

## Research Article

# Ecological Divergence and the Origins of Intrinsic Postmating Isolation with Gene Flow

Aneil F. Agrawal,<sup>1</sup> Jeffrey L. Feder,<sup>2,3</sup> and Patrik Nosil<sup>3,4</sup>

<sup>1</sup> Department of Ecology & Evolutionary Biology, University of Toronto, Toronto, ON, Canada M5S 3B2

<sup>2</sup> Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA

<sup>3</sup> Institute for Advanced Study, Wissenschaftskolleg, 14193 Berlin, Germany

<sup>4</sup> Department of Ecology and Evolutionary Biology, University of Boulder, Boulder, CO 80309, USA

Correspondence should be addressed to Aneil F. Agrawal, a.agrawal@utoronto.ca

Received 25 April 2011; Revised 14 June 2011; Accepted 21 June 2011

Academic Editor: Zachariah Gompert

Copyright © 2011 Aneil F. Agrawal et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The evolution of intrinsic postmating isolation has received much attention, both historically and in recent studies of speciation genes. Intrinsic isolation often stems from between-locus genetic incompatibilities, where alleles that function well within species are incompatible with one another when brought together in the genome of a hybrid. It can be difficult for such incompatibilities to originate when populations diverge with gene flow, because deleterious genotypic combinations will be created and then purged by selection. However, it has been argued that if genes underlying incompatibilities are themselves subject to divergent selection, then they might overcome gene flow to diverge between populations, resulting in the origin of incompatibilities. Nonetheless, there has been little explicit mathematical exploration of such scenarios for the origin of intrinsic incompatibilities during ecological speciation with gene flow. Here we explore theoretical models for the origin of intrinsic isolation where genes subject to divergent natural selection also affect intrinsic isolation, either directly or via linkage disequilibrium with other loci. Such genes indeed overcome gene flow, diverge between populations, and thus result in the evolution of intrinsic isolation. We also examine barriers to neutral gene flow. Surprisingly, we find that intrinsic isolation sometimes weakens this barrier, by impeding differentiation via ecologically based divergent selection.

## 1. Introduction

Speciation is a central topic in evolutionary biology and occurs as inherent barriers to gene flow evolve between formerly interbreeding populations. As data accumulate, it is becoming more apparent that many different forms of barriers to gene exchange (reproductive isolation) can be involved in speciation [1–4]. One barrier that has received much attention, both historically and in recent work on the genes causing reproductive isolation, that is, “speciation genes”, is low hybrid fitness due to genetic incompatibilities (intrinsic postmating isolation) [5–9]. The evolution of intrinsic reproductive isolation has received so much attention because it has been one of the key mysteries of speciation. How is it possible, if hybrids are unfit, to evolve from the high-fitness genotype of one species to the high-fitness genotype of the other species without passing through a low-fitness intermediate?

The most famous answer to this question is attributed to the insights of Bateson, Dobzhansky, and Muller [10–13]. Specifically, intrinsic isolation often arises from between-locus genetic incompatibilities: alleles that function well within species are incompatible with one another when brought together for the first time in the genome of a hybrid (see [14] for review). Genetic architectures of this form are known as Dobzhansky-Muller (DM) incompatibilities (illustrated in Table 1) and, in modern parlance, allow populations to evolve reproductive isolation without passage through a fitness valley that would be opposed by selection.

The accumulation and origin of genetic incompatibilities, as well as their effects on levels of gene flow upon secondary contact, have received much attention (Table 2 for review). Along these lines, it can be difficult for intrinsic postzygotic barriers to evolve in the face of gene flow (i.e., in parapatry or sympatry). The logic is that in interbreeding populations, deleterious genotypic combinations will

TABLE 1: Simple Dobzhansky-Muller incompatibility. With this type of genetic architecture, it is possible to evolve from one high fitness genotype ( $AB$ ) to another ( $ab$ ) without passing through any fitness valleys (by  $AB \rightarrow aB \rightarrow ab$ ). Matings between two high fitness types yield unfit hybrids (i.e.,  $AB \times ab$  produces the unfit type  $Ab$ ).

	A	A
B	Fit	Fit
b	Unfit	Fit

be created and then quickly eliminated by selection [3]. However, divergent selection in general might overcome gene flow to cause population divergence [15–20]. Thus, one might imagine scenarios for the origin of intrinsic incompatibilities in the face of gene flow where divergent selection drives the evolution of incompatibilities, perhaps in the face of gene flow. Such scenarios represent the process of “ecological speciation”, where reproductive isolation evolves as populations adapt to different ecological environments, as a simple and incidental by-product of adaptive genetic divergence [21–26].

Ecological speciation is expected to result in reduced fitness of immigrants into foreign environments and hybrids (i.e., extrinsic isolation), due to an ecological mismatch between the phenotype of immigrants and hybrids and local environmental conditions [4, 42]. However, the evolution of any type of reproductive barrier might evolve as a by-product of divergent adaptation. Thus, the negative gene-by-gene interactions that cause intrinsic postmating isolation might also arise via adaptive genetic divergence if, for example, the different alleles favored in divergent habitats are incompatible with one another. Along these lines, intrinsic isolation may evolve in the face of gene flow, as long as the diversifying effects of divergent selection can overcome the homogenizing effects of gene flow. Verbal arguments along these lines were made quite some time ago, for example, in the writings of Sir Ronald Fisher [43] and in more recent reviews [44].

Indeed, there is empirical evidence that ecological divergence can drive the evolution of intrinsic postmating isolation. Consider the classic case of ecological adaptation and intrinsic hybrid sterility in *Mimulus* monkey-flowers adapted to different soil types [45]. In this example, crossing experiments revealed that a yet unidentified gene conferring tolerance to copper in the soil (or a gene tightly physically linked to it) interacts with a small number of genes in another population to generate intrinsic hybrid incompatibility. Thus, genes conferring adaptation to copper contribute to postmating isolation. Evidence for a role for divergent adaptation in the evolution of intrinsic reproductive isolation also stems from broad comparative studies in fish [46] and insects (fruit flies and moths/butterflies [25]). These studies considered multiple species pairs and examined the relationship between indices of the degree of ecological divergence between species pairs and the extent of intrinsic postmating isolation between species pairs. Both studies report positive correlations between ecological

divergence and intrinsic isolation, independent from time (i.e., after controlling for time since species divergence using genetic distance between species pairs) and report positive correlations between ecological divergence and intrinsic isolation. Finally, experimental and genetic studies in yeast also implicate divergent adaptation in promoting the evolution of intrinsic incompatibilities [47, 48].

Although these arguments and observations about how divergent selection might overcome gene flow to generate intrinsic incompatibilities have been around for some time and make intuitive sense, there has been little formal theoretical exploration of the specific conditions under which genetic incompatibilities might originate during ecological speciation with gene flow. As noted by Coyne and Orr [3], “we currently lack any theory telling us how much intrinsic isolation can evolve under a given level of hybridization” (page 177). Here we aim to help fill this theoretical gap by exploring the evolution of intrinsic isolation in the face of gene flow using a combination of analytical and simulation models. We examine how migration, selection, and recombination interact to affect this process. We also examine how the evolution of intrinsic isolation affects rates of neutral gene flow. The results help clarify that ecological adaptation can generate intrinsic incompatibilities.

Our treatment differs from and extends important past models of DM incompatibilities in several ways. For example, classic theoretical treatments of intrinsic incompatibilities (e.g., [27, 28, 33]) generally did not consider the evolutionary processes driving genetic divergence in the underlying genes involved; allopatry is explicitly assumed or implied, and incompatibilities evolve by drift or some unspecified form of selection. Newer models include selection, but it is usually not divergent between environments (see Table 2 for consideration of past models). Our focus on the origin of incompatibilities via divergent selection and its interaction with gene flow sets our treatment apart from most previous models. While many past studies involving selection were completely simulation based, our model helps to answer the call for analytical treatments of speciation models [49].

Our models are extensions of classical theoretical work on geographic clines by Clarke [16] and Endler [18] and of more recent models by Gavrillets [14], all of which did consider divergent selection in some context. Indeed, the hybrid zone literature in general discussed how geographic clines generated by divergent selection could be steepened by the spread of modifier loci that also contribute to DM incompatibilities [19, 20]. Our work is most closely related to that of Clarke [16], though our goals are quite different. He was interested in understanding how clines of key morphological characters (controlled by a single locus) could be sharpened by epistatic modifiers. To our knowledge, he was the first to explicitly model how epistatic modifiers can build upon differences established by divergent selection, though he did not explicitly discuss reproductive isolation or measure barriers to gene flow in other parts of the genome.

Our work thus extends previous models by considering (1) the joint evolution of extrinsic versus intrinsic isolation during ecological divergence, (2) the explicit effects of divergent ecological selection on epistatically interacting loci,

TABLE 2: In chronological order, a summary of some major previous models of Dobzhansky-Muller Incompatibilities (DMIs). Uniform selection refers to scenarios where selection occurs (i.e., some alleles or genotypic combinations are more fit than others), but selection is not divergent between different ecological environments (i.e., no environmental variation). Uniform selection thus fits the “mutation-order speciation” model recently advanced by Schluter [26]. In contrast, divergent selection fits the ecological speciation model. IBD: Isolation-by-distance.

Model	Process driving divergence	Migration and gene flow	Discrete patches (populations)	Method	Notes
[16]	Divergent selection	Yes	Clinal divergence rather than discrete patches	Analytical	Divergent selection is on major locus, modifier locus affects the fitness of alleles at the major locus, but is not under divergent selection itself, barrier to neutral gene flow not examined. Also, focus of the model is how the modifier, once fixed, alters the shape of the cline, rather than on how the modifier fixes in the first place.
[18]	Divergent selection	Yes	Clinal divergence rather than discrete patches	Analytical	Divergent selection is on major locus, modifier locus causes DMI but is not under divergent selection itself, barrier to neutral gene flow not examined.
[27]	Stochastic (drift)	No	Yes	Simulation	Drift fixed new mutations in small populations; two loci considered.
[28]	Unspecified	No	Yes	Analytical	The classic paper on the accumulation of DMIs. No assumptions are made about evolutionary causes of substitutions driving DMIs.
[29, 30]	Uniform selection, not explicitly divergent	Yes	Yes	Analytical	Migration can allow the spread of a gene combination whose component alleles are not individually favored, but the conditions for this are restricted (low migration, strong advantage for particular genotypic combinations).
[31]	Uniform selection or drift	No	Yes	Analytical and simulation	Population subdivision has no effect on time to speciation by drift. Under selection, speciation occurs more rapidly between two large populations than between many small ones.
[32]	Unspecified; populations initially fixed for different incompatible alleles	Yes	Yes, stepping stone	Analytical	Focused on the effect of intrinsic incompatibilities on the barrier to neutral gene flow.
[33]	Unspecified	No	Yes	Analytical	Adds stochasticity of molecular evolution to the results of Orr [28].
[34]	Uniform selection	Yes	Yes, between neighboring patches	Simulation	Speciation in subdivided populations occurs most rapidly when there is some migration (which spreads incompatible alleles); many (250) loci are considered.
[35]	Uniform selection	Yes	Yes	Simulation	Strong selection counters the inhibitory effects of gene flow, permitting the evolution of intrinsic postmating isolation.
[36]	Uniform selection	Yes	Yes,	Analytical	Migration occurs through a single, spatially structured population and affects the accumulation of DMIs.
[37]	Uniform selection	Yes	Yes	Analytical	Loci causing DMIs themselves are under selection (i.e., new alleles at these loci can be favored), but selection is uniform. Barrier to gene flow is considered, including the effect of chromosomal inversions.
(Gavrilets [14] and previous models by same author (e.g., [38, 39])	Divergent selection, uniform selection, or drift	Yes	Yes	Analytical	Incompatible alleles advantageous in one environment but neutral in other (thus wider range of scenarios examined in current study), barrier to gene flow examined.
[40]	Uniform selection	Yes	No, clinal with IBD	Simulation	IBD and outbreeding depression were sufficient to drive speciation, in the absence of ecological differences.

TABLE 2: Continued.

Model	Process driving divergence	Migration and gene flow	Discrete patches (populations)	Method	Notes
[41]	Uniform selection	Yes	Yes	Simulation	Focused on the origin of DMIs due to different, incompatible mutations diverging between populations adapting to identical selection regimes. Gene flow strongly impeded strong divergence via such a “mutation-order” process (c.f., Schluter [26])
Current study	Divergent selection	Yes	Yes	Analytical and simulation	Focus on the origination of incompatibilities in the face of gene flow. Examined scenarios with and without direct, divergent selection on modifier loci which also contribute to intrinsic isolation. Examined barrier to neutral gene flow caused by evolution of intrinsic isolation.

(3) the examination of the barrier to neutral gene flow caused by intrinsic isolation, and (4) a focus on the initiation of divergence.

Our model is similar to that of Gavrillets [32] in its examination of barriers and intrinsic isolation (point (3) above). However, Gavrillets [32] assumes divergence in allopatry (via some unspecified process) and begins his analysis after intrinsic isolation already exists. In contrast, our model focuses on how divergent ecological selection causes intrinsic isolation to *originate*: ecological divergence in our models establishes the initial genetic divergence upon which incompatibilities can be built. As discussed below, the models yield some novel and counterintuitive predictions. For example, we find that the evolution of intrinsic isolation need not always strengthen the barrier to neutral gene flow; whether this occurs is dependent on the effect of intrinsic isolation on the genetic differentiation causing extrinsic isolation.

## 2. Model

If the level of migration is low compared to the strength of divergent selection, populations inhabiting different environments can quickly become differentiated with respect to genes underlying key ecological traits [3, 50, 51]. Intrinsic isolation can then arise between taxa because (1) the genes underlying key ecological traits also generate incompatibilities themselves, or (2) subsequent mutations that are not necessarily integral to ecology are incompatible with the altered genetic backgrounds generated by the initial adaptive changes. Attempting to model the evolution of reproductive isolation in a spatially heterogeneous environment with multilocus genotypes can quickly lead to an overly complex model with numerous parameters. To minimize these difficulties, we initially focus on the simplest possible model that allows us to illustrate some key points: a two-locus haploid model with two populations connected by migration. We then show through simulation that the major results apply in diploid systems.

We are interested in exploring the following scenario. Consider a species in which the  $AB$  haplotype is both fit and common in Population 1 (Figure 1). A second population of this species (Population 2) becomes established in a new habitat in which there is ecological selection favoring an

TABLE 3: Fitnesses of the four haplotypes in the full model. For Case 1,  $t_1 = t_2 = 0$ .

	Population 1		Population 2	
	$A$	$a$	$A$	$a$
$B$	1	$1 - s_1$	$(1 - s_2)(1 - t_2)$	$1 - t_2$
$b$	$b(1 - t_1)(1 - \epsilon_A)$	$(1 - t_1)(1 - s_1)(1 + \epsilon_a)$	$(1 - s_2)(1 - \epsilon_A)$	$1 + \epsilon_a$

$a$  allele. Assuming migration levels are low relative to the strength of divergent selection,  $aB$  becomes the common genotype in Population 2. Imagine that the  $B$  locus is an epistatic modifier of the  $A$  locus. While the  $B$  allele works well with the  $A$  allele, the  $b$  allele improves the fitness of individuals carrying an  $a$  allele. For example, the  $a$  allele may be a gene that confers tolerance to high levels of copper, as in the monkey-flower example, but also incurs some pleiotropic cost. The  $b$  allele might mitigate this cost, thus improving the intrinsic fitness of individuals carrying an  $a$  allele. The  $b$  allele, however, could interact negatively with the  $A$  allele, regardless of the environment, uncovering deleterious consequences of the  $A$  allele normally masked by the  $B$  allele. If  $AB$  is common in Population 1 and  $ab$  is common in Population 2, then recombinants between them will be less fit on average, even in the absence of copper. This would be considered intrinsic isolation.

Specifically, we assume that the  $A$  locus controls a key ecological trait and is subject to divergent ecological selection. In Population 1, ecological selection reduces the fitness of individuals with an  $a$  allele by an amount  $s_1$ ; in Population 2, ecological selection reduces the fitness of individuals with an  $A$  allele by an amount  $s_2$ . The  $B$  locus is an epistatic modifier of the  $A$  locus. Specifically, the  $b$  allele improves the intrinsic fitness of the  $a$  genotype by an amount  $\epsilon_a$ , perhaps by mitigating some ecologically independent pleiotropic cost associated with  $a$ . The  $b$  allele interacts negatively with the  $A$  allele, reducing intrinsic fitness by  $\epsilon_a$ . There may also be direct ecological selection on the  $B$  locus. If so,  $b$  genotypes have a fitness disadvantage of  $t_1$  in Population 1 whereas  $B$  genotypes have a disadvantage of  $t_2$  in Population 2. These fitness effects are shown explicitly in Table 3.

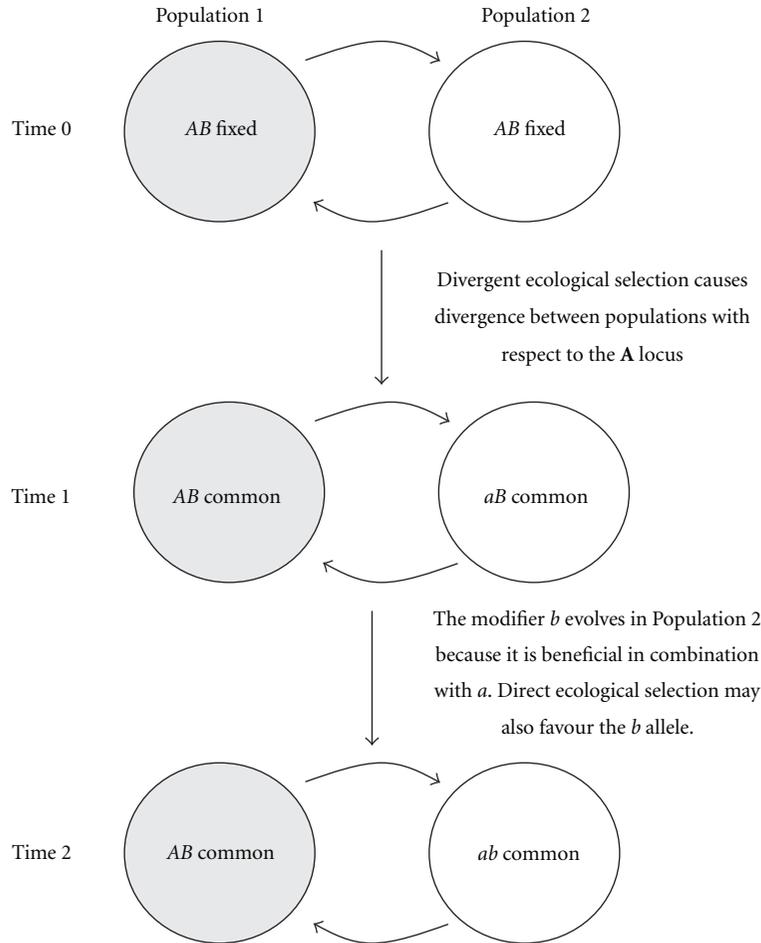


FIGURE 1: Divergent selection and the evolution of intrinsic isolation. Time 0: Two populations, connected by migration inhabit different environments. Both are initially fixed for the AB genotype. Time 1: Divergent ecological selection has caused divergence at the A locus, with the *a* allele evolving to high frequency in Population 2. Time 2: The modifier *b* evolves in Population 2, either because it ameliorates a pleiotropic cost of the *a* allele (Model 1) or because the B locus also experiences divergent ecological selection (Model 2).

We assume that migration occurs following selection within each population. The migration rate of individuals into Population 1 is  $m_1$ , and the rate into Population 2 is  $m_2$ . Technically, this is accomplished as follows. Both populations contribute to the migrant pool, but they may do so differently. A fraction  $c$  of individuals in the migrant pool comes from Population 1, while the remainder  $(1 - c)$  comes from Population 2. In each population, a fraction  $m$  of all individuals come from the migrant pool with the remainder being residents. This means  $m_1 = m(1 - c)$  and  $m_2 = cm$ . A number of ecological scenarios can be described by the appropriate choice of  $m$  and  $c$ . For example, if Population 1 is much larger than Population 2 such that migration is primarily from Population 1 into Population 2, then  $c$  would be close to 1. Throughout, we have assumed soft selection (i.e., a deme's contribution to the migrant pool is independent of its genetic composition/fitness).

Following migration, mating occurs randomly within each population. The rate of recombination between the two loci is  $r$ . The populations are assumed to be sufficiently large

that drift can be ignored. Recursion equations describing this life cycle (selection, migration, reproduction) were created in the standard way and are available as supplementary material in the form of a *Mathematica* notebook (see Supplementary Material available online at doi:10.1155/2011/435357).

A number of assumptions were made in order to make analytical progress. We assume that both ecological and intrinsic selections are weak in an absolute sense but strong relative to migration. Specifically, the selection parameters  $s_1, s_2, t_1, t_2, \varepsilon_A$ , and  $\varepsilon_a$  are  $O(\xi)$ , and  $m$  is  $O(\xi^2)$  where  $\xi \ll 1$ . We performed the analysis with both tight and loose linkage (i.e., assuming  $r$  is  $O(\xi)$  or  $O(\xi^0)$ ). The results represented assume tight linkage unless otherwise stated. Differences between the analyses are noted when they are important (see supplementary material for details). We also conducted numerical simulations for comparison with our analytical results. These simulations indicate that the analytical results are reasonably robust to these assumptions provided that migration does not become too strong relative to selection (see Results).

### 3. Results

Our primary results are most easily seen in a simplified case in which there is no ecological selection on the epistatic modifier. We present this case below. The full model including both ecological selection on the modifier and the complete range of epistatic effects is discussed in the supplementary material and presented in the numerical results below.

*Case 1 (Epistatic Trade-off).* In this case, the epistatic modifier allele  $b$  improves the fitness of individuals carrying the  $a$  allele by an amount  $\varepsilon_a$  but reduces the fitness of  $A$  individuals by an amount  $\varepsilon_A$ . There is no divergent ecological selection acting on the  $B$  locus,  $t_1 = t_2 = 0$ .

*3.1. Equilibrium Allele Frequencies and Differentiation.* Divergent ecological selection causes the populations to differentiate with respect to the ecologically important locus  $A$ . Assuming the  $b$  allele is initially absent, the equilibrium frequencies for  $A$  at migration-selection balance in Populations 1 and 2, respectively, are given by

$$\begin{aligned} p_{A,1} &= 1 - \frac{m_1}{s_1} + O(\xi^2), \\ p_{A,2} &= \frac{m_2}{s_2} + O(\xi^2). \end{aligned} \quad (1)$$

We can quantify the differentiation between the two populations at this locus as

$$\delta_{A|B \text{ fixed}} = p_{A,1} - p_{A,2} = 1 - \frac{m_1}{s_1} - \frac{m_2}{s_2} + O(\xi^2). \quad (2)$$

The  $b$  allele can invade the system by first spreading in Population 2, where  $b$  is favored if the  $a$  allele is sufficiently common there. This will be the case provided that migration into this population is not too high,

$$m_2 < \frac{s_2(r + s_2 + \varepsilon_A + \varepsilon_a)\varepsilon_a}{r(s_2 + \varepsilon_A + \varepsilon_a)}. \quad (3)$$

Following invasion, the equilibrium frequencies of the  $B$  allele in the two populations are

$$\begin{aligned} p_{B,1} &= 1 - \frac{m_1(r + \varepsilon_A)}{\varepsilon_A(r + s_1 - \varepsilon_a)} + O(\xi^2), \\ p_{B,2} &= \frac{m_2(r + \varepsilon_a)}{\varepsilon_a(r + s_2 + \varepsilon_a)} + O(\xi^2). \end{aligned} \quad (4)$$

At this equilibrium,  $b$  is common in Population 2. Because  $b$  enhances the fitness advantage of  $a$  over  $A$ , the equilibrium frequencies of  $A$  change due to the invasion of  $b$ . The leading order approximations for the equilibrium frequencies of  $A$  at this new equilibrium are

$$\begin{aligned} p_{A,1} &= 1 - \frac{m_1(r + s_1)}{s_1(r + s_1 - \varepsilon_a)} + O(\xi^2), \\ p_{A,2} &= \frac{m_2(r + s_2 + \varepsilon_A + \varepsilon_a)}{(s_2 + \varepsilon_A + \varepsilon_a)(r + s_2 + \varepsilon_a)} + O(\xi^2). \end{aligned} \quad (5)$$

The invasion of  $b$  alters the equilibrium frequency of  $A$  in Population 2, thereby changing the degree of differentiation between the two populations at the ecologically important locus. The change in differentiation is quantified by

$$\begin{aligned} \Delta\delta &= \delta_{A|b \text{ present}} - \delta_{A|B \text{ fixed}} \\ &= -m_1 \frac{\varepsilon_a}{s_1(r + s_1 - \varepsilon_a)} \\ &\quad - m_2 \left( \frac{r + s_2 + \varepsilon_A + \varepsilon_a}{(s_2 + \varepsilon_A + \varepsilon_a)(r + s_2 - \varepsilon_a)} - \frac{1}{s_2} \right) + O(\xi^2). \end{aligned} \quad (6)$$

When migration is weak relative to selection, there is strong divergence at the  $A$  locus and the evolution of the  $b$  locus has very little effect on the magnitude of divergence (i.e.,  $|\Delta\delta| \ll \delta$ ). The expression above tends to be negative, implying that the evolution of an epistatic modifier reduces differentiation at the  $A$  locus. This provides our first hint that the evolution of incompatibilities can reduce barriers to gene flow (see below).

The equilibrium described above is stable under the assumption that migration is weak relative to selection. When linkage is loose ( $r \gg s$ ), analysis of the largest eigenvalues of the Jacobian matrix indicates that stability requires the following conditions

$$\begin{aligned} m_1 &< \min \left[ \frac{s_1 \varepsilon_A}{2\varepsilon_A + \varepsilon_a}, \frac{s_1 \varepsilon_A}{\varepsilon_A + \varepsilon_a + s_1} \right], \\ m_2 &< \frac{(\varepsilon_A + \varepsilon_a + s_2)\varepsilon_a}{2(\varepsilon_A + \varepsilon_a) + s_2}. \end{aligned} \quad (7)$$

The reasons for instability are fairly intuitive. With high migration, the  $b$  allele will often find itself in the wrong population and with the wrong (i.e.,  $A$ ) allele. Its fate is determined by its epistatic interactions. When  $\varepsilon_A$  is low,  $b$  fixes in both populations because the disadvantage of  $Ab$  relative to  $AB$  in Population 1 is small compared to the advantage of  $ab$  relative to  $aB$  in Population 2. If  $\varepsilon_a$  is high,  $ab$  can fix in both populations because  $ab$  is more fit than  $AB$  across the metapopulation (the average fitness over both environments is the relevant fitness measure if populations are well mixed). If  $\varepsilon_a$  is low and  $\varepsilon_A$  is high, then  $B$  fixes in both populations because the disadvantage of  $Ab$  relative to  $AB$  in Population 1 is large compared to disadvantage of  $aB$  relative to  $ab$  in Population 2. While the stability conditions above serve as a useful heuristic, they should be regarded with caution as they are obtained from analyses assuming that migration is weak relative to selection (i.e., instability occurs outside the region where the model assumptions are valid).

The stability conditions outlined above assume loose linkage. Stability is also influenced by the recombination rate  $r$  when linkage is tight (see *Mathematica* file provided as online supplementary material). In the simulations described below, we see that tight linkage tends to destabilize the two-locus polymorphism. In the remaining sections we focus on parameter regions where the polymorphic equilibrium is stable.

*3.2. Intrinsic Isolation.* Intrinsic isolation is usually measured empirically by comparing the fitness of hybrids to the average

fitness of parental types. Typically, such experiments are done in the lab where ecological selection is typically absent, or at least different than in either of the environments from which the parents are collected. Even in the lab, the ecological trait controlled by the *A* locus may affect fitness. To account for this, we model extrinsic selection for or against the *a* allele in the lab by the variable  $s_{\text{lab}}$ , where the value of  $s_{\text{lab}}$  may be negative, positive, or zero. Assuming that the intrinsic fitness effects represented by  $\varepsilon_A$  and  $\varepsilon_a$  are truly “intrinsic” (i.e., environment independent), the fitness of the four haplotypes when measured in the lab are given by  $w_{AB} = (1 + s_{\text{lab}}/2)$ ,  $w_{aB} = (1 - s_{\text{lab}}/2)$ ,  $w_{Ab} = w_{AB}(1 - \varepsilon_A)$ , and  $w_{ab} = w_{aB}(1 + \varepsilon_a)$ .

To calculate the mean fitness of population  $x$  (where  $x$  may represent one of the source populations or a hybrid cross such as  $F_1$  and  $F_2$ ), we must know the genotype frequencies. These genotype frequencies can be calculated as  $f_{AB,x} = p_{A,x}p_{B,x} + D_x$ ,  $f_{Ab,x} = p_{A,x}(1 - p_{B,x}) - D_x$ ,  $f_{aB,x} = (1 - p_{A,x})p_{B,x} - D_x$ , and  $f_{ab,x} = (1 - p_{A,x})(1 - p_{B,x}) + D_x$ , where  $D_x$  is the linkage disequilibrium in population  $x$ . These values for the source populations are given above and in the supplementary material. For the  $F_1$  hybrid population,  $p_{A,F_1} = 1/2(p_{A,1} + p_{A,2})$  and  $p_{B,F_1} = 1/2(p_{B,1} + p_{B,2})$ . The linkage disequilibrium in the  $F_1$  population is

$$D_{F_1} = \frac{1}{4} \left( 1 - r \left( 2 + \frac{m_1(s_1 + \varepsilon_A)}{s_1 \varepsilon_A (r + s_1 - \varepsilon_a)} + \frac{m_2(s_2 + \varepsilon_A + 2\varepsilon_a)}{\varepsilon_a (r + s_2 + \varepsilon_a) (\varepsilon_A + \varepsilon_a + s_2)} \right) \right) + O(\xi^2). \quad (8)$$

Assuming future hybrid generations can be produced without selection, the allele frequencies of the  $F_n$  hybrid generation remain the same as those in the  $F_1$ . The disequilibrium of later generation hybrids is given by  $D_{F_n} = D_{F_1}(1 - r)^{n-1}$ .

Intrinsic isolation is quantified by comparing the mean fitness of  $F_n$  hybrids to the average of the mean fitnesses of individuals from each of the two parental populations:

$$I_{F_n} = \frac{1}{2}(E[w_{P1}] + E[w_{P2}]) - E[w_{F_n}], \quad (9)$$

where  $E[w_x] = \sum f_{i,x} w_{i,x}$  is the mean fitness of population  $x$ . Evaluating this expression with the equilibrium genotype frequencies gives

$$I_{F_n} = (\varepsilon_A + \varepsilon_a) \left( \frac{1}{4} - D_{F_n} \right) + O(\xi^2). \quad (10)$$

This is always positive because  $0 \leq D_{F_n} < 1/4$  indicating that hybrids are less fit than parental types. As expected, the degree of isolation depends on the intrinsic epistatic effects of the modifier,  $\varepsilon_A$  and  $\varepsilon_a$ . The modifier allele *b* causes the evolution of intrinsic isolation;  $I_{F_n} = 0$  if *b* is absent from the system, but  $I_{F_n} > 0$  after *b* invades. To leading order, this measure of intrinsic isolation is largely independent of any inadvertent ecological selection in the lab,  $s_{\text{lab}}$ , though higher order terms include interactions between ecological and intrinsic selection (see supplementary material). Similarly, the true ecological selection that initiated divergence ( $s_1, s_2$ ) does not appear in (10). These parameters do have

a weak effect on  $I_{F_n}$  (see higher order terms shown in supplementary material) because they affect the degree of divergence between parental populations and thus the magnitude of disequilibrium in hybrids. While ecological selection does not have an important quantitative direct effect on the degree of intrinsic isolation, as measured in the lab, it does have a critical role. Ecological selection initiates the divergence between populations that allows the invasion of *b* and the establishment of a stable equilibrium in which both loci are polymorphic. In principle, the effects of ecology could be discerned if lab rearing conditions were to properly emulate the divergent environmental selection pressures experienced by the parental populations in nature.

**3.3. The Barrier to Gene Flow.** In the previous section, we quantified the extent of intrinsic isolation as might be measured empirically. Intuitively, we would expect the evolution of intrinsic isolation to serve as a barrier to gene flow and, as such, contribute to speciation [14, 32]. Here we show that this is not always true.

We quantify the barrier to gene flow using the metric proposed by Zhivotovsky and Christiansen [52]. This measure reflects the disadvantage experienced by an unlinked neutral marker that enters a focal population through migration from the other population. In the absence of differentiation between the populations, a neutral marker that migrates into a population would have no disadvantage and the barrier strength would be zero. If the two populations are completely incompatible such that all hybrids die, then the barrier strength is infinite. The barrier is quantified as

$$B = -\ln \left( \prod_{t=t_0}^{\infty} \frac{\bar{w}_{m,t}}{\bar{w}_{r,t}} \right), \quad (11)$$

where  $\bar{w}_{m,t}$  is the average fitness, in generation  $t$ , of the descendants of migrants that entered the population in generation  $t_0$ .  $\bar{w}_{r,t}$  is defined analogously for a corresponding set of resident descendants. This measure of barrier strength is closely related to that suggested by Bengtsson [53] provided that the focal neutral locus is unlinked to the genes responsible for the barrier [54].

The analytical approximations given below are based on the first five terms in the infinite series in (11). This provides a reasonably good approximation because the effectiveness of the barrier quickly declines with each additional generation postmigration as the unlinked neutral marker disassociates from migrant genetic background in which it initially immigrated. The results presented below assume that recombination is strong relative to selection. Details on the calculation of the barrier, including results that allow for tight linkage, are provided in the supplementary material.

Taking Population 1 as our focal population, we find that the barrier to gene flow of alleles from Population 2 into Population 1 is

$$B_1 = \frac{31}{32}(s_1 - \varepsilon_a) + \frac{1}{32}r(57 - 42r + 22r^2 - 7r^3 + r^4) \times (\varepsilon_A + \varepsilon_a) + O(\xi^2). \quad (12)$$

Several important points emerge from this result. Both ecological selection ( $s_1$ ) and intrinsic selection ( $\varepsilon_A, \varepsilon_a$ ) affect the strength of the barrier. Whereas ecological selection clearly has a positive effect on the barrier, the same cannot be said of intrinsic epistatic selection. The second term in (12) indicates a positive effect of intrinsic epistatic selection on barrier strength but the first term shows a negative effect. These conflicting effects can be understood as follows.

Consider an unlinked neutral allele  $n$  entering Population 1 from Population 2 in an  $ab$  genotype (the common genotype in Population 2). For as long as the marker allele  $n$  remains associated with the  $a$  allele, it suffers from ecological selection against  $a$  in Population 1. The  $b$  allele provides an intrinsic advantage to  $a$ -bearing individuals ( $\varepsilon_a$ ), thereby reducing the net selection against the  $ab$  genotype. By ameliorating the selection against  $a$ , the  $b$  allele makes it easier for the neutral marker  $n$  to successfully migrate into the new population, thus reducing the barrier (see first term of (12)). However, recombination can separate the  $a$  and  $b$  alleles from each other such that the marker allele  $n$  may then find itself in the  $Ab$  genotype, which suffers negative intrinsic epistatic selection ( $w_{ab} - w_{Ab} \approx \varepsilon_A + \varepsilon_a$ ). This breakdown of the coadapted gene complex makes it more difficult for the neutral marker to successfully introgress into the new population, thereby increasing the barrier strength (see second term of (12)).

The extent to which the marker will suffer from negative effects of the coadapted gene complex breakdown depends on the rate of recombination,  $r$ , between the **A** and **B** loci. If the **A** and **B** loci are tightly linked (i.e.,  $r$  is low), the unlinked marker will have left the  $ab$  haplotype before this haplotype is destroyed by recombination. Conversely, when  $r$  is high, the marker has a good chance of still being on the  $ab$  haplotype when recombination occurs between the **A** and **B** loci. This effect of recombination is reflected in the second term of (12) which shows that  $r$  mediates the negative effects of intrinsic epistatic effects on gene flow (or, equivalently, the positive effects on barrier strength). Tightly linked coadapted complexes serve as weaker barriers to gene flow at unlinked loci than loosely linked complexes. Gavrillets [32] previously showed that barrier strength increases with recombination for DM incompatibilities (see also Bengtsson [53]).

In the opposite direction, the leading order approximation for the barrier entering Population 2 from Population 1 is

$$B_2 = \frac{31}{32}(s_2 + \varepsilon_a) + \frac{1}{32}r(57 - 42r + 22r^2 - 7r^3 + r^4) \times (\varepsilon_A + \varepsilon_a) + O(\xi^2). \quad (13)$$

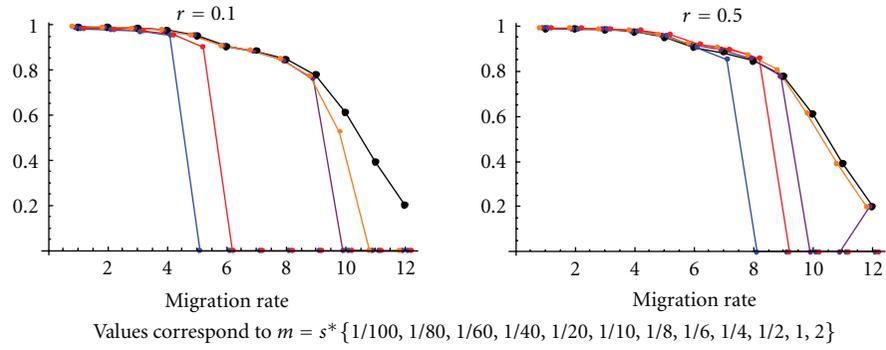
This equation is similar to (12) except the first term which is different in two ways. First, the relevant ecological selection is that which occurs in Population 2 (i.e.,  $s_2$  rather than  $s_1$ ). Second, the intrinsic epistatic effect is positive in (13) rather than being negative as in (12). Just as the presence of  $b$  ameliorated the ecological disadvantage of  $a$ -bearing migrants competing against  $A$ -bearing residents in Population 1 (causing a negative effect on  $B_1$ ), the  $b$  allele exaggerates the ecological advantage of  $a$ -bearing residents

in Population 2 competing against  $A$ -bearing migrants (i.e.,  $w_{AB}$  versus  $w_{ab}$  in Population 2). Combined with the second term, which reflects the consequences of breaking down the coadapted gene complex, the intrinsic epistatic effects always increase the barrier in Population 2.

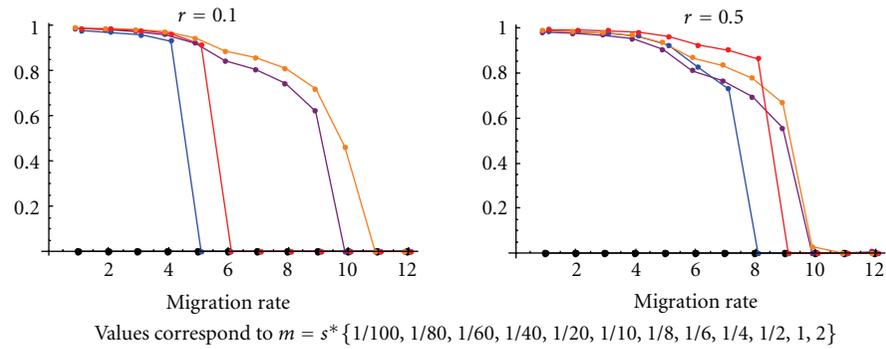
The reason that Population 1, rather than Population 2, experiences conflicting effects of intrinsic isolation is because of the perspective we have taken here. In discussing the effect of intrinsic isolation on the barrier strength, we have implicitly compared the barriers when  $b$  is absent (and there is no intrinsic isolation) to the barriers after  $b$  invades. The strength of the barriers when  $b$  is absent is given by (12) and (13) with  $\varepsilon_A = \varepsilon_a = 0$ . The effects of intrinsic isolation on each population should be interpreted with respect to that population's current genetic composition relative to its composition prior to the evolution of intrinsic isolation. Initially, Population 1 had the appropriate epistatic allele ( $B$ ), whereas Population 2 did not. The evolution of intrinsic isolation increases the barrier strength for the population in which the new modifier allele is common (Population 2). However, the effect on the barrier in the other population (Population 1) may be positive or negative, depending on recombination.

**3.4. Numerical Simulations.** Numerical simulations were performed to test several of the major conclusions from analytical approximations presented above. The results are presented in Figures 2 and 3. Divergence at the **A** and **B** loci between the two populations is shown in Figures 2(a), 2(b), 3(a), and 3(b) for Cases 1 and 2 (full model), respectively. In each figure, the black line represents the level of divergence before the evolution of the  $b$  allele whereas the colored lines show divergence after the introduction of this allele. For both models, divergence between populations is only possible when migration is weak relative to selection (Figures 2(a), 2(b), 3(a) and 3(b)). Divergence at the **B** locus (i.e., evolution of the epistatic modifier  $b$ ) typically has little quantitative effect on divergence at the **A** locus (compare colored lines to black lines in Figures 2(a) and 3(a)), provided polymorphism persists. However, as discussed in our analytical section above, the introduction of the epistatic modifier can completely destabilize the polymorphism, causing loss of divergence at both loci. This tends to occur when the epistatic modifier confers a large benefit to the  $a$  allele (i.e.,  $\varepsilon_a$  large relative to  $s$ ) and there is linkage between the **A** and **B** loci (compare left [ $r = 0.1$ ] and right [ $r = 0.5$ ] columns in Figures 2(a) and 2(b)). Direct ecological selection on the **B** locus helps to maintain divergence at both loci (compare Figures 2(a) and 2(b) with Figures 3(a) and 3(b)).

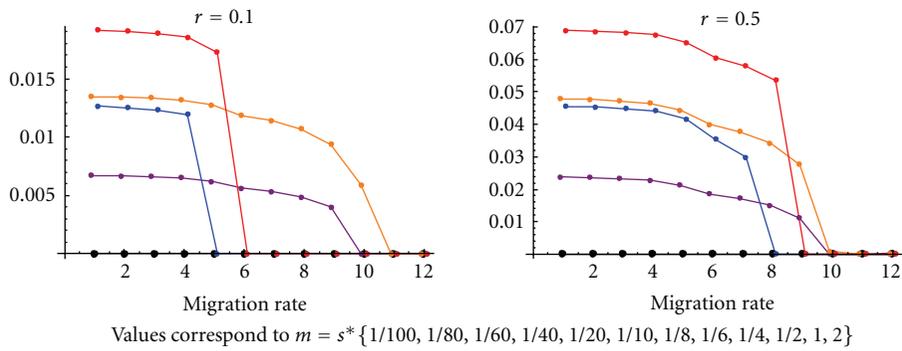
The strength of intrinsic isolation is shown in Figures 2(c) and 3(c) for Cases 1 and 2, respectively. As indicated by the points in black, there is no intrinsic isolation prior to the evolution of the  $b$  allele. However, the introduction of the epistatic modifier causes the evolution of intrinsic isolation in the face of gene flow. The strength of intrinsic isolation increases with the magnitude of epistatic effects on either background ( $\varepsilon_A$  and  $\varepsilon_a$ ) (see colored lines in Figures 2(c) and 3(c)). Additional simulations (not shown) confirm that divergent ecological selection on the **B** locus has little



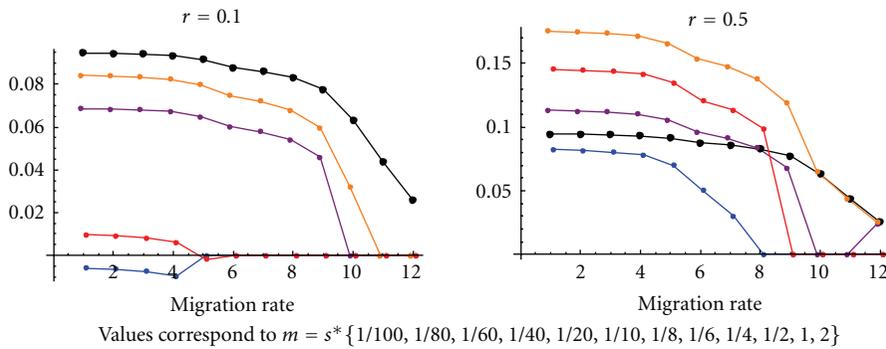
(a) Divergence at A locus



(b) Divergence at B locus



(c) Intrinsic Isolation (based on  $F_2$  hybrids)



— Before  $b$  introduced  
 —  $\epsilon_A = 0.5s, \epsilon_a = 0.5s$   
 —  $\epsilon_A = 1.5s, \epsilon_a = 0.5s$   
 —  $\epsilon_A = 0.5s, \epsilon_a = 1.5s$   
 —  $\epsilon_A = 1.5s, \epsilon_a = 1.5s$

(d) Barrier into Population 1

FIGURE 2: Continued.

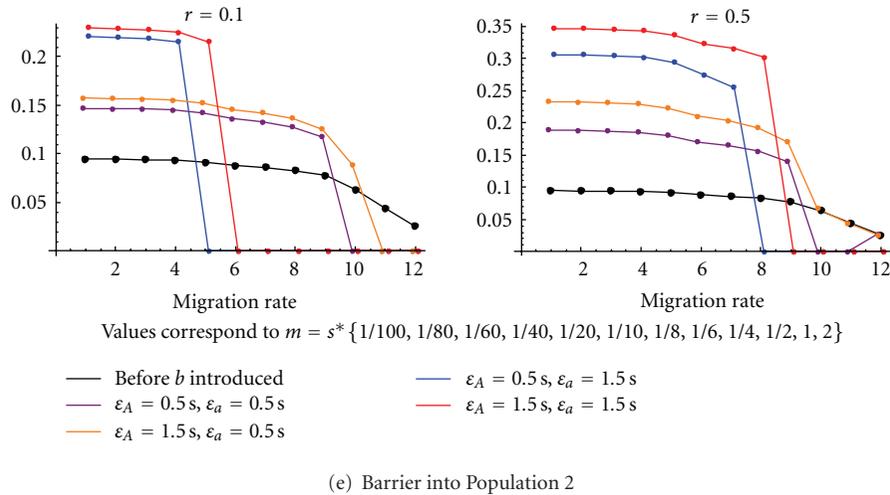


FIGURE 2: Numerical simulation results for the epistatic trade-off model (Case 1). Various properties are shown as a function of the migration rate, expressed relative to the strength of ecological selection on the **A** locus,  $s$ . The migration rates numbered 1–12 correspond to  $m = s^* \{1/100, 1/80, 1/60, 1/40, 1/20, 1/10, 1/8, 1/6, 1/4, 1/2, 1, 2\}$ ; migration is symmetric, that is,  $m_1 = m_2 = 1/2m$  ( $c = 1/2$ ). The points in black represent values prior to the evolution of the modifier (i.e., assuming both populations fixed for the **B** allele as shown in Time 1 in Figure 1). Colored points represent values after the evolution of the modifier; different colors correspond to modifiers of varying effects (see legend). Left column: moderate linkage  $r = 0.1$ ; right column: no linkage:  $r = 0.5$ . (a) Divergence at **A** locus. (b) Divergence at **B** locus. (c) Intrinsic isolation as measured by the average fitness individuals from Populations 1 and 2 compared to the average fitness of  $F_2$  hybrids. (d) Strength of the barrier into Population 1. (e) Strength of barrier into Population 2. Parameter values used:  $s = s_1 = s_2 = 0.1$ ;  $t_1 = t_2 = 0$ . See text for details.

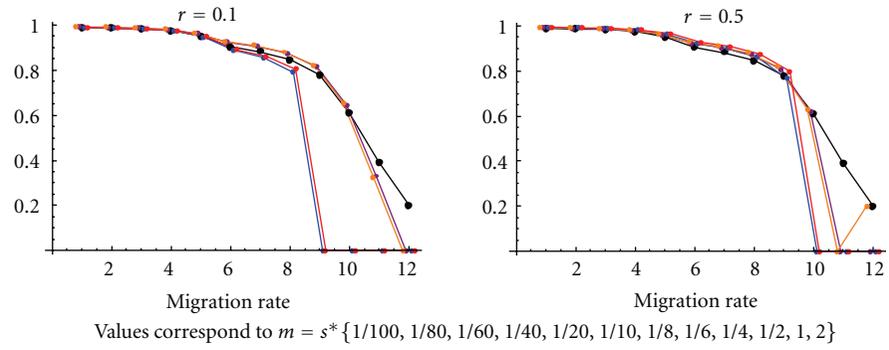
quantitative effect on strength of intrinsic isolation, provided polymorphism persists. However, ecological selection can have a critical qualitative effect by allowing divergence to occur despite gene flow; intrinsic isolation is limited by the genetic differentiation at the interacting genes.

When measured by using early generation hybrids (e.g.,  $F_2$ ), intrinsic isolation is larger if **A** and **B** are unlinked rather than linked (compare left and right columns of Figures 2(c) and 3(c), noting scales differ). This is because coadapted gene complexes are more likely to be broken down in early generation hybrids if recombination is high. However, if intrinsic isolation is measured using very late generation hybrids, say  $F_{100}$ , then linkage has a smaller effect because linkage disequilibrium has been mostly eliminated in late generation hybrids even if recombination is low (not shown).

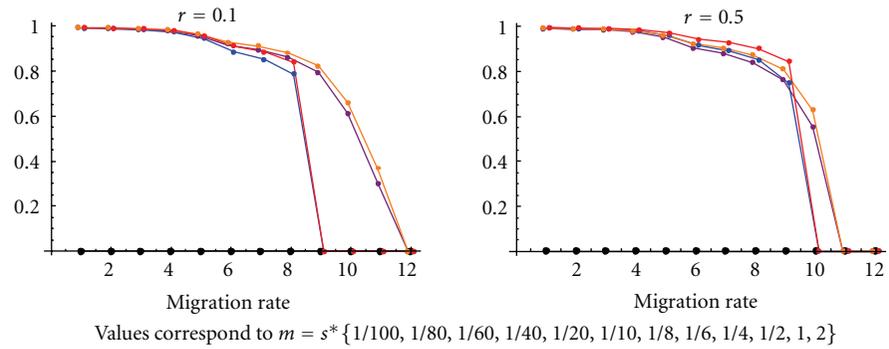
The strength of the barrier to gene flow into each population is shown in Figures 2(d), 2(e), 3(d), and 3(e) for Cases 1 and 2, respectively. The black line represents the barrier strength prior to the introduction of the  $b$  allele. Differentiation at the **A** locus alone creates a barrier to gene flow because of divergent ecological selection on this gene. The evolution of  $b$  always increases the strength of Barrier 2 (i.e., colored lines lie above black line). Both beneficial epistatic effects on the  $a$ -allele background ( $\epsilon_a$ ) and deleterious epistatic effects on the  $A$ -allele background ( $\epsilon_A$ ) strengthen the barrier (see Figure 2(e)). As predicted from the analytical approximations, the  $b$  allele can either increase or decrease the strength of Barrier 1;  $\epsilon_A$  increases this barrier, whereas  $\epsilon_a$  decreases this barrier. For example, high values of  $\epsilon_a$  greatly reduce the barrier (red and blue lines in Figure 2(d)), though these negative effects occur

primarily when there is linkage between genes. As predicted analytically, epistatic effects increase barriers to a greater extent when there is loose linkage (compare left and right columns in Figures 2(d) and 2(e)). If the interacting genes (**A** and **B**) are closely linked, a neutral marker allele has a better chance of escaping a migrant chromosome before an adaptive complex (such as  $ab$ ) is converted into a maladaptive complex (such as  $Ab$ ). Thus, linkage between the interacting genes makes it easier for neutral markers to move between populations, as observed for past models that focused on barriers to gene flow per se [14, 32]. Finally, and unsurprisingly, it is clear that direct ecological selection on the **B** locus strengthens the barrier to gene flow (compare Figures 2(d) and 2(e) with 3(d) and 3(e)).

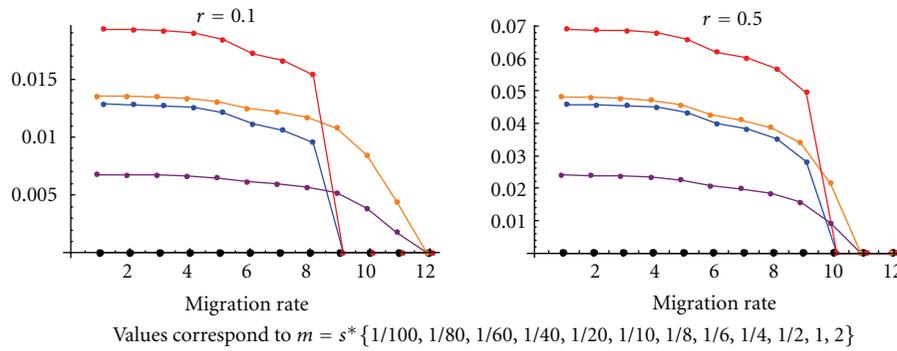
A comparison of the analytical approximations to simulation results is presented in the supplementary material (Figures S1 and S2). As expected, the analytical results provide good approximations under the conditions under which they were derived ( $m \ll s$ ). The analytical approximations tend to underestimate the magnitude of intrinsic isolation and overestimate the barrier strengths. The deviations of the analytical approximations from the simulation results increase as migration becomes of the same magnitude as selection. The most obvious discrepancy occurs when the analytical results wrongly predict the two-locus polymorphism is stable. Instability is predicted to occur when migration is strong relative to selection. However, the analysis is based on a weak migration assumption, so it is not surprising that our analytically derived stability conditions are not very good at quantitatively predicting instability.



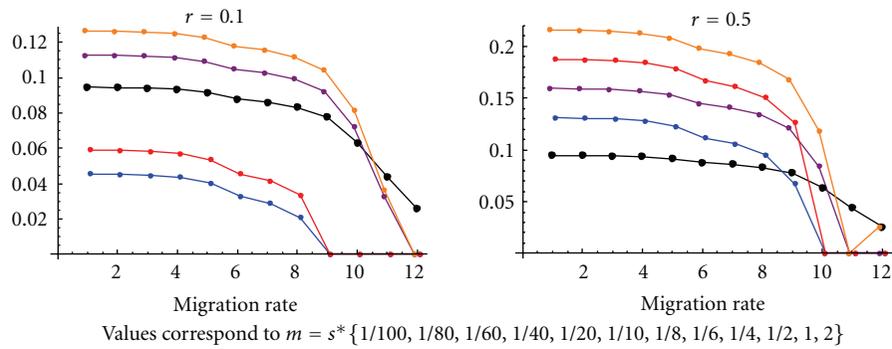
(a) Divergence at **A** locus



(b) Divergence at **B** locus



(c) Intrinsic Isolation (based on  $F_2$  hybrids)



— Before  $b$  introduced  
 —  $\epsilon_A = 0.5s, \epsilon_a = 0.5s$   
 —  $\epsilon_A = 1.5s, \epsilon_a = 0.5s$   
 —  $\epsilon_A = 0.5s, \epsilon_a = 1.5s$   
 —  $\epsilon_A = 1.5s, \epsilon_a = 1.5s$

(d) Barrier into Population 1

FIGURE 3: Continued.

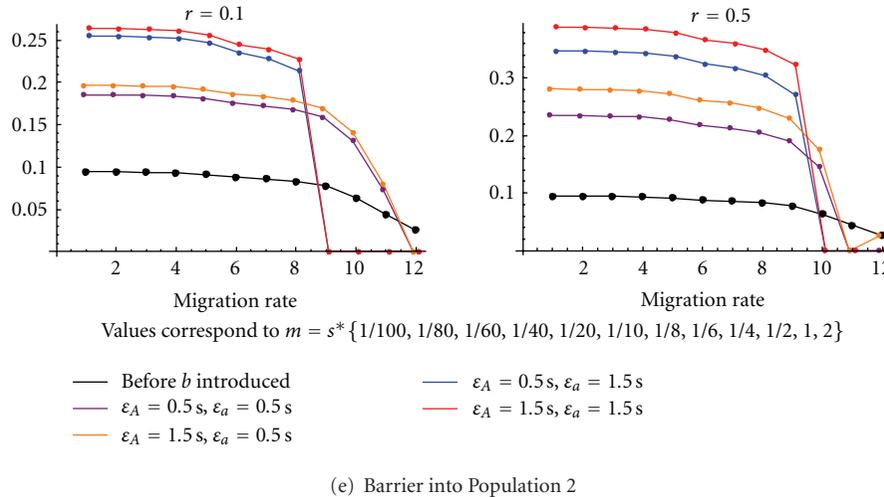


FIGURE 3: Numerical simulation results for the model with modifier with both epistatic trade-off and direct, divergent selection (Case 2, full model). Parameter values used:  $s = s_1 = s_2 = 0.1$ ;  $t_1 = t_2 = 0.5$  s. See Figure 2 for other details.

We have also tested whether our results are sensitive to the assumption of haploidy by performing a series of simulations for diploids. Results for these simulations are presented in the supplementary material (Figures S(3)–S(8)). Qualitatively, all of our major results apply to diploids, including the counter intuitive finding that the evolution of intrinsic isolation can reduce barrier strength. This is most likely to occur in diploid systems when negative epistatic effects are recessive (compare Figures S3(d), S6(d), and S9(d)). This type of gene action underlies the dominance theory of Haldane’s Rule [55] and is common in *Drosophila* incompatibilities [3, 7, 56].

## 4. Discussion

**4.1. The Evolution of Intrinsic Postmating Isolation.** Here, we presented a theoretical analysis of the role that divergent natural selection can play in generating intrinsic reproductive isolation during divergence with gene flow. Our objective was to assess the extent to which intrinsic reproductive isolation must be a by-product of genetic divergence in allopatry.

The typical DM incompatibility is often depicted as shown in Table 1 where fitnesses are not ecologically dependent. This representation has the characteristic property that two high-fitness genotypes ( $AB$  and  $ab$ ) are connected by a ridge of high-fitness genotypes, so it is possible to move from one high-fitness genotype to the other by a series of more or less neutral substitutions. This type of representation is common in text books (Barton et al. [57, page 643]) and speciation models (see [14, 27, 32, 49], see also Table 2). Truly allopatric populations can diverge along neutral ridges in the adaptive landscape. However, such divergence will be severely impeded if there is gene flow both because migration directly homogenizes populations and because selection will oppose alleles that are incompatible with other alleles found in some parts of the range. Even if divergence occurs via drift in allopatry, the two-locus polymorphism will collapse

once contact is reestablished in a two-patch model [32, 53]. Such DM incompatibilities only become stable when more complicated geographies, such as stepping-stone type models, are invoked (e.g., see [32]).

Alternatively, high fitness ridges may not be completely flat so that selection can drive the local fixation of alleles that have the potential to be involved in incompatibilities. However, when there is gene flow, such alleles are expected to quickly spread across a population as they are initially beneficial everywhere if there is no divergent ecological selection. In the face of gene flow, incompatibilities can only evolve if mutations at different interacting loci spread simultaneously in different parts of the metapopulation [36] or if new alleles interact with previously established incompatibilities [37].

Although it is intuitively obvious that the addition of divergent ecological selection will make it easier to establish incompatibilities in the face of gene flow, this possibility has been largely ignored in terms of formal models (but see [18]). Our simplest finding was that maintaining differentiation is the primary contribution of divergent selection to the evolution of intrinsic isolation. In particular, establishing genetic differentiation at ecologically important loci provides the appropriate setting for the subsequent evolution of epistatically interacting modifiers that can pleiotropically cause intrinsic reproductive isolation. This can happen in two ways. Either divergent ecological selection can act directly on the modifier locus, or not.

Divergent ecological selection acting directly on a modifier locus (Case 2) is the easiest way to maintain high levels of differentiation, even with strong epistatic effects. In the case where ecological selection does not act directly on the modifier locus (Case 1), there are limits to the strength of intrinsic isolation that can evolve. It is difficult for epistatic modifiers with strong effects to remain differentiated between populations, unless there is the right balance between the modifier’s positive effect on the

$a$ -allele background ( $\epsilon_a$ ) and its negative effect on the  $A$ -allele background ( $\epsilon_A$ ). Without the appropriate balance, strong modifiers are either lost in both populations or fixed in both, in the latter case often eliminating polymorphism at the ecologically differentiated locus ( $A$ ) in the process. Loose linkage makes it somewhat easier to maintain a high level of differentiation for strong modifiers, thus allowing for the evolution of stronger intrinsic isolation (compare left and right columns of Figure 2).

In sum, divergent ecological selection allows intrinsic isolation to evolve through alleles that would otherwise not contribute to reproductive barriers in the presence of gene flow. For example, consider a modifier that did not experience ecological selection and was neutral on one background but had intrinsic epistatic effects on the other. It would either spread to fixation in both populations or be completely lost, thus not contributing to reproductive isolation. However, when the modifier also experiences divergent ecological selection, populations can differentiate with respect to allele frequencies for this gene, allowing reproductive isolation to evolve.

**4.2. The Barrier to Neutral Gene Flow.** An interesting counterintuitive and somewhat unexpected result from our analysis is that the same genes that cause intrinsic isolation can *reduce* the barrier to gene flow. By improving the intrinsic fitness of an ecologically maladaptive allele ( $a$ ), an epistatic modifier ( $b$ ) reduces the net selection against foreign genotypes, potentially making it easier for foreign alleles to enter the focal population. For a related reason, the epistatic modifiers typically have asymmetric effects on the barriers to gene flow. The barrier into the population containing the derived modifier allele ( $b$ ) tends to be stronger than the barrier into the other population. The derived modifier allele's beneficial epistatic effect on its preferred background makes it easier for genotypes from its population to invade the other population while at the same time making it more difficult for the reverse direction of gene flow.

Our result about the asymmetric effects of epistatic modifiers on barrier strengths seems related to asymmetric effects with respect to intrinsic isolation in classic DM incompatibilities. We found that the barrier from the population containing the derived modifier allele into the population with the ancestral allele tends to be weaker than the reverse situation, whereas Orr [28] showed that derived alleles are more likely to involve (cause) incompatibilities. However, the underlying reasons for these two types of "asymmetries" are very different. Our result is driven by the derived modifier improving the intrinsic fitness of one genotype, thereby mitigating its ecological disadvantage in one habitat and augmenting its ecological advantage in the other. In contrast, Orr's result does not involve ecological dependence and is driven by negative effects of the derived allele in (hybrid) combinations that have gone untested by selection in allopatry.

**4.3. Implications for Empirical Studies of "Speciation Genes".** Much empirical work has been dedicated to finding "speciation genes" underlying reproductive isolation (for reviews

see [3, 5, 6, 58–61]). Most studies have centered on genes underlying intrinsic postmating isolation, particularly in hybrids between species of *Drosophila*. A major discovery emerging from this work, particularly in animal systems, is that genes underlying intrinsic postzygotic isolation often show molecular signatures of positive selection [3, 9, 62–64]. In one case, the gene "Overdrive", the mechanism of selection appears to be intragenomic conflict rather than environmental, ecologically based selection [8], and a similar mechanism is suspected in other cases [7, 65]. In contrast, genetic studies in yeast implicate ecological adaptation as the driver of postmating isolation [47, 48]. Nevertheless, in most cases of putative speciation genes, the mechanisms of selection involved are not yet known. As discussed in the introduction, there is empirical evidence from other types of studies that ecological divergence can indeed drive the evolution of intrinsic postmating isolation. Collectively, these studies suggest that the theoretical scenarios treated in our paper are plausible and that studies of speciation genes might consider whether the genes involved evolve via divergent selection, perhaps in the face of gene flow.

## 5. Conclusions

Postzygotic isolation is often dichotomized as "intrinsic" or "ecological/extrinsic". However, many cases of the speciation process may involve both [3]. Divergent ecological selection is well documented and can allow for genetic differentiation despite gene flow [14, 18]. Moreover, adaptation to particular environmental challenges may often come at some pleiotropic cost. This creates the potential for epistatic modifiers that ameliorate such costs but that may be incompatible on the original genetic background. Under such conditions, divergent ecological selection can be considered the catalyst for the evolution of intrinsic isolation. However, the genes that contribute to intrinsic isolation may have mixed effects on barrier to gene flow. Thus, although divergent selection can thus promote the evolution of intrinsic isolation, the overall importance of this for the process of ecological speciation will depend on (1) the degree to which this actually reduces gene flow and (2) the extent and rate to which this occurs relative to commonly documented scenarios where divergent selection drives the evolution of premating or extrinsic barriers to gene flow [3, 24, 25, 66].

## Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant) to A. F. Agrawal, as well as NSF and USDA support to J. L. Feder. During a period of working on this project, P. Nosil and J. L. Feder were hosted by the Institute for Advanced Study, *Wissenschaftskolleg*, Berlin. The authors thank S. Gavrillets and other anonymous reviewers for comments on previous versions of this manuscript.

## References

- [1] J. A. Coyne and H. A. Orr, "Patterns of speciation in *Drosophila*," *Evolution*, vol. 43, pp. 362–381, 1989.

- [2] J. Ramsey, H. D. Bradshaw, and D. W. Schemske, "Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae)," *Evolution*, vol. 57, no. 7, pp. 1520–1534, 2003.
- [3] J. A. Coyne and H. A. Orr, *Speciation*, Sinauer Associates, Sunderland, Mass, USA, 2004.
- [4] P. Nosil, T. H. Vines, and D. J. Funk, "Perspective: reproductive isolation caused by natural selection against immigrants from divergent habitats," *Evolution*, vol. 59, no. 4, pp. 705–719, 2005.
- [5] C. I. Wu and C. T. Ting, "Genes and speciation," *Nature Reviews Genetics*, vol. 5, no. 2, pp. 114–122, 2004.
- [6] H. A. Orr, "The genetic basis of reproductive isolation: insights from *Drosophila*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 1, pp. 6522–6526, 2005.
- [7] D. C. Presgraves, "Speciation genetics: epistasis, conflict and the origin of species," *Current Biology*, vol. 17, no. 4, pp. R125–R127, 2007.
- [8] N. Phadnis and H. A. Orr, "A single gene causes both male sterility and segregation distortion in *Drosophila* hybrids," *Science*, vol. 323, no. 5912, pp. 376–379, 2009.
- [9] S. Tang and D. C. Presgraves, "Evolution of the *Drosophila* nuclear pore complex results in multiple hybrid incompatibilities," *Science*, vol. 323, no. 5915, pp. 779–782, 2009.
- [10] W. Bateson, "Heredity and variation in modern lights," in *Darwin and Modern Science*, A. C. Seward, Ed., pp. 85–101, Cambridge University Press, Cambridge, UK, 1909.
- [11] T. Dobzhansky, *Genetics and the Origin of Species*, Columbia University Press, New York, NY, USA, 1st edition, 1973.
- [12] H. J. Muller, "Bearing of the *Drosophila* work on systematics," in *The New Systematics*, J. Huxley, Ed., pp. 185–268, Oxford University Press, Oxford, UK, 1940.
- [13] H. J. Muller, "Isolating mechanisms, evolution, and temperature," *Biology Symposium*, vol. 6, pp. 71–125, 1942.
- [14] S. Gavrillets, *Fitness Landscapes and the Origin of Species*, Princeton University Press, Princeton, NJ, USA, 2004.
- [15] J. B. S. Haldane, "The theory of a cline," *Journal of Genetics*, vol. 48, no. 3, pp. 277–284, 1948.
- [16] B. Clarke, "Evolution of morph-ratio clines," *American Naturalist*, vol. 100, pp. 389–402, 1966.
- [17] J. A. Endler, "Gene flow and population differentiation," *Science*, vol. 179, no. 4070, p. 243, 1973.
- [18] J. A. Endler, *Geographic Variation, Speciation, and Clines*, Princeton University Press, Princeton, NJ, USA, 1977.
- [19] N. H. Barton and G. M. Hewitt, "Analysis of hybrid zones," *Annual review of ecology and systematics*, vol. 16, pp. 113–148, 1985.
- [20] N. H. Barton and G. M. Hewitt, "Adaptation, speciation and hybrid zones," *Nature*, vol. 341, no. 6242, pp. 497–503, 1989.
- [21] D. J. Funk, "Isolating a role for natural selection in speciation: host adaptation and sexual isolation in *Neochlamisus bebbianae* leaf beetles," *Evolution*, vol. 52, no. 6, pp. 1744–1759, 1998.
- [22] D. Schluter, "Ecological causes of speciation," in *Endless Forms: Species and Speciation*, D. J. Howard and S. H. Berlocher, Eds., pp. 114–129, Oxford University Press, Oxford, UK, 1998.
- [23] D. Schluter, "Ecology and the origin of species," *Trends in Ecology & Evolution*, vol. 16, no. 7, pp. 372–380, 2001.
- [24] H. D. Rundle and P. Nosil, "Ecological speciation," *Ecology Letters*, vol. 8, no. 3, pp. 336–352, 2005.
- [25] D. J. Funk, P. Nosil, and W. J. Etges, "Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 9, pp. 3209–3213, 2006.
- [26] D. Schluter, "Evidence for ecological speciation and its alternative," *Science*, vol. 323, no. 5915, pp. 737–741, 2009.
- [27] M. Nei, T. Maruyama, and C. Wu, "Models of evolution of reproductive isolation," *Genetics*, vol. 103, no. 3, pp. 557–579, 1983.
- [28] H. A. Orr, "The population genetics of speciation: the evolution of hybrid incompatibilities," *Genetics*, vol. 139, no. 4, pp. 1805–1813, 1995.
- [29] M. Slatkin, "Epistatic selection opposed by immigration in multiple locus genetic systems," *Journal of Evolutionary Biology*, vol. 8, no. 5, pp. 623–633, 1995.
- [30] N. H. Barton, "The role of hybridization in evolution," *Molecular Ecology*, vol. 10, no. 3, pp. 551–568, 2001.
- [31] H. A. Orr and L. H. Orr, "Waiting for speciation: the effect of population subdivision on the time to speciation," *Evolution*, vol. 50, no. 5, pp. 1742–1749, 1996.
- [32] S. Gavrillets, "Hybrid zones with Dobzhansky-type epistatic selection," *Evolution*, vol. 51, no. 4, pp. 1027–1035, 1997.
- [33] H. A. Orr and M. Turelli, "The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities," *Evolution*, vol. 55, no. 6, pp. 1085–1094, 2001.
- [34] S. A. Church and D. R. Taylor, "The evolution of reproductive isolation in spatially structured populations," *Evolution*, vol. 56, no. 9, pp. 1859–1862, 2002.
- [35] A. H. Porter and N. A. Johnson, "Speciation despite gene flow when developmental pathways evolve," *Evolution*, vol. 56, no. 11, pp. 2103–2111, 2002.
- [36] A. S. Kondrashov, "Accumulation of Dobzhansky-Muller incompatibilities within a spatially structured population," *Evolution*, vol. 57, no. 1, pp. 151–153, 2003.
- [37] A. Navarro and N. H. Barton, "Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation," *Evolution*, vol. 57, no. 3, pp. 447–459, 2003.
- [38] S. Gavrillets and A. Hastings, "Founder effect speciation: a theoretical reassessment," *American Naturalist*, vol. 147, no. 3, pp. 466–491, 1996.
- [39] S. Gavrillets, "Waiting time to parapatric speciation," *Proceedings of the Royal Society B*, vol. 267, no. 1461, pp. 2483–2492, 2000.
- [40] G. A. Hoelzer, R. Drewes, J. Meier, and R. Doursat, "Isolation-by-distance and outbreeding depression are sufficient to drive parapatric speciation in the absence of environmental influences," *PLoS Computational Biology*, vol. 4, no. 7, Article ID e1000126, 2008.
- [41] P. Nosil and S. M. Flaxman, "Conditions for mutation-order speciation," *Proceedings of the Royal Society B*, vol. 278, pp. 399–407, 2011.
- [42] H. D. Rundle and M. C. Whitlock, "A genetic interpretation of ecologically dependent isolation," *Evolution*, vol. 55, no. 1, pp. 198–201, 2001.
- [43] R. A. Fisher, *The Genetical Theory of Natural Selection*, Clarendon Press, Oxford, UK, 1930.
- [44] M. Turelli, N. H. Barton, and J. A. Coyne, "Theory and speciation," *Trends in Ecology & Evolution*, vol. 16, no. 7, pp. 330–343, 2001.
- [45] M. R. Macnair and P. Christie, "Reproductive isolation as a pleiotropic effect of copper tolerance in *Mimulus-guttatus*," *Heredity*, vol. 50, pp. 295–302, 1983.
- [46] D. I. Bolnick, T. J. Near, and P. C. Wainwright, "Body size divergence promotes post-zygotic reproductive isolation in centrarchids," *Evolutionary Ecology Research*, vol. 8, no. 5, pp. 903–913, 2006.

- [47] J. R. Dettman, C. Sirjusingh, L. M. Kohn, and J. B. Anderson, "Incipient speciation by divergent adaptation and antagonistic epistasis in yeast," *Nature*, vol. 447, no. 7144, pp. 585–588, 2007.
- [48] H. Y. Lee, J. Y. Chou, L. Cheong, N. H. Chang, S. Y. Yang, and J. Y. Leu, "Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species," *Cell*, vol. 135, no. 6, pp. 1065–1073, 2008.
- [49] S. Gavrilets, "Perspective: models of speciation: what have we learned in 40 years?" *Evolution*, vol. 57, no. 10, pp. 2197–2215, 2003.
- [50] A. P. Hendry, P. Nosil, and L. H. Rieseberg, "The speed of ecological speciation," *Functional Ecology*, vol. 21, no. 3, pp. 455–464, 2007.
- [51] P. Nosil, "Adaptive population divergence in cryptic color-pattern following a reduction in gene flow," *Evolution*, vol. 63, no. 7, pp. 1902–1912, 2009.
- [52] L. A. Zhivotovsky and F. B. Christiansen, "The selection barrier between populations subject to stabilizing selection," *Evolution*, vol. 49, no. 3, pp. 490–501, 1995.
- [53] B. O. Bengtsson, *The Flow of Genes through a Genetic Barrier*, Cambridge University Press, Cambridge, UK, 1985.
- [54] K. V. Pylkov, L. A. Zhivotovsky, and F. B. Christiansen, "The strength of the selection barrier between populations," *Genetical Research*, vol. 76, no. 2, pp. 179–185, 2000.
- [55] M. Turelli and H. A. Orr, "The dominance theory of Haldane's rule," *Genetics*, vol. 140, no. 1, pp. 389–402, 1995.
- [56] D. C. Presgraves, L. Balagopalan, S. M. Abmayr, and H. A. Orr, "Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*," *Nature*, vol. 423, no. 6941, pp. 715–719, 2003.
- [57] N. H. Barton, D. E. G. Briggs, J. A. Eisen, D. B. Goldstein, and N. H. Patel, *Evolution*, CSHL Press, Cold Spring Harbor, NY, USA, 2007.
- [58] H. A. Orr and D. C. Presgraves, "Speciation by postzygotic isolation: forces, genes and molecules," *BioEssays*, vol. 22, no. 12, pp. 1085–1094, 2000.
- [59] K. Bomblies, "Doomed lovers: mechanisms of isolation and incompatibility in plants," *Annual Review of Plant Biology*, vol. 61, pp. 109–124, 2010.
- [60] L. H. Rieseberg and B. K. Blackman, "Speciation genes in plants," *Annals of Botany*, vol. 106, no. 3, pp. 439–455, 2010.
- [61] P. Nosil and D. Schluter, "The genes underlying the process of speciation," *Trends in Ecology & Evolution*, vol. 26, pp. 160–167, 2011.
- [62] D. A. Barbash, D. F. Siino, A. M. Tarone, and J. Roote, "A rapidly evolving MYB-related protein causes species isolation in *Drosophila*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 9, pp. 5302–5307, 2003.
- [63] D. A. Barbash, P. Awadalla, and A. M. Tarone, "Functional divergence caused by ancient positive selection of a *Drosophila* hybrid incompatibility locus," *PLoS Biology*, vol. 2, no. 6, pp. 839–848, 2004.
- [64] J. Mallet, "What does *Drosophila* genetics tell us about speciation?" *Trends in Ecology & Evolution*, vol. 21, no. 7, pp. 386–393, 2006.
- [65] D. C. Presgraves, "Does genetic conflict drive rapid molecular evolution of nuclear transport genes in *Drosophila*?" *BioEssays*, vol. 29, no. 4, pp. 386–391, 2007.
- [66] T. C. Mendelson, "Sexual isolation evolves faster than hybrid inviability in a diverse and sexually dimorphic genus of fish (Percidae: *Etheostoma*)," *Evolution*, vol. 57, no. 2, pp. 317–327, 2003.