^* C_{x,y}$, where f_x^* is the frequency of genotype x after viability selection and $C_{x,y}$ is the appropriate entry from table 2.

The assortment modifier is a single locus i. The genes at this locus have additive effects in the following sense. Let n be the number of 1 alleles at the modifier locus carried by both individuals of a pair that might mate. The strength of the assortment parameter for these individuals is then $A + n\delta A$. We assume that the modifier has a weak effect. Specifically, the parameter δA, which measures the

Table 2: Mating probabilities for Gaussian and quadratic assortment

	00	01	10	11
Gaussian assortment:				
00	1	$1 - A$	$1 - A$	$(1 - A)^4$
01	$1 - A$	1	1	$1 - A$
10	$1 - A$	1	1	$1 - A$
11	$(1 - A)^4$	$1 - A$	$1 - A$	1
Quadratic assortment:				
00	1	$1 - A$	$1 - A$	$1 - 4A$
01	$1 - A$	1	1	$1 - A$
10	$1 - A$	1	1	$1 - A$
11	$1 - 4A$	$1 - A$	$1 - A$	1

strength of the modifier, is much smaller than A. We emphasize that before the modifier appears, A may be large or small, positive or negative. This allows us to study the evolution of strong assortment even when the modifier has weak effects.

We focus attention on the case where the population is initially fixed at the modifier locus and at equilibrium for the focal trait. A mutation appears at the modifier locus, and we determine whether it spreads through the population. That is done by calculating the selection coefficients (i.e., the a described in the previous section) and then substituting them into equation (2). The selection coefficients are calculated using the method described by Kirkpatrick et al. (2002) and are given in appendix C in the online edition of the *American Naturalist*. In short, this method involves writing the fitness expression in a polynomial form, as in a series expansion, and the selection coefficients are the coefficients of each term on the series. Once we know the selection coefficients for a fitness scheme, we can calculate changes in allele frequencies and associations after selection and recombination. In the next section, we analyze the four combinations of assumptions (two assortment functions and two genetic systems), assuming selective mating. The section following that one studies the outcomes when mating is nonselective.

Modifiers with Direct Selection. This section shows how modifiers evolve when they act according to the two mating functions (Gaussian and quadratic) described above and commonly used in models of assortative mating. As in the plant model of Kirkpatrick and Nuismer (2004), these rules generate sexual selection, such that not all genotypes have equal reproductive success (a similar case is $n = 1$ in the models of Gavrilets and Boake [1998] and Matessi et al. [2001]). A major finding is that for the simple genetic system we are considering, assortative mating never evolves. Instead, there is always selection to reduce assortment, and the evolutionary equilibrium is with random mating.

Under both rules, the approach is as follows. We find the rate of evolution of a modifier that alters the strength of assortment by substituting the corresponding selection coefficients (the a's) derived in appendix C into equation (2). That allows us to derive the equations for the rate of change of the modifier, scaled over its effect on the strength of assortment δA, $\Delta p_i/(\delta A p q_i)$, and the association between the modifier and the focal loci at QLE, $\tilde{D}_{ijk}$ and $\tilde{D}_{ij,j}$, which are given in appendix C. To check our analytical approach, we have done exact iterations, as described below (in "Simulations"). Note that for quadratic mating, the maximum value that assortment can take is such that $1 - 4A(A + n\delta A)$ has to remain positive, so in the absence of modifiers, A can take values of up to 0.25.

The first conclusion is that the modifier's rate of change ($\Delta p_i/pq_i$) is simply proportional to the size of its effect, δA. That is a consequence of our assumption that it has a weak effect and allows us to study the scaled rate of change ($\Delta p_i/(\delta A p q_i)$). The other parameters that determine the modifier's fate are the initial strength of assortment in the population (A), the strength of disruptive natural selection (S), and the recombination rates (r_{jk} and r_{ij}). The scaled rate of change is a convenient measurement because when it is positive, a modifier of increased assortment will invade, whereas if it is negative, only modifiers that decrease assortment can invade.

An important observation is that for some parameter values, there are two alternative stable equilibria that the population can attain before the modifier appears (Kirkpatrick and Ravigné 2002). At one equilibrium, the association between the focal positions ($\hat{D}_{jk}$ for the haploid model, $\hat{D}_{j,j}$ for the diploid model) may be small. This corresponds to a situation in which the focal trait is unimodal. At the second equilibrium, the association is large and positive. Here the focal trait is bimodal; that is, the population is close to fissioning into two species.

Figures 3 and 4 show the analytic approximation and the exact value of the scaled rate of change ($\Delta p_i/(\delta A p q_i)$) for the haploid and diploid models, respectively, for both the Gaussian and quadratic rules, assuming free recombination. As we can see from those figures, there is no regime of initial S or A that will favor a modifier of increased assortment. Selection on the modifier depends on the genetic composition of the population before the modifier arises. As shown in figures 3 and 4, the rate of change does not always vary monotonically with strength of selection. This is because there are two components that contribute to selection against the modifier. The first one is due to assortment only, which decreases the fraction of matings that any individual obtains, but it mostly affects intermediate traits. The second component is due to the combination of assortative mating with disruptive selection, which favors the production of better-adapted individuals (those with extreme traits). Therefore, if the initial composition is such that all genotypes are present in roughly the same frequency (as for small S and small A), then the modifier is strongly selected against. In other words, a modifier of increased assortment is not beneficial for the population as a whole, through its effects of individuals with genotypes 01 and 10. However, in the other extreme case where initially there are only extreme traits (00 and 11), which occurs for very large A, as in figure 3*a*, then increased assortment is neither favored nor selected against.

Simulations. As mentioned above, the analytic results are based on approximations, as derived in appendix C; these approximations are accurate in the limit of small modifier

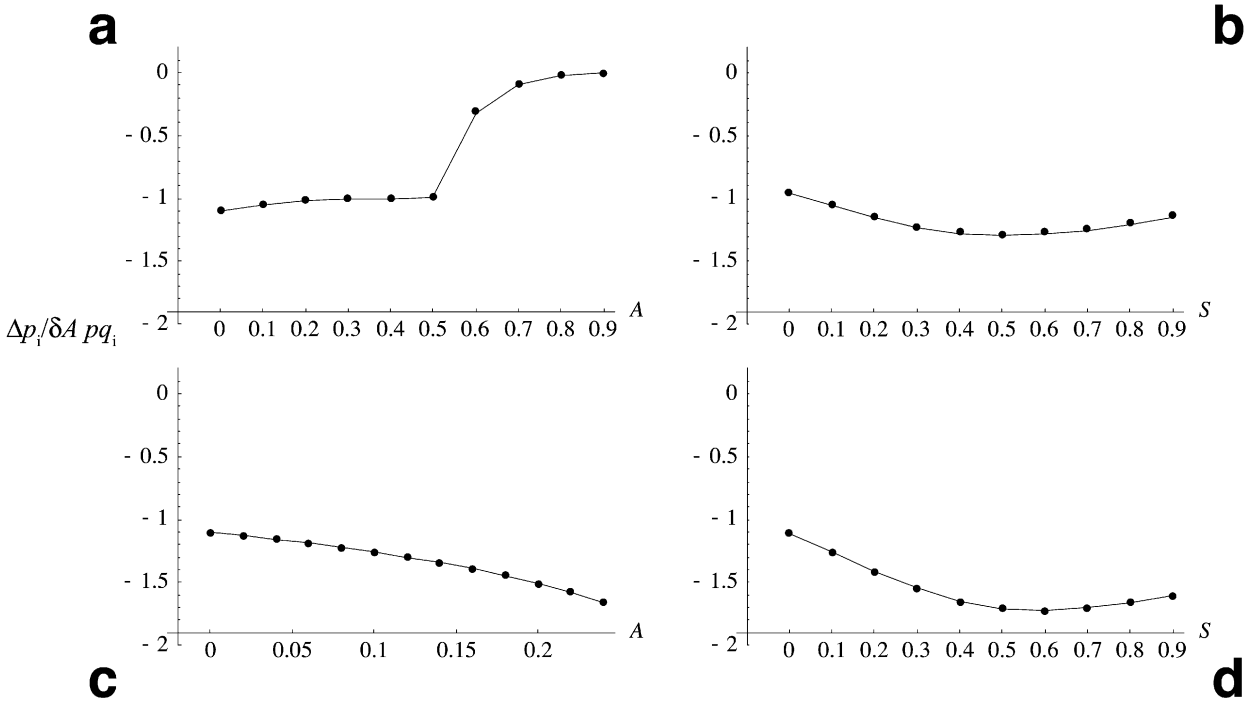

Figure 3: Scaled rate of change of a modifier of increased assortment, with direct selection, for a population of haploids. Exact results at equilibrium from numerics are shown by circles, and a quasi-linkage equilibrium approximation is shown as a line. *a*, *b*, Rate of change for individual mates following a Gaussian mating scheme and (*a*) varying the strength of initial assortment for disruptive selection, $S = 0.1$, or (*b*) varying the strength of initial disruptive selection for constant initial assortment, $A = 0.1$. *c*, *d*, Rate of change for individual mates following a quadratic mating scheme and (*c*) varying the strength of initial assortment for disruptive selection, $S = 0.1$, or (*d*) varying the strength of initial disruptive selection for constant initial assortment, $A = 0.1$.

effects. We can use numerical iterations of genotype frequencies to check the accuracy of the approximations for particular cases and to verify that there are no major errors in the calculations. We iterated genotype frequencies as follows. First, we iterated the genotype frequencies when the modifier allele 1 is absent until the system reached a stationary state (which typically takes about 50 generations). Second, we introduced the modifier allele 1 at a frequency of $p_i = 0.1$ and such that its association with the other loci (D_{ijk} in haploids, $D_{ij,j}$ in diploids) is initially 0 and let the system evolve until the scaled rate of change $\Delta p_i/(\delta A p q_i)$ reached an equilibrium. For haploids, we used a modifier with effect $\delta A = 0.01$, and for diploids we used $\delta A = 0.005$. In all our simulations, we held the allele frequencies at the focal loci fixed at 1/2.

Figures 3 and 4 show both the analytic approximation as lines, for which no δA is specified as the rate of change is scaled, and the exact results as circles. We can see that in general, the approximation works quite well. We can measure the accuracy of our approximation by measuring the relative difference between the exact result and the approximated result, such that accuracy is given by $(x_{\text{approx}} - x_{\text{exact}})/x_{\text{exact}}$. In general, the approximation works well, with accuracy of $\sim 10^{-2}$ for $\delta A = 0.01$. However, when both the approximation and the exact result are very small, the relative difference may become large (of order 10^{-1} to 1), because it is hard to compare two very small numbers close to 0.

Conclusions on modifiers with direct selection. The most striking result from the analyses is that we found no cases in which stronger assortment evolved. In all cases, modifiers that decrease assortment will invade, and so the evolutionarily stable state is random mating. The strength of selection against the modifier depends on how many individuals with intermediate traits are present in the population before the modifier invasion, because a modifier of increased assortment always decreases the fitness of these individuals and is thus selected against. When the population is composed exclusively of extreme individuals, then the modifier is neither selected against nor favored. This is simply because in this situation, there is no variation for the modifier to act on.

Pure Modifiers. In the previous section, we saw that the

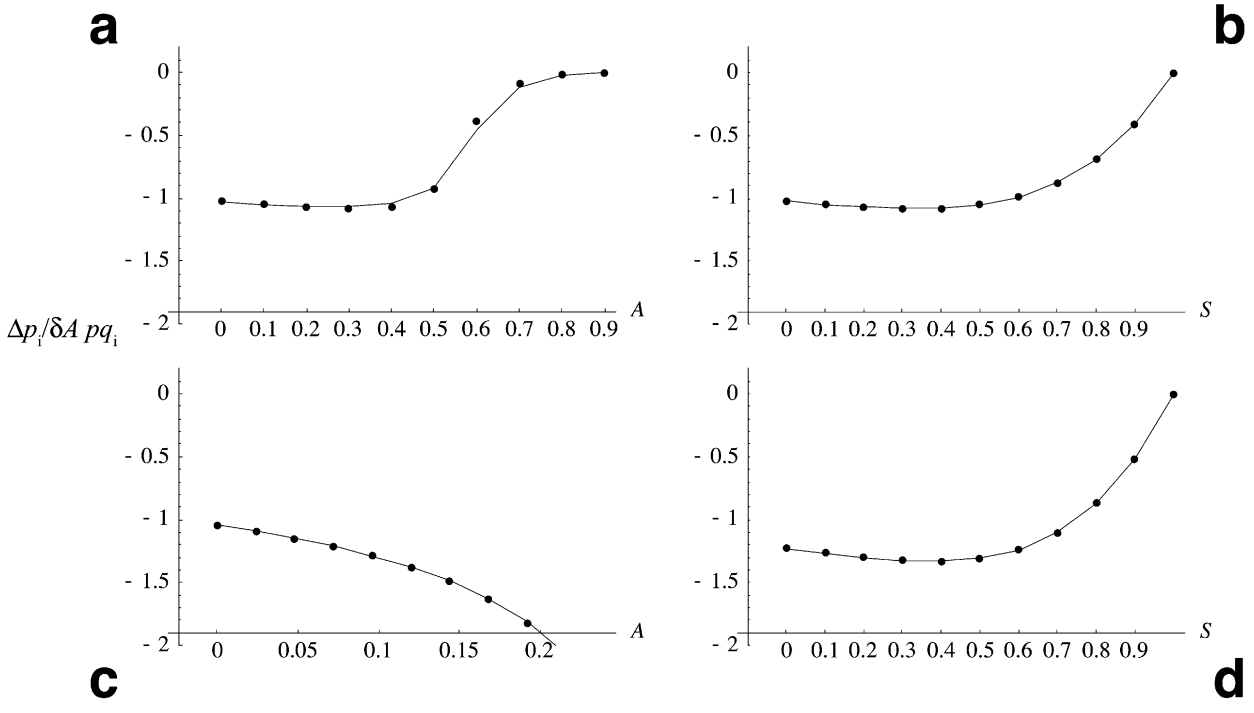

Figure 4: Scaled rate of change of a modifier of increased assortment, with direct selection, for a population of diploids. Exact results at equilibrium from numerics are shown by circles, and a quasi-linkage equilibrium approximation is shown as a continuous line. *a*, *b*, Rate of change for individual mates following a Gaussian mating scheme and (*a*) varying the strength of initial assortment for disruptive selection, $S = 0.1$, or (*b*) varying the strength of initial disruptive selection for constant initial assortment, $A = 0.1$. *c*, *d*, Rate of change for individual mates following a quadratic mating scheme and (*c*) varying the strength of initial assortment for disruptive selection, $S = 0.1$, or (*d*) varying the strength of initial disruptive selection for constant initial assortment, $A = 0.1$.

Gaussian and quadratic assortment functions can produce selective mating. If individuals prefer to mate with others that have a similar phenotype, then rare individuals may go unmated. This leads to direct selection on the assortment modifiers. To isolate the effect of indirect selection on the modifiers, we can alter the mating frequencies in such a way that all genotypes have the same reproductive success. This is the approach taken by Dieckmann and Doebeli (1999) and has been more explicitly studied by Polechova and Barton (2005). This normalization procedure is biologically very restrictive, as the only behavioral or physiological mechanism that could produce this outcome is strict monogamy. Even if restrictive or implausible, however, it is useful as a limiting case, because it is highly favorable to the evolution of assortative mating. It also makes the analysis difficult, and we are forced to partly abandon our analytic approach and turn to numerical methods (in contrast with the case presented here, Polechova and Barton [2005] were able to derive analytic approximations using the infinitesimal model). Nevertheless, we pursue this issue in order to learn more about the forces acting on assortment modifiers and to compare our results with earlier work. We emphasize that this normalization is not the same as the animal model of Kirkpatrick and Nuismer (2004) or the $n = \infty$ of Matessi et al. (2001), which has a long tradition in the sexual selection literature, where females are assured of mating but there is sexual selection on males.

The normalization is done such that the marginal fitnesses due to assortment only of all genotypes are the same. We do so by assuming that the assortment fitness of each genotype x is multiplied by a factor α_x, chosen such that

$$\sum_y \alpha_x \alpha_y f_y C_{x,y} = 1 \tag{9}$$

for each genotype x. This means that all individuals, regardless of their genotype, have the same mating success, which is consistent with the definition of nonselective mating. In contrast with the animal model of Kirkpatrick and Nuismer (2004) and other models of assortative mating (i.e., Matessi et al. 2001), our normalization does not induce sexual selection on males. The normalizing factors

α change as the population evolves. In terms of our analysis, this means that they must be calculated numerically in each generation.

We studied the evolution of the modifier by numerical simulation. The population was allowed to reach an initial equilibrium in the absence of genetic variation at the modifier locus by letting it evolve for 1,000 generations. A modifier allele that increased the strength of assortment was then introduced at a frequency of $p_i = 0.1$. We then let the population reach equilibrium by allowing it to evolve for another 100 generations, at which time the frequency of the modifier allele was increasing or decreasing at an approximately constant rate. At this point, we calculated the scaled rate of change of the modifier $(\Delta p_i/(\delta A p q_i))$.

Figures 5*a*–5*b* and 5*c*–5*d* show typical results for the Gaussian and the quadratic assortment functions, respectively, for the haploid model in which the trait is controlled by two loci. The modifier allele that increases assortment always invades or remains at the frequency at which it is introduced. The exact rate of change of the modifier depends on the initial linkage disequilibrium between the loci that control the focal trait, such that if very large (as for very large values of A for Gaussian assortment in fig. 5*a*), the modifier is neutral and the rate of change is close to 0. The rate of change of the modifier is largest when the initial linkage disequilibrium has an intermediate value because it favors an increased association between the loci controlling the focal trait.

Similar results are shown in figure 6 for a diploid population, where the focal trait is controlled by a single locus j. The modifier locus i is also biallelic and modifier allele 1 increases assortment to $A + n\delta A$, where n is the number of modifier alleles 1 in the mating pair. As in haploids, we studied the fate of the modifier numerically, as the normalization scheme makes the analysis intractable. Initially, we let the population evolve in the absence of modifier allele 1 for 1,000 generations, after which we were certain that the population was in equilibrium. A modifier allele that increases assortment was then introduced at a frequency $p_i = 0.1$, and we let the population evolve for a further 100 generations, during which the frequency of allele 1 at the modifier locus changed at a roughly constant rate. We then measured the scaled rate of change of the

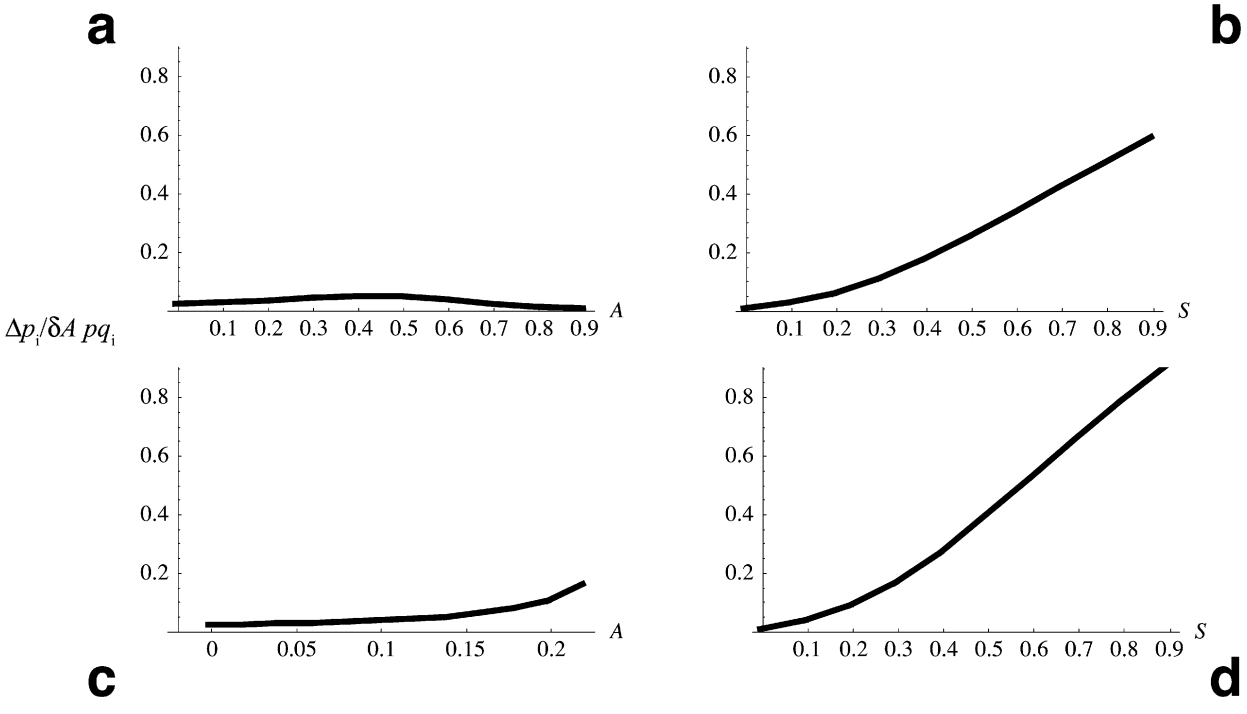

Figure 5: Scaled rate of change of a modifier of increased assortment, when mating is nonselective, for a population of haploids. *a*, *b*, Rate of change for individual mates following a Gaussian mating scheme, normalized such that all individuals have the same marginal fitness and (*a*) varying the strength of initial assortment for disruptive selection, $S = 0.1$, or (*b*) varying strength of initial disruptive selection for constant initial assortment, $A = 0.1$. *c*, *d*, Rate of change for individual mates following a quadratic mating scheme, normalized such that all individuals have the same marginal fitness and (*c*) varying the strength of initial assortment for disruptive selection, $S = 0.1$, or (*d*) varying strength of initial disruptive selection for constant initial assortment, $A = 0.1$.

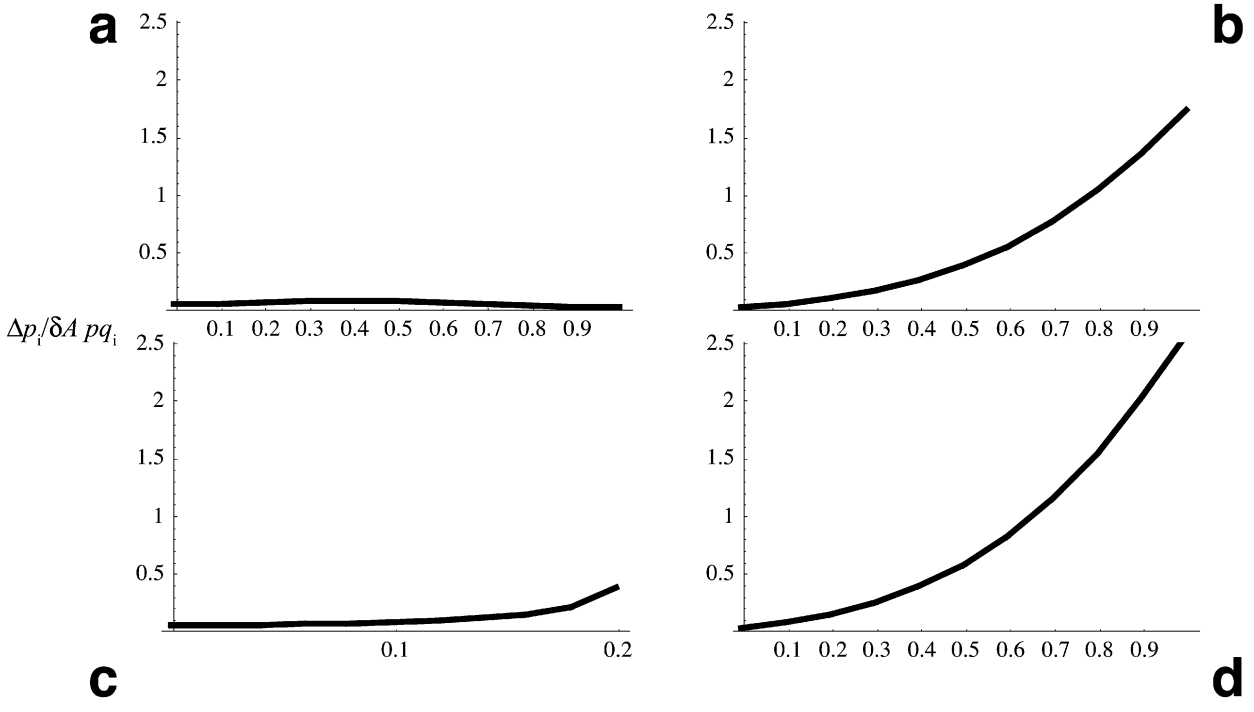

Figure 6: Scaled rate of change of a modifier of increased assortment, when mating is nonselective, for a population of diploids. *a*, *b*, Rate of change for individuals mate following a Gaussian mating scheme, normalized such that all individuals have the same marginal fitness and (*a*) varying the strength of initial assortment for disruptive selection, $S = 0.1$, or (*b*) varying strength of initial disruptive selection for constant initial assortment, $A = 0.1$. *c*, *d*, Rate of change for individuals mate following a quadratic mating scheme, normalized such that all individuals have the same marginal fitness and (*c*) varying the strength of initial assortment for disruptive selection, $S = 0.1$, or (*d*) varying strength of initial disruptive selection for constant initial assortment, $A = 0.1$.

modifier allele 1, $\Delta p_i/(\delta A p q_i)$. The results are qualitatively similar to those for the haploid scenario. Figure 6*a*–6*b* illustrates the scaled rate of change for the Gaussian mating scheme, and figure 6*c*–6*d* does so for the quadratic mating scheme. The modifier allele 1 always invades or remains at a constant frequency. The exact rate of change depends on the initial association between positions controlling the focal trait $D_{j,j}$, such that when this association has a very large value (close to 0.25), then the rate of change is close to 0 (as in fig. 6*a* for large values of A). The maximum rate of change is achieved when the association $D_{j,j}$ has initially intermediate values (0.10–0.15), as the modifier favors an increased association between the positions controlling the focal trait.

In summary, a modifier that increases assortment spreads whenever the focal trait is under disruptive viability selection (positive S). The only exception to this pattern occurs when the population has already fissioned into two modes, that is, when the initial disequilibrium between loci or positions controlling the focal trait is maximal (0.25). In this case, a modifier that increases assortment is neutral because there is no fitness variation at the focal trait that can cause the modifier to evolve by indirect selection. We have checked the robustness of our results by using the assumption of the animal model, that is, that all females mate but not all males. The cost induced by males not mating is so large that increased assortment will not invade (results not shown).

Discussion

The evolution of prezygotic isolation is the key process in sympatric speciation and reinforcement. However, conditions for increased assortment to evolve have not been clearly defined. We have used a novel method (Kirkpatrick et al. 2002) to describe the evolution of a modifier that controls the strength of assortment on one focal trait. We have derived the conditions for this modifier to invade and its final effect on the genetic composition of the population.

We began by deriving the conditions for increased assortment to evolve under rather general assumptions. In our haploid model, epistasis for fitness is necessary for increased assortment to evolve. By epistasis for fitness, we

mean that fitness is not the sum of the marginal fitnesses of each locus that determines the trait.

We then examined some specific examples. We found that selection for increased assortment is weak, and so in most models of assortative mating, it is easily overwhelmed by the cost of increased assortment. Thus, in our framework, modifiers of increased assortment will invade only if mating is nonselective. There may be an intermediate case between selective and nonselective mating where a modifier may invade, but we have not found it.

We emphasize that only a few studies have looked at the evolution of assortative mating explicitly. This contrasts with the abundant literature on assortative mating itself and its importance for speciation and reinforcement. The latter class of articles has focused on the consequences of assortative mating for the population via allelic frequencies, linkage disequilibrium, and correlations between traits (Fisher 1918) and whether it can lead to reproductive isolation (Kirkpatrick and Ravigné 2002). In this article, we have asked how assortment itself evolves.

Comparison with Previous Literature on Evolution of Assortment: Classification of Models

In order to compare our study with existing literature on evolution of assortment, it is helpful to outline and separate the different assumptions underlying different models (for a more detailed review, see chap. 10 of Gavrilets 2004). First is a useful reminder on the meaning of assortative mating as studied here: assortative mating is nonrandom mating that generates a correlation between the phenotypes of individuals in the mating pair. We are thus not dealing with nonrandom mating generated by sexual selection, where females have a preference for a male trait or there is competition for mates. Second, note that in our classification of models, we include only those where assortment can explicitly evolve, to avoid confusion with the vast literature that deals with fixed strength of assortment and its consequences for the genetic composition of the population and reproductive isolation.

In general, there can be up to three sets of loci, and these may or may not overlap. The first one, F, affects a focal trait that is under natural and sexual selection. The second set of loci, A, is under assortative mating. Finally, a third set, M, consists of the modifiers that control the strength of assortment on set A. Our distinction between models of assortment relies on the overlap between these sets. We follow Fry's (2003) classification of models as one- and two-trait models, as mentioned in the introduction to this article. Table 3 includes all models of evolution of assortative mating according to whether they are one- or two-trait models.

We call two-trait models those in which the loci under assortative mating are not under natural and sexual selection ($A \cap F = \varnothing$). Probably the simplest and best-known two-trait model is that proposed by Felsenstein (1981). In his study, two loci are under disruptive selection (F), and one locus controls assortative mating (A). He first studied the consequences of disruptive selection and assortative mating on speciation and found that for speciation to occur, assortment had to be strong. He then introduced a fourth locus, which controlled the strength of assortment (set M). If the increased-assortment allele grows in frequency, assortment will evolve. Felsenstein found that selection for increased assortment was weak.

We call one-trait models those in which the loci under assortative mating are also under natural and sexual selection ($A \subseteq F$). To our knowledge, they were proposed by Maynard Smith (1966).

The main difference between one- and two-trait models of evolution of assortment is that in the first, linkage disequilibrium emerges naturally, while in the second kind, selection or assortment generally has to be strong for linkage disequilibrium to increase in magnitude. This is true both for the study of the consequences of assortative mating on speciation and reinforcement and for assortment itself to evolve.

One-trait models of evolution of assortment have differed in their focus (see Gavrilets 2004, pp. 367–396). With the exception of those of Sawyer and Hartl (1981) and Matessi et al. (2001), all the one-trait models mentioned in table 3 have focused on the consequences of assortative mating and its evolution on speciation and on reinforcement after secondary contact. The overall conclusion from these models is that assortment can evolve but that it generally requires strong selection against hybrids/heter-

Table 3: Classification of models of evolution of assortative mating

	Assumptions	Models of this kind
Two-trait models	$F \cap A = \varnothing$, $A \cap M = \varnothing$	Maynard Smith 1966; Felsenstein 1981; Dieckmann and Doebeli 1999
One-trait models	$A \subseteq F$, $A \cap M = \varnothing$	Endler 1977; Sanderson 1989; Doebeli 1996; Cain et al. 1999; Dieckmann and Doebeli 1999; Servedio 2000; Matessi et al. 2001

Note. F = loci under natural and sexual selection; A = loci under assortative mating; M = modifiers that control the strength of assortment on the set A.

ozygotes. Furthermore, when assortment is not cost free (as studied by Sanderson [1989]), the conditions become more stringent. In general, selection for increased assortment is weak.

In contrast with these overall conclusions, the study of Dieckmann and Doebeli (1999) suggested that evolution for increased assortment occurred easily. Although they did not quantify the selection for increased assortment, it seemed that, since it occurred so easily, it should be strong and such that assortment would evolve rapidly. However, as their study relied on simulations, it is difficult to know the source of selection for increased assortment and whether it really occurred over a broad range of parameters.

The study closest to ours is the one-trait model of Matessi et al. (2001), who studied not only whether increased assortment is favored but also whether it will increase indefinitely to fixation. Their motivation and idea are similar to ours, but their approach and some of their assumptions are rather different. Their study assumed one focal trait under disruptive selection, controlled by one diploid locus. Mating is by similarity, where females have a preference to mate with males that are like them. Matessi et al. (2001) explored mating schemes that induced sexual selection in both sexes, as in the plant model of Kirkpatrick and Nuismer (2004), but most of the analysis was done for a mating scheme in which females were assured of mating but sexual selection acts on males, as in the animal model of Kirkpatrick and Nuismer (2004). This contrasts with our nonselective mating scheme, where all individuals have the same number of offspring and mating does not induce selection in females or males. The main conclusion of their article is that, under some conditions, increased assortment can invade and go to fixation. However, in contrast with our results, Matessi et al. found that there can be an interval in the values of the initial assortment, such that modifiers of increased assortment will not invade. In that situation, assortment would evolve toward an intermediate value. In their study, Matessi et al. assumed that the focal trait was always at an evolutionarily stable state. For the invasion analysis, Matessi et al. assumed Gaussian assortment and a modifier of weak effects, and for the fixation analysis, they assumed partial dominance. Because of this latter assumption and because there is always sexual selection against males, the modifier does not necessarily go to fixation. So for a modifier of increased assortment to invade and go to fixation, it has to be nonselective for all individuals.

Two recent studies on this topic have come to our attention. The first one, by Pennings et al. (2008), follows the line of research of Dieckmann and Doebeli (1999) and Matessi et al. (2001) in order to find the evolutionary stable level of assortment and whether it leads to complete reproductive isolation. Pennings et al. found that assortment does not always evolve and that when it does, it may remain at a level where isolation is not complete. The second study, by S. P. Otto, M. R. Servedio, and S. L. Nuismer (unpublished manuscript), is quite complementary to the one presented here, and their conclusions are very similar to ours (except that they focused on one focal trait controlled by one diploid locus). The main difference between our study and those of Pennings et al. and Otto et al. is that we are able to find only two extreme results, assortment evolving indefinitely and assortment never evolving, while Pennings et al. and Otto et al. have been able by different means to find intermediate costs of assortment where it evolves up to a point. This difference seems to come from the explicit frequency-dependent selection that both these models include.

Significance of Our Results

As mentioned above, most previous results on the evolution of assortment have relied on numerical simulations. Thus, it is not clear why the study of Dieckmann and Doebeli (1999) comes to conclusions different from those of previous studies of a similar scenario (Sawyer and Hartl 1981; Sanderson 1989; Cain et al. 1999; Servedio 2000; Matessi et al. 2001). Our study is the closest to a full analytical answer to when increased assortment will evolve. Furthermore, for part of our analysis, we did not assume any specific fitness form, which cannot be said of previous studies. This approach, although less explicit, helps in understanding the structure of previous models. This has allowed us to provide a clear picture of when assortment can evolve.

As expected, we have found that when selection acts on two haploid loci, epistasis for fitness is crucial for increased assortment to invade. Similarly, underdominance at one selected diploid locus would be required for assortment to evolve. This is because positive epistasis, or underdominance, brings jointly favored genes together, which, in turn, favors assortment between these gene combinations. One source of epistasis is disruptive natural selection on an additive trait. Since disruptive selection is relatively frequent in natural populations (Kingsolver et al. 2001), evolution for increased assortment may be more frequent than previously thought. Furthermore, there are many other sources of epistasis, so our conclusions extend further than only disruptive selection.

As we have seen, the selection coefficient favoring assortment is proportional to the increase in fitness of offspring that are produced assortatively (eq. [4])—a quantity that is, in principle, measurable. Whether assortment will evolve depends on the counterbalancing selection that may arise through effects of the modifier on sexual selection.

By analyzing specific fitness schemes, we have found that selection for increased assortment is weak unless disruptive selection is initially very strong. We have seen that even in the most favorable conditions, when assortative mating is initially strong and there is epistasis via disruptive selection, increased assortment will evolve extremely slowly. For instance, when mating is nonselective and follows a Gaussian scheme, for initial strong assortment $A = 0.4$ and disruptive selection $S = 0.1$, we may need about 100,000 generations for assortment to be complete. However, in agreement with our previous result that epistasis for fitness is crucial, for much weaker initial assortment but strong disruptive selection ($S = 0.9$, which effectively means that 90% of hybrids die before mating), we need only about 100 generations for assortment to be complete.

Crucially, when assortative mating is selective, that is, when not all individuals have the same number of offspring, we have not found any example under which assortment evolves. Although this conclusion is biologically intuitive, we examine it in more detail. Several articles on the consequences and evolution of assortative mating have made strong assumptions about the mating process, to the effect that the mating process was nonselective. For instance, the Dieckmann and Doebeli study of speciation via evolution of assortative mating implied that all individuals, regardless of their sex and genotype, had the same mating success (Dieckmann and Doebeli 1999, p. 357). That is, assortative mating was neutral and nonselective. Thus, their conclusion that assortment would evolve easily follows from the fact that they used nonselective assortment, and a relaxation of such an assumption would likely not allow assortment to evolve so easily (Bolnick 2004).

Limitations and Outlook

We have focused our study on the evolution of assortative mating in one-trait models. To make progress, we have made some strong assumptions. First, our results are limited to the case where the allelic frequencies of loci controlling the main trait are 1/2 and in equilibrium. This case implies initial large variance in the phenotype, and consequently, it favors the evolution of assortment. However, it is known that such an equilibrium is unstable under many fitness regimes, and we require some stabilizing factor involving frequency-dependent selection. One such factor is competition between individuals. Previous studies of competitive speciation (Bürger and Schneider 2006; Bürger et al. 2006) have shown that even with the stabilizing component of competition, the polymorphism of the main trait is lost for intermediate values of assortment. Thus, in general, the evolution of assortment may be even harder than we have shown.

Second, we have investigated only the scenario where the change in assortment will be small compared to other selective pressures. With that assumption, we have looked at two extreme cases, as mentioned above: nonselective mating, where assortment becomes complete, and selective mating, where assortment never invades. In the favorable case of nonselective mating, selection for increased assortment to evolve is small (at most, $\Delta p_i/pq_i \sim 10^{-2}$). Therefore, there may be an intermediate case where increased assortment evolves even if slightly selective. Recently, Schneider and Bürger (2005) showed that slight costs can still allow speciation to occur. Whether this conclusion implies that assortment can evolve with slight costs remains to be shown. Certain factors, such as some form of dominance in diploids and whether the change toward increased assortment is of the order of other selective pressures, may allow selective assortment to evolve. Perhaps, as occurs for modifiers of selfing, a costly modifier of assortment could invade if it has a very strong effect and produces excellent offspring (Charlesworth et al. 1990). More likely, those scenarios would have to be investigated numerically and are out of the scope of this article.

One-trait models are an extremely useful starting point in the study of reproductive isolation. However, it is known (Fry 2003; Gavrilets 2004) that these models provide the most favorable conditions for sympatric speciation to occur. It would then be of extreme interest to study the evolution of assortative mating and, most generally, the evolution of reproductive isolation while relaxing the assumption of the same set of loci under selection and assortment. Thus, two-trait models, where assortative mating (host choice) and viability selection are controlled by different loci, would be an interesting comparison. Habitat choice, as proposed by Diehl and Bush (1989), is the ideal candidate for future work on evolution of assortment.

Acknowledgments

We thank R. Bürger, A. Gardner, T. Johnson, A. Kalinka, D. Roze, and J. Taylor for discussion and comments on the manuscript. We are especially grateful to J. Hermisson and S. Otto for sharing their unpublished manuscripts with us. This work was funded by Natural Environment Research Council grant NER/A/S/2002/00857 and National Science Foundation grants DEB-9973221 and EF-0328594.

Literature Cited

Barton, N. H. 1995. Evolution of recombination. Genetical Research 65:123–145.

Barton, N. H., and M. Turelli. 1991. Natural and sexual selection on many loci. Genetics 127:229–255.

Bolnick, D. I. 2004. Waiting for sympatric speciation. Evolution 58: 895–899.

Bürger, R., and K. A. Schneider. 2006. Intraspecific competitive divergence and convergence under assortative mating. American Naturalist 167:190–205.

Bürger, R., K. A. Schneider, and M. Willensdorfer. 2006. The conditions for speciation through intraspecific competition. Evolution 60:2185–2206.

Cain, M., V. Andreasen, and D. Howard. 1999. Reinforcing selection is effective under a relatively broad set of conditions in a mosaic hybrid zone. Evolution 53:1343–1353.

Charlesworth, D., M. Morgan, and B. Charlesworth. 1990. Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. Evolution 44:1469–1489.

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

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

Diehl, S., and G. Bush. 1989. The role of habitat preference on adaptation and speciation. Pages 345–365 *in* D. Otte and J. Endler, eds. Speciation and its consequences. Sinauer, Sunderland, MA.

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

Doebeli, M., and U. Dieckmann. 2000. Evolutionary branching and sympatric speciation caused by different types of ecological interactions. American Naturalist 156(suppl.):S77–S101.

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

Endler, J. A. 1977. Geographic variation, speciation and clines. Princeton University Press, Princeton, NJ.

Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? Evolution 35:124–138.

Fisher, R. 1918. The correlation between relatives on the supposition of Mendelian inheritance. Transactions of the Royal Society of Edinburgh 52:399–433.

Fry, J. 2003. Multilocus models of sympatric speciation: Bush versus Rice versus Felsenstein. Evolution 57:1735–1746.

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

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

Karlin, S., and J. McGregor. 1974. Towards a theory of the evolution of modifier genes. Theoretical Population Biology 5:59–103.

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

Kirkpatrick, M. 2000. Reinforcement and divergence under assortative mating. Proceedings of the Royal Society B: Biological Sciences 267:1649–1655.

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

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

Kirkpatrick, M., T. Johnson, and N. H. Barton. 2002. General models of multilocus evolution. Genetics 161:1727–1750.

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

Kondrashov, A., and M. Shpak. 1998. On the origin of species by means of assortative mating. Proceedings of the Royal Society B: Biological Sciences 265:2273–2278.

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

Maynard Smith, J. 1966. Sympatric speciation. American Naturalist 100:637–650.

Pennings, P. S., M. Kopp, G. Meszéna, U. Dieckmann, and J. Hermisson. 2008. An analytically tractable model for competitive speciation. American Naturalist 171:E44–E71.

Polechova, J., and N. Barton. 2005. Speciation through competition: a critical reevaluation. Evolution 59:1194–1210.

Rice, W. 1984. Disruptive selection on habitat preference and the evolution of reproductive isolation: a simulation study. Evolution 38:1251–1260.

Sanderson, N. 1989. Can gene flow prevent reinforcement? Evolution 43:1223–1235.

Sawyer, S., and D. Hartl. 1981. On the evolution of behavioral reproductive isolation: the Wallace effect. Theoretical Population Biology 19:261–273.

Schneider, K. A., and R. Bürger. 2005. Does competitive divergence occur if assortative mating is costly? Journal of Evolutionary Biology 19:570–588.

Servedio, M. R. 2000. Reinforcement and the genetics of nonrandom mating. Evolution 54:21–29.

Associate Editor: Thomas Lenormand
Editor: Michael C. Whitlock