

CAN REINFORCEMENT COMPLETE SPECIATION?

Claudia Bank,^{1,2,3} Joachim Hermisson,³ and Mark Kirkpatrick⁴

¹ *Institut für Populationsgenetik, Vetmeduni Vienna, Vienna, Austria*

² *E-mail: claudia.bank@vetmeduni.ac.at*

³ *Mathematics and Biosciences Group, Department of Mathematics and Max F. Perutz Laboratories, University of Vienna, Vienna, Austria*

⁴ *Section of Integrative Biology C-0930, University of Texas, Austin, Texas 78712*

Received April 21, 2011

Accepted July 8, 2011

Data Archived: Dryad:doi:10.5061/dryad.cb506

Hybridization is common in nature, even between “good” species. This observation poses the question of why reinforcement is not always successful in leading to the evolution of complete reproductive isolation. To study this question, we developed a new “quasi-linkage disequilibrium” (QLD) approximation to obtain the first analytic results for the evolution of modifiers that increase mate discrimination against hybrids and heterospecifics. When such modifiers have small effects, they evolve more readily under a one-allele than a two-allele mechanism (*sensu* Felsenstein 1981). The strength of selection on the modifier decreases as hybrids decrease in frequency, and so further reinforcement may not occur once hybridization is sufficiently rare. The outcome is qualitatively different when modifiers have large effects, however, for example, when a single mutation can cause complete reproductive isolation. In this case, modifiers in a two-allele mechanism can be selected as or more strongly than those in a one-allele mechanism. Furthermore, they can spread under quite general conditions. Thus, whether complete closure of genetic introgression by reinforcement occurs may depend on the size of effects that mutations have on the sensory systems used in mate choice.

KEY WORDS: Linkage disequilibrium, mathematical model, quasi-linkage equilibrium, reproductive isolation, sexual selection.

Reinforcement is the evolution of increased prezygotic isolation that results from selection against interspecific matings (Dobzhansky 1940; Noor 1999; Kirkpatrick and Ravigné 2002; Servedio and Noor 2003; Coyne and Orr 2004). There is a long history of debate about whether reinforcement could work in principle and whether it actually happens in nature (Noor 1999; Servedio and Noor 2003; Coyne and Orr 2004). Support for reinforcement comes from two sources. Empirical studies have found evidence in various species (reviewed in Coyne and Orr 1989; Ortíz-Barrientos et al. 2009). Second, theoretical studies have established that increased premating isolation can evolve under a broad range of conditions (Kirkpatrick and Servedio 1999; de Cara et al. 2008; Otto et al. 2008; Pennings et al. 2008; Matessi et al. 2001; Barton and de Cara 2009).

A critical question that remains unresolved is whether reinforcement is able to complete speciation, that is, cause hybridization rates to evolve to zero. If so, we might not expect hybridization to persist in regions where species ranges overlap. There are, however, numerous reports of ongoing hybridization in nature. A meta-analysis concluded that “at least 25% of plant species and 10% of animal species hybridize” (Mallet 2005). Furthermore, cases are known in which hybridization has persisted for thousands of generations (Turner 1971). Several authors have argued on intuitive grounds that reinforcement cannot finish speciation definitively (reviewed by Howard 1993). Despite the long history of interest in this topic, we still lack a firm theoretical answer to the question of whether reinforcement is incapable of stopping hybridization, or whether it continues because other factors (such as constraints on sensory systems or opposing

pressures of gene flow) intervene. To determine if and when reinforcement is capable of finishing off speciation, this paper develops the simplest possible analytic models.

A key aspect of this problem is the distinction between one- and two-allele mechanisms of speciation made by Felsenstein (1981). In a one-allele mechanism, a single allele increases prezygotic isolation regardless of the species or genetic background that it occurs in. An example is an allele that increases the precision of sensory discrimination used in mate choice, as was recently reported from populations of *Drosophila* (Ortíz-Barrientos and Noor 2005). In a two-allele mechanism, an allele that enhances isolation in one species decreases it in the other. This situation occurs, for example, when males of one species are blue and the other species are red, and reinforcement occurs when a female preference allele for blue spreads in the first species while a preference allele for red spreads in the second. Felsenstein pointed out that increased premating isolation should evolve under much more general conditions under a one-allele mechanism than under a two-allele mechanism. That is because under the one-allele mechanism, the same allele unconditionally increases isolation and so can spread unopposed through both nascent species. With a two-allele mechanism, by contrast, recombination breaks down the association between the modifier alleles that increase reproductive isolation and traits that they act on (e.g., a male display). This raises the important question of whether the potential for reinforcement to complete speciation depends critically on whether a one- or two-allele mechanism is involved. Simulation results support this intuition (Servedio 2000, and a review of studies of either one- or two-allele models in Servedio and Noor 2003; Coyne and Orr 2004; Servedio 2009), but to date no analytic results are available.

Our goal here is to find simple analytical approximations to study three questions: How strong is selection for reinforcement in populations that hybridize rarely? What is the difference between the outcomes of reinforcement under one- and two-allele mechanisms? And: Can reinforcement complete speciation? Our approach is to use haploid four- and five-locus models that correspond, respectively, to the one- and two-allele mechanisms. We begin by laying out the assumptions and notation of these models. We then describe the methods used to analyze them. These include a new “quasi-linkage disequilibrium” (QLD) approximation that is adapted to cases where two species are highly but not completely diverged and there is ongoing hybridization between them. We then develop results for the invasion rates of modifiers that enhance prezygotic isolation.

The Models

We study a population that is in an advanced state of divergence toward two reproductively isolated species. We do not make any

assumptions on how this state was reached. Scenarios include an allopatric phase with accumulation of incompatibilities and divergent selection on preferences, or sympatric speciation via adaptation to different resources. Our approach visualizes the two incipient species as a single population that has very strong associations (linkage disequilibria). In the absence of hybridization, the two species would be fixed for alternative alleles, resulting in maximal pairwise linkage disequilibria. Allele frequencies across the entire population (i.e., both species combined) therefore correspond to the frequencies of the two species.

Individuals are haploid and have unlinked autosomal loci. (The generality of these assumptions is discussed below.) Two alleles, denoted 0 and 1, segregate at each locus. Following birth, viability selection acts on males. Mating then occurs. Free recombination and fusion of gametes produces the zygotes to begin the next generation.

Natural and sexual selection act on a male display trait that is not expressed in females. The phenotype z of a male is given by the sum of the allelic values at two loci, t_0 and t_1 . Thus, a male can have phenotypic values of $z = 0, 1$, or 2 . Viability selection acts on the male trait, which could result from maladaptation of intermediate phenotypes to the environment or intrinsic male-specific genetic incompatibilities. Males with extreme phenotypes ($z = 0$ or 2) have equal viability normalized to 1, whereas intermediate males ($z = 1$) have viability $1 - S$. Although our main interest is in cases where hybrids have reduced survival ($S > 0$), the model does allow for the possibility of hybrid vigor ($S < 0$).

The mating system is polygynous, with females choosing their mates. A female's mating preference for a male trait is determined by the allele she carries at locus c . Mating preferences are symmetric and are scaled such that the maximum value is 1. This is the preference value that a female carrying allele i at locus c has for a male carrying the matching allele at locus t_0 and t_1 (i.e., a female with preference allele 0 prefers males with $z = 0$, and a female with preference allele 1 prefers males with $z = 2$). The preference value for intermediate males (i.e., hybrids with $z = 1$) is P_h , and the preference value for extreme (i.e., heterospecific) males is P_e . The values of P_h and P_e are less than 1, corresponding to preferences that discriminate against hybrids and heterospecifics. Thus, premating isolation increases when P_e and P_h approach 0. Because P_e and P_h are independent parameters, the model allows for any type of mating preference function. For example, an exponential preference function is implemented by setting $P_e = P_h^2$, whereas for a linear preference function $P_e = 2P_h - 1$ (see Fig. 1). The shape of the preference function affects the relative frequency with which females mate hybrids and heterospecifics, and it also affects the lifetime fitness of hybrid males. With a linear preference function, hybrid males have the same mating success as pure males, whereas with an exponential preference function, hybrid males have reduced mating

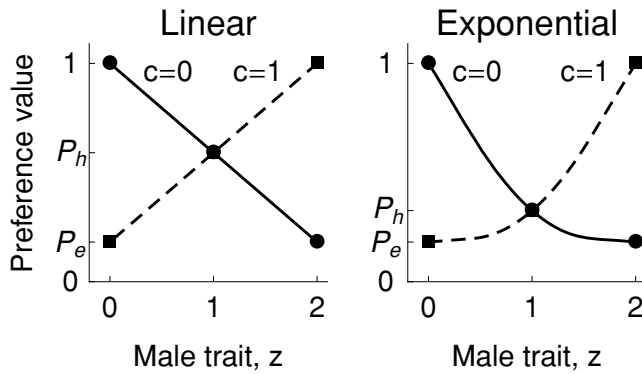


Figure 1. Two types of mating preference functions. With linear preferences (left), all males have equal mating success, whereas with exponential preferences (right), hybrids have decreased mating rates.

success. The probability that a female mates with a given male phenotype is proportional to her preference value for that phenotype and its frequency among juveniles. Matings are normalized so that all females have equal fitness (Kirkpatrick 1982), and so the preference does not entail any costs of choosiness.

This mating system with an existing strong preference provides an initial state of the population on the verge of speciation without the need of explicit assumptions on the geographical population structure. In contrast to other studies that have considered the evolution of a preference from an initial state of random mating in a geographically structured population (e.g., Servedio 2000), the initial degree of reproductive isolation in our model is achieved via existing mating discrimination and disruptive selection. Here, we choose an underlying two-allele mechanism to provide a strong degree of divergence as initial state of our analysis. Similarly, and presumably without qualitative differences in the results, one could assume that the initial amount of isolation is obtained via an underlying one-allele mechanism, like in de Cara et al. (2008), who studied a one-allele modifier on top of a one-allele assortment mechanism. Given this initial state, further isolation can be achieved by either a one-allele or a two-allele mechanism, and we study both possibilities below.

Our analysis is vastly simplified by assuming a constant population size for both (incipient) species. This occurs, for example, if each species has its own niche in which there is strong density regulation. This assumption allows us to fix allele frequencies at the c , t_0 , and t_1 loci at a constant value. In the main body of this paper, we describe our methods and results for the case of two equally sized populations, corresponding to allele frequencies fixed at $1/2$. A general analysis for asymmetric population sizes and small departures of the individual allele frequencies from the equilibrium is carried out in the Appendices.

Our approach is to determine the fate of mutations that modify female mating preferences. We analyze two main models.

Table 1. Preference values according to the one-allele model. The modifier allele increases the existing preference.

Female haplotype at		Male phenotype		
m	c (preference)	0	1	2
0	0	1	: P_h	: P_e
1	0	1	: $(1 - \delta_h)P_h$: $(1 - \delta_e)P_e$
0	1	P_e	: P_h	: 1
1	1	$(1 - \delta_e)P_e$: $(1 - \delta_h)P_h$: 1

In our “one-allele model,” a modifier increases the strength of a preference regardless of what male the female prefers (see Table 1). (An example of such a modifier is an allele that increases a female’s visual acuity of males, regardless of their color.) Modifier allele 1 at locus m increases reproductive isolation by reducing the preference for extreme (heterospecific) males from P_e to $P_e(1 - \delta_e)$ and the preference for intermediate (hybrid) males from P_h to $P_h(1 - \delta_h)$. Thus, δ_e and δ_h measure the increased discrimination that females with the modifier have against mating with extreme and hybrid males. A modifier with $\delta_e = \delta_h = 1$ generates complete reproductive isolation between the species, whereas a modifier with $\delta_e = \delta_h = 0$ has no effect at all. Note that the generality of the definition of the modifier strength can cause a modifier to have conflicting effects. For instance, a modifier with $\delta_h = 0$ and $\delta_e > 0$ will shift the probability of mating not only towards mating with conspecifics, but also toward matings with hybrids.

In our “two-allele model,” a modifier unconditionally increases mating discrimination against a particular male phenotype (see Table 2). (An example of this situation is an allele that decreases a female’s response to males of a particular color.) We implemented this model in two ways. The first is a five-locus model, in which we simultaneously introduce a modifier allele at each of two loci, m_0 and m_1 . At locus m_0 , allele 1 increases the preference for $z = 0$ males. This is achieved by reducing the

Table 2. Preference values according to the five-locus two-allele model. Each modifier allele conducts an increase in preference for a certain phenotype, independent of the existing preference.

Female haplotype at			Male phenotype		
m_0	m_1	c	0	1	2
0	0	$x \in \{0, 1\}$	P_e^x	: P_h	: P_e^{1-x}
0	1	x	$(1 - \delta_e)P_e^x$: $(1 - \delta_h)P_h$: P_e^{1-x}
1	0	x	P_e^x	: $(1 - \delta_h)P_h$: $(1 - \delta_e)P_e^{1-x}$
1	1	x	$(1 - \delta_e)P_e^x$: $(1 - \delta_h)^2 P_h$: $(1 - \delta_e)P_e^{1-x}$

preference value for hybrid males from P_h to $P_h(1 - \delta_h)$, and reducing the preference value for $z = 2$ males from P_e to $P_e(1 - \delta_e)$ (if the female carries a preference for $z = 0$ at locus c), or from 1 to $1 - \delta_e$ (if the female carries a preference for $z = 2$ at locus c). At modifier locus m_1 , allele 1 has the converse effects: it enhances the preference for $z = 2$ males. Thus, each modifier can either increase or decrease a female's overall preference, depending on her genotype at locus c . We assume equal frequencies of the modifiers at loci m_0 and m_1 . This symmetry greatly decreases the number of dynamic variables that need to be considered and ensures that allele frequencies at loci t_0 , t_1 , and c remain at their initial equilibrium of 1/2.

The second way to implement the two-allele model is a four-locus model, in which we introduce just one modifier allele that enhances the preference for $z = 2$ males in the way described above, and ignore the asymmetry effects. In the Appendices, we show that both of these models yield similar results as long as the modifier frequency is small such that associations between the two modifier alleles in the five-locus model are negligible.

The frequency of the modifier(s) at the zygote stage in both the one- and two-allele models is denoted p_m , and its change in frequency across one generation is written Δp_m .

Analysis and Results

Our first goal is to find the state of the population before the modifiers appear. Here, the population segregates at the three loci c , t_0 , and t_1 . In general, seven quantities are needed to describe the state of such a population. These are the allele frequencies at each of the loci, p_c , p_{t_0} , and p_{t_1} , three pairwise associations (linkage disequilibria) between pairs of loci, D_{ct_0} , D_{ct_1} , and $D_{t_0t_1}$, and a three-way association $D_{ct_0t_1}$ that characterizes an excess of four of the eight haplotypes.

Given the symmetry assumption of allele frequencies fixed at 1/2, one can show that $D_{ct_0t_1} = 0$ (cf. Appendix S2). Thus, the state of the population can be completely described by just two quantities. The first is D_{tt} , the association (or linkage disequilibrium) between the two male trait loci. This can also be interpreted as a measure of the frequency of hybrids in the population, which is given by $0.5 - 2D_{tt}$. Thus, $D_{tt} = 0$ corresponds to panmixia (50% hybrids), and $D_{tt} = 0.25$ corresponds to complete reproductive isolation between the two extreme phenotypes (0% hybrids). The second nonzero association is D_{ct} , the pairwise association between the preference locus, and one of the male trait loci. (Because the trait loci are interchangeable, the association between c and t_0 is equal to the association between c and t_1 .) With random mating and no viability selection, D_{ct} and D_{tt} decay to 0. At the other extreme, with complete isolation between the two species (i.e., in the absence of any hybrids), these two associations take their maximum possible value of 1/4.

We developed exact recursions for the dynamics of these associations using the multilocus framework developed by Kirkpatrick et al. (2002), which is based on the work of Barton and Turelli (1991). Those results are presented in Appendix S1. One result that emerges from the recursions is that P_h , the female preference value for intermediate (hybrid) males with $z = 1$, and the quantity $(1 - S)$, which is the viability of those males, always appear together as a product. (This makes sense as viability selection and mating happen successively in the life cycle and both selection processes act only on males.) We therefore decrease the number of parameters in the model by defining $W_h = P_h(1 - S)$, which can be interpreted as the lifetime fitness of intermediate males relative to males of the pure species. This finding shows that natural selection (acting on viability) and sexual selection (via mate discrimination) acting against hybrids make equivalent contributions to reinforcement (Kirkpatrick and Servedio 1999).

THE QLD APPROXIMATION

A major motivation for using the multilocus framework is that it lends itself readily to developing simple approximations. When the strengths of selection and nonrandom mating are weak (relative to recombination), associations between loci converge to a “quasi-linkage equilibrium” (QLE) at which the values of the associations are of order ϵ (Kimura 1965; Nagylaki 1974; Turelli and Barton 1990). Approximations for the dynamics can then be found by neglecting terms that involve products of two or more associations. We generalized this basic idea to the present study in which two species are almost completely isolated due to strong selection. In this case, the associations D_{ct} and D_{tt} are near their equilibrium values for completed speciation. We therefore developed a “quasi-linkage disequilibrium” approximation that applies when selection is so strong that deviations of the large associations from their respective values at complete speciation are small (of order ϵ). Note that natural selection against hybrids alone (without an existing mechanism of mate discrimination) is not sufficient to obtain strong enough associations to satisfy this condition, as was pointed out in a toy model by Kirkpatrick and Ravigné (2002).

The approximations for the dynamics are found by neglecting higher order terms (of order ϵ^2). The approximation is expected to be most accurate for high existing levels of reproductive isolation, which are achieved by strong premating isolation and strong selection against hybrids (i.e., $P_e, W_h \ll 1$). In the symmetric case of equal population sizes, the QLD approximation for this initial equilibrium before any modifier appears is

$$\begin{aligned}\hat{D}_{ct} &= \frac{1}{4} - \frac{P_e(2 + 3P_e - W_h)}{2(1 + 5P_e + 6P_e^2 - (3 + P_e)W_h)}, \\ \hat{D}_{tt} &= \frac{1}{4} - \frac{P_e(1 + P_e)}{1 + 5P_e + 6P_e^2 - (3 + P_e)W_h}.\end{aligned}\quad (1)$$

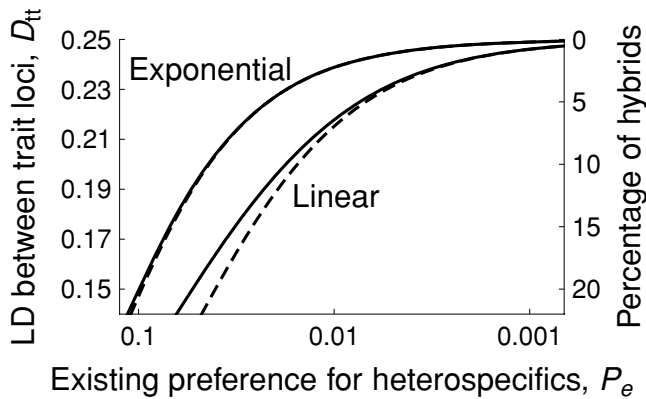


Figure 2. Buildup of pairwise linkage disequilibrium between the trait loci with decreasing preference value P_e (corresponding to increasing mating discrimination). Note that the preference value decreases from left to right, corresponding to an increase in pre-mating isolation and thus progress toward speciation. Solid curves represent results from the exact recursions, whereas dashed curves show the QLD approximation from equation (1) for a linear and exponential preference function with $S = 0.5$. The association between the trait loci, D_{tt} can be interpreted as a measure of population divergence that reflects the percentage of hybrids in the population, with $D_{tt} = 0$ corresponding to panmixia (50% hybrids), and $D_{tt} = 0.25$ corresponding to complete reproductive isolation between the two extreme phenotypes (0% hybrids).

Figure 2 shows how these associations vary as a function of the strength of preferences for linear and exponential preference functions. The exponential function builds up disequilibria more rapidly because it leads to stronger sexual selection against hybrid males and because it leads to a stronger correlation between the preference and the male trait in mated pairs. All derivations and an extension of the QLD approximation to unequal population sizes for both species are given in Appendix S2.

SELECTION FOR REINFORCEMENT

We are now in a position to study the evolution of modifiers that further increase prezygotic isolation. We introduce these modifiers into a population at the initial equilibrium. Our symmetry and equilibrium assumptions reduce the number of dynamic variables in the one-allele model from 19 to 3, and in the two-allele model from 36 (for the five-locus model) or 19 (for the four-locus model) to 6. Details are given in Appendix S3.

A natural measure of the strength of selection for reinforcement is the relative rate of increase in the frequency of a rare modifier: $\lambda = \lim_{p_m \rightarrow 0} (\Delta p_m / p_m)$. This quantity is the eigenvalue corresponding to the invasion of the modifier at the initial equilibrium with $p_m = 0$. It can also be interpreted as the strength of selection (i.e., the effective selection coefficient) for the modifier when it is rare. We calculated λ in two ways. The first uses exact recursions. These were developed using the multilocus framework of

Kirkpatrick et al. (2002) and were checked against recursions based on the haplotype frequencies. Second, we calculated approximations for λ . Here, we used a combined QLD/QLE approach, taking the QLD approximation described above as basis for a QLE approximation of the associations between the invading modifier and the other loci. This assumes that the existing mating discrimination against heterospecifics is strong and that the effects of modifiers are small ($P_e, \delta_e, \delta_h \ll 1$ of order ϵ , and ignoring terms of order ϵ^2 and higher). This approximation leads to simple expressions that can be interpreted. In the next three subsections, we present results for modifiers with weak effects, and then turn to the consequences of modifiers with strong effect.

A ONE-ALLELE MECHANISM

Under a one-allele mechanism, the modifier increases the strength of a female's mating preference, regardless of what type of male she prefers. Our QLD approximation for the rate of invasion of the modifier, which is our measure for the strength of selection for reinforcement, yields:

$$\lambda_1 = 11W_hAP_e\delta_h + 3(1 - 3W_h)AP_e\delta_e, \quad (2)$$

where

$$A := \frac{(1 - 2W_h)}{(15 - 56W_h + 33W_h^2)}. \quad (3)$$

Figure 3 compares the approximation from equation (2) to numerical results.

The first term on the right-hand side of equation (2) reflects the effect of the modifier on preferences for intermediate (i.e., hybrid) males. The second term represents the effect on preferences for extreme (heterospecific) males. These two effects make independent contributions to the strength of selection for the modifier. Modifiers that increase discrimination against intermediate males (larger δ_h) and against extreme males (larger δ_e) are favored. (Our QLD approximation requires that $W_h < 1/3$ to assure that A and the second term on the right of equation (2) are positive, but this requirement is already implicit in our earlier assumption that the two populations are highly diverged.)

Equation (2) further shows that the strength of selection for the modifier is proportional to the existing preference value for extreme (heterospecific) males, P_e . That is because a higher preference value for heterospecific males leads to larger numbers of hybrids, increasing the opportunity for reinforcement to act. Furthermore, the left term of equation (2), reflecting the effect of the modifier on matings with hybrids, is proportional to W_h . This makes sense, because if no hybrids survived ($W_h = 0$), we would expect no selection for further mating discrimination against those. The second term in the approximation for the invasion rate (2) can be understood as follows: a reduction of heterospecific matings (by δ_e) is favored most strongly if these

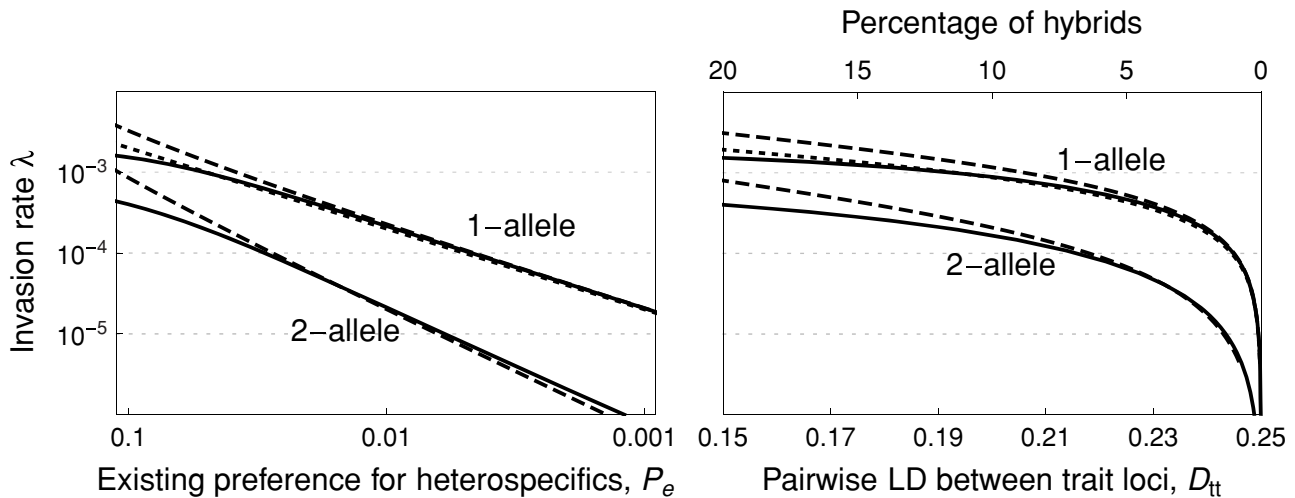


Figure 3. Invasion rates of modifiers of small effect ($\delta_e = \delta_h = 0.1$) with decreasing existing preference value P_e (left), corresponding to increasing discrimination against heterospecifics, and increasing existing population divergence, measured by D_{tt} , the pairwise linkage disequilibrium between the trait loci (right). The preference function is exponential, and $S = 0.5$. Solid curves represent results from the exact recursions, whereas dashed curves show the QLD approximations from equations (2) and (5). The dotted curve represents the less conservative approximation λ_1^* (eq. 4) for the one-allele model.

matings are still frequent (proportional to P_e) and if the fitness of the resulting hybrids is low (reflected by the term $1 - 3W_h$).

The result from equation (2) simplifies substantially in the case when hybrid males have low lifetime fitness ($W_h \ll 1$). In this case, we anticipate that invasion will not depend (at least to leading order) on how the modifier affects preferences for hybrid males because they are very rare. Indeed, we find that the rate of invasion is approximately

$$\lambda_1^* = \frac{P_e \delta_e}{5}, \quad (4)$$

visualized in Figure 3. The simple conclusion is that the strength of selection for the modifier is simply proportional to the existing preference value for extreme males, which determines the frequency of hybrids against which selection can act, and the amount by which the modifier enhances mate discrimination.

A TWO-ALLELE MECHANISM

With a two-allele mechanism, a modifier strengthens the preferences of females for conspecific males in one species, whereas it increases the probability of heterospecific matings for females of the other species. Thus, reinforcement requires that a modifier becomes associated with the correct preference allele.

The strength of selection for a modifier is now:

$$\lambda_2 = 11W_h A P_e \delta_h - \frac{11}{2} W_h A P_e \delta_e. \quad (5)$$

This expression gives important insights about the relation between one- and the two-allele mechanisms (see Fig. 3 for a visualization of the results for the one- and two-allele mechanism). In

both, the strength of selection for reinforcement is proportional to the existing preference for extreme (heterospecific) males. Furthermore, increased discrimination against intermediate (hybrid) males, δ_h , contributes equally to the spread of modifiers in the one- and two-allele mechanisms (cf. the first terms on the right side of eqs. 2 and 5). This makes sense, because both the one- and the two-allele mechanism increase mating discrimination against hybrids in the same way, independent of the existing female preference.

The two mechanisms differ dramatically, however, regarding the impact of how the modifier alters preferences for extreme males (cf. the second terms on the right of eqs. 2 and 5). In a one-allele mechanism, decreasing the preference for extreme (heterospecific) males ($\delta_e > 0$) favors spread of a modifier. The opposite occurs in a two-allele mechanism: larger values of δ_e inhibit the spread of the modifier. In fact, this effect can override the advantage of increased discrimination against hybrid males and prevent the modifier from spreading at all.

There is a simple intuitive explanation for these results. With both one-allele and two-allele mechanisms, females that avoid mating hybrid males will produce fewer hybrid offspring. Consequently, increased discrimination against intermediate males is favored in both cases. With a one-allele mechanism, increased discrimination against extreme males always translates into fewer heterospecific matings and so favors the spread of the modifier.

That is not true, however, for a two-allele mechanism. Here, the modifier acts independent of the existing preference (indeed, it can be interpreted as an independent different mating discrimination mechanism that acts on the same male trait). Therefore, a

modifier that increases the preference for males of one species can be carried by a female of the other species, leading to an increased hybridization rate. Hence, the invasion rate can only be positive if it affects matings with hybrids, which requires $\delta_h > 0$ and $W_h > 0$. (Note that the two-allele modifier would be neutral if hybrid males were dead ($W_h = 0$), whereas the one-allele modifier would still promote increased premating isolation.) Furthermore, $\delta_h > \delta_e/2$ is a necessary condition for positive invasion rates in the two-allele model. This is because $\delta_e > 0$ shifts matings not only toward conspecifics, but also toward hybrids. To make up for this effect, we need a sufficient effect of the modifier on matings with hybrids.

A 1.5-ALLELE MECHANISM

The assumptions behind the one- and two-allele models developed above are not the only ways that we can imagine modifiers to act. We developed a third model that represents an intermediate case. Here, two modifier alleles are introduced at m_0 and m_1 , which increase the preference if they occur with the corresponding existing preference, but are neutral if they occur in the background of the opposite existing preference. This model represents a two-allele model in the sense of Felsenstein (1981), because two alleles are needed to spread to finish speciation. At the same time, it is similar to a one-allele model, because the modifier works only on the background of the existing preference and is therefore independent of the buildup of associations. Hence, this example shows that there is a smooth transition between one- and two-allele mechanisms, depending on how strongly the action of the modifier is conditioned to the existing preference. A QLD approximation for modifiers of small effect yields

$$\lambda_{1.5} = \frac{\lambda_1}{2}. \quad (6)$$

This is exactly what we expect because the 1.5-allele-mechanism works essentially like the one-allele mechanism, but only in one of the two populations/species. A comparison of the exact and approximated invasion rates for all three models is pictured in Appendix S4.

STRONG MODIFIERS AND THE COMPLETION OF SPECIATION

When modifiers that increase reproductive isolation become established, the preference value for heterospecifics (P_e) decreases (recall that this means a stronger preference for conspecifics). Equations (2) and (5) then imply that the strength of selection for further reinforcement diminishes. That is simply because there are fewer and fewer hybrids for selection to act against. This observation carries an important consequence. If reinforcement proceeds by the fixation of a series of modifiers with weak effects, then the two nascent species will evolve to a point where selection for further isolation is too weak to overcome other evolutionary

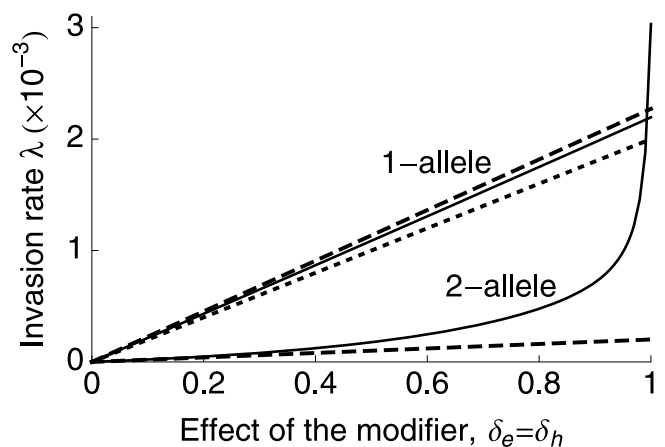


Figure 4. Invasion rates for increasing effect of the modifier, assuming similar effects on matings with intermediate and heterospecific males ($\delta_e = \delta_h$), for an exponential preference function with $S = 0.5$. Solid curves represent results from the exact recursions, whereas dashed curves show the QLD approximations from equations (2) and (5). The dotted curve represents the approximation from equation (4).

forces, such as drift. (As an example, follow one of the dashed lines indicating the size of the invasion rates in Fig. 3. For a given population size N reinforcement can be considered to be stronger than drift if $N\lambda > 1$, thus if the invasion rates for the given parameters are above $1/N$. Obviously, reinforcement loses its power earliest in a speciation process driven by a two-allele mechanism with small step size in a small population.) The result will be that the species will continue to hybridize indefinitely, unless other forces than reinforcement complete the speciation process. The strength of reinforcement would be further reduced if there was a cost of choosiness. That would cause the effect of reinforcement to cease at an even earlier stage of the speciation process.

But what happens if a mutation arises that has a large effect on reproductive isolation? The analyses to this point assume that modifiers have weak effect ($\delta_e, \delta_h \ll 1$), which means that the proportional change in the preference conferred by the modifier was assumed to be small. To answer this question, we used the exact expressions for λ (see Appendix S3). With the one-allele model, we find that the weak modifier approximations hold very well over the entire range of possible values for δ_h and δ_e (see Fig. 4). From equation (2), we see that the strength of selection for the modifier is proportional to its effect. Hence, a modifier that causes complete speciation ($\delta_h = \delta_e = 1$) will be selected most strongly of all. The linear relationship between the invasion rate and the effect of the modifier is not too surprising. However, combined with the linear relationship of the invasion rate to the preference value P_e , this yields a particularly strong advantage for modifiers of large effect that arise in a population with only an intermediate level of existing mating discrimination.

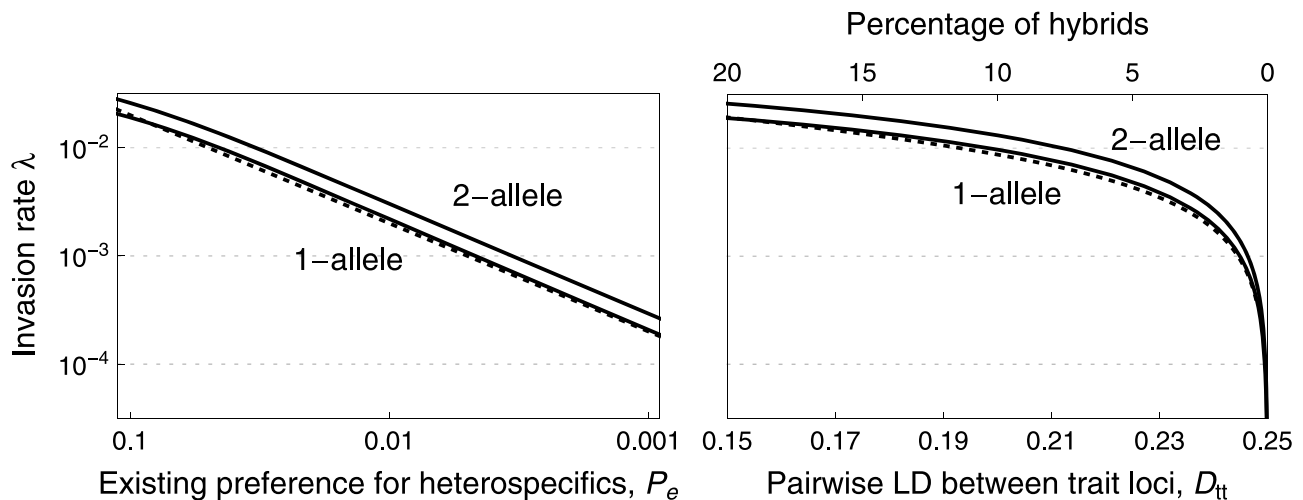


Figure 5. Invasion rates of modifiers that conduct complete speciation ($\delta_e = \delta_h = 1$) with decreasing existing preference value P_e (left), corresponding to increasing discrimination against heterospecifics, and increasing existing population divergence, measured by D_{tt} , the pairwise linkage disequilibrium between the trait loci (right). The preference function is exponential, and $S = 0.5$. The solid curves represent results from the exact recursions, whereas the dotted curve shows the QLD approximation from equation (4) for the one-allele model.

In the two-allele model, on the other hand, the weak modifier approximations do not apply when the modifier has strong effects. Although the approximation holds well for small and intermediate values of δ , there is a sudden increase of the invasion rates close to $\delta_e = \delta_h = 1$. Figure 4 shows that this increase is dramatic, and in fact can cause λ to reach even larger values than it does in a one-allele mechanism (see also Fig. 5). This occurs because a very strong modifier under a two-allele mechanism causes each individual carrying the modifier to have a totally new preference that is completely independent of the ancestral preference.

A major conclusion follows. Over a broad range of conditions, selection will favor the spread of a mutation that arises causing complete isolation in a single step. Thus, a major factor that determines whether reinforcement can finish off speciation is whether such mutations are within the evolutionary repertoire of the species.

Discussion

This paper studies the question of when reinforcement will and will not be able to cause the evolution of complete reproductive isolation. Our strategy is to work with the simplest possible models so that simple analytic results can be obtained. To do so, we generalized the “quasi-linkage equilibrium” approximation (Turelli and Barton 1990; Barton and Turelli 1991; Kirkpatrick et al. 2002) to allow for very strong selection. Some of our findings confirm and generalize conclusions suggested by earlier simulation studies, whereas other findings are apparently new.

Our analysis identifies two factors that play key roles in determining the outcome of reinforcement. The first is the effect

size of mutations that alter the female mating preference. If these modifiers have small, quantitative effects, then complete cessation of hybridization is generally not expected to evolve. That is because the strength of selection for a new modifier is inversely proportional to the strength of existing mating preferences. With the fixation of each successive modifier, the frequency of hybrids and hence the opportunity for reinforcement decline. Ultimately, the strength of selection favoring the spread of another modifier becomes so feeble that even weak evolutionary forces (e.g., drift) will dominate their evolution. This result confirms an argument made by Moore (1957), who pointed out that selection for further reinforcement must decline as hybrids become more rare. A very different outcome is expected when there are mutations that have large, qualitative effects on intermediately strong mating preferences. A modifier that completely eliminates hybrid matings will spread under a broad range of conditions, leading to the complete closure of gene flow between the nascent species.

This theoretical observation leads to the suggestion that the outcome of reinforcement may depend critically on the biological details of the mechanisms causing prezygotic isolation. Consider, for example, mate choice based on a pheromone. One can imagine single amino acid changes in a receptor that would render it sensitive to one pheromone but completely insensitive to another. In that case, a single mutation could completely eliminate hybrid matings and cause the closure of speciation by reinforcement. The situation could, however, be quite different for systems of mate choice based on acoustic cues. There one can imagine mutations that make quantitative changes to the responsiveness of females to particular frequencies, but perhaps no single mutation can make females responsive to some signals but completely

insensitive to others. In this case, reinforcement may never be able to completely eliminate hybridization. Hence, if modifiers to complete isolation and modifiers with a weaker effect both exist, the long-term outcome of reinforcement evolution may depend on historical contingency. Whereas direct fixation of the large modifier will complete the speciation process, fixation of smaller modifiers may lead to a dead end with residual gene flow. Similar results on the importance of mutational step size have been obtained for the evolution of selfing, where a gene causing complete selfing spreads most easily, because it has the potential to keep high fitness genes associated (Lande and Schemske 1985).

A second key factor is whether mutations that alter mating preferences involve a one-allele or a two-allele mechanism. This important distinction was first identified by Felsenstein (1981), but until now analytic results to compare these mechanisms have not been available. Again the outcome depends on whether modifiers have weak or strong effects. Consider first the evolution of modifiers with weak effects. The strength of selection favoring a modifier is generally stronger under a one-allele rather than a two-allele mechanism, confirming the conclusions from simulations drawn by Felsenstein (1981) and later workers (e.g., Liou and Price 1994; Servedio 2000, 2007, reviews in Servedio and Noor 2003; Coyne and Orr 2004; Servedio 2009). Increasing discrimination against mating with hybrid males favors invasion of a modifier under both one-allele and two-allele mechanisms. With a one-allele mechanism, decreasing the preference value for extreme (heterospecific) males always favors spread of the modifier. With a two-allele system, in contrast, a decreased preference value for extreme males can inhibit the spread of modifier that has a weak effect, but favor the modifier if it has a large effect. This is because the two-allele modifier represents an independent new preference, which disturbs the existing preference if its effect is small, but has the potential to override the existing preference if strong.

There are several recent studies that have analyzed the strength of reinforcement analytically. Otto et al. (2008) provide QLE approximations and a local stability analysis of a diploid one-allele model of group- and preference-based assortment. In concordance with our present study, they find that sexual selection against hybrids alone is sufficient to produce positive reinforcement. Further related analytical studies of reinforcement according to a one-allele mechanism include Matessi et al. (2001), de Cara et al. (2008), and Pennings et al. (2008), whose basic differences are summarized in Table 1 of Otto et al. (2008), all of which only consider one-allele mechanisms and keep their focus on the initialization of premating isolation. Among those, Otto et al. (2008) and Pennings et al. (2008) are the only ones to consider strong selection and modifiers of large effect. Whereas Otto et al. (2008) claims that there are no qualitative differences between weak and strong modifiers, Pennings et al. (2008) points

out that the mutational step size is an important factor in predicting whether the evolution of complete isolation is possible. This question was elaborated in more detail in Rettelbach et al. (2011). If the step size was smaller, partial isolation and thus incomplete speciation were a frequent outcome of their simulations. This points to a general problem of previous studies that have investigated the probability of speciation in the adaptive dynamics modeling framework: the assumption of mutations of small effect might obscure mechanisms that become important in large sized mutations.

de Cara et al. (2008) use the multilocus framework to study a model for the evolution of assortment. They investigate a population in which divergence has been initiated via assortment based on similarity (an underlying one-allele model with two loci) and study whether a modifier that acts as a one-allele mechanism of reinforcement will spread. Because their assortment function confers a huge selective advantage to hybrids, their analytical results show only negative invasion rates. Numerical results that have been corrected for this effect show a decrease of the invasion rate of the modifier close to complete assortment. This is an indication that our results should be transferable (at least qualitatively) to an underlying one-allele mechanism. Theoretical studies have also addressed the buildup of strong associations before. Barton (1983) developed a “strong coupling” limit for a continuous cline that holds when selection is much stronger than recombination. A recent publication by Barton and de Cara (2009) reports how incompatibilities can become associated, also considering limits of maximal selection, tight linkage, and strong assortment and weak epistasis. The approach taken in this paper is more general in some ways as it can accommodate any combination of evolutionary forces so long as they lead to very strong associations. Barton and de Cara (2009) also provide a generalization of two-allele models to account for coupling of both pre- and postzygotic incompatibilities at arbitrarily many loci. They give a verbal argument that recombination modifiers—which can be viewed as a one-allele mechanism—are expected to evolve more readily than associations. They do however not provide an analytical comparison. Probably, the first analytical results for a two-allele model have been reported by Kirkpatrick and Servedio (1999), although this study accounts only for weak selection against hybrids, and the evolution of a female preference from random mating.

Furthermore, there have been many simulation studies investigating the increase of premating isolation via one- or two-allele mechanisms (see also reviews in Servedio and Noor 2003; Coyne and Orr 2004; Servedio 2009). Most of these concentrate on just one of the possible mechanisms though, thereby making an explicit comparison impossible. Servedio (2000) provides a simulation study of reinforcement according to one- and two-allele mechanisms in scenarios of secondary contact in two-island and continent-island models. She finds that reinforcement occurs most

easily under a one-allele mechanism in a two-island model. On the other hand, in a continent-island model a two-allele mechanism was shown to evolve under more general conditions than a one-allele mechanism. This observation is intuitively clear because a one-allele mechanism gives an advantage to both the resident and the immigrant phenotypes, whereas a two-allele mechanism only promotes resident phenotypes. An indication of this effect can also be seen in our results for arbitrary relative population sizes in Appendix S4. Servedio's study is, however, restricted to the evolution of premating isolation from random mating upon secondary contact and does not provide any analytical results.

How general are our conclusions? We suspect that the qualitative conclusions apply to diploid as well as haploid systems. One reason for that intuition is the finding that the strength of linkage disequilibrium that develops between a preference locus and a male trait locus is the same in haploid and diploid systems (Kirkpatrick et al. 2002). Because this disequilibrium is a critical component that drives evolution of modifiers in these models, it seems plausible that our conclusions will carry over to diploids, at least qualitatively. A second major assumption in our models is free recombination. Results from previous models of reinforcement suggest that linkage between loci will tend to make spread of a modifier easier (Servedio and Kirkpatrick 1997). A third issue is our simplifying assumptions of symmetry. Further analyses suggest that violating the symmetry assumptions changes our conclusions quantitatively but not qualitatively for most of the parameter range (see Appendix S4). Fourth, we assume that modifier alleles have no direct effects on survival or fecundity and evolve only via indirect selection. One could imagine that modifiers that increase discrimination against heterospecifics will suffer search costs. If so, that could prevent reinforcement from proceeding. On the other hand, modifiers might have positive pleiotropic fitness effects and spread as a result. Finally, we assume that the display trait experiences natural and sexual selection only in males. This is biologically motivated by the fact that many such traits are limited in expression to males. In the event that the trait is also expressed in females, our intuition is that disruptive natural selection will further decrease the frequency of hybrids and therefore the opportunity for reinforcement. We expect, however, that the qualitative results from our model will be unchanged.

Perhaps the most general conclusion emerging from our result is that the genetic and physiological architecture of mate preference mechanisms can play a decisive role in deciding how speciation proceeds. In some organisms, preference modifiers might alter the color or sonic pitch to which females are most sensitive, for example, but still allow them to respond to a broad range of male displays. In these cases, reinforcement might never be able to complete speciation. Even though there is no absolute genetic constraint to the evolution of complete reproductive isolation, closure is never reached because selection for further

reinforcement becomes so weak. Complete cessation of gene flow would require some other mechanism, for example, selection on pleiotropic effects of preference genes, to do that job. In other species, single mutations may occur that allow females to discriminate with complete precision between conspecifics and heterospecifics, for example, by allowing them to exploit a new sensory modality. If these kinds of modifiers are available, then reinforcement can cause elimination of all gene flow between hybridizing species. Thus, the outcome of reinforcement may depend on how much individual mutations affect the sensory systems used in mate choice.

ACKNOWLEDGMENTS

We thank N. Barton, S. van Doorn, M. Servedio, an anonymous reviewer, and the Kirkpatrick lab for discussion and/or helpful comments on the manuscript. C.B. and J.H. acknowledge financial support by the Vienna Science and Technology fund (WWTF). MK is grateful for support from NSF grant DEB-0819901.

LITERATURE CITED

- Barton, N. 1983. Multilocus clines. *Evolution* 37:454–471.
- Barton, N. H., and M. A. R. de Cara. 2009. The evolution of strong reproductive isolation. *Evolution* 63:1171–1190.
- Barton, N. H., and M. Turelli. 1991. Natural and sexual selection on many loci. *Genetics* 127:229–255.
- Coyne, J., and H. A. Orr. 1989. Patterns of speciation in *Drosophila*. *Evolution* 43:362–381.
- . 2004. *Speciation*. Sinauer Associates, Sunderland, MA.
- de Cara, M. A. R., N. H. Barton, and M. Kirkpatrick. 2008. A model for the evolution of assortative mating. *Am. Nat.* 171:580–596.
- Dobzhansky, T. 1940. Speciation as a stage in evolutionary divergence. *Am. Nat.* 171:580–596.
- Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* 35:124–138.
- Howard, D. J. 1993. Reinforcement: origin, dynamics, and fate of an evolutionary hypothesis. Pp. 46–69 in R. G. Harrison ed. *Hybrid zones and the evolutionary process*. Oxford Univ. Press, Oxford, UK.
- Kimura, M. 1965. Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection. *Genetics* 52:875–890.
- Kirkpatrick, M. 1982. Sexual selection and the evolution of female choice. *Evolution* 36:1–12.
- Kirkpatrick, M., and M. Servedio. 1999. The reinforcement of mating preferences on an island. *Genetics* 151:865–884.
- Kirkpatrick, M., and V. Ravigné. 2002. Speciation by natural and sexual selection: models and experiments. *Am. Nat.* 159:22–35.
- Kirkpatrick, M., T. Johnson, and N. Barton. 2002. General models of multi-locus evolution. *Genetics* 161:1727–1750.
- Lande, R., and D. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution* 39:24–40.
- Liou, L., and T. Price. 1994. Speciation by reinforcement of premating isolation. *Evolution* 48:1451–1459.
- Mallet, J. 2005. Hybridization as an invasion of the genome. *Trends Ecol. Evol.* 20:229–237.
- 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.

- Moore, J. A. 1957. An embryologist's view of the species concept. Pp. 325–338 in E. Mayr ed. *The species problem*. American Association for the Advancement of Science, Washington, DC.
- Nagylaki, T. 1974. Quasilinear equilibrium and the evolution of two-locus systems. *Proc. Natl. Acad. Sci.* 71:526–530.
- Noor, M. A. 1999. Reinforcement and other consequences of sympatry. *Heredity* 83:503–508.
- Ortíz-Barrientos, D., and M. Noor. 2005. Evidence for a one-allele assortative mating locus. *Science* 310:1467–1467.
- Ortíz-Barrientos, D., A. Grealy, and P. Nosil. 2009. The genetics and ecology of reinforcement. Implications for the evolution of prezygotic isolation in sympatry and beyond. *Ann. NY Acad. Sci.* 1168:156–182.
- Otto, S. P., M. R. Servedio, and S. L. Nuismer. 2008. Frequency-dependent selection and the evolution of assortative mating. *Genetics* 179:2091–2112.
- Pennings, P. S., M. Kopp, G. Meszéna, U. Dieckmann, and J. Hermisson. 2008. An analytically tractable model for competitive speciation. *Am. Nat.* 171:E44–E71.
- Rettelbach, A., J. Hermisson, U. Dieckmann, and M. Kopp. 2011. Effects of genetic architecture on the evolution of assortative mating under frequency-dependent disruptive selection. *Theor. Popul. Biol.* 79:82–96.
- Servedio, M. 2000. Reinforcement and the genetics of nonrandom mating. *Evolution* 54:21–29.
- . 2007. Male versus female mate choice: sexual selection and the evolution of species recognition via reinforcement. *Evolution* 61:2772–2789.
- . 2009. The role of linkage disequilibrium in the evolution of premating isolation. *Heredity* 102:51–56.
- Servedio, M., and M. Kirkpatrick. 1997. The effects of gene flow on reinforcement. *Evolution* 51:1764–1772.
- Servedio, M., and M. Noor. 2003. The role of reinforcement in speciation: theory and data. *Ann. Rev. Ecol. Evol. Syst.* 34:339–364.
- Turelli, M., and N. H. Barton. 1990. Dynamics of polygenic characters under selection. *Theor. Popul. Biol.* 38:1–57.
- Turner, J. 1971. Two thousand generations of hybridisation in a *Heliconius* butterfly. *Evolution* 25:471–482.

Associate Editor: N. Barton

Supporting Information

The following supporting information is available for this article:

Appendix S1. Derivation of recursions.

Appendix S2. QLD approximations for the 3-locus model.

Appendix S3. Combining QLD & QLE approximations to obtain invasion rates.

Appendix S4. Visualizations.

Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.