The Limits to Parapatric Speciation: Dobzhansky–Muller Incompatibilities in a Continent–Island Model

Claudia Bank,[*,†,1] **Reinhard Bürger,**[‡] **and Joachim Hermisson**[†,‡]
[*]Institut für Populationsgenetik, Vetmeduni Vienna, 1210 Vienna, Austria, [†]Mathematics and Biosciences Group, Max F. Perutz Laboratories, 1030 Vienna, Austria, and [‡]Department of Mathematics, University of Vienna, 1090 Vienna, Austria

---

ABSTRACT How much gene flow is needed to inhibit speciation by the accumulation of Dobzhansky–Muller incompatibilities (DMIs) in a structured population? Here, we derive these limits in a classical migration–selection model with two haploid or diploid loci and unidirectional gene flow from a continent to an island. We discuss the dependence of the maximum gene-flow rate on ecological factors (exogeneous selection), genetic factors (epistasis, recombination), and the evolutionary history. Extensive analytical and numerical results show the following: (1) The maximum rate of gene flow is limited by exogeneous selection. In particular, maintenance of neutral DMIs is impossible with gene flow. (2) There are two distinct mechanisms that drive DMI evolution in parapatry, selection against immigrants in a heterogeneous environment and selection against hybrids due to the incompatibility. (3) Depending on the mechanism, opposite predictions result concerning the genetic architecture that maximizes the rate of gene flow a DMI can sustain. Selection against immigrants favors evolution of tightly linked DMIs of arbitrary strength, whereas selection against hybrids promotes the evolution of strong unlinked DMIs. In diploids, the fitness of the double heterozygotes is the decisive factor to predict the pattern of DMI stability.

THE (Bateson–) Dobzhansky–Muller model (DMM) (Bateson 1909; Dobzhansky 1936; Muller 1942) is the standard model to explain the evolution of intrinsic postzygotic isolation. A variety of theoretical studies show the plausibility of this model (reviewed in Coyne and Orr 2004; Gavrilets 2004), and numerous empirical studies report Dobzhansky–Muller incompatibilities (DMIs) across several species of animals and plants (see reviews by Lowry *et al.* 2008 and Presgraves 2010).

The appeal of the DMM is its generality: Speciation happens as a by-product of divergence without the need of special selection scenarios or complex adaptations to cross fitness valleys (*cf*. Orr 1995). The sole, but crucial, assumption is sufficient evolutionary time to let the process unfold and the spatial separation of the incipient species during this phase. The plausibility of the DMM as a mechanism for speciation and the widespread belief in the prevalence of allopatric speciation (Mayr 1942; Coyne and Orr 2004) go hand in hand. However, given that even tiny amounts of gene flow (of the order of a single migrant per generation) can have substantial effects on divergence (Slatkin 1987), it is clear that the assumption of strict allopatry is an idealization. This poses the question of the relevance of the DMM in parapatry and thus whether DMIs can originate or be maintained in the presence of gene flow. It has been suggested that the accumulation of DMIs in parapatry is possible (*e.g.*, Gavrilets 1997, 2004; Porter and Johnson 2002; Kondrashov 2003; Agrawal *et al.* 2011; Nosil and Flaxman 2011). However, usually only the limit of weak migration has been studied. In contrast, we focus on the maximum rates of gene flow that admit the maintenance of a DMI.

A key factor in speciation via the DMM is the mechanism that drives divergence. Although the original DMM is consistent with neutral substitutions that spread solely by genetic drift, all previous results indicate that selection as a driving force is needed in the presence of gene flow (*e.g.*, Gavrilets 1997). Recently, a distinction of two categories was brought forward to describe the selection scenarios that can promote speciation through DMIs in a structured population

---

doi: 10.1534/genetics.111.137513
Manuscript received February 19, 2012; accepted for publication April 20, 2012
Supporting information is available online at http://www.genetics.org/content/suppl/2012/04/27/genetics.111.137513.DC1.
[1] Corresponding author: School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), EPFL SV-IBI, Station 15, CH-1015 Lausanne, Switzerland. E-mail: claudia.bank@epfl.ch

(Schluter 2009; but see also Turelli *et al.* 2001; Rundell and Price 2009): Whereas "ecological speciation" refers to the case where divergent selection results in the spread of (potentially incompatible) locally adapted substitutions, "mutation-order speciation" is characterized by fixation of different globally advantageous alleles in populations adapting to similar selection pressures. The latter mechanism is termed "mutation order" because it is assumed to be critically dependent on the order in which substitutions arise (Mani and Clarke 1990).

Whereas speciation due to ecological divergence is supported by considerable empirical evidence, only a few cases of mutation-order speciation have been identified (reviewed in Schluter 2009). Furthermore, simulation studies have strengthened verbal arguments claiming that mutation-order speciation is expected to be less likely than ecological speciation (Unckless and Orr 2009; Nosil and Flaxman 2011). However, there is a lack of analytical work to justify this nomenclature and to quantify the conditions for the evolution of DMIs according to these mechanisms.

In this study, we address these issues in two minimal models for the evolution of a DMI on an island that receives migrants from a continent. We are able to provide an almost complete characterization of the conditions and limits for the origination and maintenance of DMIs in the presence of gene flow. Models, results, and biological interpretations are presented in the main body of the article. Our methods combine classical approaches (Karlin and McGregor 1972a; Lande 1979; Rutschman 1994) with more advanced tools from dynamical systems theory (Hofbauer 1990; Nagylaki *et al.* 1999). Since our analytical results require elaborate derivations, all details about methods and proofs are collected in the Supporting Information.

Models

To determine the conditions for the origin and maintenance of a DMI, we use classical migration–selection models in continuous time. We assume that a population is divided into two panmictic subpopulations, continent and island, each of infinite size. There is unidirectional gene flow from the continent to the island at rate m. Individuals may be haploid or diploid; both cases are considered in separate sections. Selection acts on two diallelic loci, $\mathcal{A}$ and $\mathcal{B}$, with alleles a, A and b, B, respectively. Lowercase letters denote the compatible alleles and uppercase letters the incompatible alleles. The fitness schemes are given below. The loci may be linked and recombine at rate r.

We consider the fate of individual substitutions at both loci, but ignore recurrent mutation. A DMI corresponds to a (stable) two-locus polymorphism. It can evolve from a homogeneous ancestral state either by one single-locus substitution in each of the two subpopulations (a so-called derived-derived DMI) or by two subsequent substitutions in the same subpopulation (a derived-ancestral DMI).

For a derived-derived DMI, we assume that the entire population is initially fixed for the "wild-type haplotype" ab. Subsequently, the mutant alleles A and B may invade on the island and the continent, respectively. We assume that on the continent the "continental haplotype" aB is superior to the wild type, whereas on the island the fitness of the "island haplotype" Ab is at least as high as that of the wild type.

We distinguish the following evolutionary scenarios:

1. *Continent–island*. Allele B invades first, goes to fixation on the continent, and reaches migration–selection equilibrium on the island. Subsequently, allele A appears on the island.
2. *Island–continent*. Allele A invades first on the island. Subsequently, allele B invades and becomes fixed on the continent.
3. *Secondary contact*. Both populations are in secondary contact after an allopatric phase. During allopatry, the continent has become fixed for allele B and the island for allele A.
4. *Island–island*. Here, aB is the ancestral state. Subsequently, first allele b and then allele A invade on the island.
5. *Continent–continent*. Here, Ab is the ancestral state. Allele a and then allele B invade on the continent.

Evolution on the continent is independent of the population composition on the island. We assume that every equilibrium on the continent is monomorphic and any substitution occurs quickly. In all scenarios, the continent is eventually fixed for the continental haplotype aB. We can then treat the continent simply as a source of aB migrants and need only to keep track of the allele frequencies on the island. The different evolutionary scenarios introduced above correspond to different initial conditions for the haplotype frequencies on the island (see *Appendix*, *Initial conditions*).

Below, we analyze how the equilibrium structure on the island depends on the model parameters for selection, recombination, and migration. In particular, we are interested in the parameter ranges where the island haplotype Ab can be maintained. Since the alleles a and b will always be present due to immigration, a stable DMI corresponds to an asymptotically stable, fully polymorphic equilibrium on the island. For sufficiently strong migration, the island will always be swamped by the continental haplotype aB (implying loss of A). Therefore, our focus is on the determination of the maximum migration rate $m_{\max}$ that allows for a fully polymorphic equilibrium. If this equilibrium is not globally stable, *i.e.*, if convergence depends on the initial condition, we will investigate under which of the scenarios it can be reached.

The haploid model

Let x_1, x_2, x_3, and x_4 denote the frequencies of the four haplotypes ab, aB, Ab, and AB on the island. They satisfy $x_i \geq 0$ for every i and $\sum_{i=1}^{4} x_i = 1$. We assume that selection acts on individuals during the haploid phase of their life cycle according to the scheme for Malthusian fitness values given in Table 1.

Table 1 Haplotype fitnesses and frequencies in the haploid model

Haplotype	*ab:* wild type	*aB:* continental	*Ab:* island	*AB:* recombinant
Fitness	$w_1 = 0$	$w_2 = \beta$	$w_3 = \alpha$	$w_4 = \alpha + \beta - \gamma$
Frequency	x_1	x_2	x_3	x_4

Table 2 Genotype fitnesses in the diploid model

	aa	*aA*	*AA*
bb	0	α	2α
bB	β	$\alpha + \beta - \gamma_1$	$2\alpha + \beta - \gamma_3$
BB	2β	$\alpha + 2\beta - \gamma_2$	$2\alpha + 2\beta - \gamma_4.$

This scheme is entirely general. The fitness of the wild-type haplotype is arbitrarily normalized to 0. The parameters α and β measure a potential selective advantage of the island and continental haplotypes, respectively, on the island (exogeneous selection). They can be positive or negative. However, below we assume that $\alpha > 0$ because, as will be shown, otherwise a DMI cannot exist. Finally, the epistasis paramter γ measures the strength of the incompatibility among the *A* and *B* alleles (endogeneous selection). We assume $\gamma \geq 0$, such that epistasis is negative or absent ($\gamma = 0$).

Assuming weak evolutionary forces in continuous time, the haplotype dynamics read

$$\dot{x}_1 = x_1\,(w_1 - \bar{w} - m) - rD, \quad \dot{x}_2 = x_2\,(w_2 - \bar{w} - m) + rD + m,$$
$$\dot{x}_3 = x_3\,(w_3 - \bar{w} - m) + rD, \quad \dot{x}_4 = x_4\,(w_4 - \bar{w} - m) - rD. \tag{1}$$

Here, the "dot" denotes the derivative with respect to time, $D = x_1x_4 - x_2x_3$ is the measure of linkage disequilibrium, and

$$\bar{w} = \beta x_2 + \alpha x_3 + (\alpha + \beta - \gamma)x_4$$

is the mean Malthusian fitness. Alternatively, the dynamics can be written in terms of the allele frequencies and the linkage-disequilibrium coefficient D (see *Appendix*, Haploid Model).

The diploid model

In the diploid case, selection acts on the 16 different combinations of two-locus haplotypes. Neglecting position effects and assuming no dominance for the *A* and *B* alleles in the wild-type (*ab*) background, we arrive at the fitness scheme with six parameters given in Table 2.

We thus allow for a quadruple of four independent epistasis parameters, $\Gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)$, for the four different incompatibility genotypes. These genotypes differ in the number of *A*/*B* conflicts, one (*aAbB*), two (*AAbB*, *aABB*), or four (*AABB*), and correspond to the H_0, H_1, or H_2 incompatibilities in Turelli and Orr (2000). As in the haploid case, we assume $\alpha,\ \gamma_i \geq 0$. In the most plausible case, epistasis is nondecreasing with the number of conflicts, *i.e.*, $\gamma_1 \leq \min[\gamma_2, \gamma_3]$, $\max[\gamma_2, \gamma_3] \leq \gamma_4$. Assuming Hardy–Weinberg equilibrium throughout, we can write the dynamics entirely in terms of the haplotype frequencies. These dynamics take the same functional form as in the haploid case (Equation 1), but with the fitness values replaced by the corresponding marginal fitnesses and the diploid mean fitness (File S2, Equations S.101 and S.102).

Results: Haploid Model

We first treat the haploid model for which the exact conditions for the existence of a stable (parapatric) DMI can be derived. These results are based on the analysis of the equilibrium structure of the model, which is described in the *Appendix*. All proofs and derivations, which require advanced mathematical methods, are given in the supporting information. File S1 contains a complete analysis of the haploid model, which, as a by-product, extends and complements previous studies of haploid two-locus selection (and migration) models. File S2 contains the derivations of the results for the diploid model and File S4 contains a *Mathematica* notebook with additional supporting material.

We are interested in the maximum rate of gene flow that allows for the evolutionary origin and the maintenance of a DMI in parapatry. In general, this rate depends on the initial conditions that correspond to the different evolutionary scenarios. As it turns out, however, we need to distinguish only two maximum rates, which we denote as $m_{\max}^-$ and $m_{\max}^+$. A globally asymptotically stable internal equilibrium exists if and only if $m < m_{\max}^-$. As a consequence, a stable DMI will evolve from every initial condition (hence, under all evolutionary scenarios). A locally stable internal equilibrium exists if and only if $m < m_{\max}^+$. Local stability is a necessary and sufficient condition for the long-term maintenance of a DMI. If $m_{\max}^+ = 0$, no DMI can exist under any level of gene flow. In addition to $m_{\max}^\pm$, we consider the equilibrium frequency $\hat{x}_3$ of the island haplotype at a DMI as a measure of two-locus population differentiation that is achieved for a given set of parameters. With this definition, differentiation is positive if and only if a DMI can be maintained.

Both maximum rates, $m_{\max}^+$ and $m_{\max}^-$, are functions of the selection coefficients α, β, and γ and the recombination rate r. At most one stable internal equilibrium can exist, and extensive numerical calculations did not produce complex attractors such as limit cycles. If a DMI exists for a given set of parameters, the allele frequencies at this equilibrium are thus unique. A (globally) stable DMI for some migration rate $\tilde{m} > 0$ implies the existence of a (globally) stable DMI for every positive $m < \tilde{m}$. Hence, the rates $m_{\max}^\pm$ are the unique boundaries of well-defined parameter regimes. Obviously, $m_{\max}^- \leq m_{\max}^+$. If $m_{\max}^- < m < m_{\max}^+$, the evolutionary dynamics are bistable and the DMI will not be reached from every initial condition. Extensive numerical simulations (and a proof for the special case of independent loci) show that a DMI will evolve for such a bistable case if and only if the second substitution enters the population in a deme where the incompatible allele is absent. Clearly, this is the case for secondary contact, but also if the first substitution occurs on the island and the second substitution on the continent or if both substitutions occur on the continent. In contrast, a stable DMI will never be reached for $m^-_{\max} < m < m^+_{\max}$ if the first substitution is on the continent and the second on the island or if both substitutions occur on the island. In both these latter scenarios, migrants from the continent carrying the incompatible allele hamper the establishment of the second substitution on the island.

Necessary conditions for stable DMIs

Fully polymorphic equilibria (or complex attractors) cannot exist if all evolutionary trajectories approach the boundary of the allele-frequency space. In File S1, section S.2, we establish necessary conditions for the existence of a DMI by means of Lyapunov functions. In particular, a DMI can exist only if

$$\alpha > 0 \quad \text{and} \quad \gamma > \beta; \tag{2}$$

i.e., the island haplotype *Ab* must be on a local peak of the fitness landscape, relative to both its single-step mutational neighbors, *ab* and *AB*. Most significantly, the condition excludes the evolution of a DMI by a neutral process even in the secondary contact scenario: The substitution of the island allele *A* must be adaptive, $\alpha > 0$. We henceforth assume (2) in addition to $\gamma \geq 0$ for all our derivations and results.

This condition does not require that the island haplotype occupies an absolute fitness maximum. The fitness of the continental type *aB* can be larger; *i.e.*, $\beta > \alpha$. If this is the case, however, we can show that a DMI can exist only if

$$r > \beta - \alpha, \tag{3}$$

i.e., if continental immigrants are broken up by sufficiently strong recombination. Conditions (2) and (3) constrain the parameter space for a parapatric DMI under any level of gene flow, *i.e.*, even under arbitrarily weak migration. For higher gene-flow rates we find additional conditions, which can be expressed as upper bounds for the maximum gene-flow rate $m^+_{\max}$,

$$m^+_{\max} \leq \alpha - \beta + r, \tag{4a}$$

$$m^+_{\max} \leq \max\left[\alpha - \beta, \ \frac{\alpha}{4}\right], \tag{4b}$$

$$m^+_{\max} \leq \max\left[\alpha - \beta, \ \frac{(\gamma - \beta)}{4}\right]. \tag{4c}$$

Condition (4a) implies and extends (3). For gene-flow rates above any of these bounds, we obtain global convergence to one of the boundary equilibria. As detailed in the *Appendix*, *Boundary equilibria and global stability*, at most three boundary equilibria can exist if $m > 0$. In addition to the monomorphic equilibrium corresponding to fixation of the continental haplotype (denoted by $\mathbf{M}_2$), there are two equilibria with a single-locus polymorphism: An equilibrium $\mathbf{S}_{\mathcal{B}}$ with locus $\mathcal{B}$ polymorphic and locus $\mathcal{A}$ fixed for allele *a* exists if the continental haplotype *aB* is maladaptive on the island and the ancestral type *ab* resists the immigration pressure (*i.e.*, if $m < -\beta$). Similarly, an equilibrium $\mathbf{S}_{\mathcal{A}}$ exists (with locus $\mathcal{A}$ polymorphic and locus $\mathcal{B}$ fixed for allele *B*) whenever the recombinant haplotype *AB* can resist the immigration pressure (*i.e.*, if $m < \alpha - \gamma$).

Limiting cases

Our results for the haploid model are based on a complete analytical characterization of the maximum rates of gene flow admitting a DMI. However, since the analytical expressions for $m^-_{\max}$ and $m^+_{\max}$ require multiple case distinctions, the statement of the complete, explicit results is deferred to the *Appendix*, *Boundary equilibria and global stability* and *Internal equilibria and local stability.* Below, we focus on several limiting cases for which the results take a simple form. In particular, we treat the cases of weak migration (small *m*), complete linkage ($r = 0$), and independent loci ($D = 0$). Discussion of these limits already allows us to draw the most important conclusions about DMI stability.

Weak migration: For $m = 0$ the haploid two-locus model was already studied by Rutschman (1994); see also Feldman (1971). In this case, the only boundary equilibria are the monomorphic equilibria $\mathbf{M}_i$ (where $x_i = 1$), and at least one of them is stable. Rutschman's results together with those in File S1, section S.5 imply that if an internal equilibrium exists, it is unstable. As a consequence, if $m > 0$, any stable internal equilibrium must be a perturbation of the monomorphic equilibrium $\mathbf{M}_3$ (fixation of the island haplotype) that is pushed into the interior of the state space due to immigration of *aB* haplotypes. The stability properties of the perturbed equilibrium are the same as for the unperturbed equilibrium (Karlin and McGregor 1972a; Bürger 2009).

We therefore obtain the necessary and sufficient conditions for a stable DMI under weak migration from the stability analysis of $\mathbf{M}_3$ if $m = 0$. Expressed in terms of $m^{\pm}_{\max}$,

$$\begin{aligned} &m^+_{\max} > 0 \quad \text{if and only if} \\ &\alpha > 0 \quad \text{and} \quad \gamma > \beta \quad \text{and} \quad r > \beta - \alpha, \end{aligned} \tag{5a}$$

and, if $m^+_{\max} > 0$,

$$\begin{aligned} &m^-_{\max} > 0 \quad \text{if and only if} \\ &\beta < 0 \quad \text{or} \quad \gamma < \alpha \quad \text{or} \quad r < \alpha - \beta. \end{aligned} \tag{5b}$$

From (5a) we conclude that (2) and (3) are not only necessary, but also sufficient conditions for an asymptotically stable internal equilibrium if migration is sufficiently weak.

A perturbation analysis yields the approximate coordinates of the perturbed internal equilibrium near $\mathbf{M}_3$. In particular, the equilibrium frequency of the island haplotype *Ab*, which we use as our measure of population differentiation, is

$$\hat{x}_3 = 1 - \frac{\alpha - \beta}{\alpha - \beta + r} \cdot \frac{m}{\alpha - \beta} - \frac{r}{\alpha - \beta + r}\left(\frac{m}{\alpha} + \frac{m}{\gamma - \beta}\right). \tag{6}$$

The second term describes the loss of *Ab* due to direct competition with the immigrating continental haplotype *aB*, whereas the third term results from competition of *Ab* with the recombination products *ab* and *AB*. Equation 6 shows that, for small m, population differentiation increases with the recombination rate r if and only if breaking up the immigrating continental type reduces the competition experienced by the island type. This is the case if either $\beta > \alpha$ or $\alpha^{-1} + (\gamma - \beta)^{-1} < (\alpha - \beta)^{-1}$. These conditions can be summarized as

$$\frac{\partial \hat{x}_3}{\partial r} \gtrless 0 \Leftrightarrow \beta(\gamma - \alpha) \gtrless (\alpha - \beta)^2. \tag{7}$$

Tight linkage: If the two incompatibility loci are completely linked ($r = 0$), at most two haplotypes can coexist at an equilibrium (*i.e.*, one of *ab*, *Ab*, or *AB*, in addition to the continental haplotype *aB*); see File S1, section S.8.1. A fully polymorphic DMI, with haplotypes *Ab* and *aB* present, exists and is globally asymptotically stable if

$$m < m^{\pm}_{\max} = \alpha - \beta. \tag{8}$$

Thus, a bistable regime does not exist for $r = 0$. The DMI simply leaves the state space through the monomorphic equilibrium $\mathbf{M}_2$ as m increases beyond $\alpha - \beta$. As a consequence, the population differentiation decreases to zero as m approaches $m^{\pm}_{\max}$. Indeed, we find

$$\hat{x}_3 = 1 - \frac{m}{\alpha - \beta}. \tag{9}$$

Perturbation results guarantee that the same qualitative behavior [*i.e.*, global stability of a DMI ($m^{+}_{\max} = m^{-}_{\max}$) and continuous decrease of $\hat{x}_3$ to zero] holds for recombination rates below a threshold $r^* > 0$.

Independent loci: If recombination is sufficiently strong relative to selection and migration, we can assume that the loci are in linkage equilibrium. In this case, the evolutionary dynamics reduce from three to two dimensions, and we can determine the exact position and the stability of all internal equilibria (see File S1, section S.8.2). We can exclude cycling behavior and show that at most two internal equilibria can exist, one of which is asymptotically stable whenever it exists. The maximum rate of gene flow admitting a globally stable DMI can be obtained from the formulas for general r in the *Appendix*, *Internal equilibria and local stability*, or from Equations S.93 and S.94 in File S1. It simplifies to

$$m < m^{-}_{\max} = \begin{cases} \dfrac{(\alpha - \gamma)(\gamma - \beta)}{\alpha} & \text{if } \beta \le \gamma \le \alpha + \beta, \\ \dfrac{-\alpha\beta}{\gamma - \beta} & \text{if } \gamma > \alpha + \beta, \end{cases} \tag{10}$$

where negative values imply that no DMI can exist. For local stability, we find the following condition:

$$m < m^{+}_{\max} = \begin{cases} \dfrac{(\alpha - \gamma)(\gamma - \beta)}{\alpha} \; (= m^{-}_{\max}) & \text{if } \beta \le \gamma \le \min\left[\dfrac{1}{2}\alpha, \alpha + \beta\right], \\ \dfrac{-\alpha\beta}{\gamma - \beta} \; (= m^{-}_{\max}) & \text{if } \alpha + \beta < \gamma \le -\beta, \\ \dfrac{\alpha(\gamma - \beta)}{4\gamma} \; (> m^{-}_{\max}) & \text{if } \gamma > \max\left[\dfrac{1}{2}\alpha, \; |\beta|\right]. \end{cases} \tag{11}$$

Thus, for sufficiently large γ, *i.e.*, if $\gamma > \max[|\beta|, \alpha/2]$, we obtain a bistable regime in which evolution of a DMI depends on initial conditions.

An explicit (but lengthy) expression for the level of population differentiation $\hat{x}_3$ is given in File S1, Equation S.97. An illuminating result is obtained if we study $\hat{x}_3$ in the limit of strong epistasis, $\gamma \to \infty$. Here, we find

$$\hat{x}_3(\gamma = \infty) = \frac{1}{2}\left(1 + \sqrt{1 - \frac{4m}{\alpha}}\right). \tag{12}$$

From (11), we note that $m^{+}_{\max} = \alpha/4$ in this case. Thus, the amount of population differentiation at the DMI is independent of β for strong epistasis and never falls below 1/2. Compared to tight linkage (Equation 9), this different behavior of $\hat{x}_3$ results from the fact that, here, the stable internal equilibrium does not leave the state space through one of the boundary equilibria at $m = m^{+}_{\max}$, but disappears upon merging with an unstable internal equilibrium.

Stability of parapatric DMIs in the haploid model

We can draw several conclusions concerning the conditions and mechanisms for the evolution of parapatric DMIs that summarize our main findings for the haploid model.

First, condition (4b) defines a sharp upper bound on the rate of gene flow that admits a (two-locus) DMI. Because this bound depends only on the exogenous selection coefficients α and β, a DMI is impossible for higher migration rates, irrespective of the choice of γ or r. It is sharp, in the sense that DMIs with an appropriate genetic architecture (in terms of γ and r) can evolve for every rate of gene flow below it. Second, depending on which of the terms on the right-hand side of (4a) is relevant for the bound, the DMIs that remain stable under the highest gene-flow rates may have widely diverging architectures: On the one hand, if $m^{+}_{\max} = \alpha - \beta > \alpha/4$, then a tightly linked DMI of arbitrary strength (provided $\gamma > \beta$) will sustain the maximum level of gene flow. On the other hand, if $\alpha/4 \ge \alpha - \beta$, Equation 11 shows that a strong DMI among unlinked loci (r, $\gamma \to \infty$) can sustain the maximum level of gene flow.

We can understand this result in terms of two different mechanisms that enable the evolution and the maintenance of a parapatric DMI: Whereas the first, the "direct" mechanism is based on selection against maladapted immigrants in a spatially heterogeneous environment, the second, the "indirect" mechanism acts through selection against hybrids and requires epistasis and recombination. The two mechanisms act in different regions of the parameter space, corresponding to different shapes of the fitness landscape on the island (see Figure 1–Figure 3 and the analytical results in the *Appendix*).

Slope-type fitness: Selection against immigrants: See Figure 1A, Figure 3A1 and area with light shading in Figure 2. A slope-type fitness landscape results under two (nonexclusive) scenarios: (i) local adaptation ($\beta < 0$, *i.e.*, the continental haplotype *aB* is maladaptive on the island; see Figure 1A); and (ii) weak incompatibility ($\gamma < \alpha$; hence, *aB* has lower fitness than the recombinant genotype *AB*). In this case, a stable DMI can be maintained by selection against immigrants, *i.e.*, if (and only if) the fitness advantage of the island genotype over the invading continental type is sufficiently large to compensate the migration pressure. For sufficiently small m, the island haplotype is nearly fixed at the unique stable equilibrium. Since we assume that the haplotype *aB* is fixed on the continent, selection against immigrants requires heterogeneous selection across the continent and the island.

For slope-type fitness, DMIs are usually, but not always, globally stable (Figure 3A1). For low recombination, global stability extends from $m = 0$ to the maximum rate $m^+_{\max} = m^-_{\max}$. Bistability occurs if gene-flow rates are not much lower than $m^+_{\max}$ and rare island haplotypes are efficiently broken up by recombination, such that a stable boundary equilibrium can be maintained. This is the case, for instance, if both r and γ are large.

From our analytical and numerical results and their graphical representations (see Figures 1A and 3A1 and the interactive visualization in File S4), we conclude the following characteristic properties for the maximum gene-flow rates $m^{\pm}_{\max}$ and the degree $\hat{x}_3$ of population differentiation:

1. The maximum migration rate satisfies $m^+_{\max} \leq \alpha - \beta$ [this follows from (4b) and (4c) because $\beta < 0$ or $\gamma < \alpha$]. This maximum is achieved for tightly linked incompatibility loci ($r = 0$). With increasing r, both $m^+_{\max}$ and $m^-_{\max}$ are strictly decreasing (see Figure 1A). This can be understood as follows: As we show in File S1, section S.2.4, $D < 0$ holds at every internal equilibrium if $r < \infty$. Thus, recombination increases the frequencies of the ancestral and recombinant types (*ab* and *AB*) relative to the continental and island types (*aB* and *Ab*). Since at least one of the genotypes *ab* or *AB* has a higher fitness than the continental type *aB*, recombination increases the marginal fitness of at least one of the alleles *a* or *B* that enters the island through gene flow. This facilitates the fixation of the respective allele, leading to loss of the DMI.
2. As might be expected, both maximum rates, $m^+_{\max}$ and $m^-_{\max}$, increase with a steeper slope of the landscape (higher α or lower β; see, *e.g.*, the interactive visualization in File S4). In contrast, Figure 3 shows that the strength of the incompatibility is far less important (and irrelevant if $r = 0$).
3. Finally, also the level of differentiation $\hat{x}_3$ is strictly decreasing if the number of hybrids is increased by increasing r (*cf.* Figure 1A and Equation 7).

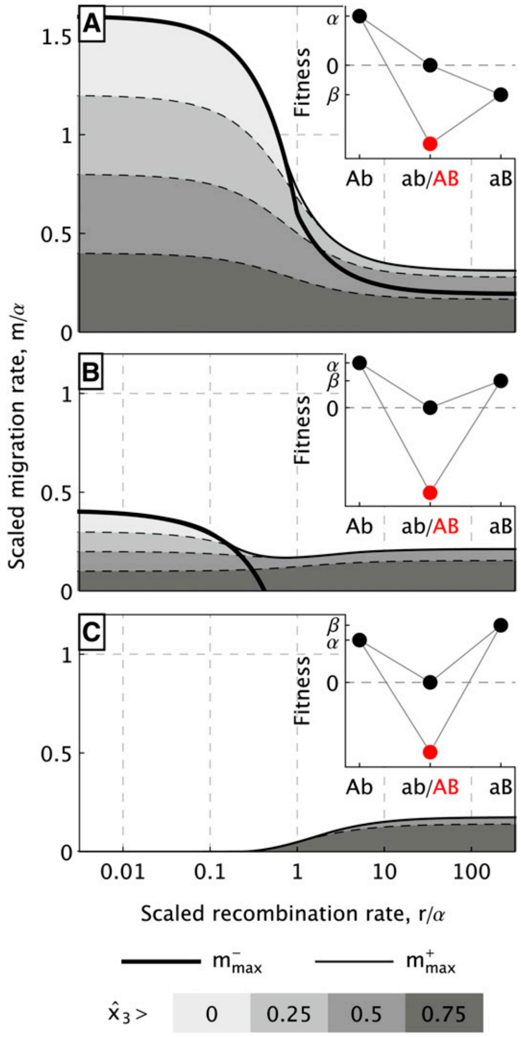

Figure 1 Maximum migration rates (black solid lines) and degree of population differentiation (gray areas) as functions of the (scaled) recombination rate r. We identify three regimes. (A) Selection against immigrants for slope-type fitness. Here, $\beta/\alpha = -0.6$, $\gamma/\alpha = 2.5$. (C) Selection against hybrids in a homogeneous environment. Here, $\beta/\alpha = 1.2$, $\gamma/\alpha = 4$. (B) Combination of both for intermediate fitness combinations. Here, $\beta/\alpha = 0.6$, $\gamma/\alpha = 4$. The corresponding type of fitness landscapes is displayed in each respective top right corner.

Double-peak fitness with superior continental type: Selection against hybrids: See Figure 1C, Figure 3A3, and area with dark shading in Figure 2. If $\beta > \alpha$, the invading continental type is the fittest genotype on the island (fitness scheme in Figure 1C.) Nevertheless, the island genotype can be maintained in a locally stable DMI if both the ancestral and the recombinant genotypes have relatively low fitness. The condition we need is $\gamma > \beta > \alpha$, thus, a strong incompatibility. In addition, sufficiently strong recombination is required to protect the island genotype *Ab* by decomposing the invading continental types *aB* into its constituent alleles (*cf.* Equation 3). Due to the low fitness of the recombination products *ab* and *AB*, both *a* and *B* are kept at low frequencies. Thus, recombination creates a drainage through which the invading alleles are discharged.

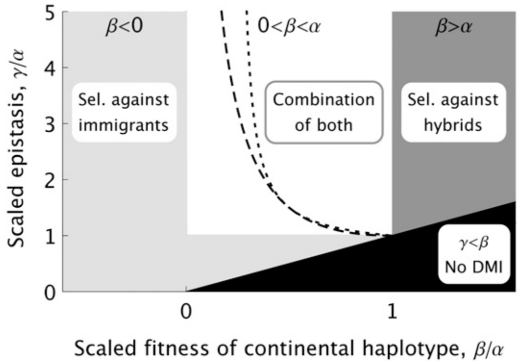

Figure 2 Parameter ranges corresponding to two different mechanisms for DMI evolution in the haploid model as a function of the heterogeneity of the environment (measured as β/α) and the strength of the incompatibility (measured as γ/α). In the intermediate parameter range (open area), both mechanisms contribute and we see a stepwise transition for various characteristic properties: (1) decrease/increase of population differentiation with increasing recombination rate for weak migration (left/right of dashed line) and (2) decrease/U-shape or increase of $m^+_{\max}$ with increasing recombination rate (left/right of dotted line).

This second, indirect mechanism for maintaining a stable DMI does not depend on ecological differentiation between the continent and the island; *i.e.*, exogenous selection can be homogenous. Indeed, under several scenarios (substitution on the island first, both on the continent, or secondary contact), evolution of a stable DMI can occur with a homogeneous fitness function across both subpopulations. The resulting DMI is always only locally stable. Also the dependence of $m^\pm_{\max}$ and $\hat{x}_3$ on the model parameters differs distinctly from the one under slope-type fitness (Figure 1C and Figure 3A3):

1. Since, in haploids, hybrid genotypes are formed only by recombination, selection against hybrids is most efficient for large r. Both $m^+_{\max}$ and $\hat{x}_3$ increase with r (*cf.* Figure 1C).
2. The efficiency of this mechanism is enhanced by strong epistasis that leads to a deep fitness valley (large α and large $\gamma - \beta$), as evidenced by the dependence of $m^+_{\max}$ on these parameters (*cf.* Figure 3A3). From (4b), we infer that a DMI can exist only if $m < (1/4)\alpha$. That this upper bound is approached for high r and γ has the following explanation. If all continental migrants recombine with resident island types (when $r \to \infty$), and if all recombinant types *AB* die (when $\gamma \to \infty$), only island and wild-type genotypes remain on the island. Then the dynamics of the island types are $\dot{x}_2 = \alpha x_2(1-x_2)-m$, which have a stable equilibrium with $x_2 > 0$ if $m < (1/4)\alpha$.

Double-peak fitness with superior island type: Combination of both mechanisms: See Figures 1B, Figure 3A2, and open area in Figure 2. If $0 < \beta < \alpha < \gamma$, we obtain a double-peak fitness landscape on the island with the island type on the higher peak. In this parameter region, both mechanisms described above interact and contribute to the maximum rate of gene flow in a complex way. There is no simple sharp boundary that separates the parameter regions where either mechanism dominates, but we can illustrate the transition for several characteristic properties.

1. If recombination is weak relative to the fitness difference of the island and the continental type, $r \ll \alpha - \beta$, the haplotypes compete mostly as entire units and selection against immigrants dominates. The resulting DMI is globally stable for small m, and heterogeneous selection is needed to maintain a parapatric DMI. The maximum rate $m^-_{\max}$ for a globally stable DMI decreases monotonically with r and is zero if $r \geq \alpha - \beta$ (*cf.* Figure 1B). For higher recombination rates, a locally stable DMI can exist for spatially homogeneous selection; it can evolve from favorable initial conditions.
2. The dependence of $m^+_{\max}$ on the recombination rate changes as follows. Whereas $m^+_{\max}$ is monotone decreasing in r if $\gamma(4\beta - \alpha) < 3\alpha\beta$ (see Equation S.38 in File S1), otherwise we obtain a U-shaped dependence with a minimum at an intermediate recombination rate; *cf.* Figure 1B. If $\beta > (3/4)\alpha$, the highest migration rate can be tolerated if the loci are independent because (4b), (4c), and (11) imply $m^+_{\max} = (1/4)\alpha$ if $\gamma \to \infty$ and $r \to \infty$. $m^+_{\max}$ increases monotonically with γ, indicating that both mechanisms are more efficient for strong DMIs in this parameter range (*cf.* Figure 3A2). This can be understood since selection against the *B* allele (lower fitness of *AB*) also leads to reduced marginal fitness of the allele (since *ab* haplotypes are less fit than *aB* haplotypes if $\beta > 0$).
3. Alternatively, we can ask whether formation of more hybrids (by an elevated r) increases or decreases population differentiation. The answer depends on the level of gene flow. For low m, differentiation decreases with higher r below the threshold implicitly given in Equation 7 and increases monotonically above this threshold (dashed curve in Figure 2; Figure 1B represents a case in which $m^+_{\max}$ is highest for tight linkage although population differentiation increases with recombination for weak migration). For high gene-flow rates near the maximum values admitting a stable DMI, the dependence of $\hat{x}_3$ on r can be U-shaped, similar to the dependence of $m^+_{\max}$ on r.

Results: Diploid Model

Although a comprehensive analytical treatment of the general diploid model seems out of reach, progress can be made for particular cases. For most of our results, we focus on two choices of the epistasis coefficients $\Gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)$: (1) a recessive model with no fitness costs for double heterozygotes, $\Gamma_0 := (0, \gamma, \gamma, 2\gamma)$; and (2) a codominant model with costs proportional to the number of *A/B* conflicts, $\Gamma_1 := (\gamma/2, \gamma, \gamma, 2\gamma)$.

As we will see, these two models represent two main types of DMI structure. Two additional properties make them analytically more tractable than the general case. First, there is no dominance at locus $\mathcal{A}$ or $\mathcal{B}$ if the other locus is monomorphic. In particular, this excludes stable boundary equilibria caused by overdominance: If $m = 0$, the only equilibria at the boundary are the monomorphic equilibria $\mathbf{M}_1$–$\mathbf{M}_4$. Second, the condition $\gamma_4 = 2\gamma_2 = 2\gamma_3$ preserves the

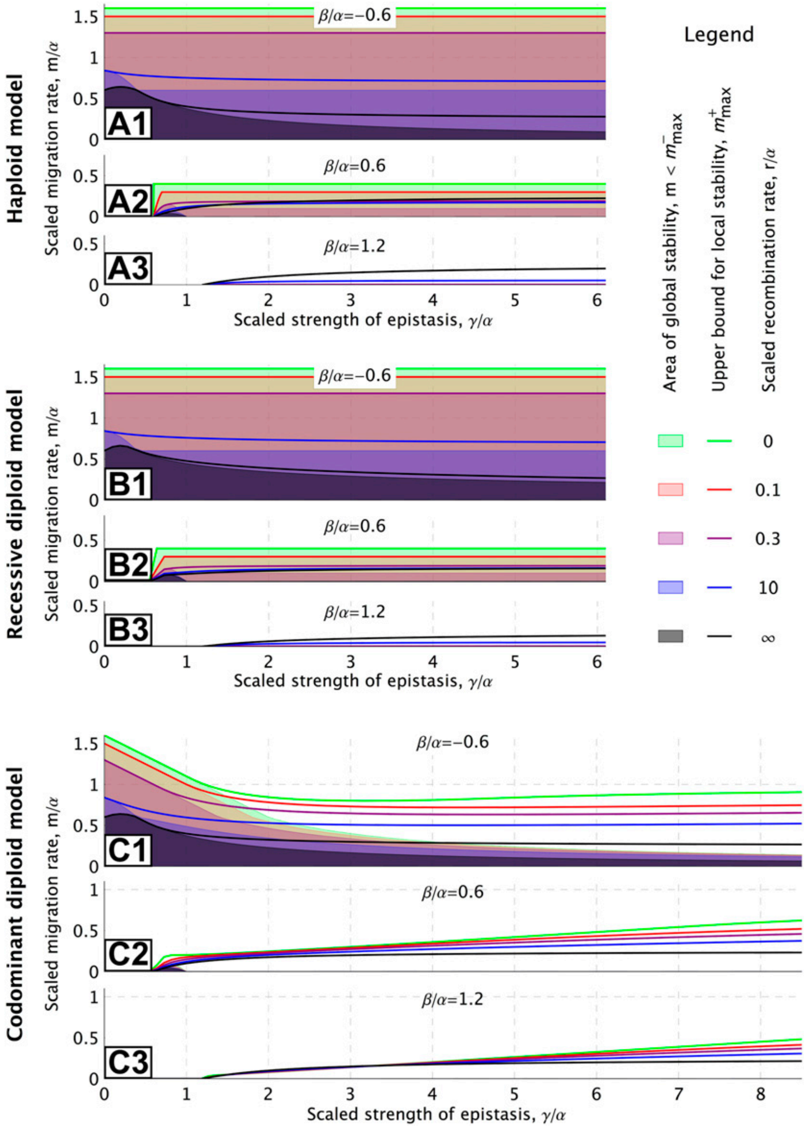

Figure 3 Maximum migration rates as functions of the (scaled) strength of epistasis for the haploid (A1–A3), recessive diploid (B1–B3), and codominant diploid (C1–C3) models and for various recombination rates. Each panel represents a different environmental scenario determined by β. For the haploid model, all curves are determined analytically using the formulas in *Appendix*, *Internal equilibria and local stability*. See main text for a detailed discussion of this figure and File S4 for interactive visualization of other parameter values.

symmetry of the haploid model. Thus, the equilibrium structure is symmetric under the transformation $\alpha \to \gamma - \beta$ and $\beta \to \gamma - \alpha$, which corresponds to an exchange of the ancestral haplotype *ab* and the recombinant type *AB*.

As in the haploid model, if $m > 0$, necessary and sufficient conditions for a globally stable DMI can be derived from a stability analysis of the boundary equilibria. For our special models, the same boundary equilibria as in the haploid model exist under the same conditions: The monomorphic equilibrium $\mathbf{M}_2$ exists always, and the single-locus polymorphisms $\mathbf{S}_{\mathcal{A}}$ or $\mathbf{S}_{\mathcal{B}}$ exist if $m < \alpha - \gamma$ or $m < -\beta$, respectively (*cf.* Figure A1 in the *Appendix*). In general, however, the stability conditions in the haploid and diploid case differ.

Necessary conditions for recessive and codominant DMIs

Necessary conditions for a stable DMI are derived in Sections S.10 and S.12 of File S2. In terms of the maximum rates of gene flow, we find for the epistasis scheme $\Gamma = (\gamma_1, \gamma, \gamma, 2\gamma)$, which comprises both our special models,

$$\begin{aligned} &m^+_{\max} > 0 \quad \text{if and only if} \\ &\alpha > 0 \quad \text{and} \quad \gamma > \beta \quad \text{and} \quad r > \beta - \alpha - \gamma_1, \end{aligned} \tag{13a}$$

and, if $m^+_{\max} > 0$,

$$\begin{aligned} &m^-_{\max} > 0 \quad \text{if and only if} \\ &\beta < 0 \quad \text{or} \quad \gamma < \alpha \quad \text{or} \quad r < \alpha - \beta - \gamma_1. \end{aligned} \tag{13b}$$

There is a striking similarity with the corresponding Equations 5a and 5b for the haploid model. For both models, the conditions $\alpha > 0$ and $\gamma > \beta$ are required for a parapatric DMI, thus excluding neutral DMIs and setting a lower bound for the strength of the incompatibility. For the rest of this section, we therefore assume $\alpha > 0$ and $\gamma > \beta$.

For the recessive model ($\gamma_1 = 0$), the conditions (13a) and (13b) coincide with their haploid counterparts. In the codominant model ($\gamma_1 = \gamma/2$), selection against double heterozygotes deepens the fitness valley between the island and the continental genotypes. For $m = 0$, this leads to an increased range of simultaneous stability of the monomorphic equilibria $\mathbf{M}_2$ and $\mathbf{M}_3$. This is reflected by the conditions (13a) and (13b): whereas the constraints for a locally stable DMI get weaker for larger γ_1, the conditions for global stability become more stringent.

For weak migration, the amount of population differentiation, as measured by the equilibrium frequency $\hat{x}_3$ of the island haplotype, is calculated to

$$\begin{aligned}\hat{x}_3 = 1 - \frac{\alpha - \beta + \gamma_1}{\alpha - \beta + \gamma_1 + r} \cdot \frac{m}{\alpha - \beta + \gamma_1} \\ - \frac{r}{\alpha - \beta + \gamma_1 + r}\left(\frac{m}{\alpha} + \frac{m}{\gamma - \beta}\right).\end{aligned} \tag{14}$$

Differentiation increases or decreases with r according to

$$\frac{\partial \hat{x}_3}{\partial r} \gtrless 0 \Leftrightarrow (\beta - \gamma_1)(\gamma - \alpha - \gamma_1) \gtrless (\alpha - \beta + \gamma_1)^2. \tag{15}$$

In the context of the haploid result (Equation 7), we can understand this result as follows: For weak migration, invading alleles appear on the island almost exclusively in heterozygous genotypes containing the island haplotype (*i.e.*, in *aBAb*, *ABAb*, or *abAb*). The dynamics of these three genotypes, together with the homozygous island type *AbAb*, essentially follow those of a haploid model if the following substitutions are made: $\beta_{\text{hap}} \rightarrow \beta_{\text{dip}} - \gamma_1$ and $\gamma_{\text{hap}} \rightarrow \gamma_{\text{dip}} - \gamma_1$.

Finally, for $\Gamma = (\gamma_1, \gamma, \gamma, 2\gamma)$, we can also derive the following upper bound for $m^+_{\max}$ (Equation S.111 in File S2):

$$m^+_{\max} \leq \max\left[\frac{1}{4}\alpha, \alpha - \beta + \gamma_1\right]. \tag{16}$$

For $m^-_{\max}$, complete analytical results can, in principle, be derived from a local stability analysis of each boundary equilibrium (File S2, section S.13). Since they are quite complex and involve a large number of case distinctions, we discuss only the limiting cases of weak migration, tight linkage ($r = 0$), and independent loci ($D = 0$) in more detail.

Recessive DMI model

Tight linkage: If $r = 0$, there are only two equilibria that can be stable for $m > 0$. One is the monomorphic equilibrium $\mathbf{M}_2$, which is globally stable for $m > \alpha - \beta$, but unstable otherwise. The other equilibrium, at which only the haplotypes *Ab* and *aB* are present, gives rise to the DMI. Hence, we recover the result $m^{\pm}_{\max} = \alpha - \beta$ (Equation 8) from the haploid model. Note, however, that this does not imply that the dynamics of the diploid and haploid models are the same.

Independent loci: In linkage equilibrium, the maximum migration rate $m^-_{\max}$ is determined from the stability conditions of the single-locus polymorphisms $\mathbf{S}_{\mathcal{A}}$ and $\mathbf{S}_{\mathcal{B}}$ (*cf.* File S4, section 2.3). We obtain

$$\begin{aligned} m^-_{\max} \\ = \begin{cases} \dfrac{\alpha - \gamma}{2\gamma}\left(\sqrt{(\alpha - \gamma)^2 + 4\gamma(\gamma - \beta)} - (\alpha - \gamma)\right) & \text{if } \beta < \gamma \leq \alpha + \beta, \\ \dfrac{-\beta}{2\gamma}\left(\sqrt{\beta^2 + 4\alpha\gamma} + \beta\right) & \text{if } \gamma > \alpha + \beta. \end{cases}\end{aligned} \tag{17}$$

This expression deviates from the corresponding haploid result (Equation 10). In particular, we obtain a slower decline to zero ($\sim\gamma^{-1/2}$) as $\gamma \rightarrow \infty$. For local stability, we find $m^+_{\max} < \alpha$ as an upper bound in addition to (16). However, we have not been able to derive an exact analytical expression for $m^+_{\max}$. Numerical results are discussed below.

Codominant DMI model

Tight linkage: If $r = 0$, local stability analysis of the boundary equilibria yields

$$m^-_{\max} = \begin{cases} \dfrac{2(\gamma - \beta)(\alpha - \gamma)}{\gamma} & \text{if } \gamma < 2\beta, \\ \dfrac{-2\alpha\beta}{\gamma} & \text{if } \gamma > 2\alpha, \\ \alpha - \beta - \dfrac{\gamma}{2} & \text{otherwise.} \end{cases} \tag{18}$$

As discussed below, this deviates strongly from the haploid and the recessive case. We deduce the following formulas for $m^+_{\max}$ (*cf.* File S4, section 2.4). If $\alpha > 4\beta$,

$$m^+_{\max} = \begin{cases} \dfrac{(\gamma - \beta)}{\gamma}(2\alpha - 2\beta - 3\gamma) & \text{if } \gamma < 2\beta, \\ \alpha - \beta - \dfrac{\gamma}{2} & \text{if } 2\beta < \gamma < \min\left[2\alpha, \dfrac{2}{3}(\alpha - \beta)\right], \\ \dfrac{(2\alpha - 2\beta + \gamma)^2}{16\gamma} & \text{if } \min\left[2\alpha, \dfrac{2}{3}(\alpha - \beta)\right] \\ & \quad < \gamma < \max[2\alpha, 6\alpha + 2\beta], \\ \dfrac{\alpha}{\gamma}(\gamma - 2\alpha - 2\beta) & \text{if } \gamma > \max[2\alpha, 6\alpha + 2\beta]; \end{cases} \tag{19}$$

and if $\alpha \leq 4\beta$ (which implies $\beta > 0$),

$$m^+_{\max} = \begin{cases} \dfrac{(\gamma - \beta)}{\gamma}(2\alpha - 2\beta - 3\gamma) & \text{if } \gamma < \dfrac{2}{7}(\alpha + 3\beta), \\ \dfrac{(2\alpha - 2\beta + \gamma)^2}{16\gamma} & \text{if } \max\left[2(\beta - \alpha), \dfrac{2}{7}(\alpha + 3\beta)\right] \\ & \quad < \gamma < 6\alpha + 2\beta, \\ \dfrac{\alpha}{\gamma}(\gamma - 2\alpha - 2\beta) & \text{if } \gamma > 6\alpha + 2\beta. \end{cases} \tag{20}$$

Independent loci: Under the assumption of linkage equilibrium, the dynamics of allele frequencies in the codominant model map precisely to those of the corresponding haploid model (*cf.* File S4). Intuitively, this can be seen as follows:

If $r \to \infty$, alleles are combined randomly. To recover the dynamics of the haploid model, we assign to each diploid genotype the average fitness of the corresponding allele combination in the haploid model; *e.g.*, $w_d(aAbB) = 1/2\,(w_h(ab) + w_h(AB) + w_h(Ab) + w_h(aB)) = \alpha + \beta - \gamma/2$, which indeed is the fitness in the codominant model. We obtain the same existence and stability results as in the haploid case, namely (10) and (11). All equations reported above are visualized in interactive *Mathematica* plots in File S4, section 2.4.

Analytical results and conjectures for general epistasis coefficients

In File S4, we analyze how our results for the special epistasis schemes generalize to a diploid model with an arbitrary two-locus incompatibility. Throughout our analysis, we assume that the strength of the incompatibility is nondecreasing with the number of *A/B* conflicts, *i.e.*, that the epistasis parameters fulfill the condition $\gamma_4 \geq \{\gamma_3, \gamma_2\} \geq \gamma_1$. Additional (technical) assumptions are needed for some of our formal proofs in File S2. When this is the case, our conjectures are confirmed by perturbation arguments and by numerical simulations. We find the following necessary conditions for a stable parapatric DMI (*cf.* Equation S.125 in File S2):

$$\alpha > 0 \quad \text{and} \quad \beta < \max[\gamma_3, \gamma_4 - \gamma_3]. \tag{21}$$

The latter condition simplifies to $\beta < \gamma_3$ if we exclude overdominance at locus $\mathcal{B}$, which requires $\gamma_3 < \beta < \gamma_4 - \gamma_3$. Since overdominance at locus $\mathcal{B}$ stabilizes the island genotype, the condition is slightly relaxed in the general case. Without overdominance, we also retain the lower bound on the recombination rate of the codominant model, $r > \beta - \alpha - \gamma_1$.

Finally, we find that

$$m^{+}_{\max} \leq \max[\alpha - \beta, \alpha] \tag{22}$$

provides an upper bound for the maintenance of any (structurally stable) DMI in the diploid model, independent of epistasis parameters or recombination rates. This bound is sharp and is already attained for our two special models: if $m < \alpha - \beta$, a recessive DMI ($\gamma_1 = 0$) exists for any $\gamma > \beta$ and complete linkage ($r = 0$); if $m < \alpha$, a codominant DMI exists for $\gamma_1 \to \infty$ and $r = 0$.

An important consequence of (21) and (22) is that a stable neutral DMI (or any DMI with $\alpha \leq 0$) is impossible in the diploid continent–island model. A proof of this fact for the general diploid DMI model is given in File S2, section S10. We note that this is not a trivial finding: In particular, our general assumption of nondecreasing γ_i with the number of conflicts is essential, and counterexamples can be found if $\gamma_1 \gg \max[\gamma_3, \gamma_4]$ (*cf.* File S4). A special case of potential biological relevance is the scenario of tightly linked loci ($r = 0$) in the secondary-contact scenario (*i.e.*, in the complete absence of *ab* and *AB* haplotypes). In this case, the model reduces to a one-locus two-allele problem and we readily obtain the condition

$$m^{\pm}_{\max} = \begin{cases} \alpha - \beta - \gamma_1 & \text{if } \gamma_1 \leq \dfrac{\alpha - \beta}{3}, \\[2ex] \dfrac{(\alpha - \beta + \gamma_1)^2}{8\gamma_1} & \text{if } \gamma_1 > \dfrac{\alpha - \beta}{3}. \end{cases} \tag{23}$$

Thus, a neutral incompatibility can be maintained if $\gamma_1 > 0$, corresponding to a previous result by Lande (1979). Indeed, the maximum rate of gene flow even increases toward ∞ with growing γ_1. As we will see below, this behavior deviates strongly from any other scenario where no haplotype is excluded from the dynamics. Note that the equilibrium is unstable in the full parameter space and thus not a true DMI according to our definition. Indeed, the incompatibility is quickly lost if the ancestral haplotype *ab* is initially present at a low frequency or if recombination deviates (however slightly) from zero.

Stability of parapatric DMIs in the diploid model

For various parameter combinations, Figure 3B and C display $m^{-}_{\max}$ and $m^{+}_{\max}$ as functions of the strength of epistasis for the recessive and the codominant model, respectively. Corresponding figures for other choices of epistasis schemes are provided in File S3. As in the haploid model, we find that at most a single stable DMI can exist for given parameters. Locally stable DMIs refer to bistable cases, where a DMI coexists with one (or several) stable boundary equilibria. In these cases, evolution depends on the initial conditions. As in the case of the haploid model, all our numerical results show convergence to the DMI if the second substitution occurs on the continent and for secondary contact. In contrast, a DMI never evolves in a bistable case for the *continent–island* and *island–island* scenarios. We find two main types of equilibrium patterns, which are well represented by the recessive or the codominant DMI model.

Recessive incompatibilities: Comparison of Figure 3B with 3A shows that the maximum gene-flow rates $m^{\pm}_{\max}$ for the recessive diploid model and for the haploid model are very similar. This holds, in particular, for low recombination rates, but even for independent loci the differences are modest. We obtain qualitatively similar results for other choices of the epistasis parameters that do not impose selection against the double heterozygotes, *i.e.*, if $\gamma_1 = 0$ [see File S3, Figure S4 for $\Gamma = (0, \gamma, 0, 2\gamma)$, Figure S5 for $\Gamma = (0, 0, \gamma, 2\gamma)$, and Figure S6 for $\Gamma = (0, 0, 0, 2\gamma)$].

Codominant incompatibilities: Comparison of Figure 3C with 3A shows that the maximum rates $m^{\pm}_{\max}$ for the codominant diploid model are very similar to those for the haploid model for high recombination rates [for $\Gamma = (\gamma/2, \gamma, \gamma, 2\gamma)$ they are even identical in the limit $r \to \infty$]. However, significant differences emerge if recombination is weak relative to selection, $r \leq \min[\alpha, \beta]$. On the one hand, the maximum rate for global stability, $m^{-}_{\max}$, is severely reduced relative to that in the haploid model. Irrespective of the other parameters, we find that $m^{-}_{\max} \to 0$ as $\gamma_1 \to \infty$. On the other hand, the maximum rate for a locally stable DMI, $m^+_{\max}$, can be much larger, in particular for large γ. We thus find large parameter regions with bistable dynamics for the codominant model. Again, models with different choices of the epistasis parameters, but $\gamma_1 > 0$, show a similar behavior [see Figure S7 in File S3 for the case $\Gamma = (2\gamma, 2\gamma, 2\gamma, 2\gamma)$].

General incompatibilities

For an epistasis scheme that is intermediate between the two special patterns studied above, *e.g.*, $\Gamma = (\gamma/50, \gamma/5, \gamma/5, 2\gamma)$, we observe a transition between the recessive and the codominant DMI pattern with increasing γ (in File S3, compare Figure S8 with Figures S2 and S3; see also Figure S9). The maximum gene-flow rates are very similar to those in the recessive DMI model as long as the incompatibility is weak and $\gamma_1 \ll \max[\alpha, \beta]$. However, when observed on a larger scale in terms of γ, such that the fitness costs of the double heterozygote are of the order of α and β, the pattern resembles that of the codominant model (File S3, Figure S9).

As in the haploid case, we can understand these results in terms of the fitness landscape and the corresponding mechanisms for the evolution and maintenance of DMIs: heterogeneous ecological selection pressures and selection against hybrids.

1. For a slope-type fitness landscape, due to either local adaptation ($\beta < 0$) or a weak incompatibility ($\alpha > \max[\gamma_2, \gamma_4 - \gamma_2]$ and $\beta < \min[\gamma_3, \gamma_4 - \gamma_3]$), the DMI is maintained by heterogeneous selection. There are no hybrids that are so unfit that they could pose a selection barrier. If gene flow is sufficiently low, we obtain global stability of the DMI even for high recombination (as in the haploid case, $m^-_{\max}$ always decreases with r but, here, $m^-_{\max} > 0$ for $r \rightarrow \infty$).
2. The most significant differences between the haploid and the diploid model occur for a double-peak fitness landscape. In diploids, any selection against the double heterozygote F_1 hybrids ($\gamma_1 > 0$) deepens the fitness valley between the peaks and generates a direct, recombination-independent fitness cost for hybrids. In the haploid model, costly hybrids can be produced only by recombination. As a consequence, the difference between both models is strongest for small r. Since the fitness valley protects both local fitness maxima, the evolution of the DMI (which is close to the "island peak") will always depend on initial conditions. The maximum gene-flow rate $m^+_{\max}$ depends on the depth of the valley and on its shape. If $w_{AaBa} \gg \max[w_{aabb}, w_{AABB}]$ (small γ_1), the barrier is most shallow in the center, *i.e.*, at the *AaBb* hybrids. In this case, selection against hybrids is weakest (and $m^+_{\max}$ smallest) for low recombination (*cf.* Figure 4, recessive model). In the opposite case (large γ_1) the barrier is most shallow at the recombinant hybrids *aabb* or *AABB* of island and continental individuals (*cf.* Figure 4, codominant model). As a consequence, $m^+_{\max}$ increases with *r*.
3. The individual effects of the epistasis coefficients γ_2, γ_3, and γ_4, are further explored in File S3. Compared to γ_1, they are relatively minor provided certain weak conditions are satisfied (File S2). The parameters γ_2 (selection against *aABB* individuals) and γ_3 (selection against *AABb* types) have opposite effects. In particular, reducing γ_2 below $\gamma_4/2$ leads to overdominance at the $\mathcal{A}$ locus for $p_B = 1$ and thus stabilizes a boundary equilibrium without the island haplotype *Ab*. In contrast, $\gamma_3 < \gamma_4/2$ entails overdominance at the $\mathcal{B}$ locus for $p_A = 1$ and thus contributes to the stability of a DMI with $x_3 > 0$.

Discussion

The standard model to explain the evolution of intrinsic postzygotic isolation in allopatry, the Dobzhansky–Muller model, has repeatedly been described as a plausible speciation scenario, also in the presence of gene flow (*e.g.*, Gavrilets 2004; Wu and Ting 2004). However, little is known about the limits up to which the mechanism could work. As phrased by Turelli *et al.* (2001, p. 341), "we would like to understand how much gene flow is needed to inhibit the accumulation of DMIs." The origin (and maintenance) of the first incompatibility in a genetically homogeneous ancestral population is the crucial step of this process, since the evolution of any additional DMI can only be easier.

In our study, we address this question for a minimal model with two loci and two alleles. We focus on the particular case of unidirectional gene flow, *e.g.*, from a continent to an island population. For this model, we obtain a complete analytical characterization of the maximum rates of gene flow that permit DMI evolution. The model accounts for exogeneous and endogeneous selection and allows for different temporal–spatial patterns in the origination of the substitutions. We thus can disentangle the relative importance of the external environment, the genetic architecture of the incompatibility, and the evolutionary history for the evolution and maintenance of parapatric DMIs.

Two mechanisms can drive the evolution of a parapatric DMI. In a heterogeneous environment, a DMI can emerge as a by-product of selection against maladapted immigrants. In addition, selection against unfit hybrids can maintain a parapatric DMI even in homogeneous environments. Both mechanisms rely on exogenous selection, but lead to opposite predictions concerning the genetic architecture of the DMI and its dependence on the evolutionary history. Our main results are the following:

1. The maximum rate of gene flow can never exceed a bound set by exogenous selection pressures. This implies that the adaptive advantage of the single substitutions rather than the strength of the incompatibility is the most important factor in determining whether a DMI can evolve or be maintained in the presence of gene flow. In particular, neutral DMIs cannot exist in parapatry.
2. In the haploid model, selection against immigrants is most effective in driving the evolution of parapatric DMIs if both incompatibility loci are tightly linked. In contrast, selection against hybrids works best for strong, but loosely linked incompatibilities. The results for the diploid model

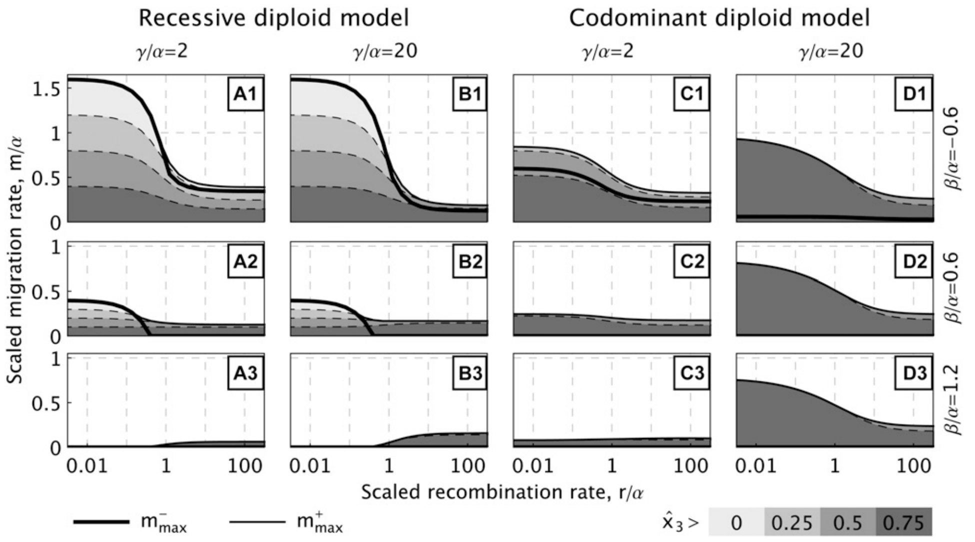

Figure 4 Maximum migration rate (solid lines) and population differentiation (shaded areas) in the recessive and codominant diploid models as a function of the recombination rate γ. For each model, two choices for the strength γ of epistasis (weak/strong) and three choices for β, describing different types of selection scenarios (corresponding to local adaptation/heterogeneous/homogeneous environments), are represented. Similar to the haploid model (*cf.* Figure 1), the results in the recessive model are almost independent of the strength of epistasis (compare A1–A3 and B1–B3), but strongly dependent on the environment (compare rows 1, 2, and 3). In contrast, for the codominant model, $m^+_{\max}$ and $m^-_{\max}$ depend only weakly on the environment if epistasis is strong (D1–D3). Generally, the codominant model produces stronger population differentiation than the recessive model (dark shading dominating), although the overall maximum migration rate in scenarios of local adaptation is lower; compare A1 and B1 with C1 and D1.

are very similar to those for the haploid model if the incompatibility is recessive, *i.e.*, if double-heterozygous F_1 hybrids are not affected by the incompatibility. In contrast, for codominant incompatibilities, the fitness of double heterozygotes becomes a decisive factor to determine DMI evolution.

3. In many cases, a DMI will evolve only if the two incompatible substitutions appear in the correct order on the continent and on the island. In particular, if selection acts mainly against hybrids in a homogeneous environment, the outcome of the evolutionary process is dependent on historic contingency.

In the following, we discuss the most important implications that arise from these results and relate them to existing empirical and theoretical work.

The importance of exogenous selection

Our results highlight the importance of exogeneous selection (parameters α and β in the model) and of adaptive evolution during parapatric speciation. At least one of the substitutions involved in a stable parapatric DMI must have a selective advantage on the island relative to the ancestral type ($\alpha > 0$ for the "island substitution"). This confirms previous claims that neutral postzygotic speciation is impossible in the presence of gene flow (Barton and Bengtsson 1986; Gavrilets 1997; Kondrashov 2003; Lemmon and Kirkpatrick 2006; but see Gavrilets 2004). Although the formal proof of this fact is quite involved for the diploid model (see File S2), there is an intuitive explanation: If both substitutions are neutral, the compatible ancestral haplotype has a marginal advantage over all haplotypes that carry a substitution. Thus, whenever the ancestral type is produced by recombination (or if its initial frequency is positive), it can rise in frequency until the island substitution (the allele that is not present on the continent) is lost. These results also exclude the maintenance of a neutral DMI after secondary contact, even if such a DMI may have originated in a phase of strict allopatry. The only exception is a DMI that is effectively just a single underdominant diploid locus (a single recombinational unit). This may occur, *e.g.*, if both substitutions are joined in an inversion. However, as has previously been shown in a simulation study, even tiny recombination rates preclude the maintenance of a neutral DMI upon secondary contact (Feder and Nosil 2009).

These findings generalize easily to pairwise DMIs in the background of other genetic barriers to gene flow by substituting an effective migration rate (Barton and Bengtsson 1986). Since the maximum rate of gene flow for the maintenance of a neutral DMI is zero, such an incompatibility will be maintained only if isolation is already complete. It thus cannot contribute to the speciation process, even at a late stage. These results are consistent with empirical evidence for signatures of positive selection, which have been observed for most incompatibility genes that have been identified so far (Presgraves *et al.* 2003; Johnson 2010).

The maximum rates of gene flow for locally or globally stable DMIs increase with a higher selective advantage of the island substitution (higher α) and with higher disadvantage of the incompatible continental substitution (lower β). Strong local adaptation (with $\alpha > 0$ and $\beta < 0$) can enable the accumulation of DMIs under strong gene flow ($m^+_{\max} = \alpha - \beta$; *e.g.*, Figure 3, A1 and B1). Since empirical studies have indeed identified cases of very strong divergent selection (*e.g.*, Mullen and Hoekstra 2008), this implies that the evolution of DMIs as a by-product of local adaptation might be possible in nearly sympatric scenarios. However, the level of population differentiation that is maintained by a DMI at high gene-flow rates will become vanishingly low (Figure 1A and Figure 4, A1 and B1).

Stability of a parapatric DMI does not require local adaptation, but is also possible in a homogeneous environment if there is sufficiently strong selection against hybrids (*i.e.*, if $\gamma > \beta$ in the haploid model). It is even possible that the "island" substitution is less beneficial on the island than the invading, incompatible "continental" substitution ($\beta > \alpha > 0$; *e.g.*, Figure 3, A3–C3). Although the maximum rates of gene flow are lower than for local adaptation ($m^{+}_{\max} \leq \alpha/4$ for haploids, $m^{+}_{\max} \leq \alpha$ for diploids), the amount of population differentiation that a single DMI can achieve, even close to the maximum migration rate, is still high (Figure 1C and Figure 4, A3–D3). This is in line with classical arguments about hybrid zones by Barton and Hewitt (1989), who claimed that local adaptation might be important to generate variation in the first place, but selection against hybrids is supposed to be more important for maintaining a hybrid zone.

The importance of the genetic architecture of the DMI

Reliable inferences about the genetic architecture of a DMI (its strength γ and its linkage r) must take into account under which mechanism it has evolved. If a DMI is maintained simply as a by-product of local adaptation or direct selection against immigrants in a heterogeneous environment, the strength of the incompatibility (measured by the epistasis parameters) is of minor importance. The maximum admissible migration rates often even decrease with stronger epistasis (*e.g.*, Figure 3, A1–C1). In the selection-against-immigrants scenario, a DMI on the island is most effectively protected from gene flow if tight linkage (low r) of the incompatibility loci maintains the immigrant haplotype *aB* (*e.g.*, Figure 1A). In contrast, strong recombination will quickly decompose the immigrant type, even if it does not suffer a direct fitness cost. Instead, hybrid haplotypes (*ab* and *AB*) are produced. DMIs on the island can withstand the immigration pressure if and only if sufficiently strong selection acts against these hybrids.

In haploid populations, hybrid individuals suffering from an incompatibility are produced only by recombination. This also holds for recessive diploid incompatibilities ($\gamma_1 = 0$), where double-heterozygous F_1 individuals pay no cost. If selection against hybrids is the driving mechanism of DMI evolution, we therefore find that DMI stability against gene flow increases with recombination (*cf.* Figure 1C). This is different for codominant DMIs, where the incompatibility affects double heterozygotes (H_0 incompatibility in the terminology of Turelli and Orr 2000). Here, selection against F_1 hybrids ($\gamma_1 > 0$) can protect the island genotypes from swamping by continental types even if recombination is low or absent. Indeed, the maximum rate of gene flow permitting stable DMIs decreases with recombination if selection against F_1 hybrids is strong (*cf.* Figure 4C3 and D3). Our results conform to related findings by Gavrilets (1997), who noted the importance of F_1 fitness for the strength of a barrier to gene flow induced by a DMI in a model of a genetic cline.

The importance of the evolutionary history

The two incompatible substitutions can originate in various temporal–spatial patterns, either on the continent or on the island. They can occur in the presence of gene flow or in a temporary allopatric phase. We find that, with respect to the stability of the DMI, all evolutionary histories can be categorized in two classes. If the second substitution arises on the continent (and the first substitution either on the island or on the continent), or if both alleles originate in an allopatric phase, they are initially protected from incompatibility selection. A DMI will evolve under these conditions whenever the model parameters allow for a stable incompatibility, *i.e.*, independent of initial conditions. In contrast, if the second substitution occurs on the island, a DMI will evolve if and only if the corresponding equilibrium is globally stable (*i.e.*, there is no alternative stable evolutionary endpoint). The intuitive reason is that the second allele faces competition from its incompatible counterpart, which is maintained on the island either due to local selection or due to recurrent gene flow from the continent.

As expected, a favorable evolutionary history is always decisive in a homogeneous environment. The dependence on the history is also increased by high recombination rates and by selection against double heterozygotes in the codominant diploid model. We thus find that historic contingency becomes important in all scenarios in which DMIs primarily evolve due to selection against hybrids. In contrast, DMIs are usually (but not always) globally stable if selection acts directly against immigrants.

The strong dependence on the evolutionary history that we observe is partially a consequence of the migration asymmetry that is inherent to the continent–island model. We can extend our results to include very weak back migration from the island to the continent under the following assumptions. Back migration is much weaker than both selection and migration from the continent, such that the equilibria derived for unidirectional gene flow are only slightly perturbed. Back migration nevertheless is much stronger than recurrent new mutation on the continent, such that the first substitution will reach equilibrium in both demes, independent of its deme of origin. The impact of weak back migration of this type depends on the fitness landscape on the continent. If fixation of the continental haplotype (*aB*) is the only stable equilibrium on the continent, we retain all predictions from the model with strict unidirectional migration. If the fitness function on the continent is double peaked, with local optima for fixation of either the continental or the island haplotype (*aB* or *Ab*), the continental substitution *B* must occur first, since otherwise both demes fix for the island haplotype. In this case, however, the island substitution *A* faces competition from the continental *B* allele and will be able to invade only if the DMI in the unidirectional model is globally stable.

An important consequence of this result is that heterogeneous selection across both demes is a prerequisite for the evolution a parapatric DMI if there is (even minimal) back migration. The same line of arguments also shows that evolution of a codominant DMI may be difficult in the presence of bidirectional gene flow, since strong selection against double heterozygotes will plausibly lead to a double-peak fitness landscape on the continent and on the island. Indeed, most hybrid incompatibility loci that have been identified so far are recessive (*cf*. Presgraves 2003). This is expected particularly for DMIs that have evolved via gene duplications or selfish elements (Johnson 2010; Presgraves 2010).

Ecological vs. mutation-order speciation

According to Schluter (2009), adaptive postzygotic isolation with gene flow falls into two broad categories. *Ecological speciation* refers to the case that each substitution of a DMI pair is beneficial in one subpopulation, but deleterious in the other. Speciation is then a by-product of local adaptation. In contrast, the case in which both substitutions are beneficial in both demes is referred to as *mutation-order speciation* (Mani and Clarke 1990). Characteristic for this second scenario is that speciation depends on the "chance occurrence and fixation of different alleles between populations adapting to similar selection pressures" (Schluter 2009, p. 737).

The present study provides a framework for an adjustment and formal validation of this verbal classification. In our model, local adaptation corresponds to a deleterious effect of the continental substitution on the island ($\beta < 0$). "Mutation order" expresses dependence on the evolutionary history, which is the case in all regimes with a locally, but not globally, stable DMI. We find that these types are not exclusive, nor do they cover the entire parameter space where a DMI can evolve: On the one hand, the potential for "pure" mutation-order speciation with (strictly) homogeneous selection across both demes is relatively limited in the continent–island model and becomes impossible if there is back migration. On the other hand, the correct mutation order is often essential in addition to local adaptation to explain the evolution of a DMI, in particular for codominant DMIs in the diploid model. Finally, there is a large parameter range where weak DMIs (with $\gamma < \alpha$) evolve under heterogeneous selection, but without the need of local adaptation and independently of the evolutionary history.

From our model, we do not find a clean way to classify the DMIs themselves into categories. Instead, we distinguish two mechanisms that can drive the evolution of parapatric DMIs (*cf*. Figure 2). This distinction retains some flavor of Schluter's classification: Selection against immigrants relies on ecological selection due to a heterogeneous environment, whereas selection against hybrids in a homogeneous environment must rely on mutation order. For each mechanism there are parameter regions in which one dominates the other. This has characteristic consequences for predictions concerning the genetic architecture or historic contingency. However, and in contrast to Schluter's scheme, the two mechanisms are not exclusive and reliable predictions become more complex in the broad transition region.

The limited potential of mutation-order speciation has previously been reported by Nosil and Flaxman (2011), who performed a simulation study in a two-island model. Furthermore, Agrawal *et al.* (2011) studied the evolution of a DMI in the context of ecological speciation and examined the resulting barriers to neutral gene flow.

Limitations of the model

Our continent–island model of DMI evolution represents a minimal model to investigate under which conditions of the genetics and the environment a single two-locus DMI can evolve. We have shown that in this simple case a comprehensive analytical understanding of these conditions can be achieved. Naturally, several further issues emerge from our study. Maybe the most pressing question is how our results depend on our assumptions about population structure. Since the continent–island model represents an extreme case of asymmetric mutation, it should be complemented by a study of a two-island model with bidirectional migration. As outlined above, we generally expect that DMIs are more difficult to evolve with bidirectional migration, where the second mutation always experiences immediate incompatibility selection. In contrast, the maintenance of a DMI may be possible under even higher migration rates, if part of the gene flow is due to back migration of advantageous types. These intuitive expectations are supported by preliminary results and previous studies (*e.g.*, Karlin and McGregor 1972a; Lande 1979). Further extensions could address DMI evolution in a stepping-stone model or in continuous space, building on the preliminary results by Gavrilets (1997) and Kondrashov (2003), respectively.

As a technical point, we have used continuous-time dynamics, implicitly assuming weak selection. However, numerical checks against a model with discrete generations did not show qualitative changes even for fairly large selection coefficients (results not shown). More significantly, our derivations do not account for genetic drift. Thus, we implicitly assume that drift is much weaker than selection and migration. Following heuristic arguments by Yeaman and Otto (2011) suggests that the stability properties of DMIs should remain unaltered under drift if selection and migration parameters are $\gg 1/2N$. Whereas the regime where drift is much stronger than migration (such that the population is almost always monomorphic) was analyzed by Gavrilets (2000), the case in which all three forces are of similar size has yet to be studied.

Also our assumptions on the genetics of the model are restrictive and call for further study. A valuable (and relatively easy) next step would be to extend the model to nonautosomal incompatibilities, since empirical studies report a growing number of DMIs that involve the sex chromosomes or mitochondria (Presgraves 2003; Chou and Leu 2010). Since we consider only a two-locus DMI, extensions to higher-order incompatibilities would be welcome. We expect, however, that the analysis will be much more complex and results will have to rely on extensive numerical work. Finally, our model considers the origin and maintenance of only a single, first DMI in a previously homogeneous population. Although this is an important question, also with regard to reports of segregating incompatibilities within populations (Cutter 2012), in the context of parapatric speciation the entire dynamics of DMI accumulation up to complete isolation need to be addressed. In the light of a growing number of whole-genome scans for incompatibilities (*e.g.*, Kao *et al.* 2010), it would be particularly interesting to predict how distribution patterns of DMIs along the genome depend on environmental and genetic conditions.

Our results are informative for an accumulation process in various ways. Primarily, they characterize the conditions for the onset of parapatric speciation via accumulating DMIs. Because the first single DMI will establish population differentiation, this will reduce the effective gene flow between the populations and hence simplify the accumulation of additional incompatibilities. Moreover, if one interprets the maximum migration rate as a maximum rate of effective gene flow, our model can be readily extended to two populations at an advanced stage of the speciation process. In particular, we can also assume that barriers to gene flow have been built by prezygotic mechanisms, such as reinforcement (*e.g.*, Nosil *et al.* 2003). In this case, our model predicts the maximum residual amount of gene flow that permits DMIs to accumulate and finish the speciation process when other evolutionary forces, such as reinforcement, have become too weak to shut off gene flow (*cf.* Bank *et al.* 2012).

Predictions and conclusions

What can we conclude about the possibility of evolution of postzygotic reproductive isolation in the presence of gene flow? We have seen that the substitutions involved in a parapatric DMI must be adaptive. Also, at least some environmental heterogeneity is usually necessary, in particular if there is gene flow from the island back to the continent. In this case, recessive incompatibilities are far easier to evolve than codominant incompatibilities. If selection on the island acts most strongly against maladapted immigrants from the continent, the emergence of recessive DMIs is promoted by tight linkage between loci. Consequently, we expect an advantage for incompatibilities that evolve in regions of low recombination, such as chromosomal rearrangements (Bürger and Akerman 2011; Yeaman and Whitlock 2011). This supports a "genomic islands of speciation" model (Wu and Ting 2004), where DMIs accumulate in clusters on the genome. Incompatibilities are just a by-product of strong diversifying selection in this case, and the evolutionary outcome is largely independent of the chronological order of the substitutions. On the contrary, if selection on the island acts primarily against hybrids, the correct mutation order is decisive. In stark contrast to the "genomic island" picture, incompatibility pairs are rather expected to be loosely linked or unlinked in this case. Also upon secondary contact, populations that have undergone initial divergence under similar selection pressures in allopatry are more likely to maintain incompatibilities that are unlinked.

Although there are still many open questions, the present study provides theoretical groundwork for a characterization of the conditions for the evolution of differentiation via DMIs in parapatry. In particular, our analytical results show that progress on these issues is not necessarily limited to simulation studies. We hope that our methods can serve as a basis for future work on the evolution of reproductive isolation in parapatry and shed light on the ongoing discussion about the plausibility of ecological and mutation-order speciation.

Acknowledgments

We thank N. Barton, J. Hofbauer, M. Servedio, and the members of the Doktoratskolleg Populationsgenetik for helpful discussions or comments on the manuscript. C.B. and J.H. are members of the Mathematics and BioSciences Group at the University of Vienna, which is funded by the Vienna Science and Technology Fund through project MA06-01. R.B. is grateful for support by grant P21305-N13 of the Austrian Science Fund.

Literature Cited

Agrawal, A. F., J. L. Feder, and P. Nosil, 2011 Ecological divergence and the origins of intrinsic postmating isolation with gene flow. Int. J. Ecol. 2011: 1–15.

Bank, C., J. Hermisson, and M. Kirkpatrick, 2012 Can reinforcement complete speciation? Evolution 66: 229–239.

Barton, N., and B. O. Bengtsson, 1986 The barrier to genetic exchange between hybridising populations. Heredity 57: 357–376.

Barton, N. H., and G. M. Hewitt, 1989 Adaptation, speciation and hybrid zones. Nature 341: 497–503.

Bateson, W., 1909 Heredity and variation in modern lights, pp. 85–101 in *Darwin and Modern Science*, edited by A. C. Seward. Cambridge University Press, Cambridge, UK.

Bürger, R., 2009 Multilocus selection in subdivided populations I. Convergence properties for weak or strong migration. J. Math. Biol. 58: 939–978.

Bürger, R., and A. Akerman, 2011 The effects of linkage and gene flow on local adaptation: a two-locus continent-island model. Theor. Popul. Biol. 80: 272–288.

Chou, J.-Y., and J.-Y. Leu, 2010 Speciation through cytonuclear incompatibility: insights from yeast and implications for higher eukaryotes. BioEssays 32: 401–411.

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

Cutter, A. D., 2012 The polymorphic prelude to Bateson-Dobzhansky-Muller incompatibilities. Trends Ecol. Evol. 27: 209–218.

Dobzhansky, T., 1936 Studies on hybrid sterility. II. Localization of sterility factors in Drosophila pseudoobscura hybrids. Genetics 21: 113–135.

Feder, J. L., and P. Nosil, 2009 Chromosomal inversions and species differences: When are genes affecting adaptive divergence and reproductive isolation expected to reside within inversions? Evolution 63: 3061–3075.

Feldman, M. W., 1971 Equilibrium studies of two locus haploid populations with recombination. Theor. Popul. Biol. 2: 299–318.

Gavrilets, S., 1997 Hybrid zones with Dobzhansky-type epistatic selection. Evolution 51: 1027–1035.

Gavrilets, S., 2000 Waiting time to parapatric speciation. Proc. Biol. Sci. 267: 2483.

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

Hofbauer, J., 1990 An index theorem for dissipative semiflows. Rocky Mountain J. Math. 20: 1017–1031.

Johnson, N. A., 2010 Hybrid incompatibility genes: Remnants of a genomic battlefield? Trends Genet. 26: 317–325.

Kao, K. C., K. Schwartz, and G. Sherlock, 2010 A genome-wide analysis reveals no nuclear Dobzhansky-Muller pairs of determinants of speciation between S. cerevisiae and S. paradoxus, but suggests more complex incompatibilities. PLoS Genet. 6: e1001038.

Karlin, S., and J. McGregor, 1972a Application of method of small parameters to multi-niche population genetic models. Theor. Popul. Biol. 3: 186–209.

Karlin, S., and J. McGregor, 1972b Polymorphisms for genetic and ecological systems with weak coupling. Theor. Popul. Biol. 3: 210–238.

Kondrashov, A. S., 2003 Accumulation of Dobzhansky-Muller incompatibilities within a spatially structured population. Evolution 57: 151–153.

Lande, R., 1979 Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement. Evolution 33: 234–251.

Lemmon, A. R., and M. Kirkpatrick, 2006 Reinforcement and the genetics of hybrid incompatibilities. Genetics 173: 1145–1155.

Lowry, D. B., J. L. Modliszewski, K. M. Wright, C. A. Wu, and J. H. Willis, 2008 The strength and genetic basis of reproductive isolating barriers in flowering plants. Philos. Trans. R. Soc. B 363: 3009–3021.

Mani, G. S., and B. C. Clarke, 1990 Mutational order: a major stochastic process in evolution. Proc. R. Soc. Lond. B Biol. Sci. 240: 29–37.

Mayr, E., 1942 *Systematics and the Origin of Species, from the Viewpoint of a Zoologist*. Harvard University Press, Cambridge, MA.

Mullen, L. M., and H. E. Hoekstra, 2008 Natural selection along an environmental gradient: a classic cline in mouse pigmentation. Evolution 62: 1555–1570.

Muller, H., 1942 Isolating mechanisms, evolution and temperature. Biol. Symp. 6: 71–125.

Nagylaki, T., J. Hofbauer, and P. Brunovský, 1999 Convergence of multilocus systems under weak epistasis or weak selection. J. Math. Biol. 38: 103–133.

Nosil, P., and S. M. Flaxman, 2011 Conditions for mutation-order speciation. Proc. Biol. Sci. 278: 399–407.

Nosil, P., B. J. Crespi, and C. P. Sandoval, 2003 Reproductive isolation driven by the combined effects of ecological adaptation and reinforcement. Proc. Biol. Sci. 270: 1911–1918.

Orr, H. A., 1995 The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139: 1805–1813.

Porter, A. H., and N. A. Johnson, 2002 Speciation despite gene flow when developmental pathways evolve. Evolution 56: 2103–2111.

Presgraves, D., 2003 A fine-scale genetic analysis of hybrid incompatibilities in Drosophila. Genetics 163: 955–972.

Presgraves, D. C., 2010 The molecular evolutionary basis of species formation. Nat. Rev. Genet. 11: 175–180.

Presgraves, D. C., L. Balagopalan, S. M. Abmayr, and H. A. Orr, 2003 Adaptive evolution drives divergence of a hybrid inviability gene between two species of Drosophila. Nature 423: 715–719.

Rundell, R. J., and T. D. Price, 2009 Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. Trends Ecol. Evol. 24: 394–399.

Rutschman, D., 1994 Dynamics of the two-locus haploid model. Theor. Popul. Biol. 45: 167–176.

Schluter, D., 2009 Evidence for ecological speciation and its alternative. Science 323: 737–741.

Slatkin, M., 1987 Gene flow and the geographic structure of natural populations. Science 236: 787–792.

Turelli, M., and H. A. Orr, 2000 Dominance, epistasis and the genetics of postzygotic isolation. Genetics 154: 1663–1679.

Turelli, M., N. H. Barton, and J. A. Coyne, 2001 Theory and speciation. Trends Ecol. Evol. 16: 330–343.

Unckless, R. L., and H. A. Orr, 2009 Dobzhansky-Muller incompatibilities and adaptation to a shared environment. Heredity 102: 214–217.

Wu, C.-I., and C.-T. Ting, 2004 Genes and speciation. Nat. Rev. Genet. 5: 114–122.

Yeaman, S., and S. P. Otto, 2011 Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. Evolution 65: 2123–2129.

Yeaman, S., and M. C. Whitlock, 2011 The genetic architecture of adaptation under migration-selection balance. Evolution 65: 1897–1911.

Communicating editor: L. M. Wahl

Appendix

Haploid Model

For the haploid model, we determined the equilibrium structure completely and obtained explicit analytical expressions for the maximum rates of gene flow below which locally or globally stable DMIs can be maintained. Since the proofs are complex and require advanced mathematical arguments, they are given in File S4. File S4 provides a self-contained analytical treatment of the haploid model. For dedicated readers interested in additional results, we also provide a fully annotated *Mathematica* notebook in File S4. In this *Appendix*, we give a summary of our results. In particular, we state all formulas for the maximum rates of gene flow $m_{\max}^{\pm}$ that are used in the main body of the article.

For many purposes, it is convenient to write the dynamics in terms of the frequencies p and q of the incompatible alleles (A and B) and the linkage-disequilibrium coefficient D:

$$\dot{p} = \alpha p(1-p) - \gamma(1-p)(pq+D) + \beta D - mp, \tag{A1a}$$

$$\dot{q} = \beta q(1-q) - \gamma(1-q)(pq+D) + \alpha D + m(1-q), \tag{A1b}$$

$$\begin{aligned}\dot{D} = {} & [\alpha(1-2p) + \beta(1-2q)]D - \gamma[(1-p)(1-q) - D](pq+D) \\ & - rD - m[p(1-q) + D].\end{aligned} \tag{A1c}$$

We note that D has to satisfy

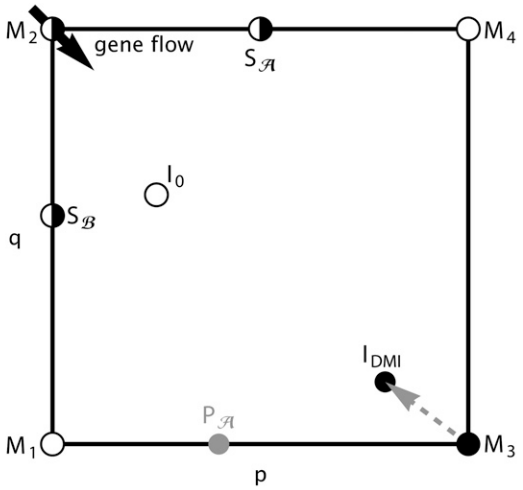

Figure A1 State-space diagram, projected onto the plane spanned by the allele frequencies p and q, showing the possible equilibria and initial states of the haploid dynamics. Immigration of continental haplotypes moves the monomorphic equilibrium $\mathbf{M}_3$ (which exists for $m = 0$) into the interior of the state space, thus yielding a stable DMI. Open circles represent instable equilibria, whereas half-filled circles represent equilibria which can be stable or instable. The gray circle represents a potential polymorphic initial state.

$$-\min[pq,\ (1-p)(1-q)] \le D \le \min[p(1-q),\ (1-p)q]\ . \tag{A2}$$

Equilibrium structure

Figure A1 sketches the positions in state space of all equilibria that can exist for the haploid model with $m \geq 0$. These are the four monomorphic equilibria $\mathbf{M}_i$, where the corresponding haplotype is fixed on the island ($x_i = 1$). Two boundary equilibria with a single polymorphic locus may exist on the edges $q = 1$ or $p = 0$. We denote them by $\mathbf{S}_{\mathcal{A}}$ and $\mathbf{S}_{\mathcal{B}}$ (locus $\mathcal{A}$, $\mathcal{B}$ polymorphic), respectively. There is at most one asymptotically stable internal equilibrium, denoted by $\mathbf{I}_{\mathrm{DMI}}$. If it exists we say that a parapatric DMI is maintained. Finally, an unstable internal equilibrium, $\mathbf{I}_0$, may exist.

In the absence of migration ($m = 0$), all four monomorphic equilibria $\mathbf{M}_i$ exist. In the relevant parameter space ($\alpha > 0$ and $\gamma > \beta$, see Equation 2), only $\mathbf{M}_2$ (the continental haplotype *aB* is fixed) and $\mathbf{M}_3$ (the island haplotype *Ab* is fixed) can be stable. An unstable internal equilibrium ($\mathbf{I}_0$) exists if and only if $\mathbf{M}_2$ and $\mathbf{M}_3$ are both stable. No further equilibrium can exist in this case. For sufficiently small $m > 0$, $\mathbf{I}_{\mathrm{DMI}}$ exists if and only if the island equilibrium $\mathbf{M}_3$ is stable for $m = 0$ [leading to conditions (2) and (3)]. Indeed, $\mathbf{I}_{\mathrm{DMI}}$ is a perturbation of $\mathbf{M}_3$ and enters the state space when m becomes positive. $\mathbf{I}_{\mathrm{DMI}}$ is globally asymptotically stable if and only if it is the only internal equilibrium. If the unstable internal equilibrium $\mathbf{I}_0$ exists, the dynamics are bistable, and $\mathbf{I}_{\mathrm{DMI}}$ and one of the boundary equilibria ($\mathbf{M}_2$, $\mathbf{S}_{\mathcal{A}}$, or $\mathbf{S}_{\mathcal{B}}$) are locally stable attractors.

As m increases, there are three different scenarios for the fate of $\mathbf{I}_{\mathrm{DMI}}$ (File S1, section S.7; note that these patterns in the dynamics do not easily correspond to the mechanisms of DMI evolution that we identify):

1. If the unstable internal equilibrium $\mathbf{I}_0$ exists already at $m = 0$, we have $m^-_{\max} = 0$. Both internal equilibria merge and disappear at $m = m^+_{\max} > 0$, and the previously locally stable boundary equilibrium becomes globally stable.
2. The unstable equilibrium $\mathbf{I}_0$ enters the state space through one of the boundary equilibria at $m = m^-_{\max} > 0$. Both internal equilibria merge and disappear at $m = m^+_{\max} > m^-_{\max}$.
3. The stable equilibrium $\mathbf{I}_{\mathrm{DMI}}$ leaves the state space through one of the three boundary equilibria at $m = m^+_{\max} = m^-_{\max}$, which then becomes globally stable for larger m. Thus, an unstable internal equilibrium never exists.

We note that a DMI always exhibits negative linkage disequilibrium unless independent loci are assumed (File S1, section S.2.4). This is not at variance with the usual finding that migration induces positive linkage disequilibrium because in our model the frequencies of the "extreme" gametes (*Ab* and *aB*) are x_2 and x_3.

Boundary equilibria and global stability

At a boundary equilibrium, at least one haplotype frequency x_i is zero. For recombination rates $r > 0$, this is possible only if one of the two loci is monomorphic (we deal with the case $r = 0$ separately in File S1, section S.8.1). Due to constant immigration of the continental haplotype *aB*, there are no equilibria with $p = 1$ or $q = 0$. Thus, boundary equilibria satisfy $p = 0$ or $q = 1$, and hence, $D = 0$. If $m > 0$, $\mathbf{M}_2$ is the only monomorphic equilibrium. A linear stability analysis reveals that $\mathbf{M}_2$ is asymptotically stable if

$$m > \max[-\beta, \alpha - \gamma, \alpha - \beta - r]. \tag{A3}$$

The three lower bounds correspond to the invasion criteria for the other three haplotypes. If $m < -\beta$, the selective advantage of the wild-type *ab* over *aB* is sufficiently large to exceed the growth of the continental type due to migration. Similarly, if $m < \alpha - \gamma$, the recombinant-type *AB* outperforms the continental type near $\mathbf{M}_2$. Note that recombination of *aB* haplotypes with rare *ab* or *AB* types does not affect the haplotype frequencies. This is different if rare island haplotypes *Ab* invade the monomorphic equilibrium $\mathbf{M}_2$. Then recombination between *aB* and *Ab* reduces the growth rate of *Ab* in proportion to its frequency. We therefore obtain the criterion $m < \alpha - \beta - r$ for the invasion of $\mathbf{M}_2$ by the island haplotype. We define the following critical migration rates:

$$m_{\mathcal{A}} = \frac{(\alpha-\gamma)(\gamma-\beta)}{\alpha}\left(1 + \frac{\alpha+\beta-\gamma}{r}\right), \tag{A4a}$$

$$m_{\mathcal{B}} = \frac{-\beta\alpha}{\gamma-\beta}\left(1 + \frac{\gamma-\beta-\alpha}{r}\right), \tag{A4b}$$

$$m_2 = \alpha - \beta - r. \tag{A4c}$$

The equilibrium $\mathbf{S}_{\mathcal{A}}$ has the coordinates $(p,q,D) = (1-m/(\alpha-\gamma),\ 1,\ 0)$ and exists if $m \leq \alpha - \gamma$. It is asymptotically stable if

$$-\beta < \alpha - \gamma \quad \text{and} \quad m_{\mathcal{A}} < m < \alpha - \gamma, \tag{A5}$$

which requires $r > \gamma - \beta$. Similarly, the equilibrium $\mathbf{S}_{\mathcal{B}}$ has coordinates $(0,\ -m/\beta,\ 0)$, exists if $m \leq -\beta$, and is asymptotically stable if

$$\alpha - \gamma < -\beta \quad \text{and} \quad m_{\mathcal{B}} < m < -\beta, \tag{A6}$$

which requires $r > \alpha$. The coordinates and stability conditions of $\mathbf{S}_{\mathcal{A}}$ and $\mathbf{S}_{\mathcal{B}}$ are related by the substitutions $\alpha \to \gamma - \beta$ and $\beta \to \gamma - \alpha$. This substitution exchanges the roles of the wild-type *ab* and the recombinant haplotype *AB* in the fitness scheme and expresses an important symmetry of the haploid model. From conditions (A3), (A5), and (A6), we deduce at once that at most one of the boundary equilibria can be stable for given parameters.

From the stability properties of the boundary equilibria, we can derive conditions for a globally stable DMI (and thus for $m^{-}_{\max}$). In particular, $\mathbf{I}_{\mathrm{DMI}}$ exists and is globally stable if and only if the boundary of the state space is repelling, *i.e.*, if genotypes that are initially absent can invade every boundary equilibrium. From (A3), (A5), and (A6) we deduce the following values for $m^{-}_{\max}$:

$$m_{\max-} = \begin{cases} m_{\mathcal{A}} & \text{if } \gamma - \alpha < 0 < \gamma - \beta < \alpha \text{ and } r > \gamma - \beta, \\ \mathrm{m}_{\mathcal{B}} & \text{if } \beta < 0 < \alpha < \gamma - \beta \quad \text{and } r > \alpha, \\ m_2 & \text{if } r \leq \min[\alpha, \alpha - \beta, \gamma - \beta]. \end{cases}$$

(A7a) (A7b) (A7c)

Condition (A7a) delineates the parameter regime with a globally stable DMI ($m < m^{-}_{\max} = m_{\mathcal{A}}$) from that where the boundary equilibrium $\mathbf{S}_{\mathcal{A}}$ is locally or globally stable. Conditions (A7b) and (A7c) provide analogous delineations for the regions where $\mathbf{S}_{\mathcal{B}}$ and $\mathbf{M}_2$ are stable. In each case, there are two possibilities: If $\mathbf{I}_{\mathrm{DMI}}$ leaves the state space at $m = m^{-}_{\max}$, the boundary equilibrium involved in this bifurcation becomes globally stable. If the unstable equilibrium $\mathbf{I}_0$ enters the state space at $m = m^{-}_{\max}$, the corresponding boundary equilibrium becomes locally stable.

Internal equilibria and local stability

The possible internal equilibria are determined in File S1, section S.3. The coordinate p of an internal equilibrium can be expressed as a root of a third-order polynomial. Given p, the coordinate q is the root of a quadratic polynomial. Finally, D is a simple rational function of p and q. Only solutions that fall inside the state space [with $p,\ q \in [0,1]$ and D has to satisfy (A2)] correspond to equilibria of the model. The explicit solutions are complicated and several case distinctions are necessary to ensure that a triple $(p,\ q,\ D)$ is in the admissible range. Therefore, we give explicit formulas for the coordinates of $\mathbf{I}_{\mathrm{DMI}}$ and $\mathbf{I}_0$ only in special cases and rather focus on the conditions for the existence of these equilibria.

As explained above, there are two possibilities how a stable DMI can vanish from the state space as m increases. The equilibrium can leave the state space through one of the boundary equilibria at $m^{-}_{\max}$. Alternatively, the internal equilibria $\mathbf{I}_{\mathrm{DMI}}$ and $\mathbf{I}_0$ can merge and disappear at a bifurcation point $m = m^{*}$ that is given by (File S1, section S.4)

$$\begin{aligned} m^{*} = \frac{1}{27r\alpha\gamma^2(\gamma-\beta)}\Big(& -2\alpha^2(\gamma-\beta)^2[8\beta(\gamma-\alpha)+\gamma(\beta-\alpha)] \\ & -3\alpha\gamma(\gamma-\beta)[2\beta(\gamma-\alpha)+\gamma(\beta-\alpha)]r \\ & -3\alpha\gamma^2(\gamma-\beta)r^2 - 2\gamma^3r^3 \\ & +2\sqrt{\left[\alpha^2\left(4\beta^2-5\beta\gamma+\gamma^2\right)+\gamma^2r^2-\alpha(\beta-\gamma)\gamma(3\beta+r)\right]^3}\Big). \end{aligned} \tag{A8}$$

This migration rate corresponds to a bifurcation point of $\mathbf{I}_{\mathrm{DMI}}$ and $\mathbf{I}_0$ if and only if $\mathbf{I}_{\mathrm{DMI}}$ does not leave the state space if $m < m^{*}$. Since $\mathbf{I}_{\mathrm{DMI}}$ can leave the state space only at $m = m^{-}_{\max}$, the boundaries of the regime with a bifurcation of $\mathbf{I}_{\mathrm{DMI}}$ and $\mathbf{I}_0$ are obtained by equating m^{*} (Equation A8) with $m^{-}_{\max}$ (Equation A7). At this boundary, the recombination rate r can take one of the following values:

$$r_{\mathcal{A}} = (\gamma-\alpha)\frac{3(\gamma-\beta)-\alpha}{2\gamma-\alpha}, \tag{A9}$$

$$r_{\mathcal{B}} = \beta\ \frac{3\alpha+\beta-\gamma}{\beta+\gamma}, \tag{A10}$$

$$r_2 = \frac{3\alpha(\gamma-\beta)-\sqrt{\alpha(\gamma-\beta)(4\gamma\beta+5\alpha\gamma-9\alpha\beta)}}{2\gamma}. \tag{A11}$$

We further define

$$r_2^{*} = \begin{cases} r_2 & \text{if } r_2 \text{ is real,} \\ \min[\alpha, \gamma-\beta] & \text{otherwise.} \end{cases} \tag{A12}$$

To provide expressions for the upper bound $m^{+}_{\max}$ below which $\mathbf{I}_{\mathrm{DMI}}$ exists, we need to distinguish several cases (*cf.* Theorem S.5 in File S1). They correspond to the mechanisms maintaining or establishing a DMI that result from the different types of fitness landscapes discussed in the main text.

1. Selection scenario 1: $0 < \alpha < \beta < \gamma$ (double-peak landscape with maximum at the continental type). This selection scenario represents the parameter regime in which "selection against hybrids" is driving DMI evolution (area with dark shading in Figure 2). With this fitness landscape, a globally stable internal equilibrium cannot exist; *i.e.*,

   $$m^{-}_{\max} = 0. \tag{A13}$$

   If $r \leq \beta - \alpha$, there is no internal equilibrium and the continental equilibrium $\mathbf{M}_2$ is globally stable for any m. If $r > \beta - \alpha$, $\mathbf{I}_{\mathrm{DMI}}$ exists if $m \leq m^{*}$,

   $$m^{+}_{\max} = \begin{cases} 0 & \left(=m^{-}_{\max}\right) & \text{if} & r \leq \beta - \alpha, \\ m^{*} & \left(>m^{-}_{\max}\right) & \text{if} & r > \beta - \alpha. \end{cases} \tag{A14}$$

If $m > m_{\max}^{+}$, $\mathbf{M}_2$ is globally stable.

2. Selection scenario 2: $0 < \alpha < \beta < \gamma$ (double-peak landscape with maximum at the island type). This selection scenario represents the intermediate parameter regime (open area in Figure 2) in which both mechanisms interfere. In this case, $\mathbf{I}_{\text{DMI}}$ exists always for small m. If $r < \alpha - \beta$, it is globally stable for m up to

   $$m_{\max}^{-} = m_2. \tag{A15}$$

   Thus, $\mathbf{I}_{\text{DMI}}$ exists and is stable up to

   $$m_{\max}^{+} = \begin{cases} m_{\max}^{-} & \text{if } r \leq r_2, \\ m^* \quad (> m_{\max}^{-}) & \text{if } r > r_2. \end{cases} \tag{A16}$$

   If $r < r_2$, the monomorphic equilibrium $\mathbf{M}_2$ is globally stable for $m \geq m_{\max}^{-} = m_{\max}^{+}$. If $r_2 < r < \alpha - \beta$, $\mathbf{M}_2$ becomes locally stable at $m = m_{\max}^{-}$ (where $\mathbf{I}_0$ enters the state space) and globally stable at $m = m_{\max}^{+} = m^*$ (where $\mathbf{I}_{\text{DMI}}$ and $\mathbf{I}_0$ merge). If $r \geq \alpha - \beta$, $\mathbf{M}_2$ is already locally stable at $m = 0$ and becomes the unique globally stable equilibrium at $m = m_{\max}^{+} = m^*$.
3. Selection scenario 3: $\gamma - \alpha < 0 < \gamma - \beta < \alpha$ (slope-type landscape with weak DMI $\gamma < \alpha$). This selection scenario represents a part of the parameter regime in which "selection against immigrants" drives DMI evolution. In Figure 2, it is represented by the area with light shading for $0 < \beta/\alpha < 1$. With this landscape, $\mathbf{I}_{\text{DMI}}$ is globally stable for small m, and we have

   $$m_{\max}^{-} = \begin{cases} m_2 & \text{if} \quad r \leq \gamma - \beta, \\ m_{\mathcal{A}} & \text{if} \quad r > \gamma - \beta. \end{cases} \tag{A17}$$

   For the existence of $\mathbf{I}_{\text{DMI}}$, we need to distinguish two subcases:
   i. $\gamma > \alpha/2$:

   $$m_{\max}^{+} = \begin{cases} m_{\max}^{-} & \text{if} \quad r \leq \min[\gamma - \beta,\ r_2^*] \ \text{ or } \ \gamma - \beta < r \leq r_{\mathcal{A}}, \\ m^* \quad (> m_{\max}^{-}) & \text{if} \quad r_2^* < r \leq \gamma - \beta \ \text{ or } \ r > \max[\gamma - \beta,\ r_{\mathcal{A}}]. \end{cases} \tag{A18}$$

   ii. $\gamma < \alpha/2$:

   $$m_{\max}^{+} = \begin{cases} m_{\max}^{-} & \text{if} \quad r \leq r_2^* \quad \text{or} \quad r \geq r_{\mathcal{A}}, \\ m^* \quad (> m_{\max}^{-}) & \text{if} \quad r_2^* < r < r_{\mathcal{A}}. \end{cases} \tag{A19}$$
4. Selection scenario 4: $\beta < 0 < \alpha < \gamma - \beta$ (slope-type with local adaptation). This selection scenario represents the remaining part of the parameter regime in which selection against immigrants drives DMI evolution. In Figure 2 it is depicted by the area with light shading for $\beta\ /\ \alpha < 0$. With this landscape, $\mathbf{I}_{\text{DMI}}$ is globally stable for small m, and we have

   $$m_{\max}^{-} = \begin{cases} m_2 & \text{if} \quad r \leq \alpha, \\ m_{\mathcal{B}} & \text{if} \quad r > \alpha. \end{cases} \tag{A20}$$

   For the existence of $\mathbf{I}_{\text{DMI}}$, we distinguish two subcases:
   i. $\gamma + \beta > 0$:

   $$m_{\max}^{+} = \begin{cases} m_{\max}^{-} & \text{if} \quad r \leq \min[\alpha, r_2^*] \quad \text{or} \quad \alpha < r \leq r_{\mathcal{B}}, \\ m^* \quad (> m_{\max}^{-}) & \text{if} \quad r_2^* < r \leq \alpha \quad \text{or} \quad r > \max[\alpha, r_{\mathcal{B}}]. \end{cases} \tag{A21}$$

   ii. $\gamma + \beta < 0$:

   $$m_{\max}^{+} = \begin{cases} m_{\max}^{-} & \text{if} \quad r \leq r_2^* \quad \text{or} \quad r \geq r_{\mathcal{B}}, \\ m^* \quad (> m_{\max}^{-}) & \text{if} \quad r_2^* < r < r_{\mathcal{B}}. \end{cases} \tag{A22}$$

Initial conditions

The above results show that for sufficiently large values of r, we have $m_{\max}^{+} > m_{\max}^{-}$ in parts of the parameter space. This leads to bistable regimes, where maintenance of a DMI depends on the initial conditions. Depending on the order of substitution events in the evolutionary history of the two populations, in the *Models* section of the main text we have outlined five scenarios for the evolution of a parapatric DMI. Here, we show how they correspond to three different initial conditions that are illustrated in Figure A1.

1. For both the continent–island and island–island scenarios, the island allele A enters the island only after the allele frequencies at the locus $\mathcal{B}$ have reached equilibrium. Locus $\mathcal{B}$ is initially at the polymorphic $\mathbf{S}_{\mathcal{B}}$ equilibrium, whenever this exists (*i.e.*, if $m \leq -\beta$). Otherwise, the population starts in the vicinity of the monomorphic equilibrium $\mathbf{M}_2$. Here, we assume that the ancestral b allele is not completely lost, or reintroduced at very low frequencies, such that the dynamics are not artificially stuck to the $q = 1$ boundary if this boundary is unstable.
2. For the island–continent and continent–continent scenarios, A mutants are present on the island while the continent is still fixed for the ab haplotype. A stable single-locus polymorphism with coordinates $\boldsymbol{p}_{\mathcal{A}} : (p, q, D) = (1 - m/\alpha, 0, 0)$ results whenever the island allele can invade, which is the case if $m < \alpha$. Otherwise, the population is initially in the vicinity of the monomorphic equilibrium $\mathbf{M}_1$ at which the ancestral haplotype is fixed.
3. For secondary contact, the island population is, by definition, initially fixed for the island haplotype, *i.e.*, at $\mathbf{M}_3$.

For the case of independent loci, we show in File S1 (section S.8.2) that, in a bistable regime, a locally stable DMI will always evolve under the scenarios corresponding to the second or the third initial condition (island–continent, continent–continent, and secondary contact), but never under the scenarios corresponding to the first initial condition (island–island and continent–island). This holds independently of the size of the domain of attraction of the DMI. Extensive numerical calculations show that this finding extends to the general model with linkage.

GENETICS

Supporting Information
http://www.genetics.org/content/suppl/2012/04/27/genetics.111.137513.DC1

The Limits to Parapatric Speciation: Dobzhansky–Muller Incompatibilities in a Continent–Island Model

Claudia Bank, Reinhard Bürger, and Joachim Hermisson

DOI: 10.1534/genetics.111.137513

File S1: Analysis of the haploid model

Here, we derive the possible equilibrium and bifurcation structures of the haploid model (Section *The haploid model*). Our main results are formulated and proved in Section S.7. The supplementary *Mathematica* notebook will be useful in checking several of the results presented below. It also contains interactive gadgets to visualize these results.

The results derived in this Online Supplement are not merely auxiliary to those presented in the main text, but they complement and extend previous studies of two-locus haploid models. Theorem S.2 characterizes the equilibrium structure of the diallelic haploid two-locus model with arbitrary selection (and no migration) for the cases that are not covered by the analyses of Feldman (1971) and Rutschman (1994). These are the cases when an internal equilibrium exists and two boundary equilibria are asymptotically stable. However, in contrast to their analyses, ours is performed for the continuous-time model. This not only facilitates some computations but in particular enables the application of the index theory of continuous dynamical systems. Index theorems are a powerful tool that, to our knowledge, has not been used in the population genetic literature so far. Except for non-generic, degenerate cases, Theorems S.4 and S.5 characterize all equilibrium and bifurcation patterns for the haploid two-locus continent-island model with negative epistatic selection. In particular, they extend the main mathematical result of Bürger and Akerman (2011), who derived these patterns for nonepistatic selection.

For easier reference, we state the basic differential equations describing the haplotype dynamics:

$$\dot{x}_1 = x_1[-\alpha(x_3+x_4) - \beta(x_2+x_4) + \gamma x_4] - rD - mx_1\,, \tag{S.1a}$$
$$\dot{x}_2 = x_2[-\alpha(x_3+x_4) + \beta(x_1+x_3) + \gamma x_4] + rD + m(1-x_2)\,, \tag{S.1b}$$
$$\dot{x}_3 = x_3[\alpha(x_1+x_2) - \beta(x_2+x_4) + \gamma x_4] + rD - mx_3\,, \tag{S.1c}$$
$$\dot{x}_4 = x_4[\alpha(x_1+x_2) + \beta(x_1+x_3) - \gamma(x_1+x_2+x_3)] - rD - mx_4\,. \tag{S.1d}$$

This is a dynamical system on the simplex $S_4 = \{(x_1,x_2,x_3,x_4) : x_i \geq 0\ \forall i \text{ and } \sum_i x_i = 1\}$ which constitutes our state space. We always assume $m \geq 0$, $r \geq 0$, and $\gamma \geq 0$.

For many purposes, it will be convenient to describe the dynamics in terms of the allele frequencies $p_A = x_3 + x_4$, $p_B = x_2 + x_4$, and the measure D of linkage disequilibrium (LD). Deviating from the main text, we use the notation $p = p_A$ and $q = p_B$ throughout this Online Supplement. Then the dynamical equations read

$$\begin{aligned}\dot{p} &= \alpha p(1-p) - \gamma(1-p)(pq+D) + \beta D - mp\,, && \text{(S.2a)}\\ \dot{q} &= \beta q(1-q) - \gamma(1-q)(pq+D) + \alpha D + m(1-q)\,, && \text{(S.2b)}\\ \dot{D} &= [\alpha(1-2p) + \beta(1-2q)]D - \gamma[(1-p)(1-q) - D](pq+D) \\ &\quad - rD - m[p(1-q) + D]\,. && \text{(S.2c)}\end{aligned}$$

We recall that D has to satisfy

$$-\min[pq, (1-p)(1-q)] \le D \le \min[p(1-q), (1-p)q]\,. \tag{S.3}$$

S.1 Existence and linear stability of boundary equilibria

We denote the monomorphic equilibria $x_i = 1$ by $\mathbf{M}_i$. If $m = 0$, then all monomorphic equilibria exist. However, if $\alpha > 0$ and $\gamma > \beta$, the conditions most relevant for this investigation (see (S.24) below), $\mathbf{M}_1$ and $\mathbf{M}_4$ are always unstable.

If $m > 0$, then only $\mathbf{M}_2$ (fixation of the continental type) is an equilibrium. The eigenvalues of (the Jacobian of (S.1) at) $\mathbf{M}_2$ are

$$-\beta - m\,,\ \alpha - \gamma - m\,,\ \alpha - \beta - m - r\,. \tag{S.4}$$

Hence, $\mathbf{M}_2$ is asymptotically stable if

$$m > \max[-\beta, \alpha - \gamma, \alpha - \beta - r]\,, \tag{S.5}$$

i.e., if (A.3) holds.

If $m = 0$, the eigenvalues of $\mathbf{M}_3$ are

$$-\alpha\,,\ \beta - \gamma\,,\ -\alpha + \beta - r\,. \tag{S.6}$$

Hence, $\mathbf{M}_3$ is asymptotically stable if $\alpha > 0$, $\gamma > \beta$, and $r > \beta - \alpha$.

Next, there may exist two equilibria at which one locus is polymorphic and the other is fixed. The equilibrium $\mathbf{S}_{\mathcal{A}}$ has the coordinates $(p, q, D) = \left(1 - \frac{m}{\alpha-\gamma}, 1, 0\right)$ and is admissible if and only if $m < \alpha - \gamma$. Its eigenvalues are

$$-\alpha + \gamma + m\,,\ \frac{1}{2}\left[-(\alpha + r) + 2(\gamma - \beta) \pm \sqrt{(\alpha + r)^2 + \frac{4\alpha r m}{\gamma - \alpha}}\right]\,. \tag{S.7}$$

Hence, $\mathbf{S}_{\mathcal{A}}$ is asymptotically stable if

$$\gamma < \alpha + \beta \quad \text{and} \quad \frac{(\alpha - \gamma)(\gamma - \beta)}{\alpha}\left(1 + \frac{\alpha + \beta - \gamma}{r}\right) < m < \alpha - \gamma\,, \tag{S.8}$$

which requires $r > \gamma - \beta$.

The equilibrium $\mathbf{S}_{\mathcal{B}}$ has coordinates the $\left(0, -\frac{m}{\beta}, 0\right)$ and is admissible if and only if $m < -\beta$. Its eigenvalues are

$$\beta + m\,,\ \frac{1}{2}\left[-(\gamma - \beta + r) + 2\alpha \pm \sqrt{(\gamma - \beta + r)^2 + \frac{4(\gamma - \beta)rm}{\beta}}\,\right]. \tag{S.9}$$

Hence, $\mathbf{S}_{\mathcal{B}}$ is asymptotically stable if

$$\gamma > \alpha + \beta \quad \text{and} \quad \frac{-\beta\alpha}{\gamma - \beta}\left(1 + \frac{\gamma - \beta - \alpha}{r}\right) < m < -\beta\,, \tag{S.10}$$

which requires $r > \alpha$.

In the following lemma, we collect a few simple but important observations that follow from the above analysis.

Lemma S.1. *1. For given $m > 0$, at most one of the boundary equilibria $\mathbf{M}_2$, $\mathbf{S}_{\mathcal{A}}$, or $\mathbf{S}_{\mathcal{B}}$ can be stable. Their asymptotic stability conditions are given by* (S.5), (S.8), *or* (S.10), *respectively.*

2. $\mathbf{M}_2$ is asymptotically stable if m is sufficiently large. If $\mathbf{M}_2$ is asymptotically stable, then $\mathbf{S}_{\mathcal{A}}$ and $\mathbf{S}_{\mathcal{B}}$ are not admissible.

3. As a function of m, boundary equilibria change stability at most once. If a change in stability occurs, then it is from unstable to stable (as m increases).

Finally, if $r = 0$, there is a fully polymorphic equilibrium $\mathbf{R}_0$ on the edge $x_1 = x_4 = 0$ of S_4. Thus, only the island and the continental haplotypes are present. It satisfies $p = x_3$, $q = x_2$, $D = -pq$, and $p + q = 1$. The coordinates (p, q, D) of $\mathbf{R}_0$ are

$$\left(1 - \frac{m}{\alpha - \beta}, \frac{m}{\alpha - \beta}, \frac{m^2}{(\alpha - \beta)^2} - \frac{m}{\alpha - \beta}\right). \tag{S.11}$$

The eigenvalues are

$$-\alpha\,,\ \beta - \gamma\,,\ -\alpha + \beta + m\,. \tag{S.12}$$

This equilibrium exists and is asymptotically stable if and only if

$$m < \alpha - \beta \tag{S.13}$$

holds; cf. (8).

S.2 Global stability properties of boundary equilibria

In the following we prove global asymptotic stability of boundary equilibria for various sets of parameters by applying the theory of Lyapunov functions (e.g. LaSalle 1976, in particular, Theorem 6.4 and Corollary 6.5). By global asymptotic stability of an equilibrium we mean that every trajectory, such that initially all alleles are present, converges to this equilibrium. By (S.5), (S.8), and (S.10), there is at most one asymptotically stable boundary equilibrium for any given set of parameters. Hence, convergence of all trajectories to the boundary is sufficient for demonstrating global stability. Because global convergence to the boundary precludes the existence of an internal equilibrium, these results yield the necessary conditions for a stable DMI (i.e., a stable internal equilibrium) in the section *Necessary conditions for stable DMIs.*

S.2.1 Proof of the necessary conditions $\alpha > 0$ and $\beta < \gamma$

We define

$$Y = \frac{x_1 + x_3}{x_3 + x_4} = \frac{1-q}{p}\,, \tag{S.14}$$

where we assume $x_3 + x_4 > 0$. Differentiating Y with respect to t and using (S.1), we obtain

$$\dot{Y} = -\frac{(\beta - \gamma)x_4(x_1 + x_3) + \alpha x_1(x_3 + x_4)}{(x_3 + x_4)^2}\,. \tag{S.15}$$

We infer immediately that

$$\begin{aligned} \dot{Y} &\le 0 \quad \text{if} \quad \alpha > 0 \text{ and } \beta \ge \gamma\,, && \text{(S.16a)} \\ \dot{Y} &\ge 0 \quad \text{if} \quad \alpha \le 0 \text{ and } \beta < \gamma\,, && \text{(S.16b)} \end{aligned}$$

and the inequalities for $\dot{Y}$ are strict in the interior of S_4. Therefore, $q(t) \to 1$ as $t \to \infty$ if (S.16a) applies. It is an easy, but not necessary, exercise to show that $\mathbf{S}_{\mathcal{A}}$ is globally asymptotically stable if $\alpha > \gamma$ and $\mathbf{M}_2$ is globally asymptotically stable if $\alpha \le \gamma$ (see also below). If (S.16b) applies, then $p(t) \to 0$, and either $\mathbf{M}_2$ (if (S.5) holds) or $\mathbf{S}_{\mathcal{B}}$ (otherwise) is globally asymptotically stable.

Next, we define

$$X = \frac{x_1 + x_3}{x_1 + x_2} = \frac{1-q}{1-p}\,, \tag{S.17}$$

where $x_1 + x_2 > 0$, and obtain

$$\dot{X} = -\frac{m(x_1 + x_3) + \beta x_2(x_1 + x_3) - \alpha x_3(x_1 + x_2)}{(x_1 + x_2)^2}\,. \tag{S.18}$$

We conclude that

$$\dot{X} \leq 0 \quad \text{if} \quad \alpha < 0 \text{ and } \beta \geq 0\,, \tag{S.19}$$

and $\dot{X} < 0$ in the interior of S_4. Therefore, $q(t) \to 1$ as $t \to \infty$. Combining (S.16a), (S.16b), and (S.19) completes the proof that (2), i.e., $\alpha > 0$ and $\gamma > \beta$, is a necessary condition for the existence of a DMI.

S.2.2 Gobal convergence to $\mathbf{M}_2$

Assume $\alpha > 0$ and $\gamma > \beta$. Because

$$\begin{aligned}\dot{x}_2 &= x_2[\beta(x_1+x_3)+\gamma x_4-\alpha(x_3+x_4)-rx_3]+m(x_1+x_3+x_4)+rx_1x_4\\ &\geq x_2[(m+\beta)x_1+(m+\beta-\alpha-r)x_3+(m-\alpha+\gamma)x_4]\,,\end{aligned} \tag{S.20}$$

global convergence to $\mathbf{M}_2$ follows at once if $m > \max[-\beta, \alpha-\gamma, \alpha-\beta+r] = \alpha-\beta+r$. This implies (3) and (4a). Hence, $\mathbf{M}_2$ is globally asymptotically stable for every m if $r < \beta - \alpha$.

S.2.3 Proof of (4b) *and* (4c)

Assume $\alpha > 0$ and $\gamma > \beta$. We first prove that $m \geq \max[\alpha-\beta, \frac{1}{4}\alpha]$ implies $\dot{X} \leq 0$, i.e., $q(t) \to 1$ as $t \to \infty$. Indeed, the denominator of $-\dot{X}$ in (S.18) satisfies

$$\begin{aligned}&m(x_1+x_3)+\beta x_2(x_1+x_3)-\alpha x_3(x_1+x_2)\\ &\quad= mx_1(x_1+x_3+x_4)+mx_3(x_1+x_3+x_4)-\alpha x_1x_3+x_1x_2(m+\beta)+x_2x_3(m+\beta-\alpha)\\ &\quad\geq m(x_1+x_3)^2-\alpha x_1x_3+x_1x_2(m+\beta)+x_2x_3(m+\beta-\alpha)\geq 0\end{aligned} \tag{S.21}$$

if $m \geq \max[\alpha-\beta, \frac{1}{4}\alpha, -\beta]$. This proves $m^{+}_{\text{max}} \leq \max[\alpha-\beta, \frac{1}{4}\alpha]$ because $\alpha-\beta > -\beta$. The other inequality, (4c), follows analogously by using the Lyapunov function $(x_3+x_4)/(x_2+x_4)$ or by employing the model symmetry noted below (A.6), i.e., $\alpha \to \gamma-\beta$ and $\beta \to \gamma-\alpha$.

S.2.4 Internal equilibria exhibit negative linkage disequilibrium

We prove that every trajectory eventually enters the region $D \leq 0$ and remains there. Convergence to $D = 0$ occurs if and only if at least one allele is eventually lost. Thus, every internal equilibrium satisfies $D < 0$.

To prove these statements, we define

$$Z = \frac{x_2 x_3}{x_1 x_4}\,, \tag{S.22}$$

where $x_1 > 0$ and $x_4 > 0$ is assumed. We note that $Z = 1$ if and only if $D = 0$, and $Z < 1$ if and only if $D > 0$. Then

$$\dot{Z} = x_1 x_3 x_4 (m + \gamma x_2) + rD(x_1 x_2 x_3 + x_2 x_3 x_4 + x_1 x_2 x_4 + x_1 x_3 x_4)\,. \tag{S.23}$$

We observe that $\dot{Z} \geq 0$ holds whenever $D \geq 0$. In addition, it follows immediately that $\dot{Z} > 0$ if $rD > 0$ and $x_2 + x_3 > 0$. If $x_2 + x_3 = 0$ and $x_1 x_4 > 0$, then $\dot{x}_1 + \dot{x}_4 < 0$ if $r + m > 0$. Hence, all trajectories leave $D > 0$ if $r > 0$. If $rD = 0$, then $\dot{Z} = 0$ only if $x_3 = 0$ or if $m = 0$ and $\gamma x_2 = 0$. Thus, our result follows by investigating (i) the dynamics on $x_3 = 0$ if $r = 0$, (ii) the dynamics on $x_2 = 0$ if $m = r = 0$, and (iii) the case $m = \gamma = 0$. We leave the simple first two cases to the reader. The third case is also not difficult and follows immediately from Section 3.4.1 in Bürger and Akerman (2011)).

We conclude that at any equilibrium satisfying $D = 0$, one or both loci are fixed.

For the rest of this Online Supplement, we assume

$$\alpha > 0 \text{ and } \gamma > \beta \text{ and } r > \beta - \alpha\,, \tag{S.24}$$

because we have proved that internal equilibria can exist only if (S.24) is satisfied. We note that (S.24) holds if and only if $\mathbf{M}_3$ (island haplotype fixed) is linearly stable in the absence of migration; cf. (S.6).

S.3 Calculating the internal equilibria

We derive a cubic equation from which the coordinate p of an internal equilibrium (p, q, D) can be obtained. Given p, the coordinates q and D can be computed from relative simple explicit formulas. Using these results, we show that at most three internal equilibria can exist.

By solving $\dot{p} = 0$, we find that, for given p and q, and if $p \neq 1 - \beta/\gamma$, the value of LD at equilibrium is

$$D = D(p, q) = p\,\frac{m + (1-p)(\gamma q - \alpha)}{\beta - \gamma + \gamma p}\,. \tag{S.25}$$

Substituting this into (S.2b), assuming $\beta \neq 0$, and solving $\dot{q} = 0$ for q, we obtain

$$q_{1,2}(p) = \frac{1}{2}\left[\left(1 - \frac{m}{\beta}\right) \pm \sqrt{Q}\right] \tag{S.26}$$

where

$$Q = \left(1 + \frac{m}{\beta}\right)^2 - \frac{4\alpha m p}{\beta(\gamma - \beta)} - \frac{4\alpha(\gamma - \alpha)}{\beta(\gamma - \beta)} p(1-p) \tag{S.27}$$

needs to be nonnegative to yield an admissible equilibrium. Finally, we substitute $q = q_1(p)$ and $D = D(p, q_1(p))$ into (S.2c) and obtain that any equilibrium value p must be a solution of the equation

$$(\gamma - \beta)A(p) - \sqrt{Q}B(p) = 0\,, \tag{S.28}$$

where

$$\begin{aligned}
A(p) = {} & (\gamma-\beta)\left\{\beta[(\alpha-r)(\gamma-2\alpha)+\gamma(\alpha-\beta)] + [\beta(\gamma+\beta-2\alpha)+r(2\beta-\gamma)]m + \beta m^2\right\} \\
& + \big\{\beta\,[2\alpha(2\gamma-3\beta)(\gamma-\alpha)-\beta\gamma(\gamma-\beta)] - (2\gamma-\beta)(2\alpha-\gamma)r \\
& \quad + [2\alpha\beta(\gamma-2\beta)-\beta\gamma(\gamma-\beta)+\gamma(2\gamma-3\beta)r]m\big\}p \\
& - [2\alpha\beta(\gamma-\alpha)(\gamma-2\beta)+\beta\gamma(\gamma-2\alpha)r+\gamma^2 rm]p^2\,,
\end{aligned} \tag{S.29a}$$

$$\begin{aligned}
B(p) = {} & (\gamma-\beta)[-2\alpha\beta+\gamma(\beta+r)+\beta m] \\
& + [2\alpha\beta(\gamma-\beta)-\beta\gamma(\gamma-\beta)+r(\beta-2\gamma)]p + \gamma^2 r p^2\,.
\end{aligned} \tag{S.29b}$$

If we substitute $q = q_2(p)$ and $D = D(p, q_2(p))$ into (S.2c), we obtain

$$(\gamma - \beta)A(p) + \sqrt{Q}B(p) = 0\,, \tag{S.30}$$

instead of (S.28). If $A(p) = B(p) = 0$, then for given α, β, γ, and r, only one value of m can give rise to an internal equilibrium. We ignore this case here, but will encounter it further below. Otherwise, a solution p of (S.28) cannot be a solution of (S.30), and vice versa. Hence, for a solution p, only one of $q_1(p)$ or $q_2(p)$ can give rise to an admissible internal equilibrium. In fact,

$$\begin{gathered}
\text{if } A(p) \text{ and } B(p) \text{ have the same sign, only } q_1(p) \text{ can be admissible;} \\
\text{if } A(p) \text{ and } B(p) \text{ have opposite signs, only } q_2(p) \text{ can be admissible.}
\end{gathered} \tag{S.31}$$

Therefore, solutions p of (S.28) or (S.30) satisfy

$$\begin{aligned}
0 &= (\gamma-\beta)^2 A(p)^2 - QB(p)^2 \\
&= \frac{4\beta}{\gamma-\beta}(\beta-\gamma+\gamma p)^2[m+(\gamma-\alpha)(1-p)]P(p)\,,
\end{aligned} \tag{S.32}$$

where

$$\begin{aligned}
P(p) = {} & (\gamma-\beta)(m+r+\beta-\alpha)[\alpha\beta(\alpha+\beta-\gamma-r)-mr(\gamma-\beta)] \\
& + \big\{\alpha\beta(\alpha-\beta)(\gamma-\beta)(\alpha+\beta-\gamma)+\alpha(\gamma-\beta)[3\beta(\gamma-2\alpha)+m(4\beta-\gamma)]r \\
& \quad + [\alpha(\gamma^2+\beta\gamma-\beta^2)+\gamma(\gamma-\beta)m]r^2\big\}p \\
& - 2\alpha r[\beta(\gamma-\beta)(\gamma-2\alpha)+\gamma^2 r]p^2 + \alpha\gamma^2 r^2 p^3\,.
\end{aligned} \tag{S.33}$$

Because $p = 1 - \beta/\gamma$ never gives an equilibrium of (S.2) and $p = 1 - m/(\gamma - \alpha)$ can give rise only to a single-locus polymorphism, any internal equilibrium value p must satisfy $P(p) = 0$. It is easy to find parameter values such that P has three zeros in $(0,1)$ (e.g., $\alpha = 1$, $\beta = 1.5$, $\gamma = 4$, $r = 2$, $m = 0.4$). Apparently, only two can give rise to an equilibrium (and this will be proved further below).

The case $\beta = 0$ can be treated separately and is much simpler because $q(p)$ is uniquely determined. In particular, the value p of an internal equilibrium has to satisfy $P(p) = 0$. If $p = 1 - \frac{\beta}{\gamma}$, we necessarily have $r = \frac{\gamma(\alpha-\beta)(\alpha+\beta-\gamma)}{(\gamma-2\alpha)(\gamma-2\beta)}$. Then $q = \alpha/\gamma$ is uniquely determined and $D = -\frac{\alpha(\gamma-\alpha)(\alpha+\beta-\gamma)}{(\gamma-2\alpha)(\gamma-2\beta)}$. This equilibrium may be admissible or not.

Remark S.1. In the absence of epistasis ($\gamma = 0$) there can be at most two internal equilibria. Their coordinates are obtained from a quadratic equation in p. The admissibility conditions are given by simple formulas (Bürger and Akerman 2011).

We can summarize these findings as follows.

Theorem S.1. *The haploid dynamics* (S.2) *can have at most three internal equilibria. The coordinate* $\hat{p}$ *of an internal equilibrium* $(\hat{p}, \hat{q}, \hat{D})$ *is a zero of the polynomial* P *given by* (S.33)*. For given* $\hat{p}$ *with* $P(\hat{p}) = 0$*, only one of* $q_1(\hat{p})$ *or* $q_2(\hat{p})$ (S.26) *can yield an equilibrium value* $\hat{q}$*.* $\hat{D}$ *is calculated from* $\hat{p}$ *and* $\hat{q}$ *by* (S.25)*. This procedure yields an internal equilibrium if and only if* $0 < \hat{p} < 1$*,* $0 < \hat{q} < 1$*, and* $-\min[\hat{p}\hat{q}, (1-\hat{p})(1-\hat{q})] \le \hat{D} < 0$ *hold. If* $\gamma = 0$*, there are at most two internal equilibria.*

For several important limiting cases, explicit expressions for the internal equilibria are obtained below (see Sections S.6 and S.8).

S.4 Critical m at which two internal equilibria may bifurcate

A bifurcation of two internal equilibria can occur if and only if $P(p^*) = 0$, where $p^* \in (0,1)$ is a critical point of P, i.e., $P'(p^*) = 0$. There are at most two such critical points, and they are given by

$$p^*_{1,2} = \frac{1}{3\alpha\gamma^2 r}\left\{2\alpha[\beta(\gamma-2\alpha)(\gamma-\beta) + \gamma^2 r] \pm \sqrt{R}\right\}, \tag{S.34a}$$

where

$$\begin{aligned} R = \alpha^2\big\{&-\beta(\gamma-\beta)\left[16\alpha\beta(\gamma-\alpha)(\gamma-\beta) + \gamma^2(3\alpha^2+\beta^2) - \gamma^3(3\alpha+\beta)\right] \\ &- \beta\gamma^2(\gamma-\beta)(\gamma-2\alpha)r + \gamma^2(3\beta^2 - 3\beta\gamma + \gamma^2)r^2\big\} \\ &- 3mr\alpha\gamma^2(\gamma-\beta)[4\alpha\beta + \gamma(r-\alpha)] \end{aligned} \tag{S.34b}$$

Solving either $P(p_1^*) = 0$ or $P(p_2^*) = 0$ for m, we obtain after some straightforward manipulations that the critical value m^* must be a solution of the following quartic equation:

$$[\alpha\beta(2\beta-\gamma)(\alpha+\beta-\gamma+r)+\gamma(\gamma-\beta)rm]^2[-\psi_1+2\psi_2 rm+27\alpha\gamma^2(\gamma-\beta)r^2m^2]=0\,, \tag{S.35a}$$

where

$$\psi_1 = \alpha(\gamma-\beta)(\alpha-\beta+r)^2[4\alpha\beta(\gamma-\alpha)(\gamma-\beta)+\gamma^2r^2]\,, \tag{S.35b}$$

$$\begin{aligned}\psi_2 &= 2\alpha^2(\gamma-\beta)^2(9\beta\gamma-8\alpha\beta-\alpha\gamma)+3\alpha\gamma(\gamma-\beta)(3\beta\gamma-2\alpha\beta-\alpha\gamma)r\\ &\quad+3\alpha\gamma^2(\gamma-\beta)r^2+2\gamma^3r^3\,.\end{aligned} \tag{S.35c}$$

The zero $m = m_0$ arising from the first (linear) factor in (S.35a) does not give a valid bifurcation point for internal equilibria for the following reason. The corresponding critical value p_0 (which, depending on the parameters, can be p_1^* or p_2^*) is a double solution of $P(p)$ and has the property that $A(p_0) = B(p_0) = 0$. However, $P(p)$ was derived from (S.28) and (S.30) by squaring. Differentiating $(\gamma-\beta)A(p) \pm \sqrt{Q}B(p)$ with respect to p and evaluating at (p_0, m_0), we infer from the Implicit Function Theorem that p is uniquely determined by m in a small neighborhood of m_0 unless r assumes one of two specific values. One of them leads to boundary equilibrium (thus not to an internal bifurcation). If we denote the other by r_0 ($= \beta[\alpha(\gamma-4\beta)+\beta(\gamma-\beta)]/[\gamma(3\beta-\gamma])$, then $p_0 = p_1 = p_2$ holds at m_0. Although there is a bifurcation of three values p, only one of the resulting three branches gives rise to an admissible equilibrium.

The second (quadratic) factor in (S.35a) provides two potential solutions. However, because $\psi_1 \geq 0$, one is negative. Therefore, the critical value we are looking for is given by

$$\begin{aligned} m^* &= \frac{1}{27\alpha\gamma^2(\gamma-\beta)r}\left(-\psi_2+\sqrt{\psi_2^2+27\alpha\gamma^2(\gamma-\beta)\psi_1}\right)\\ &= \frac{1}{27\alpha\gamma^2(\gamma-\beta)r}\left\{-\psi_2+2\left[\alpha^2(\gamma-\beta)^2+3\alpha\beta(\gamma-\beta)(\gamma-\alpha)+\alpha\gamma(\gamma-\beta)r+\gamma^2r^2\right]^{3/2}\right\}.\end{aligned} \tag{S.36}$$

At this value, two equilibria with non-zero allele frequencies collide and annihilate each other. Thus, m^* is the critical value at which a saddle-node bifurcation occurs. This gives an admissible bifurcation if both equilibria are internal (hence admissible) for either $m < m^*$ or $m > m^*$.

If $p_1^* = p_2^*$, i.e., if $R = 0$, a pitchfork bifurcation could occur at m^*. As a function of α, β, and γ, the condition $p_1^* = p_2^*$ can be satisfied at m^* only for three different values of r, of which at most two can be positive. It can be shown that at each of these values, one of the emerging zeros of $P(p)$ does not give rise to an admissible equilibrium (because $D > 0$ there). Thus, only a saddle-node bifurcation can occur.

We obtain the following series expansions for m^*. If $\beta < 0$ and γ is small, then

$$m^* = \frac{(\alpha-\beta+r)^2}{8r} + \frac{\gamma(\alpha-\beta-3r)(\alpha-\beta+r)^2}{64\alpha\beta r} + O(\gamma^2)\,. \tag{S.37}$$

If $\gamma = 0$, this reduces to the expression given in Bürger and Akerman (2011).

If r is large, the following quasi-linkage-equilibrium (QLE) approximation is obtained:

$$m^* = \frac{\alpha(\gamma-\beta)}{4\gamma}\left(1 - \frac{4\beta\gamma - 3\alpha\beta - \alpha\gamma}{2\gamma r}\right) + O\left(\frac{1}{r^2}\right)\,. \tag{S.38}$$

S.5 The haploid two-locus selection model in the absence of migration

We assume $m = 0$. From Section S.3, we obtain the following properties of internal equilibria (p, q, D). The LD is given by

$$D = D(p,q) = p(1-p)\frac{\gamma q - \alpha}{\beta - \gamma + \gamma p}\,, \tag{S.39}$$

where $p \neq 1 - \beta/\gamma$; cf. (S.25). For admissibility, we need

$$\max[-pq, -(1-p)(1-q)] < D < 0\,. \tag{S.40}$$

For given p and if $\beta \neq 0$, the coordinate q of an internal equilibrium can assume only one of the following forms:

$$q_{1,2}(p) = \frac{1}{2}\left(1 \pm \sqrt{1 - \frac{4\alpha(\gamma-\alpha)p(1-p)}{\beta(\gamma-\beta)}}\right)\,. \tag{S.41}$$

By Theorem S.1, for given p, at most one of $q_1 = q_1(p)$ or $q_2 = q_2(p)$ can give rise to an equilibrium.

Our goal here is to prove the following theorem:

Theorem S.2. *Suppose* (S.24) *and* $m = 0$.

1. The haploid dynamics (S.2) *admits at most one internal equilibrium.*

2. Depending on the parameters, the internal equilibrium is given by either $(p, q_1(p), D(p, q_1(p)))$ *or* $(p, q_2(p), D(p, q_2(p)))$, *where* p *is one of* p_1 *or* p_2 *in* (S.45), *and* $q_i(p)$ *and* $D(p, q_i(p))$ *are given by* (S.41) *and* (S.39), *respectively.*

3. An internal equilibrium exists if and only if both $\mathbf{M}_2$ *and* $\mathbf{M}_3$ *are asymptotically stable. This is the case if and only if*

$$\gamma > \alpha \ \text{ and } \ \beta > 0 \ \text{ and } \ r > \alpha - \beta\,. \tag{S.42}$$

4. *The internal equilibrium is unstable whenever it exists.*

5. *If* (S.42) *does not hold, then* $\mathbf{M}_3$ *is globally asymptotically stable.*

This theorem complements the results derived by FELDMAN (1971) and RUTSCHMAN (1994) on the discrete-time dynamics of the haploid two-locus selection model. Rutschman proved global convergence to a boundary equilibrium for all parameter combinations for which no internal equilibrium exists. If transformed to the parameters used by Rutschman, condition (S.42) yields precisely the cases not covered by Rutschman's Theorem 14. Because our model is formulated in continuous time, the internal equilibrium can be determined by solving quadratic equations. This is instrumental for our proof.

We assume $\beta > 0$ because otherwise $q(t) \to 0$ as $t \to \infty$ if $m = 0$ (this follows because $\dot{q} < 0$ on $D < 0$ in this case, and we already know that all solutions eventually reach $D \leq 0$ and remain there). In addition, for $m > 0$ we showed that internal equilibria can exist only if $\alpha > 0$ and $\gamma > \beta$. Because with $m = 0$ we have the additional model symmetry of exchangeable α and β, we may assume

$$\gamma > \alpha \geq \beta > 0\,. \tag{S.43}$$

We need four lemmas to prove the above theorem.

Lemma S.2. *The coordinate p of an equilibrium is a zero of the polynomial*

$$P(p) = \gamma^2 r^2 p^2 - r[2\beta(\gamma-\beta)(\gamma-2\alpha) + r\gamma^2]p + \beta(\gamma-\beta)(\alpha-\beta-r)(\alpha+\beta-\gamma-r)\,. \tag{S.44}$$

If $r > 0$, the two zeros of P are given by

$$p_{1,2} = \frac{1}{2} + \frac{\gamma-2\alpha}{r}\frac{\beta}{\gamma}\left(1-\frac{\beta}{\gamma}\right) \pm \frac{1}{2}\left(1-\frac{2\beta}{\gamma}\right)\sqrt{1+\frac{4\alpha(\gamma-\alpha)}{r^2}\frac{\beta}{\gamma}\left(1-\frac{\beta}{\gamma}\right)} \tag{S.45}$$

and satisfy

$$|p_1 - p_2| = \left(1-\frac{2\beta}{\gamma}\right)\sqrt{1+\frac{4\alpha(\gamma-\alpha)}{r^2}\frac{\beta}{\gamma}\left(1-\frac{\beta}{\gamma}\right)} > \left|1-\frac{2\beta}{\gamma}\right|\,. \tag{S.46}$$

The minimum value of P is

$$P(p_{\min}) = \frac{-(\gamma-2\beta)^2}{4\gamma^2}[4\alpha\beta(\gamma-\alpha)(\gamma-\beta) + \gamma^2 r^2] < 0\,, \tag{S.47}$$

where

$$p_{\min} = \frac{1}{2} + \frac{1}{r}\frac{\beta}{\gamma}\left(1-\frac{\beta}{\gamma}\right)(\gamma-2\alpha)\,. \tag{S.48}$$

In addition, we have

$$P\left(\frac{\beta}{\gamma}\right) = \beta\left(1-\frac{\beta}{\gamma}\right)[(\alpha-\beta)(\alpha+\beta-\gamma)\gamma + (\gamma-2\alpha)(\gamma-2\beta)r] \tag{S.49a}$$

and

$$P\left(1-\frac{\beta}{\gamma}\right)=\beta\left(1-\frac{\beta}{\gamma}\right)[(\alpha-\beta)(\alpha+\beta-\gamma)\gamma-(\gamma-2\alpha)(\gamma-2\beta)r]\,. \tag{S.49b}$$

Proof. The expression $P(p)$ follows from (S.33) because $m=0$. The other statements are derived readily. □

An immediate consequence of (S.44) is that no internal equilibrium exists if $r=0$ and at most one if $\gamma=0$.

Lemma S.3. *(i)* $(p,q_1(p))$ *can give rise to an internal equilibrium if and only if*

$$\tilde{A}(p)=(\gamma-\beta)[(\alpha-r)(\gamma-2\alpha)+\gamma(\alpha-\beta)]+[2\alpha(\gamma-\alpha)(2\beta-\gamma)+r\gamma(\gamma-2\alpha)]p \tag{S.50}$$

and

$$\tilde{B}(p)=-(\gamma-\beta)[\beta(\gamma-2\alpha)+\gamma r]+r\gamma^2p \tag{S.51}$$

have opposite signs. Otherwise, $(p,q_2(p))$ *may give rise to an internal equilibrium.*

(ii) Assume (S.43)*. If* p_1 *and* p_2 *are the two zeros of* $P(p)$*, then* $\tilde{B}(p_1)\tilde{B}(p_2)<0$*. Hence,* $\tilde{B}(p)$ *changes sign between* p_1 *and* p_2*. Because* $\alpha>\beta$*, we have* $\tilde{A}(p_1)\tilde{A}(p_2)>0$ *for every* $r>0$ *if and only if* $\gamma\le\alpha+\beta$*. Hence, in this case,* $\tilde{A}(p)$ *does not change sign between* p_1 *and* p_2*.*

(iii) Assume (S.43)*. If* $\gamma\le\alpha+\beta$ *and* P *has two zeros in* $(0,1)$*, at most one of them can give rise to an admissible internal equilibrium* $(p,q_1,D(p,q_1))$*.*

Proof. (i) Since $m=0$, we obtain from (S.29) $A(p)=\beta(1-p)\tilde{A}(p)$ and $B(p)=-(1-p)\tilde{B}(p)$, which, on account of (S.31), proves (i).

(ii) These statements follow because

$$\tilde{A}(p_1)=\frac{\gamma-2\beta}{2\gamma^2r}\left[\phi_1+\phi_2\sqrt{4\alpha\beta(\gamma-\alpha)(\gamma-\beta)+\gamma^2r^2}\right], \tag{S.52a}$$

$$\tilde{A}(p_2)=\frac{\gamma-2\beta}{2\gamma^2r}\left[\phi_1-\phi_2\sqrt{4\alpha\beta(\gamma-\alpha)(\gamma-\beta)+\gamma^2r^2}\right], \tag{S.52b}$$

and

$$\tilde{B}(p_1)=\frac{\gamma-2\beta}{2}\left[-\gamma r+\sqrt{4\alpha\beta(\gamma-\alpha)(\gamma-\beta)+\gamma^2r^2}\right], \tag{S.53a}$$

$$\tilde{B}(p_2)=\frac{\gamma-2\beta}{2}\left[-\gamma r-\sqrt{4\alpha\beta(\gamma-\alpha)(\gamma-\beta)+\gamma^2r^2}\right], \tag{S.53b}$$

where

$$\phi_1 = 4\alpha\beta(\gamma-\beta)(\gamma-\alpha)(2\alpha-\gamma) + 2\alpha\gamma(\gamma-\alpha)(\gamma-2\beta)r + \gamma^2(2\alpha-\gamma)r^2\,, \tag{S.54a}$$
$$\phi_2 = 2\alpha(\gamma-\alpha)(2\beta-\gamma) + r\gamma(\gamma-2\alpha)\,, \tag{S.54b}$$

and

$$\begin{aligned}&\phi_1^2 - \phi_2^2[4\alpha\beta(\gamma-\alpha)(\gamma-\beta)+\gamma^2r^2]\\ &\quad= 4\alpha\beta\gamma^2(\gamma-\alpha)(\gamma-\beta)[4\alpha(\alpha-\beta)(\gamma-\alpha)(\alpha+\beta-\gamma)+(2\alpha-\gamma)^2r^2)] > 0\end{aligned} \tag{S.55}$$

if $\alpha+\beta>\gamma$.

(iii) This is an immediate consequence of (i) and (ii). □

Lemma S.4. *Under the assumption* (S.43), $(p, q_1, D(p, q_1))$ *can be an admissible internal equilibrium only if one of the following three conditions is satisfied:*

$$\gamma \ge 2\alpha \;\; \textit{and} \;\; p < 1 - \frac{\beta}{\gamma}\,, \tag{S.56a}$$

or

$$\gamma < 2\alpha \;\; \textit{and} \;\; p < \min\left[\frac{\beta}{\gamma}, 1-\frac{\beta}{\gamma}\right]\,, \tag{S.56b}$$

or

$$\gamma < 2\beta \;\; \textit{and} \;\; 1-\frac{\beta}{\gamma} < p < \frac{\beta}{\gamma}\,. \tag{S.56c}$$

Proof. We utilize the fact that an internal admissible equilibrium has to satisfy (S.40), where $D = D(p, q_1)$ is given by (S.39). We consider two cases.

Case 1. $p < 1 - \frac{\beta}{\gamma}$. Then $D < 0$ if and only if $\gamma q_1 > \alpha$, which is satisfied if and only if one of the following conditions holds:

$$\gamma \ge 2\alpha\,, \tag{S.57a}$$
$$\gamma < 2\alpha \;\; \text{and} \;\; p(1-p) < \frac{\beta}{\gamma}\left(1-\frac{\beta}{\gamma}\right)\,. \tag{S.57b}$$

Because $p < 1 - \frac{\beta}{\gamma}$, (S.57b) is equivalent to (S.56b). Thus, each of (S.56a) or (S.56b) is necessary for admissibility of an equilibrium with $p < 1 - \frac{\beta}{\gamma}$.

Case 2. $p > 1 - \frac{\beta}{\gamma}$. Then $D < 0$ if and only if $\gamma q_1 < \alpha$, which is satisfied if and only if

$$\gamma < 2\alpha \;\; \text{and} \;\; 1-\frac{\beta}{\gamma} < p < \frac{\beta}{\gamma}\,. \tag{S.58}$$

Clearly, this requires $\gamma < 2\beta$. Hence, (S.58) can be replaced by (S.56c).

We recall from Section S.3 that if $p = 1-\beta/\gamma$, there exists at most one internal equilibrium. □

Lemma S.5. *Under the assumption* (S.43), $(p, q_2, D(p, q_2))$ *can be an admissible internal equilibrium only if*

$$\gamma > 2\alpha \tag{S.59}$$

and

$$\frac{\beta}{\gamma} < p \leq 1 - \frac{\beta}{\gamma} \tag{S.60}$$

hold.

Proof. Case 1. $p < 1 - \frac{\beta}{\gamma}$. Then $D < 0$ if and only if $\gamma q_2 > \alpha$, which can be satisfied only if (S.59) and (S.60) hold.

Case 2. $p > 1 - \frac{\beta}{\gamma}$. Then $D < 0$ if and only if $\gamma q_2 < \alpha$, which can be satisfied only if $p > \max\left[\frac{\beta}{\gamma}, 1 - \frac{\beta}{\gamma}\right]$. Admissibility of an internal equilibrium also requires $D(p, q_2) > -pq_2$. A straightforward calculation shows that, because $p > 1 - \frac{\beta}{\gamma}$, this can hold only if $\alpha < \beta$, i.e., if assumption (S.43) is violated. □

Remark S.2. For convenience, we briefly recapitulate the following index theorem (Theorem 2 in Hofbauer 1990) which will play a decisive role in the proof of Theorem S.2. For every dissipative semiflow on $\mathbb{R}^n_+$ such that all fixed points are regular, the sum of their indices equals +1. To apply this theorem, we note that system (S.1) satisfies the assumptions. (Here, we assume $m \geq 0$, because this remark will also be needed below when treating the case $m > 0$.) (S.1) is obviously well defined on $\mathbb{R}^4_+$, $\mathbb{R}^4_+$ is forward invariant under (S.1) (i.e., if $x_i = 0$, then $\dot{x}_i \geq 0$), and S_4 is globally attracting in $\mathbb{R}^4_+$ under a modification of (S.1) that does not change the dynamics on S_4. To show the latter, we note that adding a sufficiently large positive constant κ to each of the gametic fitnesses w_i does not change the dynamics on S_4, but implies

$$\frac{d}{dt}\Big(\sum_i x_i\Big) = \Big(1 - \sum_i x_i\Big)[m + \kappa x_1 + (\beta+\kappa)x_2 + (\alpha+\kappa)x_3 + (\alpha+\beta-\gamma+\kappa)x_4] < 0 \tag{S.61}$$

if and only if $\sum_i x_i > 1$ and $\frac{d}{dt}\left(\sum_i x_i\right) > 0$ if and only if $\sum_i x_i < 1$.

In our case, the index of an equilibrium is $(-1)^m$, where m is the number of negative eigenvalues (they are always real). An internal equilibrium is always saturated. If it is asymptotically stable, it has index 1. Equilibria on the boundary of the simplex are saturated if and only if they are externally stable. This is the case if and only if no gamete that is missing at the equilibrium can invade. Because S_4 is attracting within $\mathbb{R}^4_+$, the index of an asymptotically stable (hence, saturated) boundary equilibrium is 1.

Now we are ready to prove the theorem.

Proof of Theorem S.2. 1. and 2. We assume (S.43). By Theorem S.1, for given p, only one of $q_1(p)$ or $q_2(p)$ can give rise to an equilibrium. By Lemma S.4, $q_1(p)$ may give rise to an internal equilibrium if either of the conditions (S.56a), (S.56b), or (S.56c) is satisfied. By Lemma S.5, $q_2(p)$ may give rise to an equilibrium if (S.59) and (S.60) are satisfied.

If $\gamma \geq 2\alpha$, (S.49b) shows that $P(1-\beta/\gamma) < 0$ and Lemma S.2 implies that one zero of P is greater than $1-\beta/\gamma$. Hence, by (S.56a), only the zero of P which satisfies $p < 1-\beta/\gamma$ can give rise to an admissible equilibrium $(p, q_1, D(p, q_1))$. By Lemma S.5, (S.59) and (S.60), only the zero $p < 1-\beta/\gamma$ can give rise to an admissible equilibrium $(p, q_2, D(p, q_2))$. However, we already know that at most one of $q_1(p)$ or $q_2(p)$ can yield an equilibrium.

If $\gamma < 2\alpha$, then Lemmas S.4 and S.5 imply that only q_1 can give rise to an equilibrium. This requires that (S.56b) or (S.56c) apply. If (S.56c) applies, (S.46) shows that at most one zero of P can satisfy this condition. If (S.56b) applies and $\gamma \leq \alpha + \beta$, we conclude from Lemma S.3 (iii) that at most one zero can give rise to an equilibrium. If $\alpha + \beta < \gamma < 2\alpha$, then (S.49a) yields $P(\frac{\beta}{\gamma}) < 0$, whence one zero of P is bigger than $\frac{\beta}{\gamma}$, in contradiction to (S.56b). This finishes the proof of the first statement and shows that an internal equilibrium can exist only if (S.42) is satisfied.

3. and 4. It remains to prove that the internal equilibrium exists if (S.42) holds and that it is unstable. This follows readily from the index theorem of HOFBAUER (1990); see Remark S.2. In our model, the only boundary equilibria are the four monomorphic states. M_1 and M_4 are never saturated because they are unstable within S_4. $\mathbf{M}_2$ and $\mathbf{M}_3$ are saturated if and only if they are asymptotically stable within S_4. Then, we have $\mathrm{ind}(\mathbf{M}_2) = \mathrm{ind}(\mathbf{M}_3) = 1$. Hence there must exist an internal equilibrium with index -1. Such an equilibrium cannot be stable.

Because $\mathbf{M}_2$ and $\mathbf{M}_3$ are both asymptotically stable if and only if (S.42) holds, statements 3 and 4 are proved.

5. Let us first assume $\beta \leq 0$. Then (S.18) implies $\dot{X} \geq 0$ if $p < 1$, and $\dot{X} = 0$ only if $x_3 = 0$. Because $\dot{x}_3 > 0$ on $x_3 = 0$ except on the invariant edges $p = 0$ or $q = 1$, we infer $p(t) \to 1$. If $p = 1$, then $\dot{x}_3 = x_3(1-x_3)(\gamma - \beta)$ yields the assertion.

Now assume $\beta > 0$. A simple calculation shows that

$$\dot{x}_3 + \dot{x}_4 = \alpha x_1 x_3 + (\alpha - \beta) x_2 x_3 + (\alpha - \gamma) x_2 x_4 + (\alpha + \beta - \gamma) x_1 x_4 \geq 0 \tag{S.62}$$

if $\alpha \geq \gamma\,(> \beta)$, and equality holds at most at edges of S_4. From this, it is an easy exercise to prove convergence of $\mathbf{M}_3$.

Finally, assume $\beta > 0$, $\gamma > \alpha$, and $r < \alpha - \beta$, which requires $\alpha > \beta$. Then global convergence to $\mathbf{M}_3$ follows from the inequality

$$\dot{x}_3 = \alpha x_1 x_3 + (\alpha - \beta) x_1 x_4 + (\gamma - \beta) x_3 x_4 - (\alpha - \beta - r) D \geq 0 \,, \tag{S.63}$$

which holds if $D \leq 0$ (as is sufficient; see Section S.2.4), and a simple investigation of the dynamics on the edges of S_4. □

S.6 Weak migration

With the aid of perturbation theoretical methods, we can derive the equilibrium and stability structure of our model for weak migration from the model without migration. Theorem 4.4 of Karlin and McGregor (1972b) implies that every equilibrium that is asymptotically stable for $m = 0$, persists for sufficiently small $m > 0$ and remains asymptotically stable. Stable equilibria on the boundary may move into the interior of the state space. Equilibria that are unstable if $m = 0$ remain unstable if $m > 0$. However, unstable equilibria at the boundary may leave the state space as m becomes positive. Application of this theorem requires that all equilibria in the unperturbed system are hyperbolic, i.e., they have no eigenvalues with zero real part. Theorem 5.4 in Bürger (2009) ensures global asymptotic stability of the perturbed equilibrium, at least if γ and m are small enough.

For our model these perturbation results in conjunction with Theorem S.2 and Lemma S.1 immediately yield the following result.

Theorem S.3. *If $m > 0$ is sufficiently small, the following equilibrium configurations can occur.*

1. If (S.42) *holds, there exists one unstable internal equilibrium, one asymptotically stable internal equilibrium (the perturbation of $\mathbf{M}_3$), and $\mathbf{M}_2$ is asymptotically stable. Neither $\mathbf{S}_{\mathcal{A}}$ nor $\mathbf{S}_{\mathcal{B}}$ is admissible.*

2. Otherwise, i.e., if $\gamma < \alpha$ or $\beta < 0$ or $r < \alpha - \beta$, the perturbation of the equilibrium $\mathbf{M}_3$ is globally asymptotically stable (at least if γ is small). The equilibrium $\mathbf{M}_2$ is unstable, and the equilibria $\mathbf{S}_{\mathcal{A}}$ and $\mathbf{S}_{\mathcal{B}}$ may be admissible. If $\mathbf{S}_{\mathcal{A}}$ or $\mathbf{S}_{\mathcal{B}}$ is admissible, it is unstable.

The coordinates of the internal equilibria can be determined explicitly to first order in m. The coordinates (p, q, D) of the stable equilibrium ($\mathbf{I}_{\mathrm{DMI}}$) are

$$\left(1 - \frac{m(\alpha + r)}{\alpha(\alpha - \beta + r)}, \frac{m(\gamma - \beta + r)}{(\gamma - \beta)(\alpha - \beta + r)}, -\frac{m}{\alpha - \beta + r}\right) . \tag{S.64}$$

S.7 The complete equilibrium and stability structure

Now we are in the position to prove our main results about the equilibrium and bifurcation structure. We continue to assume (S.24), because we have already shown that otherwise an internal equilibrium (a DMI) cannot exist and global convergence to a boundary equilibrium occurs (Section S.2). Throughout, we always consider bifurcations as a function of (increasing) m.

We define

$$m_{\mathcal{A}} = \frac{(\alpha-\gamma)(\gamma-\beta)}{\alpha}\left(1+\frac{\alpha+\beta-\gamma}{r}\right), \tag{S.65a}$$

$$m_{\mathcal{B}} = \frac{-\beta\alpha}{\gamma-\beta}\left(1+\frac{\gamma-\beta-\alpha}{r}\right), \tag{S.65b}$$

$$m_2 = \alpha-\beta-r \tag{S.65c}$$

and note that $m_{\mathcal{A}}$, $m_{\mathcal{B}}$, and m_2 are the critical values of m above which $\mathbf{S}_{\mathcal{A}}$, $\mathbf{S}_{\mathcal{B}}$, and $\mathbf{M}_2$, respectively, are asymptotically stable provided they are admissible (Section S.1). We also recall the definitions of m^* (S.36) and of

$$m_{\max}^{-} = \begin{cases} m_{\mathcal{A}} & \text{if } \gamma < \min[\alpha, \alpha+\beta, \beta+r]\,, \quad \text{(S.66a)} \\ m_{\mathcal{B}} & \text{if } \gamma \geq \alpha+\beta = \min[\alpha, \alpha+\beta, \beta+r]\,, \quad \text{(S.66b)} \\ m_2 & \text{if } \gamma \geq \alpha+\beta = \min[\alpha, \alpha+\beta, \beta+r]\,; \quad \text{(S.66c)} \end{cases}$$

cf. Appendix A.2.

Theorem S.4. *The following three types of bifurcation patterns can occur:*

Type 1.

- *If* $0 < m < m^*$*, there exist two internal equilibria; one is asymptotically stable* ($\mathbf{I}_{\mathrm{DMI}}$)*, the other* ($\mathbf{I}_0$) *is unstable. The monomorphic equilibrium* $\mathbf{M}_2$ *is asymptotically stable.*
- *At* $m = m^*$*, the two internal equilibria collide and annihilate each other by a saddle-node bifurcation.*
- *If* $m > m^*$*,* $\mathbf{M}_2$ *is the only equilibrium; it is asymptotically stable and, presumably, globally stable.*

Type 2. There exists a critical migration rate $\tilde{m}$ *satisfying* $0 < \tilde{m} < m^*$ *such that:*

- *If* $0 < m < \tilde{m}$*, there is a unique internal equilibrium* $(\mathbf{I}_{\mathrm{DMI}})$*. It is asymptotically stable and, presumably, globally stable.*
- *At* $m = \tilde{m}$*, an unstable equilibrium* $(\mathbf{I}_0)$ *enters the state space by an exchange-of-stability bifurcation with a boundary equilibrium.*
- *If* $\tilde{m} < m < m^*$*, there are two internal equilibria, one asymptotically stable* $(\mathbf{I}_{\mathrm{DMI}})$*, the other unstable* $(\mathbf{I}_0)$*, and one of the boundary equilibria is asymptotically stable.*
- *At* $m = m^*$*, the two internal equilibria merge and annihilate each other by a saddle-node bifurcation.*
- *If* $m > m^*$*, a boundary equilibrium asymptotically stable and, presumably, globally stable.*

Type 3.

- *If* $0 < m < m^-_{max}$*, a unique internal equilibrium* $(\mathbf{I}_{\mathrm{DMI}})$ *exists. It is asymptotically stable and, presumably, globally stable.*
- *At* $m = m^-_{max}$*,* $\mathbf{I}_{\mathrm{DMI}}$ *leaves the state space through a boundary equilibrium by an exchange-of-stability bifurcation.*
- *If* $m > m^-_{max}$*, a boundary equilibrium is asymptotically stable and, presumably, globally stable.*

Remark S.3. 1. Each of the bifurcation patterns of Type 2 and 3 include several subcases because the equilibria $\mathbf{S}_{\mathcal{A}}$ and $\mathbf{M}_2$ or $\mathbf{S}_{\mathcal{B}}$ and $\mathbf{M}_2$ may merge and thereby exchange stability. This may occur below or above m^*. If such a bifurcation occurs, then it is always $\mathbf{M}_2$ that becomes stable, and $\mathbf{S}_{\mathcal{A}}$ or $\mathbf{S}_{\mathcal{B}}$ leave the state space through $\mathbf{M}_2$.

2. The conjectures about global asymptotic stability of a boundary equilibrium if $m > m^*$ (Type 1 or Type 2) or if $m > m^-_{\mathrm{max}}$ (Type 3) are supported by extensive numerical iterations which provided no incidence of limit cycles or complex attractors. Global stability of $\mathbf{M}_2$ has been proved if $m > m_2$ (Section S.2.2).

3. The conjectures about global stability of the internal equilibrium in patterns of Type 2 and Type 3 are supported by extensive numerical iterations. For sufficiently small m and γ, they follow from Theorem 5.4 in Bürger (2009). For tight linkage and independent loci, limit cycles and complex attractors are excluded in Sections S.8.1 and S.8.2.

Proof of Theorem S.4. Theorem S.3 provides all equilibrium configurations for small m. Lemma S.1 provides control over the boundary equilibria. As m increases, they can vanish but not emerge. They can also become asymptotically stable as m increases. For sufficiently large m, there is always a globally asymptotically stable boundary equilibrium. By Theorem S.1, the number of internal equilibria is at most three. In addition, internal equilibria can emerge or vanish only either by a saddle-node bifurcation (Section S.4) or because an equilibrium enters or leaves S_4 through one of boundary equilibria, when an exchange of stability occurs. A bifurcation involving the two internal equilibria can occur at most at one value of m, namely at m^* (S.36). An exchange-of-stability bifurcation can occur only at the values $m_{\mathcal{A}}$, $m_{\mathcal{B}}$, or m_2. If it occurs, then the respective boundary equilibrium is asymptotically stable for every larger m for which it is admissible. By the index theorem of HOFBAUER (1990), cf. Remark S.2, the sum of the indices of all saturated equilibria equals 1.

If Case 1 of Theorem S.3 applies, then $\mathbf{M}_2$ is asymptotically stable for every $m > 0$, and it is the only boundary equilibrium. Hence, its index is 1. The index of the stable internal equilibrium is also 1. Because the sum of the indices of the internal equilibria must be 0, the index of the unstable equilibrium is -1. Because at most one bifurcation involving the two internal equilibria can occur and because for large m, $\mathbf{M}_2$ is globally asymptotically stable, the bifurcation must be of saddle-node type in which the equilibria collide and annihilate each other (but do not emerge). In principle, the internal equilibria could also leave S_4 through a boundary equilibrium (in this case, it must be $\mathbf{M}_2$). However, by the index theorem, they can do so only simultaneously. This occurs if and only if $m^* = m_2$, which is a non-generic degenerate case. Because the sum of indices of the internal equilibria must be zero, no equilibrium can enter the state space. These considerations settle the bifurcation pattern of Type 1.

If Case 2 of Theorem S.3 applies, then, for small m, the boundary equilibria are unstable, hence not saturated, and do not contribute to the sum of indices. Since then the indices of internal equilibria must sum up to 1, the only possible bifurcation that does not entail the stability of a boundary equilibrium would be a pitch-fork bifurcation of the internal equilibrium which, by Section S.4, does not occur. (Indeed, because m^* is the only value at which a bifurcation among internal equilibria can occur, and because for large m a boundary equilibrium is globally asymptotically stable, the three equilibria emerging by a pitchfork bifurcation would have to leave the state space through boundary equilibria. This, however, cannot occur, as follows easily from the results about linear stability in Section S.1.) Thus, any further bifurcations involve a boundary equilibrium. There are two possibilities.

(i) An equilibrium enters S_4 at some value $\tilde{m}$ (which can only be one of $m_{\mathcal{A}}$, $m_{\mathcal{B}}$, or m_2) through one of the unstable boundary equilibria by an exchange-of-stability bifurcation. If $m > \tilde{m}$, there is one asymptotically stable boundary equilibrium, an unstable internal equilibrium (the one that entered S_4), and one asymptotically stable internal equilibrium. Now a reasoning analogous to that applied above to Case 1 of Theorem S.3 establishes the bifurcation pattern of Type 2.

(ii) The internal equilibrium leaves S_4 by exchange of stability through one of the boundary equilibria at $m_{\max}^-$. This becomes asymptotically stable then and, presumably, globally stable. At larger values of m no equilibrium can enter S_4 through one of the (other) unstable boundary equilibria, because this would either lead to two simultaneously stable boundary equilibria, which is impossible (Lemma S.1), or this had to occur at the same value at which the hitherto stable boundary equilibrium merges with $\mathbf{M}_2$ and leaves the state space. This, too, is impossible because the sum of the indices of the new stable boundary equilibrium and the new unstable internal equilibrium would be zero. Thus, we have established the bifurcation pattern of Type 3 and excluded all other possibilities. □

Our final goal is to assign the respective parameter combinations to the three bifurcation patterns determined above. As in Appendix A.3, we distinguish four selection scenarios:

Selection scenario 1: $0 < \alpha < \beta < \gamma$

Selection scenario 2: $0 < \beta < \alpha < \gamma$

Selection scenario 3: $\gamma - \alpha < 0 < \gamma - \beta < \alpha$

Selection scenario 4: $\beta < 0 < \alpha < \gamma - \beta$

From Appendix A.2, we recall the following definitions

$$r_{\mathcal{A}} = (\gamma - \alpha)\,\frac{3(\gamma - \beta) - \alpha}{2\gamma - \alpha}\,, \tag{S.67a}$$

$$r_{\mathcal{B}} = \beta\,\frac{3\alpha + \beta - \gamma}{\beta + \gamma}\,, \tag{S.67b}$$

$$r_2 = \frac{3\alpha(\gamma - \beta) - \sqrt{\alpha(\gamma - \beta)(4\beta\gamma + 5\alpha\gamma - 9\alpha\beta)}}{2\gamma}\,. \tag{S.67c}$$

Now we define

$$f_2(r) = \alpha(\beta - \alpha)(\gamma - \beta) + 3\alpha(\gamma - \beta)r - \gamma r^2\,. \tag{S.68}$$

Then r_2 is the smaller of the two zeros of f_2 (provided they are real).

We note the following properties of f_2 and r_2:

(i) $f_2(0) < 0$ if and only if $\alpha > \beta$;
(ii) f_2 is concave and assumes its maximum at $r_{2,\max} = \frac{3}{2}\alpha(1 - \frac{\beta}{\gamma}) > 0$. The larger of the two zeros of f_2 is always greater (or equal) than α;
(iii) $f_2(r) < 0$ for every r if and only if $4\beta\gamma + 5\alpha\gamma - 9\alpha\beta < 0$;
(iv) r_2 is real if and only if $4\beta\gamma + 5\alpha\gamma - 9\alpha\beta \geq 0$, and $r_2 > 0$ if, in addition, $\alpha > \beta$.

We define

$$r_2^* = \begin{cases} r_2 & \text{if } r_2 \text{ is real}\,, \\ \min[\alpha, \gamma - \beta] & \text{otherwise}\,. \end{cases} \tag{S.69}$$

In selection scenarios 1 and 2, r_2 is always real. If r_2 is not real, we have $r_2^* = \beta - \gamma$ in selection scenario 3, and $r_2^* = \alpha$ in scenario 4. An important observation is that in selection scenario 4

$$r_2^* < \alpha \quad \text{if and only if} \quad \gamma(\alpha + \beta) > \beta(2\alpha + \beta)\,. \tag{S.70}$$

Lemma S.6. *Assume selection scenario 4, i.e.,* $\beta < 0$ *and* $\gamma > \alpha + \beta$*.*

1) An unstable internal equilibrium enters the state space through $\mathbf{M}_2$ *at* $m = m_2$ *(and exists if* $m = m_2 + \epsilon$, $\epsilon > 0$ *small) if*

$$r \le \alpha \quad \textit{and} \quad f_2(r) > 0 \tag{S.71}$$

or, equivalently, if

$$r_2^* < r \le \alpha\,. \tag{S.72}$$

2) The stable internal equilibrium leaves the state space through $\mathbf{M}_2$ *at* $m = m_2$ *if*

$$r \le \alpha \quad \textit{and} \quad f_2(r) \le 0 \tag{S.73}$$

or, equivalently, if

$$r \le \min[\alpha, r_2^*]\,. \tag{S.74}$$

Proof. The lemma follows from studying the perturbation of $\mathbf{M}_2$ near m_2. To obtain D to first order in ϵ, one has to derive p and q up to second order in ϵ.

1) If $m = m_2 + \epsilon$ ($\epsilon > 0$), to leading order in ϵ, the perturbation analysis yields

$$p(\epsilon) = \frac{(\gamma - \beta)D(\epsilon)}{r - \gamma + \beta}\,, \quad q(\epsilon) = 1 - \frac{\alpha D(\epsilon)}{r - \alpha}\,, \quad D(\epsilon) = \epsilon \frac{-(r-\alpha)(r - \gamma + \beta)}{f_2(r)}\,. \tag{S.75}$$

Because we know that $D(\epsilon) < 0$ holds at every equilibrium and because $\alpha < \gamma - \beta$, (S.71) follows. The equivalence of (S.71) and (S.72) follows easily from the properties of f_2 summarized above.

2) If $m = m_2 - \epsilon$ ($\epsilon > 0$), to leading order in ϵ, the perturbation analysis yields the same expressions for $p(\epsilon)$ and $q(\epsilon)$, and

$$D(\epsilon) = \epsilon \frac{(r-\alpha)(r - \gamma + \beta)}{f_2(r)}\,. \tag{S.76}$$

Now, (S.73) and (S.74) follow as above. If $r = r_2$, then $f_2(r) = 0$. This case is degenerate because $m_2 = m^*$ and the outgoing and incoming equilibria collide at $\mathbf{M}_2$ and annihilate each other. □

Lemma S.7. *Assume selection scenario 2, i.e.,* $0 < \beta < \alpha < \gamma$.

1) An unstable internal equilibrium enters the state space through $\mathbf{M}_2$ *at* $m = m_2$ *(and exists if* $m = m_2 + \epsilon$, $\epsilon > 0$ *small) if*

$$r_2 < r < \alpha - \beta \,. \tag{S.77}$$

2) The stable internal equilibrium leaves the state space through $\mathbf{M}_2$ *at* $m = m_2$ *if*

$$r \leq r_2 \,. \tag{S.78}$$

Proof. The proof is analogous to that of Lemma S.6 upon noting that $m_2 > 0$ if and only if $r < \alpha - \beta$, and that $\alpha - \beta < \min[\alpha, \gamma - \beta]$ holds. ☐

Next, we define

$$f_B(r) = \beta(3\alpha + \beta - \gamma) - r(\beta + \gamma) \tag{S.79}$$

and note that $f_B(r) = 0$ if and only if $r = r_{\mathcal{B}}$. In addition, we have

$$r_{\mathcal{B}} > \alpha \quad \text{if and only if} \quad \begin{cases} \gamma > -\beta \text{ and } \gamma(\alpha+\beta) < \beta(2\alpha+\beta)\,, \text{ or} \\ \gamma < -\beta \text{ and } \gamma(\alpha+\beta) > \beta(2\alpha+\beta)\,. \end{cases} \tag{S.80}$$

In selection scenario 4, we obtain

$$r_2^* < \alpha \quad \text{if and only if} \quad \begin{cases} r_{\mathcal{B}} < \alpha \text{ and } \gamma > -\beta\,, \text{ or} \\ r_{\mathcal{B}} > \alpha \text{ and } \gamma < -\beta\,. \end{cases} \tag{S.81}$$

If r_2 is real, this follows immediately from (S.70) and (S.80), or from the observation $f_2(\alpha) + \alpha f_B(\alpha) = 0$. Otherwise, it follows because $4\beta\gamma + 5\alpha\gamma - 9\alpha\beta < 0$ implies $\gamma(\alpha+\beta) < \beta(2\alpha+\beta)$.

Lemma S.8. *Assume selection scenario 4, i.e.,* $\beta < 0$ *and* $\gamma > \alpha + \beta$.

1) An unstable internal equilibrium enters the state space through $\mathbf{S}_{\mathcal{B}}$ *at* $m = m_{\mathcal{B}}$ *(and exists for* $m = m_{\mathcal{B}} + \epsilon$, $\epsilon > 0$ *small) if and only if*

$$r > \alpha \quad \textit{and} \quad f_B(r) < 0\,. \tag{S.82}$$

This holds in precisely the following cases:

$$r > \max[\alpha, r_{\mathcal{B}}] \textit{ and } \gamma > -\beta\,, \tag{S.83a}$$

$$\alpha < r < r_{\mathcal{B}} \textit{ and } \gamma < -\beta\,, \tag{S.83b}$$

$$\alpha < r < \infty \textit{ and } \gamma = -\beta < 3\alpha + \beta\,. \tag{S.83c}$$

2) The stable internal equilibrium leaves the state space through $\mathbf{S}_{\mathcal{B}}$ *at* $m = m_{\mathcal{B}}$ *if and only if*

$$r > \alpha \quad \textit{and} \quad f_B(r) \geq 0\,. \tag{S.84}$$

This holds in precisely the following cases:

$$r > \max[\alpha, r_{\mathcal{B}}] \ \textit{and} \ \gamma < -\beta\,, \tag{S.85a}$$

$$r = r_{\mathcal{B}} > \alpha\,, \tag{S.85b}$$

$$\alpha < r < r_{\mathcal{B}} \ \textit{and} \ \gamma > -\beta\,, \tag{S.85c}$$

$$\alpha < r < \infty \ \textit{and} \ \gamma = -\beta \geq 3\alpha + \beta\,. \tag{S.85d}$$

Proof. The lemma follows from studying the perturbation of $\mathbf{S}_{\mathcal{B}}$ near $m_{\mathcal{B}}$. We recall from above (S.9) that $\mathbf{S}_{\mathcal{B}}$ is admissible for $m = m_{\mathcal{B}}$ if and only if $m_{\mathcal{B}} < -\beta$. Under the assumptions of the lemma, this reduces to $r > \alpha$.

1) If $m = m_{\mathcal{B}} + \epsilon$ ($\epsilon > 0$), the perturbation analysis (for p and q terms of order ϵ^2 are required to obtain D to order ϵ) yields

$$p(\epsilon) = rD(\epsilon)\frac{\gamma - \beta}{\alpha(\alpha + \beta - \gamma)}\,, \quad D(\epsilon) = \epsilon\frac{(\alpha + \beta - \gamma)(\alpha - r)}{r f_B(r)}\,. \tag{S.86}$$

$q(\epsilon)$ is needed only to get D. Because we know $D(\epsilon) < 0$ holds at every equilibrium, condition (S.82) follows. The particular cases follow easily by studying f_B.

The proof of 2) is analogous. The difference is that the perturbation analysis yields

$$D(\epsilon) = -\epsilon\frac{(\alpha + \beta - \gamma)(\alpha - r)}{r f_B(r)}\,. \tag{S.87}$$

If $r = r_{\mathcal{B}}$, then $f_B(r) = 0$. This case is degenerate because $m_{\mathcal{B}} = m^*$ and the outgoing and incoming equilibria collide at $\mathbf{S}_{\mathcal{B}}$ and annihilate each other. □

Theorem S.5. *1. Bifurcation patterns of Type 1 occur in*

Selection scenario 1 if and only if $r > \beta - \alpha\,;$
Selection scenario 2 if and only if $r \geq \alpha - \beta\,.$

2. Bifurcation patterns of Type 2 occur in

Selection scenario 2 if and only if $r_2 < r < \alpha - \beta\,;$

Selection scenario 3 if and only if one of the following holds:

(a) $r_2^* < r \leq \gamma - \beta$,

(b) $r > \max[\gamma - \beta, r_{\mathcal{A}}]$ *and* $\gamma > \frac{1}{2}\alpha$,

(c) $\gamma - \beta < r < \infty$ *and* $\gamma = \frac{1}{2}\alpha > 3\beta$,

(d) $\gamma - \beta < r < r_{\mathcal{A}}$ *and* $\gamma < \frac{1}{2}\alpha$.

Selection scenario 4 if and only if one of the following holds:

(a) $r_2^* < r \leq \alpha$,

(b) $r > \max[\alpha, r_{\mathcal{B}}]$ *and* $\gamma > -\beta$,

(c) $\alpha < r < \infty$ *and* $\gamma = -\beta < 3\alpha + \beta$,

(d) $\alpha < r < r_{\mathcal{B}}$ *and* $\gamma < -\beta$.

3. Bifurcation patterns of Type 3 occur in

Selection scenario 2 if and only if $r \leq r_2$ *;*
Selection scenario 3 if and only if one of the following holds:

(a) $r \leq \min[\gamma - \beta, r_2^*]$,

(b) $\gamma - \beta < r \leq r_{\mathcal{A}}$ *and* $\gamma > \frac{1}{2}\alpha$,

(c) $\gamma - \beta < r < \infty$ *and* $\gamma = \frac{1}{2}\alpha < 3\beta$.

(d) $r \geq \max[\gamma - \beta, r_{\mathcal{A}}]$ *and* $\gamma < \frac{1}{2}\alpha$.

Selection scenario 4 if and only if one of the following holds:

(a) $r \leq \min[\alpha, r_2^*]$,

(b) $\alpha < r \leq r_{\mathcal{B}}$ *and* $\gamma > -\beta$,

(c) $\alpha < r < \infty$ *and* $\gamma = -\beta \geq 3\alpha + \beta$,

(d) $r \geq \max[\alpha, r_{\mathcal{B}}]$ *and* $\gamma < -\beta$.

If $\gamma = 0$ and $\beta < -\alpha < 0$, which is subsumed in selection scenario 4(a), all possible bifurcation diagrams are displayed in Figure 1 of Bürger and Akerman (2011). Each of the bifurcation patterns of Type 2 or 3 is governed by one of these bifurcation diagrams. Thus, the only fundamentally new bifurcation patterns that arise in the present, much more general model, are of Type 1.

Proof of Theorem S.5. 1. The statement about bifurcation patterns of Type 1 is an immediate consequence of condition (S.42) in Theorem S.2, (S.24), and Theorem S.4.

2. The statement about selection scenario 2 follows from Lemma S.7, that about scenario 4 from Lemmas S.6 and S.8 by employing (S.70), (S.80), and (S.81). The statements about selection scenario 3 follow from those about scenario 4 by the model symmetry noted below (A.6).

3. The statement about selection scenario 2 is a consequence of Lemma S.7, that about scenario 4 of Lemmas S.6 and S.8 together with (S.70), (S.80), and (S.81). The statements about scenario 3 follow from those about scenario 4 by the model symmetry. □

Remark S.4. If, in selection scenario 4, $\gamma(\alpha+\beta) < \beta(2\alpha+\beta)$ and $\gamma < -\beta$ hold, then for every $r > 0$ a bifurcation pattern of Type 3 occurs, whence Type 2 (or Type 1) never occurs. These two conditions are satisfied whenever $-\beta > \max[2\alpha, \gamma]$. The first assertion follows immediately from Lemmas S.6 and S.8 by recalling (S.70), (S.79), and (S.80), the second is obvious.

Therefore, in the local-adaptation scenario with a slope-type fitness function, at most one internal equilibrium (which then is globally asymptotically stable) occurs whenever the selection intensity on the two loci differs by more of a factor of two. Thus, bistable equilibrium patterns can occur in the local adaptation scenario only if the selection strength on both loci is sufficiently similar and the recombination rate is about as strong as the selection intensity.

S.8 Further limiting cases and perturbation results

S.8.1 Tight linkage

First, we treat the case $r = 0$. From Section S.1 we recall that the monomorphic equilibrium $\mathbf{M}_2$ is asymptotically stable if $m > \alpha - \beta$ (hence, whenever $\beta > \alpha$), and the polymorphic equilibrium $\mathbf{R}_0$ is asymptotically stable if $m < \alpha - \beta$. The single-locus polymorphisms $\mathbf{S}_{\mathcal{A}}$ and $\mathbf{S}_{\mathcal{B}}$ do not exist.

We infer global convergence to the (unique) asymptotically stable equilibrium because

$$V = V(x_1, x_2, x_3, x_4) = \overline{w} + m \ln x_2 \tag{S.88}$$

is a Lyapunov function on the state space S_4 (cf. PASSEKOV 1978; HOFBAUER 1985; BÜRGER and AKERMAN 2011). That V is a Lyapunov function follows from equations (2.13) - (2.16) on p. 103 in BÜRGER (2000) by taking into account that, here, selection acts on haploids and by observing that, if $r = 0$, the system (1) is formally equivalent to a one-locus selection-mutation model with four alleles and so-called house-of-cards mutation. To see this, set all mutation rates to types 1, 3, and 4 (gametes ab, Ab, AB) zero, and assume that each of types 1, 3, and 4 mutates to type 2 (aB) at rate m. Therefore (p. 103 in BÜRGER 2000), (1) is a generalized gradient system. This implies that every trajectory converges to an equilibrium.

As a consequence, for sufficiently small r, every trajectory in the perturbed system converges to an equilibrium provided all equilibria are hyperbolic if $r = 0$. Thus, in particular, (generic) global convergence to the unique stable equilibrium follows. The crucial point is that the above result for $r = 0$ implies that the chain-recurrent points of (1) are exactly the equilibria (Lemma 2.2 in NAGYLAKI *et al.* 1999). Therefore, the proof of Theorem 2.3 in NAGYLAKI *et al.* (1999) applies unaltered.

If linkage is tight, the coordinates of the (globally) stable internal equilibrium can be approximated by perturbing the equilibrium $\mathbf{R}_0$ (S.11). One obtains

$$\hat{p} = 1 - \frac{m}{\alpha-\beta}\left[1 + r\,\frac{\beta(\beta-\alpha)+m(\alpha+\beta)}{\alpha(\alpha-\beta)}\right] + O(r^2) \tag{S.89a}$$

$$\hat{q} = \frac{m}{\alpha-\beta}\left[1 - r\,\frac{(\alpha+\beta-\gamma)(m+\beta-\alpha)}{\alpha(\alpha-\beta)(\gamma-\beta)}\right] + O(r^2) \tag{S.89b}$$

$$\hat{D} = \frac{m}{\alpha-\beta}\left(1 - \frac{m}{\alpha-\beta}\right)\left[1 - r\,\frac{(\gamma-\beta)(\alpha-\beta)+m(\alpha+\beta-\gamma)}{\alpha(\alpha-\beta)(\gamma-\beta)}\right] + O(r^2)\,. \tag{S.89c}$$

S.8.2 Independent loci

Under the assumption of linkage equilibrium ($D = 0$), the dynamics (S.2) reduces to the much simpler, two-dimensional form

$$\dot{p} = p[\alpha(1-p) - \gamma(1-p)q - m]\,, \tag{S.90a}$$

$$\dot{q} = (1-q)[\beta q - \gamma pq + m]\,. \tag{S.90b}$$

The coordinates and local stability properties of the possible boundary equilibria ($\mathbf{M}_2$, $\mathbf{S}_{\mathcal{A}}$, $\mathbf{S}_{\mathcal{B}}$) are obtained from Section S.1 by letting $r \to \infty$ (if necessary). There may exist up to two internal equilibria which can be determined explicitly:

$$\hat{p}^{\pm} = 1 - \frac{\gamma-\beta}{2\gamma}(1 \mp R)\,,\ \hat{q}^{\pm} = \frac{\alpha}{2\gamma}(1 \mp R)\,, \tag{S.91}$$

where

$$R = \sqrt{1 - \frac{4m\gamma}{\alpha(\gamma - \beta)}}\,. \tag{S.92}$$

The equilibrium $(\hat{p}^+, \hat{q}^+)$ corresponds to $\mathbf{I}_{\mathrm{DMI}}$, $(\hat{p}^-, \hat{q}^-)$ to $\mathbf{I}_0$. The first one is asymptotically stable whenever it exists, the second one is unstable. It is easy to check directly, but also a straightforward consequence of Theorem S.4, that $(\hat{p}^+, \hat{q}^+)$ is admissible if and only if

$$\text{selection scenario 1 or 2 applies and } m \le m^*\,, \tag{S.93a}$$

$$\text{selection scenario 3 applies and } m \le \begin{cases} m_{\mathcal{A}} & \text{if } \gamma \le \frac{1}{2}\alpha\,, \\ m^* & \text{if } \gamma > \frac{1}{2}\alpha\,, \end{cases} \tag{S.93b}$$

$$\text{selection scenario 4 applies and } m \le \begin{cases} m_{\mathcal{B}} & \text{if } \gamma \le -\beta\,, \\ m^* & \text{if } \gamma > -\beta\,, \end{cases} \tag{S.93c}$$

where we have $m_{\mathcal{A}} = (\alpha - \gamma)(\gamma - \beta)/\alpha$, $m_{\mathcal{B}} = -\alpha\beta/(\gamma - \beta)$, and $m^* = \frac{1}{4}\alpha(1 - \beta/\gamma)$.

$(\hat{p}^-, \hat{q}^-)$ is admissible if and only if

$$\gamma > \max[\tfrac{1}{2}\alpha, -\beta] \text{ and } \max[m_{\mathcal{A}}, m_{\mathcal{B}}] < m < m^*\,. \tag{S.94}$$

Then, and only then, both internal equilibria coexist. In this case, we have $\beta/\gamma < \hat{p}^- < \hat{p}^+ < 1$.

By the studying the nullclines of (S.90a), we can exclude the existence of periodic orbits. We have $\dot{p} = 0$ if and only if $p = 0$ or $q = K_A(p)$, where

$$K_A(p) = \frac{\alpha(1-p) - m}{\gamma(1-p)}\,, \tag{S.95}$$

and we have $\dot{q} = 0$ if and only if $q = 1$ or $q = K_B(p)$, where

$$K_B(p) = \frac{m}{\gamma p - \beta}\,. \tag{S.96}$$

The intersection points of K_A and K_B in the (open) square $(0,1) \times (0,1)$ are the internal equilibria. It is easy to check that both K_A and K_B are strictly monotone decreasing in p, that K_A is convex, and K_B is concave. Whether they intersect once or twice, the (two or three) areas between them are always positively invariant. This excludes the existence of periodic orbits because a periodic orbit has to enclose an equilibrium.

The domain of attraction of $\mathbf{I}_{\mathrm{DMI}}$ contains at least all initial conditions with $p > \hat{p}^+$ and $q < \hat{q}^+$. In particular, this includes the *island – continent*, *continent – continent* and *secondary contact* scenarios. In contrast, all points in the region $p < \hat{p}^+$ and $q > \hat{q}^+$ are outside of the attraction domain and converge to a boundary equilibrium. This region includes the starting conditions for the *continent – island* and the *island – island* scenario.

From the coordinates of the equilibrium in terms of haplotype frequencies we obtain the level of population differentiation,

$$\hat{x}_3 = \frac{-\alpha\beta+\beta\gamma+\gamma^2}{2\gamma^2} + \frac{\alpha\beta-\beta\gamma+\gamma^2}{2\gamma^2}\sqrt{1+\frac{4m\gamma}{\alpha\beta-\alpha\gamma}} - 2\frac{m}{\gamma}. \tag{S.97}$$

In the limit of strong epistasis, $\gamma \to \infty$, this expression simplifies to

$$\hat{x}_3^* = \frac{1}{2}\left(1+\sqrt{1-\frac{4m}{\alpha}}\right). \tag{S.98}$$

S.8.3 Strong recombination

By assuming that both selection and migration are weak relative to recombination, we can derive a quasi-linkage-equilibrium (QLE) approximation for the internal equilibria. Formally, we fix $r > 0$ and assume $\alpha = \epsilon a$, $\beta = \epsilon b$, $\gamma = \epsilon c$, and $m = \epsilon\mu$, where a, b, c, and μ are constants and $\epsilon \to 0$. It is straightforward to verify the results stated below by using an algebraic formula manipulation program (see the supporting online *Mathematica* notebook).

To first order in $1/r$, the coordinates of the two possible internal equilibria are

$$\begin{aligned}\hat{p}_\pm &= 1 - \frac{\gamma-\beta}{2\gamma}(1\mp R) && \text{(S.99a)}\\ &\quad - \frac{\gamma-\beta}{2\gamma^2 r}\left[2\alpha\beta+\gamma(m-\beta) \pm R\alpha\frac{\beta(\gamma-\beta)(2\alpha-\gamma)+m\gamma(\gamma-5\beta)}{4m\gamma-\alpha(\gamma-\beta)}\right]\\ \hat{q}_\pm &= \frac{\alpha}{2\gamma}(1\mp R) && \text{(S.99b)}\\ &\quad + \frac{\alpha}{2\gamma^2 r}\left[(\gamma-\alpha)(\gamma-2\beta)+m\gamma \pm R(\gamma-\beta)\frac{\alpha(\gamma-\alpha)(\gamma-2\beta)+m\gamma(5\alpha-4\gamma)}{4m\gamma-\alpha(\gamma-\beta)}\right]\\ \hat{D}_\pm &= -\frac{\alpha\beta(\gamma-\alpha)(\gamma-\beta)}{2\gamma^3 r}(1\mp R)\\ &\quad - \frac{m}{r}\left[1-\frac{\beta}{\gamma}-\frac{\alpha}{2\gamma}\left(1-\frac{3\beta}{\gamma}\right) \pm \frac{\alpha}{2\gamma}\left(1-\frac{\beta}{\gamma}\right)R\right], && \text{(S.99c)}\end{aligned}$$

where we could return to the original parameters because they occur only as ratios. The stable equilibrium is $(\hat{p}_+, \hat{q}_+, \hat{D}_+)$. Because the QLE approximation is the result of a singular perturbation (Bürger 2009)), it is necessary to develop $\hat{p}_\pm$ and $\hat{q}_\pm$ to order r^{-2} to obtain $\hat{D}_\pm$ to order r^{-1}.

The QLE approximation (S.38) for m^* is readily obtained from (S.99) by solving $\hat{p}_+ = \hat{p}_-$ for m.

File S2: Analysis of the diploid model

By assuming Hardy-Weinberg proportions at every time instance, which requires sufficiently weak evolutionary forces, the structure of the dynamical equations for the diploid model becomes the same as that for the haploid model, Eq. (1),

$$\begin{aligned} \dot{x}_1 &= x_1(w_1 - \overline{w} - m) - rD\,, \quad & \dot{x}_2 &= x_2(w_2 - \overline{w} - m) + rD + m\,, \\ \dot{x}_3 &= x_3(w_3 - \overline{w} - m) + rD\,, \quad & \dot{x}_4 &= x_4(w_4 - \overline{w} - m) - rD\,, \end{aligned} \tag{S.100}$$

but with the marginal fitness values

$$\begin{aligned} w_1 &= \beta x_2 + \alpha x_3 + (\alpha + \beta - \gamma_1)x_4, \\ w_2 &= \beta + \beta x_2 + (\alpha - \gamma_1)x_3 + (\alpha + \beta - \gamma_2)x_4, \\ w_3 &= \alpha + (\beta - \gamma_1)x_2 + \alpha x_3 + (\alpha + \beta - \gamma_3)x_4, \\ w_4 &= \alpha + \beta - \gamma_1 + (\beta + \gamma_1 - \gamma_2)x_2 + (\alpha + \gamma_1 - \gamma_3)x_3 + (\alpha + \beta + \gamma_1 - \gamma_4)x_4, \end{aligned} \tag{S.101}$$

and the mean fitness

$$\overline{w} = 2\big(\beta x_2 + \alpha x_3 + (\alpha + \beta - \gamma_1)x_4 - \gamma_1 x_2 x_3 + (\gamma_1 - \gamma_2)x_2 x_4 + (\gamma_1 - \gamma_3)x_3 x_4 + (\gamma_1 - \gamma_4/2)x_4^2\big). \tag{S.102}$$

For the special case $\Gamma = (\gamma/2, \gamma, \gamma, 2\gamma)$ and in linkage equilibrium ($x_1 x_4 = x_2 x_3$), we obtain $\overline{w} = 2\big(\beta x_2 + \alpha x_3 + (\alpha + \beta - \gamma)x_4\big)$. In this case, the dynamical equations for the two-dimensional system with $p = x_3 + x_4$ and $q = x_2 + x_4$ reduces precisely to the corresponding equations for the haploid model. For all other choices of epistasis parameters, the dynamics of the diploid model is much more complex. Therefore, our analytical results are incomplete. They are presented below and, in more extensive form, in the supplementary *Mathematica* OS.

S.9 Existence and linear stability of boundary equilibria

For $m = 0$, all four monomorphic equilibria exist and stability conditions can readily be derived (see the *Mathematica* OS). We find the following conditions for asymptotic stability:

$$\mathbf{M}_1: \quad \alpha < 0,\ \beta < 0,\ \text{and } r > \alpha + \beta - \gamma_1\,, \tag{S.103a}$$

$$\mathbf{M}_2: \quad \alpha < \gamma_2,\ \beta > 0,\ \text{and } r > \alpha - \beta - \gamma_1\,, \tag{S.103b}$$

$$\mathbf{M}_3: \quad \alpha > 0,\ \gamma_3 > \beta,\ \text{and } r > \beta - \alpha - \gamma_1\,, \tag{S.103c}$$

$$\mathbf{M}_4: \quad \alpha > \gamma_4 - \gamma_2,\ \beta > \gamma_4 - \gamma_3,\ \text{and } r > \gamma_4 - \gamma_1 - \alpha - \beta\,. \tag{S.103d}$$

Compared with the haploid case, two additional boundary equilibria, denoted by $\tilde{\mathbf{S}}_{\mathcal{A}}$ and $\tilde{\mathbf{S}}_{\mathcal{B}}$, can exist for the diploid model if $m = 0$. These equilibria exist if locus $\mathcal{A}$ or $\mathcal{B}$ are over- or underdominant when the derived allele (B or A) at the other locus is fixed. They can be stable only in the case of overdominance. In particular, an overdominant equilibrium $\tilde{\mathbf{S}}_{\mathcal{A}}$ with coordinates $p = (\gamma_4 - \gamma_2 - \alpha)/(\gamma_4 - 2\gamma_2)$ and $q = 1$ exists if $\gamma_2 < \alpha < \gamma_4 - \gamma_2$. An overdominant boundary equilibrium $\tilde{\mathbf{S}}_{\mathcal{B}}$ with $p = 1$ and $q = x_4 = (1 - x_3) = (\gamma_4 - \gamma_3 - \beta)/(\gamma_4 - 2\gamma_3)$ exists if $\gamma_3 < \beta < \gamma_4 - \gamma_3$. Note that overdominance is only a necessary, but not sufficient condition for stability of $\tilde{\mathbf{S}}_{\mathcal{A}}$ or $\tilde{\mathbf{S}}_{\mathcal{B}}$ in the full model. Sufficient conditions can be derived from the Jacobian (see the *Mathematica* OS) but result in complicated cubic equations for the eigenvalues.

If $m > 0$, boundary equilibria can exist only at the edges $p = 0$ (implying $x_3 = x_4 = D = 0$) or $q = 1$ (i.e., $x_1 = x_3 = D = 0$) of the state space. The only monomorphic equilibrium to fulfill this condition is $\mathbf{M}_2$ (fixation of the continental haplotype). For an arbitrary point on the edge $x_3 = x_4 = 0$, the eigenvalues of the Jacobian are as follows:

$$\begin{aligned}\lambda_1 &= \beta - 2\beta x_2 - m \\ \lambda_{2,3} &= \frac{1}{2}\Big(\beta + 2\alpha - r - \gamma_1 - 2\beta x_2 - \gamma_2 x_2 - 2m \\ &\quad \pm \sqrt{(\beta - r - \gamma_1 + 2\gamma_1 x_2 - \gamma_2 x_2)^2 + 4 r x_2 (\beta - \gamma_1 + 2\gamma_1 x_2 - \gamma_2 x_2)}\Big)\end{aligned} \tag{S.104}$$

Setting $x_2 = 1$, we obtain the stability conditions for $\mathbf{M}_2$:

$$m > \max[-\beta, \alpha - \beta - \gamma_1 - r, \alpha - \gamma_2]. \tag{S.105}$$

As in the haploid case, these conditions correspond to the invasion criteria for the three other haplotypes, *ab*, *Ab*, and *AB*. Additional boundary equilibria, with a single polymorphic locus, can exist if $m > 0$. In Section S.13 below, we analyze the stability of these equilibria for a particular choice of the epistasis parameters.

S.10 Global stability of the boundary

Here, we derive sufficient conditions for global convergence to the boundary. Alternatively, these may be viewed as necessary conditions for the existence of a DMI (i.e., a stable internal equilibrium). The following assumptions on epistasis will be required:

$$0 \le \gamma_1 \le \min[\gamma_2, \gamma_3] \quad \text{and} \quad \max[\gamma_2, \gamma_3] \le \gamma_4\,, \tag{S.106}$$

$$\gamma_1 + \gamma_4 \ge \gamma_2 + \gamma_3\,. \tag{S.107}$$

Clearly, the recessive and the co-dominant model with $\Gamma_0 := (0, \gamma, \gamma, 2\gamma)$ and $\Gamma_1 := (\gamma/2, \gamma, \gamma, 2\gamma)$ satisfy these assumptions. In addition, we shall need the following condition on the selection parameters:

$$\beta \ge \max[\gamma_2, \gamma_3, \gamma_4 - \gamma_1]\,. \tag{S.108}$$

Throughout the following, we always assume (S.106). The results below hold whenever $m \ge 0$ and $r > 0$.

Theorem S.6. *Each of the following assumptions ensures that all trajectories converge to an equilibrium at the boundary of the state space S_4:*

(i)

$$\alpha \leq 0 \text{ and } (\text{S.107})\,; \tag{S.109}$$

(ii)

$$\alpha \geq 0 \text{ and } (\text{S.107}) \text{ and } (\text{S.108})\,; \tag{S.110}$$

(iii)

$$m \geq \max[\tfrac{1}{4}\alpha, \alpha + \gamma_1 - \beta, \gamma_2 - \gamma_3] \text{ and } (\text{S.107})\,. \tag{S.111}$$

Remark S.5. In (i), assumption (S.107) is needed only for technical reasons. Without it, internal equilibria could be excluded, but not limit cycles or other complex attractors. Importantly, the proof shows that any internal equilibrium, hence every DMI, will be in negative linkage disequilibrium.

It is not difficult to show that the assumption (S.108) in (ii) can be replaced by the weaker assumption

$$\beta \geq \max[\gamma_2 - \alpha, \gamma_3, \gamma_4 - \gamma_3, \gamma_4 - \gamma_1 - \alpha]\,. \tag{S.112}$$

The proof is based on several lemmas.

Lemma S.9. *If $m + \gamma_1 + \gamma_4 \geq \gamma_2 + \gamma_3$, then every ω-limit is contained in the set $\{D \leq 0\}$. Any internal ω-limit points have to satisfy $D < 0$. In particular, the conclusions apply if $m \geq 0$ and (S.107) hold.*

For the proof, define

$$Z = \frac{x_2 x_3}{x_1 x_4}\,. \tag{S.113}$$

Then the lemma follows from the identity

$$\begin{aligned}(x_1x_4)^2\dot{Z} &= rD(x_1x_2x_3 + x_2x_3x_4 + x_1x_2x_4 + x_1x_3x_4) \\ &\quad + x_1x_3x_4\,\{m + x_2[x_2(\gamma_2 - \gamma_1) + x_3(\gamma_3 - \gamma_1) + x_4(\gamma_1 + \gamma_4 - \gamma_2 - \gamma_3)]\}\,. \end{aligned} \tag{S.114}$$

Lemma S.10. *(i) Assume $\alpha \leq 0$ and $D \leq 0$. Then $(x_1/x_3)^{\dot{}} \geq 0$, and the inequality is strict in the interior of the state space.*

This follows immediately from the identity

$$x_3^2 \left(\frac{x_1}{x_3}\right)^{\dot{}} = -\alpha x_1 x_3 - r(x_1 + x_3)D + x_1x_3[\gamma_1 x_2 + (\gamma_3 - \gamma_1)x_4],. \tag{S.115}$$

Lemma S.11. *(i) Assume* $\alpha \leq 0$ *and* $D \geq 0$*. Then* $(x_2/x_4)^{\cdot} \geq 0$*, and the inequality is strict in the interior of the state space.*

This follows immediately from the identity

$$\begin{aligned} x_4^2 \left(\frac{x_2}{x_4}\right)^{\cdot} &= m x_4 + r(x_2+x_4)D - \alpha x_2 x_4 \\ &\quad + x_2 x_4[\gamma_1 x_1 + \gamma_2 x_2 + (\gamma_3-\gamma_1)x_3 + (\gamma_4-\gamma_2)x_4]\,. \end{aligned} \tag{S.116}$$

Lemma S.12. *Assume* $\alpha + \beta \geq \max[\gamma_2, \gamma_3, \gamma_4 - \gamma_1]$*. Then* $(x_1/x_4)^{\cdot} \leq 0$ *if* $D(x_1 - x_4) \leq 0$ *and the set* $\{x_1 \leq x_4\}$ *is forward invariant.*

The proof follows from the identity

$$\begin{aligned} x_4^2 \left(\frac{x_1}{x_4}\right)^{\cdot} &= rD(x_1-x_4) - x_1 x_4[x_1(\alpha+\beta-\gamma_1) + x_2(\alpha+\beta-\gamma_2) \\ &\quad + x_3(\alpha+\beta-\gamma_3) + x_4(\alpha+\beta-\gamma_4+\gamma_1)]\,. \end{aligned} \tag{S.117}$$

Lemma S.13. *Let*

$$Y = \frac{x_1 + x_3}{x_3 + x_4}\,. \tag{S.118}$$

Assume $\alpha > 0$ *and* (S.108)*. Then* $\dot{Y} \leq 0$ *on the set* $\{x_1 \leq x_4\}$*.*

The proof follows from the identity

$$\begin{aligned} (x_3+x_4)^2\,\dot{Y} &= -\alpha x_1(x_3+x_4) + \gamma_1 x_2 x_3(x_1-x_4) - x_4[x_1^2(\beta-\gamma_1) + x_2(x_1+x_3)(\beta-\gamma_2) \\ &\quad + x_3(2x_1+x_3)(\beta-\gamma_3) + x_1 x_4(\beta-\gamma_4+\gamma_1) + x_3 x_4(\beta-\gamma_4+\gamma_3)]\,. \end{aligned} \tag{S.119}$$

Lemma S.14. *Assume* $m \geq \max[\alpha - \beta, \gamma_2 - \gamma_3]$*. Then* $(x_2/x_3)^{\cdot} \geq 0$ *holds on the forward-invariant set* $\{x_2 \geq x_3\} \cap \{D \leq 0\}$*.*

The proof follows from the identity

$$\begin{aligned} x_3^2 \left(\frac{x_2}{x_3}\right)^{\cdot} &= -rD(x_2-x_3) + \gamma_1 x_2 x_3(x_2-x_3) + m x_3(x_1+x_3) + m x_3 x_4(x_1+x_3+x_4) \\ &\quad + (m-\alpha+\beta)x_2 x_3 + (m+\gamma_3-\gamma_2)x_2 x_3 x_4\,. \end{aligned} \tag{S.120}$$

Lemma S.15. *Let*

$$X = \frac{x_1 + x_3}{x_1 + x_2}\,. \tag{S.121}$$

Assume $\alpha > 0$ *and* $m \geq \max[\frac{1}{4}\alpha, \alpha + \gamma_1 - \beta, \gamma_2 - \gamma_3]$*. Then* $\dot{X} \leq 0$ *on the set* $\{x_3 \geq x_2\}$*.*

This statement follows from the identity

$$
\begin{aligned}
(x_1+x_2)^2\,\dot X = -m(x_1+x_3)^2 + \alpha x_1x_3 - (m-\alpha-\gamma_1x_3+\beta)x_2x_3 - (m+\beta)x_1x_2 \\
- (m-\gamma_2+\gamma_3)x_2x_4(x_1+x_3) + \gamma_1x_2^2x_3^2 + (\gamma_3-\gamma_1)x_1x_4(x_3-x_2)
\end{aligned}
\tag{S.122}
$$

because $-m(x_1+x_3)^2+\alpha x_1x_3 \le -\frac{1}{4}\alpha[(x_1+x_3)^2-4x_1x_3] = -\frac{1}{4}\alpha(x_1-x_3)^2 \le 0$.

Proof of Theorem S.6. (i) We start by noting that, because we assume $r > 0$, the only positive-invariant subsets of the boundary are the four edges of the simplex corresponding to fixation of one of the alleles. On these edges, convergence to equilibrium occurs always. Hence, it is sufficient to consider the fate of solutions starting in the interior of the state space.

The first statement follows immediately from Lemmas S.10 and S.11. For the second statement, Lemma S.9 guarantees that every ω-limit is contained in the positive-invariant set $\{D \le 0\}$, whence Lemma S.10 yields the result.

(ii) By Lemma S.9, it is sufficient to show that there are no ω-limit points satisfying $D < 0$. Because $\{D \le 0\}$ is forward invariant, Lemma S.12 shows that all trajectories in $\{x_1 \ge x_4\} \cap \{D \le 0\}$ enter the region $\{x_1 \le x_4\} \cap \{D \le 0\}$ and remain there. Now Lemma S.13 yields the result.

(iii) By Lemma S.9, it is sufficient to show that there are no ω-limit points satisfying $D < 0$. Application of Lemmas S.14 and S.15 yields the assertion. □

Remark S.6. The following simple results can also be proved:

1. If (S.107) and $\alpha+\beta \ge \max[\gamma_2,\gamma_3,\gamma_4-\gamma_1]$ hold, then any DMI (if it exists) satisfies $D<0$ and $x_1 \le x_4$.

2. If (S.107) and $\alpha+\beta \le \min[\gamma_2,\gamma_3,\gamma_4-\gamma_1]$ hold, then any DMI (if it exists) satisfies $D<0$ and $x_1 \ge x_4$.

Under the assumption of linkage equilibrium, the dynamical equation for p can be written as

$$
\begin{aligned}
\dot p = p\Big(\alpha - m - \alpha p - q(1-p)\big[2\gamma_1(1-q)(1-p) + \gamma_2 q(1-p) \\
+ 2(\gamma_3-\gamma_1)p(1-q) + (\gamma_4-\gamma_2)pq\big]\Big).
\end{aligned}
\tag{S.123}
$$

We thus see that p is a Lyapunov function and $p(t) \to 0$ as $t \to \infty$ if $\gamma_4 \ge \gamma_2$, $\gamma_3 \ge \gamma_1$, and $m > \alpha$ (here it is not necessary to assume (S.106), (S.107), or (S.108)).

S.11 Tight linkage

Arguments analogous to those in Section S.8.1 yield that, if $r = 0$,

$$V = V(x_1, x_2, x_3, x_4) = \tfrac{1}{2}\overline{w} + m \ln x_2 \tag{S.124}$$

is a Lyapunov function, whence the dynamics is gradient-like. Again, the proof of Theorem 2.3 in NAGYLAKI *et al.* (1999) applies and shows that for sufficiently small r every trajectory converges to an equilibrium point (provided every equilibrium of the unperturbed system is hyperbolic). For small r, the asymptotically stable equilibria are obtained by perturbation of the asymptotically stable equilibria for $r = 0$. In particular, if there is a unique asymptotically stable equilibrium for $r = 0$, then its perturbation is globally asymptotically stable for sufficiently small r. It is worth emphasizing that this holds for arbitary fitness assignments (such that all equilibria are hyperbolic).

S.12 Internal equilibria: Weak migration

In addition to our results from Lyapunov functions, we obtain conditions for the existence of a DMI under weak migration from perturbation arguments. On the one hand, the existence of a stable equilibrium with positive frequency of the island haplotype ($x_3 > 0$) for $m = 0$ implies the existence of a DMI for sufficiently small $m > 0$. On the other hand, if a stable equilibrium with $x_3 > 0$ does not exist if $m = 0$, a DMI cannot be maintained for small m.

Extensive computer simulations confirm that there are always uniquely defined maximum rates of gene flow, $m^{\pm}_{\text{max}}$, which separate the domains in which a locally or globally stable DMI exists (if $m < m^{\pm}_{\text{max}}$) from the region in which a DMI does not exist ($m > m^{\pm}_{\text{max}}$). We note that although this conjecture is highly plausible, we do not have a rigorous proof beyond the cases treated in the previous section. These arguments imply that $m^{+}_{\text{max}} > 0$ if there is a stable equilibrium with $x_3 > 0$ for $m = 0$, and $m^{+}_{\text{max}} = 0$, otherwise. Similarly, we can conclude that $m^{-}_{\text{max}} > 0$ if the only stable equilibrium for $m = 0$ is one with $x_3 > 0$.

Our numerical analysis of the diploid model (see the *Mathematica* OS) indicates that for $m = 0$ there are no stable equilibria in the interior of the state space. As we have seen above, two boundary equilibria with $x_3 > 0$ may exist if $m = 0$: the monomorphic equilibrium $\mathbf{M}_3$ and the single-locus polymorphism $\tilde{\mathbf{S}}_{\mathcal{B}}$. A necessary condition for the asymptotic stability of at least one of these boundary equilibria is

$$\beta < \max[\gamma_3, \gamma_4 - \gamma_3]\,. \tag{S.125}$$

Given our numerical evidence, this translates into a necessary condition for a stable DMI (hence, for $m^{+}_{\text{max}} > 0$) with weak migration. We note that this condition is slightly more stringent than the negation of (S.108). In addition, we already know from Theorem S.6 that $\alpha > 0$ is another necessary condition.

Sufficient conditions for $m^{+}_{\max} > 0$ can, in principle, be derived from a full linear stability analysis of $\mathbf{M}_3$ and $\tilde{\mathbf{S}}_{\mathcal{B}}$ for $m = 0$. In general, however, the conditions for the stability of $\tilde{\mathbf{S}}_{\mathcal{B}}$ are not transparent. For the special case of independent loci ($r \to \infty$), we find that $\tilde{\mathbf{S}}_{\mathcal{B}}$ is stable if and only if $\gamma_3 < \beta < \gamma_4 - \gamma_3$ and $\alpha > \alpha^*$, where

$$\alpha^* = \frac{(\beta - \gamma_3)[2(\gamma_3 - \gamma_1)(\gamma_4 - \gamma_3 - \beta) + (\beta - \gamma_3)(\gamma_4 - \gamma_2)]}{(\gamma_4 - 2\gamma_3)^2}\,. \tag{S.126}$$

Note that $\alpha^* > 0$ if $\gamma_3 < \beta < \gamma_4 - \gamma_3$.

If we exclude overdominance at the locus $\mathcal{B}$ (and, hence, $\tilde{\mathbf{S}}_{\mathcal{B}}$ can not be stable) further analytical results can be derived. In particular,for weak migration, any stable DMI must be a perturbation of the island equilibrium $\mathbf{M}_3$. A perturbation analysis to first order in m yields the coordinates $\{x_1, x_2, x_3\}$ of $\mathbf{PM}_3$:

$$\mathbf{PM}_3 : \left\{ \frac{mr}{\alpha\,(r + \alpha - \beta + \gamma_1)}, \frac{m}{r + \alpha - \beta + \gamma_1}, 1 - \frac{m\,(r\alpha + (r + \alpha)\,(\gamma_3 - \beta))}{\alpha\,(r + \alpha - \beta + \gamma_1)\,(\gamma_3 - \beta)} \right\}. \tag{S.127}$$

The coordinate x_3 measures the level of population differentiation which is analyzed further in the main text.

For all our analytical and numerical derivations, we have assumed that the epistasis coefficients are non-decreasing with the number of A/B conflicts, i.e., $0 \leq \gamma_1 \leq \min[\gamma_2, \gamma_3]$ and $\max[\gamma_2, \gamma_3] \leq \gamma_4$. This condition is essential for our results. Additional internal equilibria can be found for other choices of the γ_i. In particular, we found numerically that stable internal equilibria are possible for $\alpha = \beta = 0$ (and even for $\alpha < 0$) if $\gamma_1 \ll \min[\gamma_3, \gamma_4]$ (for examples see the *Mathematica* OS).

S.13 Stability conditions for the recessive and the co-dominant model

Here, we consider the diploid model with the simplified epistasis scheme $\Gamma = (\gamma_1, \gamma, \gamma, 2\gamma)$. Obviously, both the recessive model with $\Gamma_0 = (0, \gamma, \gamma, 2\gamma)$ and the co-dominant model with $\Gamma_1 = (\gamma/2, \gamma, \gamma, 2\gamma)$ follow this scheme. In the following lemma, we collect some important consequences for this choice of the epistasis parameters. They all follow from our above analysis and elementary calculations.

Lemma S.16. *Elementary facts for the diploid model with $\Gamma = (\gamma_1, \gamma, \gamma, 2\gamma)$.*

1. *Overdominance or underdominance of single loci does not occur.*
2. *The model is symmetric under the exchange of the haplotypes ab and AB. As in the haploid case, this symmetry is reflected by an invariance of the model under the transformation $\alpha \to \gamma - \beta$ and $\beta \to \gamma - \alpha$.*

3. $\alpha > 0$ and $\gamma > \beta$ are necessary conditions for the existence of a DMI.

In the absence of overdominance, we find that the boundary equilibria $\mathbf{S}_{\mathcal{A}}$ and $\mathbf{S}_{\mathcal{B}}$ exist under the same conditions and with the same coordinates as in the haploid case (Section S.1). The stability conditions are different, however. Evaluation of (S.104) at $x_2 = -m/\beta$ leads to the stability conditions for $\mathbf{S}_{\mathcal{B}}$. In analogy to (S.10), explicit bounds on the migration rate m can be derived, however, the expressions are lengthy. Due to the model symmetry, the stability conditions for the $\mathbf{S}_{\mathcal{A}}$ equilibrium can be obtained directly from the corresponding conditions for $\mathbf{S}_{\mathcal{B}}$. For explicit results, we focus on the limiting cases with $r = 0$ and $r \to \infty$. We always assume $\alpha > 0$ and $\gamma > \beta$.

If $r = 0$, the equilibrium $\mathbf{S}_{\mathcal{B}}$ is asymptotically stable if and only if

$$\begin{aligned}-\beta > m > \max\Bigg[&\frac{-\beta(\beta + 2\alpha - \gamma_1)}{\gamma},\\ &\frac{-\beta\Big(\alpha\gamma + \beta\gamma_1 - \gamma_1^2 + \sqrt{(\alpha\gamma + \beta\gamma_1 - \gamma_1^2)^2 - 4\alpha\gamma_1(\gamma - \gamma_1)(\beta + \alpha - \gamma_1)}\Big)}{2\gamma_1(\gamma - \gamma_1)}\Bigg].\end{aligned} \tag{S.128}$$

For independent loci ($r \to \infty$), and assuming $\gamma \geq \gamma_1$, the stability condition for $\mathbf{S}_{\mathcal{B}}$ can be expressed as

$$-\beta > m > \frac{-\beta\Big(\beta - 2\gamma_1 + \sqrt{(2\gamma_1 - \beta)^2 + 4\alpha(\gamma - 2\gamma_1)}\Big)}{2(\gamma - 2\gamma_1)}, \tag{S.129}$$

For the recessive model, in which $\gamma_1 = 0$, we can summarize the stability conditions for the boundary equilibria as follows. If $r = 0$, condition (S.128) can never be fulfilled and the equilibria $\mathbf{S}_{\mathcal{A}}$ and $\mathbf{S}_{\mathcal{B}}$ are never stable. The monomorphic equilibrium $\mathbf{M}_2$ is stable if and only if $m > \alpha - \beta$. For $r \to \infty$, the stability conditions for the three boundary equilibria are

$$\mathbf{M}_2: \quad m > \max[-\beta, \alpha - \gamma], \tag{S.130a}$$

$$\mathbf{S}_{\mathcal{A}}: \quad \alpha - \gamma > m > (\alpha - \gamma)\frac{\gamma - \alpha + \sqrt{(\gamma - \alpha)^2 + 4\gamma(\gamma - \beta)}}{2\gamma}, \tag{S.130b}$$

$$\mathbf{S}_{\mathcal{B}}: \quad -\beta > m > -\beta\,\frac{\beta + \sqrt{\beta^2 + 4\alpha\gamma}}{2\gamma}. \tag{S.130c}$$

For the co-dominant model, in which $\gamma_1 = \gamma/2$, we obtain for $r = 0$:

$$\mathbf{M}_2: \quad m > \max[-\beta, \alpha - \beta - \gamma/2, \alpha - \gamma], \tag{S.131a}$$

$$\mathbf{S}_{\mathcal{A}}: \quad \alpha - \gamma > m > \frac{2(\gamma - \beta)(\alpha - \gamma)}{\gamma}, \tag{S.131b}$$

$$\mathbf{S}_{\mathcal{B}}: \quad -\beta > m > \frac{-2\alpha\beta}{\gamma}. \tag{S.131c}$$

As noted above, the dynamics of the co-dominant diploid model coincides with the haploid dynamics in the limit $r \to \infty$. Thus, also the stability conditions and maximum rates of gene flow coincide and can be taken from there (Eqs. (10)–(12) in the main text). Using our results for the stability of boundary equilibria for the recessive and the co-dominant model, the analytical results for the maximum rate of gene flow $m^-_{\max}$ in the Section RESULTS: DIPLOID MODEL of the main text can easily be derived.

In addition, explicit results for $m^+_{\max}$ for the co-dominant model with tight linkage can be deduced. Their derivation is provided in the supplementary *Mathematica* OS.

SUPPLEMENTARY INFORMATION: REFERENCES

Bürger, R. 2000. The Mathematical Theory of Selection, Recombination, and Mutation. Wiley, Chichester.

Bürger, R. 2009. Multilocus selection in subdivided populations I. Convergence properties for weak or strong migration. J. Math. Biol. 58, 939-978.

Bürger, R., A. Akerman. 2011. The effects of linkage and gene flow on local adaptation: A two-locus continent-island model. Theor. Popul. Biol. 80, 272-288.

Feldman, M. 1971. Equilibrium studies of two locus haploid populations with recombination. Theor. Popul. Biol. 2, 299-318.

Hofbauer, J. 1985. The selection mutation equation. J. Math. Biol. 23, 41–53.

Hofbauer, J. 1990. An index theorem for dissipative semiflows. Rocky Mountain J. 20, 1017-1031.

LaSalle, J.P. 1976. The Stability of Dynamical Systems. SIAM, Philadelphia.

Karlin, S., McGregor, J. 1972. Polymorphism for genetic and ecological systems with weak coupling. Theor. Popul. Biol. 3, 210-238.

Nagylaki, T., Hofbauer, J., Brunovský, P. 1999. Convergence of multilocus systems under weak epistasis or weak selection. J. Math. Biol. 38, 103-133.

Passekov, V.P. 1978. On the relation between steady-state patterns of the gene frequencies' distribution and extension of the Fisher fundamental theorem for a single multi-allelic locus. Doklady MOIP, Obschaya Biologiya, 1978, pp. 135–138 (in Russian).

Rutschman, D.H. 1994. Dynamics of the two-locus haploid model. Theor. Popul. Biol. 45, 167-176.

File S3: Stability regions of DMI equilibria in the diploid model

In this section of the OS, we show the DMI pattern for several different diploid models. All plots show the scaled maximum migration rates for local stability (solid lines) and the areas of global stability (shaded) against the scaled strength of the incompatibility, γ/α – similar to figure 3, but on a slightly different scale. The data was obtained numerically using the functions in the *Mathematica* OS. For better comparison, plots of the haploid and the recessive and codominant diploid model treated in the main text are also included.

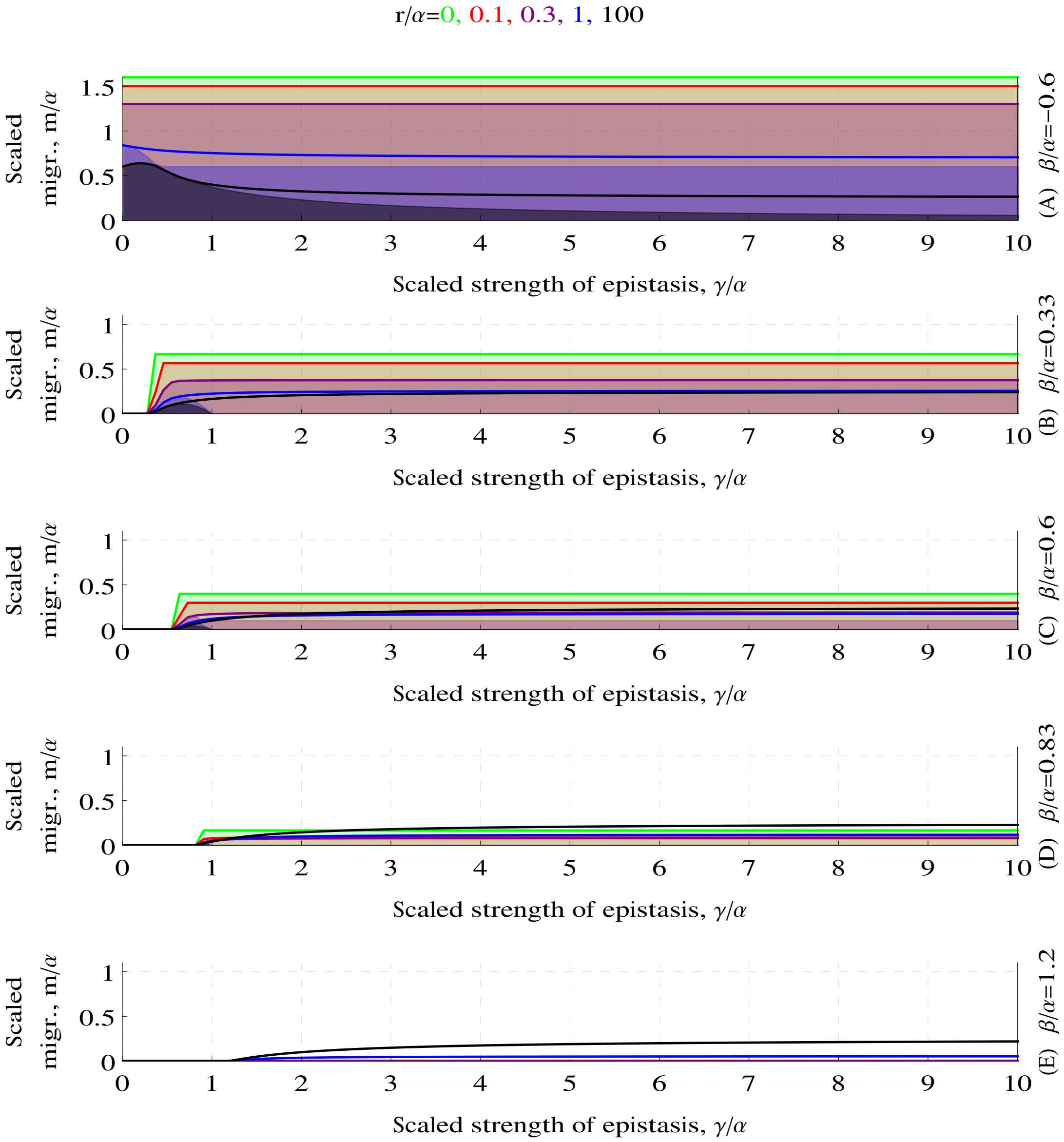

Figure S1: Haploid model

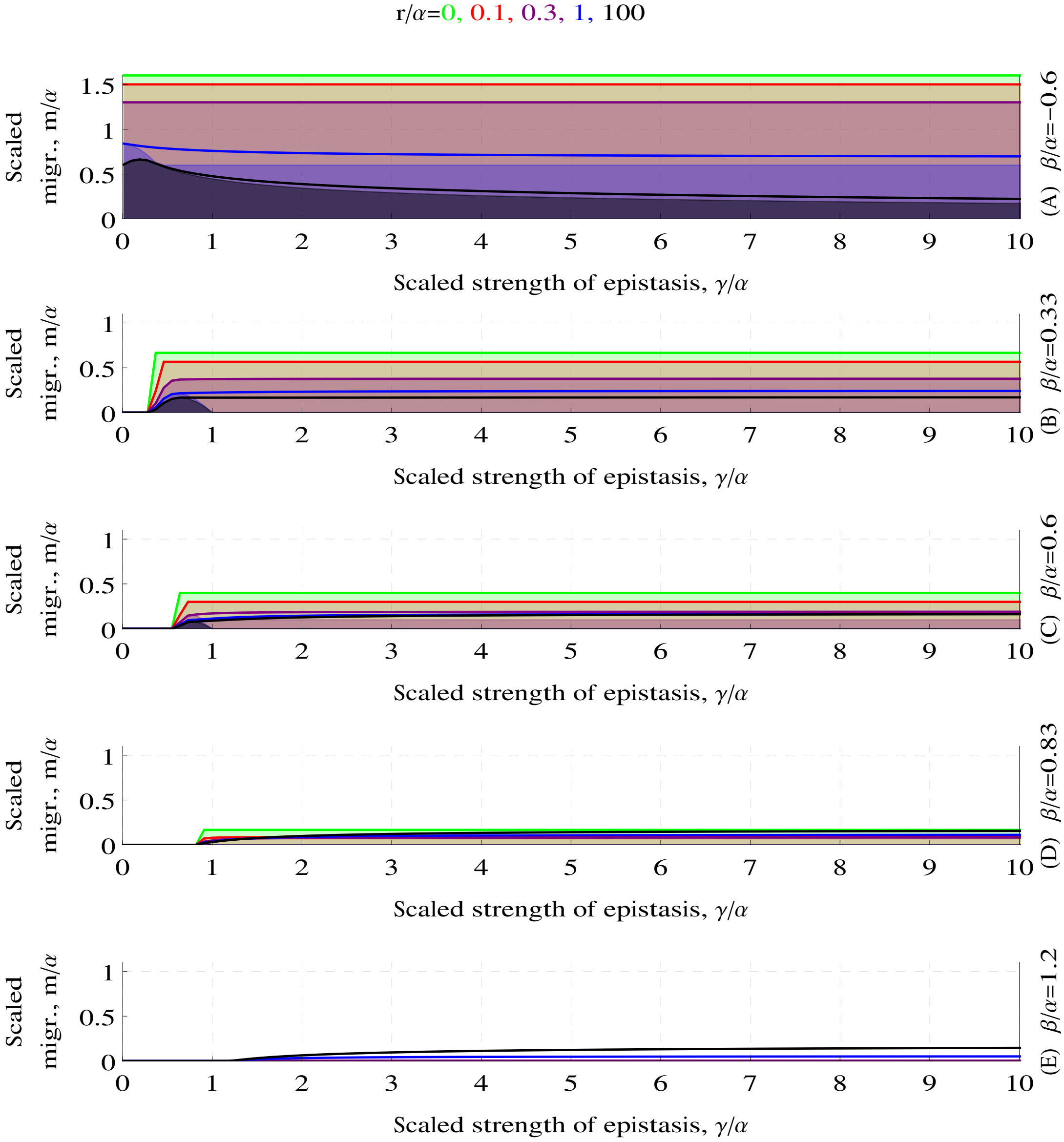

Figure S2: Recessive model, $\Gamma = (0, \gamma, \gamma, 2\gamma)$

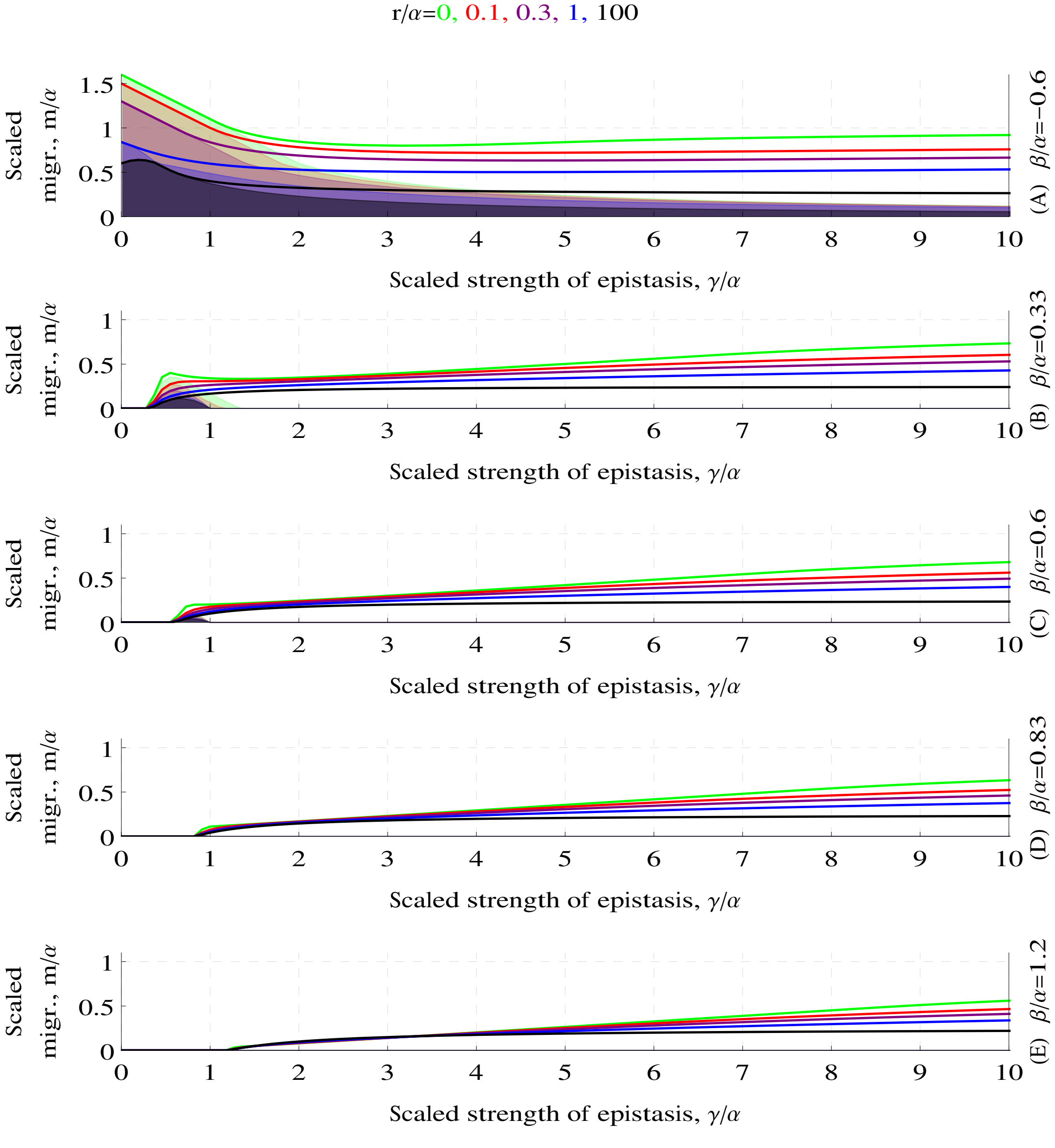

Figure S3: Codominant model, $\Gamma = (\gamma/2, \gamma, \gamma, 2\gamma)$.

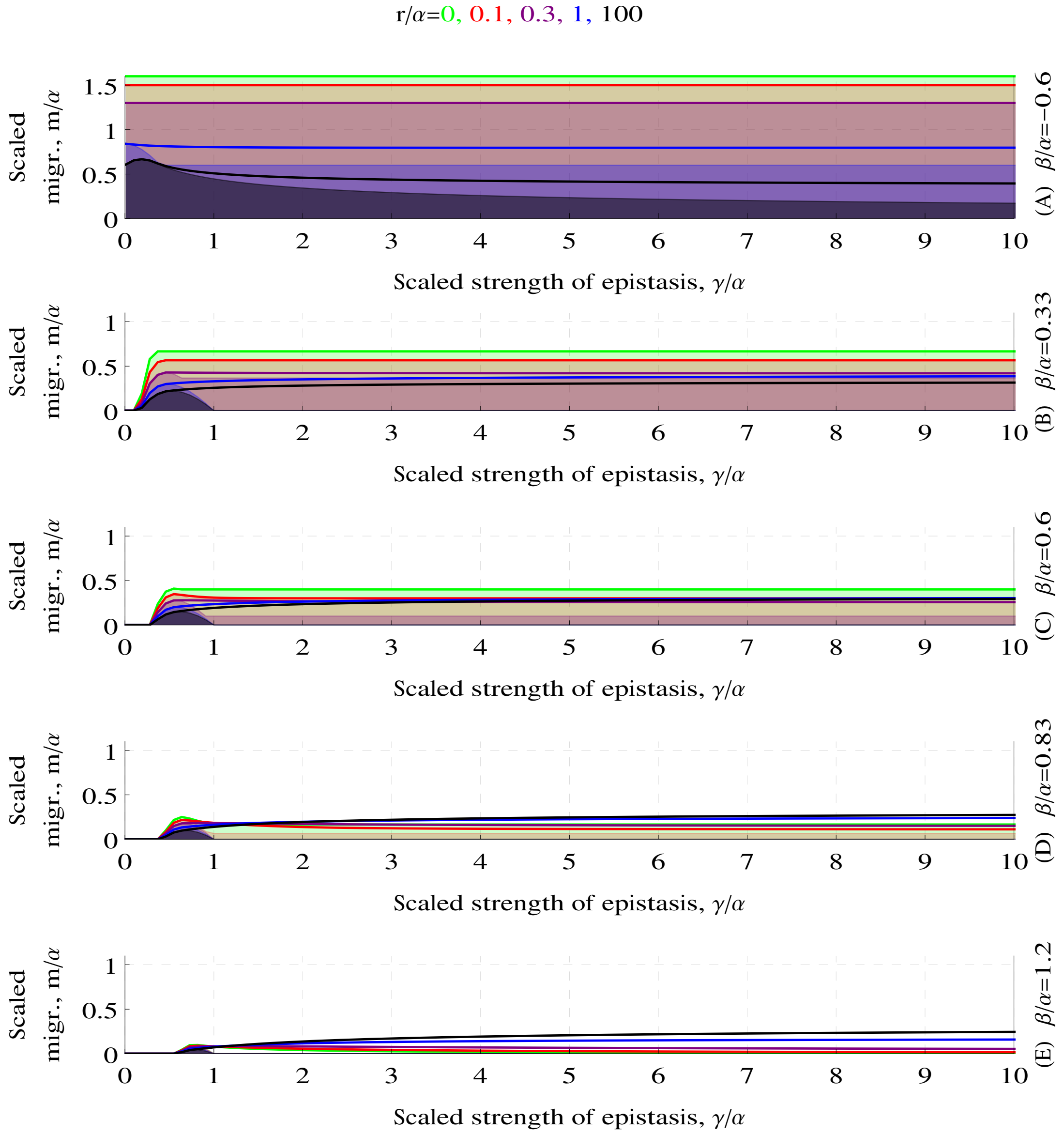

Figure S4: $\Gamma = (0, \gamma, 0, 2\gamma)$: very similar to $\Gamma = (0, \gamma, \gamma, 2\gamma)$ (see Fig. S2) and thus to the haploid case, but for $\beta > 0$ with overdominance at the $\mathcal{B}$ locus as soon as $2\gamma > \beta$. As a consequence, the minimal γ to maintain a stable DMI is reduced to $\beta/2$. Also, a globally stable DMI may exist even for $\beta > \alpha$, since the overdominant genotype $AABb$ is the fittest genotype for $\beta < 2\alpha$.

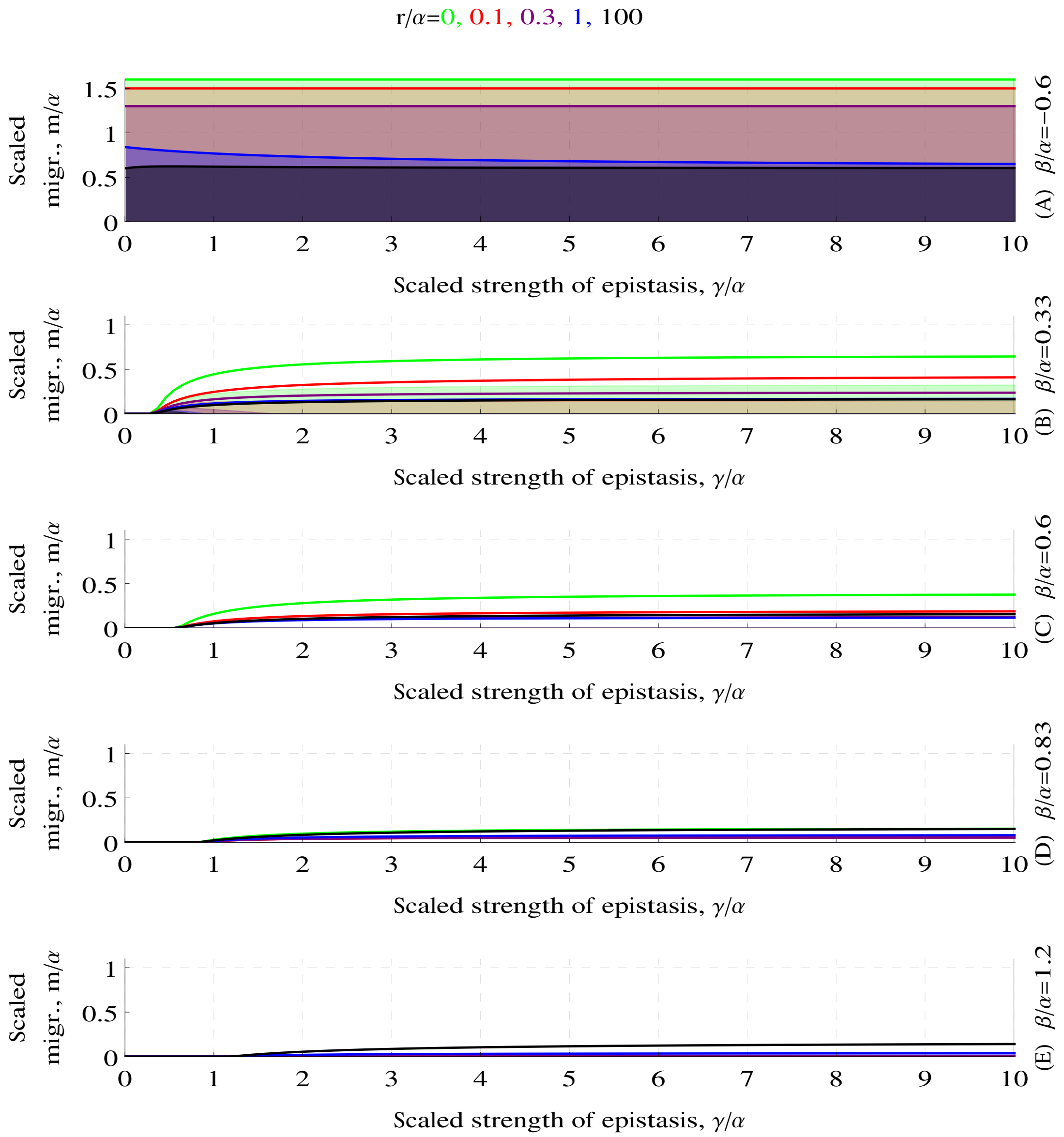

Figure S5: $\Gamma = (0, 0, \gamma, 2\gamma)$: still similar to $\Gamma = (0, \gamma, \gamma, 2\gamma)$ (see Fig. S2) and the haploid model, but now with overdominance at the $\mathcal{A}$ locus for $2\gamma > \alpha$. Due to the high fitness of the $AaBB$ genotype in this case, the stability regions for a DMI are somewhat reduced. In particular, global stability for $\beta > 0$ becomes more difficult.

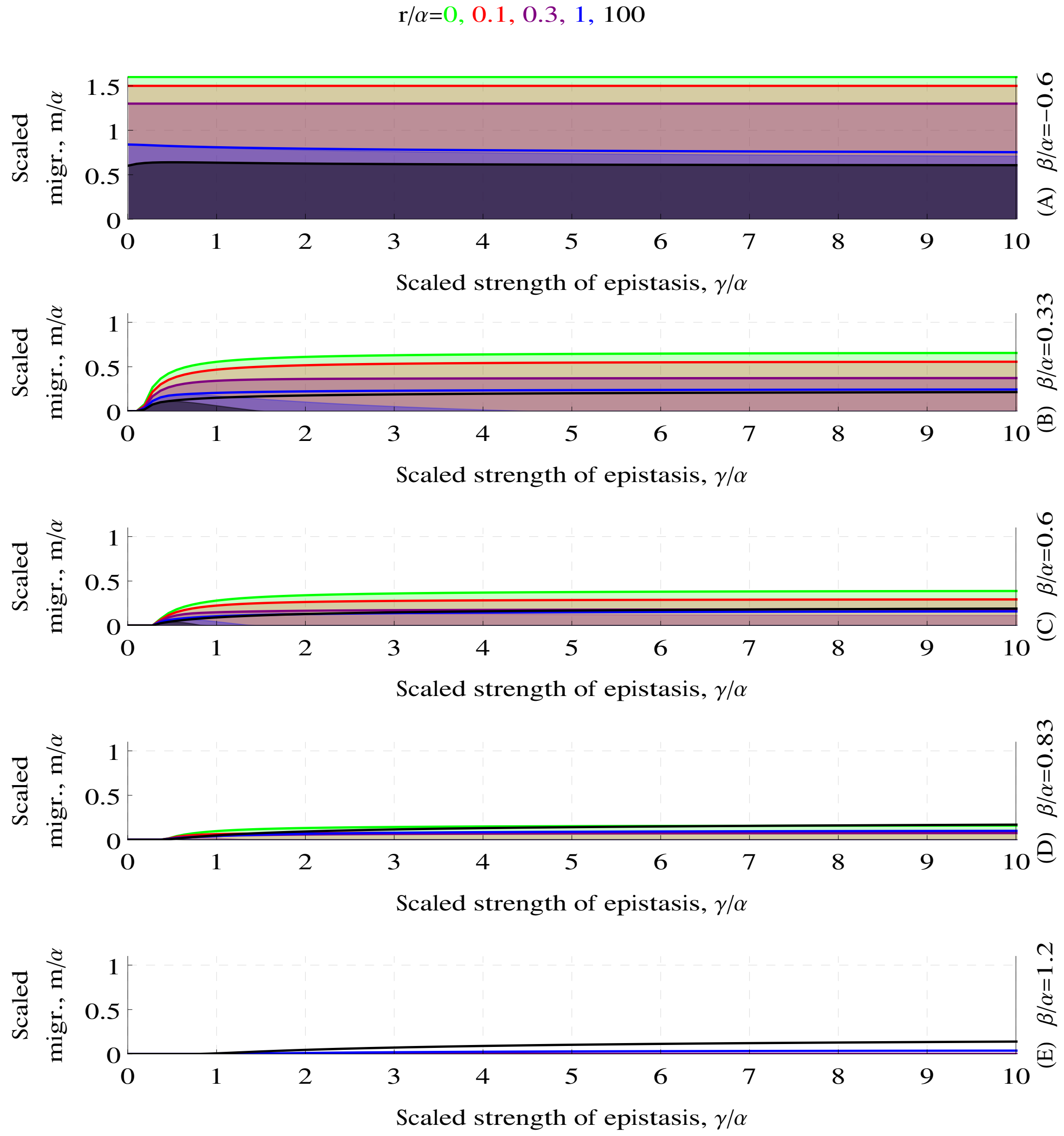

Figure S6: $\Gamma = (0, 0, 0, 2\gamma)$: the overdominance effects at the $\mathcal{A}$ locus (disfavoring DMIs) and at the $\mathcal{B}$ locus (favoring DMIs) partly cancel, resulting in stability regions that are once again similar to the haploid model (Figure S2).

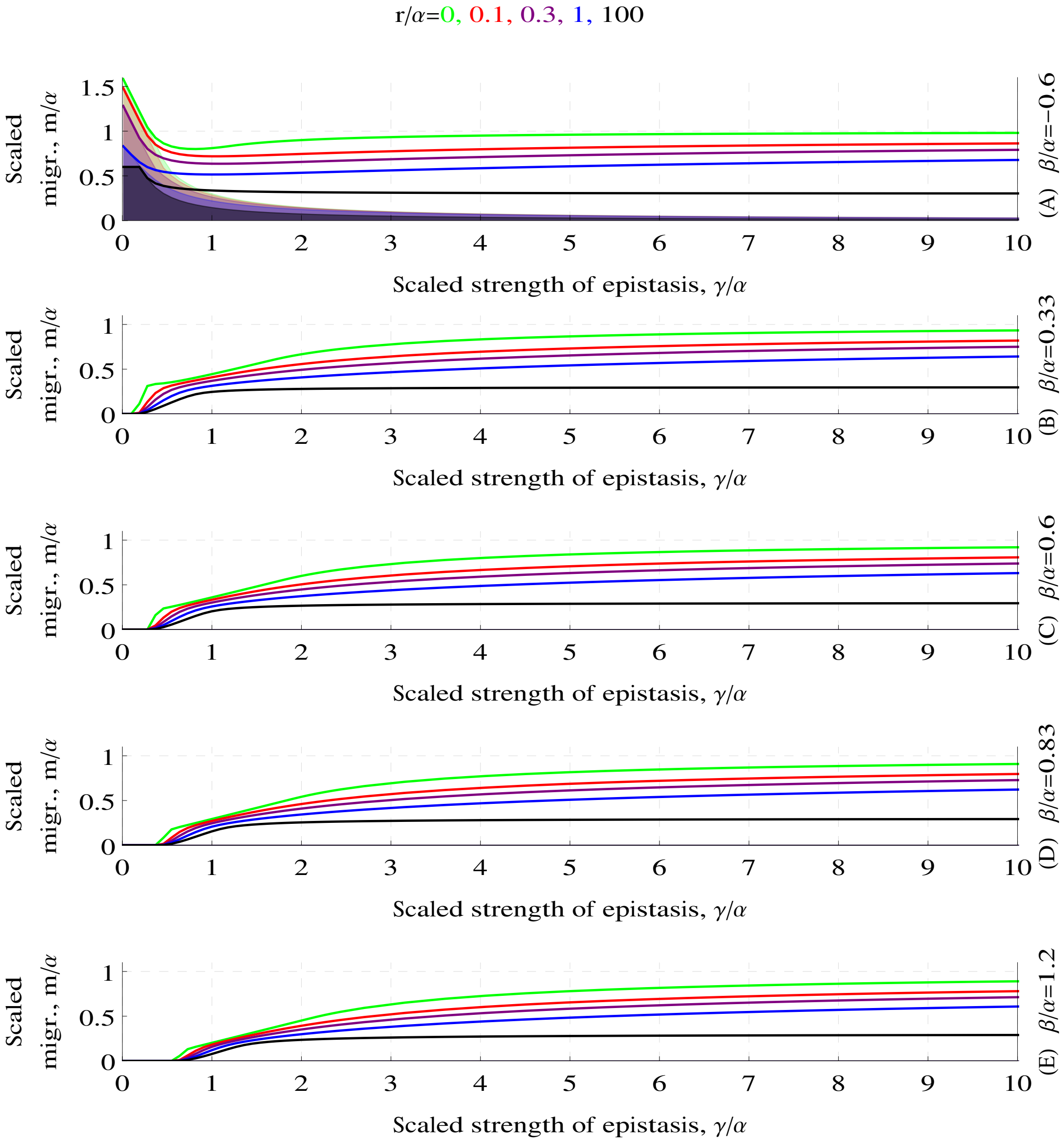

Figure S7: $\Gamma = (2\gamma, 2\gamma, 2\gamma, 2\gamma)$ (dominant DMI): very similar to the codominant case (see Fig. S3). For $\beta > 0$ and $\gamma > \beta/2$ underdominance at the $\mathcal{B}$ locus for $p_A = 1$. As a consequence, a DMI can never be globally stable for $\beta > 0$.

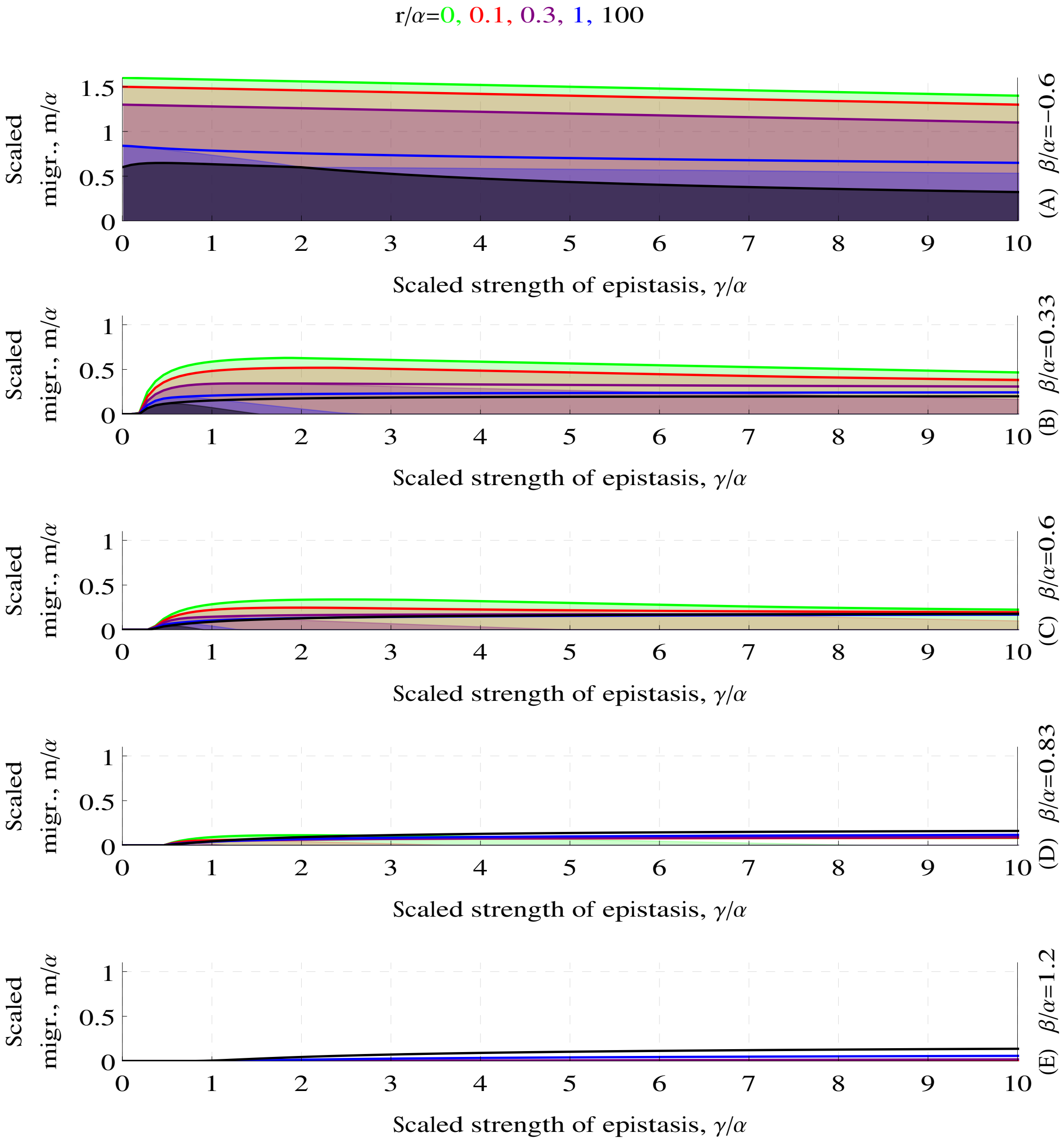

Figure S8: $\Gamma = (\gamma/50, \gamma/5, \gamma/5, 2\gamma)$: very similar to $\Gamma = (0, 0, 0, 2\gamma)$ (Fig. S6) and thus the haploid case (Figure S2).

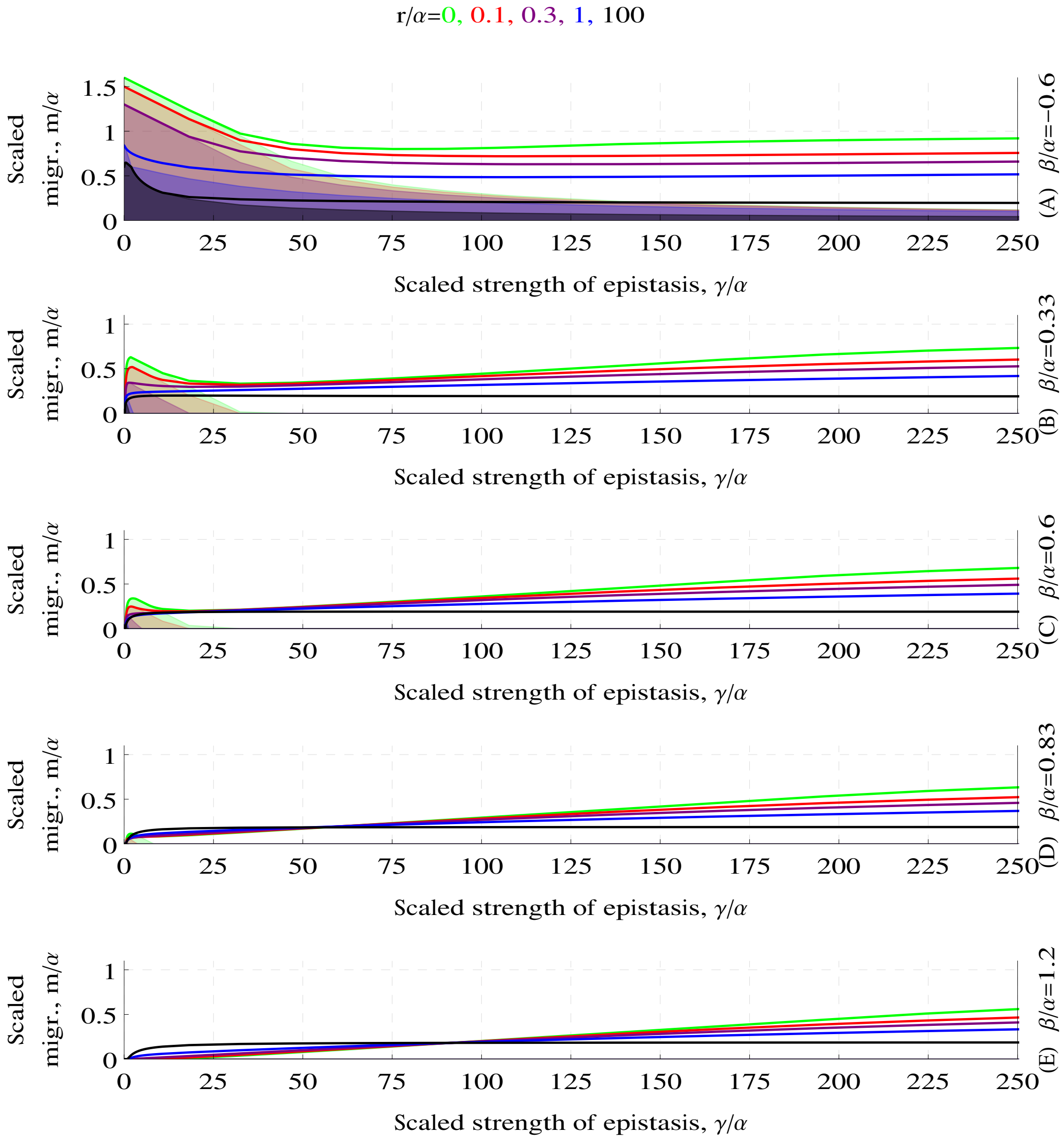

Figure S9: $\Gamma = (\gamma/50, \gamma/5, \gamma/5, 2\gamma)$ on a larger scale: Now the pattern looks very similar to the codominant model (Figure S3).