

# The Evolution of Phenotypic Switching in Subdivided Populations

Oana Carja,<sup>\*1</sup> Uri Liberman,<sup>†,‡</sup> and Marcus W. Feldman<sup>\*</sup>

<sup>\*</sup>Department of Biology, Stanford University, Stanford, California 94305-5020, <sup>†</sup>School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel 69978, and <sup>‡</sup>Efi Arazi School of Computer Science, Interdisciplinary Center, Herzliya, Israel 46150

**ABSTRACT** Stochastic switching is an example of phenotypic bet hedging, where offspring can express a phenotype different from that of their parents. Phenotypic switching is well documented in viruses, yeast, and bacteria and has been extensively studied when the selection pressures vary through time. However, there has been little work on the evolution of phenotypic switching under both spatially and temporally fluctuating selection pressures. Here we use a population genetic model to explore the interaction of temporal and spatial variation in determining the evolutionary dynamics of phenotypic switching. We find that the stable switching rate is mainly determined by the rate of environmental change and the migration rate. This stable rate is also a decreasing function of the recombination rate, although this is a weaker effect than those of either the period of environmental change or the migration rate. This study highlights the interplay of spatial and temporal environmental variability, offering new insights into how migration can influence the evolution of phenotypic switching rates, mutation rates, or other sources of phenotypic variation.

**G**ENETICALLY identical cells can show significant cell-to-cell variability in gene expression and other phenotypic characteristics. This variation in gene expression can be due to fluctuations in levels of methylation of CpG sites, in mRNA transcription, or in protein translation and may cause shifts among different regulatory states that result in bi- or multi-stability (Smits *et al.* 2006), which have also been viewed as an epigenetic switch (Lim and van Oudenaarden 2007). Moreover, these states are often heritable between generations, perhaps due to epigenetic inheritance. Examples include the lactose utilization network in *Escherichia coli*, where single cells can stochastically switch between two different states (Mettetal *et al.* 2006), or the galactose utilization network in yeast that displays bimodal patterns in the expression of *GAL* genes (Acar *et al.* 2005; Kaufmann *et al.* 2007).

In stochastic switching, individual cells can randomly switch between different phenotypes, which may be inherited. The mode of inheritance in such cases can depart from straightforward Mendelism (Bonduriansky and Day 2009; Danchin and

Wagner 2010). In particular, there can be interactions between inherited environmental conditions and epigenetic effects (see, *e.g.*, Furrow *et al.* 2011) that contribute to the statistical heritability of phenotypes.

Why do these genetically identical populations exhibit such switching behaviors? Stochastic switching is often interpreted as a bet-hedging strategy (Starrfelt and Kokko 2012) that could confer a fitness advantage in volatile environments (Thattai and van Oudenaarden 2004; Kussell and Leibler 2005). Switching increases the phenotypic diversity of the population within one generation, thereby increasing the chance of well-adapted offspring in a future environment. Studies have shown that stochastic changes can be advantageous even in comparison to environmental sensing when the switching rate is closely tuned to the rate at which the environment fluctuates (Acar *et al.* 2008). In this framework, the trade-off between the costs of maladaptation and the benefits of phenotypic variability becomes the object of mathematical analysis (Leigh 1970; Ishii *et al.* 1989; Thattai and van Oudenaarden 2004; Kussell and Leibler 2005; Leibler and Kussell 2010). It is important to note that bet hedging (increasing phenotypic variance within a generation) as a response to environmental volatility is one way to decrease the variance in fitness between generations and thus increase the geometric mean fitness of the genotype (Carja *et al.* 2013).

Copyright © 2014 by the Genetics Society of America

doi: 10.1534/genetics.114.161364

Manuscript received November 26, 2013; accepted for publication January 14, 2014; published Early Online February 4, 2014.

Supporting information is available online at <http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.161364/-/DC1>.

<sup>1</sup>Corresponding author: Department of Biology, Gilbert Hall, Stanford University, Stanford, CA 94305. E-mail: oana.carja@gmail.com; ocarja@stanford.edu

Population genetic approaches to stochastic phenotypic switching have generally been couched in terms of the dynamics of the mutation rate in a fluctuating environment. Mutation, in this context, refers to mutation between different allelic states or epiallelic states that are the cause of the phenotypic variability, and the question boils down to finding the optimal mutation rate as a function of the way the environment changes. Ishii *et al.* (1989) analyzed one of the earliest models of this kind in which alleles  $A$  and  $a$  were assigned fitnesses  $1 + s(t)$  and  $1 - s(t)$ , respectively, at time  $t$ , with  $s(t)$  allowed to fluctuate through time with mean zero. A second locus, with alleles  $B$  and  $b$ , controlled the rate of mutation between  $A$  and  $a$ . Ishii *et al.* (1989) found that the evolutionarily stable mutation rate maximized the long-term geometric mean fitness of the population. They suggested that under a wide array of conditions the uninventable mutation rate was  $1/n$  if the selection cycled with period  $2n$ .

Other studies of symmetric periodic selection models, without the mutation modifying gene as the cause of the switching, also suggested that the fitness was optimized at a mutation rate of  $1/n$  (Lachmann and Jablonka 1996). One important theoretical implication of these results is that if the mutation rate is very low, the population can be invaded by a mutator that increases this rate (up to  $1/n$ ). This violates the mutation reduction principle of Feldman and Liberman (1986), which holds in constant environments. When the selection regime does not fluctuate, the reduction principle, that zero is the uninventable mutation rate, recombination rate (Liberman and Feldman 1986), and migration rate (Balkau and Feldman 1973; Liberman and Feldman 1989), is quite general. However, in fluctuating environments, the reduction principle does not generally hold.

Recent analyses by Salathé *et al.* (2009) and Liberman *et al.* (2011) of switching and mutation modification in fluctuating environments have produced results that are much more complicated than originally derived by Ishii *et al.* (1989). If the time spent in each of two environments is a random variable, then even if the fitnesses in the two environments are symmetric, the stable switching rate can be very different from  $1/\bar{n}$ , where  $\bar{n}$  is the mean time in each environment. In addition, if the fitnesses of allele  $A$  relative to allele  $a$  in one environment and of allele  $a$  to allele  $A$  in the second environment are not very similar, then a switching-reducing modifier allele will succeed, with the result that the switching rate ultimately evolves toward zero.

In the case of symmetric selection with cycles of period 2, Liberman *et al.* (2011) showed that for any recombination rate between the switching-modifying gene and the gene under direct selection, higher switching rates are always favored, and the mean fitness at equilibrium is an increasing function of the switching rate; the opposite is true in a constant environment. Liberman *et al.* (2011) also proved that with period 4 and no recombination, the stable switching rate was  $1/2$ . Furthermore, when there is no recombination, the critical points of the mean fitness with respect to the switching rate are the same switching rates that cannot be

invaded. With asymmetric fitnesses, mathematical results have not been obtained, but numerical analysis, with and without recombination, has demonstrated that large asymmetries actually lead to a reduction of the switching rate (Salathé *et al.* 2009).

The regimes in which the benefits outweigh the costs of switching have been extensively studied both theoretically and experimentally when there is temporal heterogeneity in the selection pressures. However, organisms experience environmental heterogeneity through both time and space. The change in the selection pressure for an organism's offspring can therefore arise either from a temporal change in the fitness function or from migration to a different deme. Although the evolution of phenotypic switching has been extensively studied when the selection pressures vary temporally, there have been few studies of the evolution of stochastic switching in the case of a subdivided population, when selection varies spatially. We might expect differences in the evolutionary dynamics because in a spatially heterogeneous environment, only the migrants experience the change in the selection pressures, while in temporally varying environments the change in selection pressure is experienced by the entire population. Studies of the evolution of phenotypic variability when there is both spatial and temporal heterogeneity include that of Arnoldini *et al.* (2012), who investigated the evolution of sensed switching, finding that the relative balance of sensed and stochastic switching depended on the accuracy of the environmental stress signal. Moran (1992) explored the evolution of noninherited phenotypic variability and argued that the optimal amount of variability is zero.

Our goal here is to study the evolution of switching rates in the presence of both spatial and temporal heterogeneity in selection pressures. Do the inherent differences between temporal and spatial change in the selection regime result in different evolutionary conditions under which stochastic switching evolves? We find that the evolutionary dynamics of the system are mainly governed by the rate of environmental change. For large enough periods in the environmental fluctuations, the migration rate in the system also has strong effects on the evolution of stochastic switching: the uninventable switching rate is on the order of  $1/n$  minus the migration rate, where selection cycles with a period  $2n$ , and zero when the migration rate is  $>1/n$ . This stable switching rate is also a decreasing function of the recombination rate in the system, although this effect is weaker than those of the environmental change and migration. The evolution of stochastic switching is therefore determined by a complex interplay between all the forces that contribute to the variability under periodic change in selection pressures.

## Modeling Stochastic Switching

Consider an infinite, haploid population, divided into two demes,  $E_x$  and  $E_y$ , each with its own selection regime. Every individual in this population is defined by two biallelic loci: the major  $A/a$  locus controls the phenotype and fitness of

the individual, while the modifier  $B/b$  locus is selectively neutral and controls the switching rate between these two phenotypic states. Alleles  $B$  and  $b$  determine switching rates  $\mu_B$  and  $\mu_b$ , respectively, and our main interest is in finding the evolutionarily stable switching rate between the two phenotypes  $A$  and  $a$ , namely, a switching rate that cannot be invaded by mutants that produce different rates of switching. Toward that goal, we use an explicit population genetic model to track the allele frequencies at the modifier gene and ask under what conditions phenotypic switching is adaptive and what the role of migration is in the evolutionary dynamics of the modifier locus.

### Biological interpretation of the model

Even though the model is presented from the perspective of genetically controlled phenotypic switching (examples of which can be found in Smits *et al.* 2006; Veening *et al.* 2008; Beaumont *et al.* 2009; Rainey *et al.* 2011) between two bistable states or phenotypes (similar to Thattai and van Oudenaarden 2004; Salathé *et al.* 2009; Gaál *et al.* 2010; Liberman *et al.* 2011), it is also general enough to be applicable to a wider range of evolutionary questions. A possible mechanism for switching is epigenetic control of gene expression through variation in levels of methylation or chromatin formation. Therefore, the  $B/b$  modifier locus can be interpreted as an epigenetic locus able to influence the transition between the two phenotypes or the levels of expression of alleles  $A$  and  $a$ . Examples of such epigenetic loci are the DNMT genes that have been shown to have a role in the establishment and regulation of cytosine methylation (Bird 2002). Another possible interpretation of the model is that the switching rate is a mutation rate between the two  $A/a$  alleles at the major locus under selection. The evolution of mutation has been an area of intense research in population genetics and exceptions to the Feldman–Liberman reduction principle (Feldman and Liberman 1986) are expected in populations experiencing spatial and temporal variation. Because the  $B/b$  modifier locus may be genetic or epigenetic, we explore a wide range of switching rates.

There are four possible genotypes in the population  $AB$ ,  $Ab$ ,  $aB$ , and  $ab$ . At each generation, the population goes through recombination, selection, and phenotypic switching, in each deme separately, and then individuals may migrate between the two demes. As in general analyses of neutral modifiers, we frame the question in terms of the stability of the state of fixation in  $B$ , producing switching rate  $\mu_B$ , to invasion by allele  $b$ , which produces switching rate  $\mu_b$ . We assume these rates of switching to be the same in both phenotypic directions. This symmetry assumption makes our analysis more tractable but relaxing this assumption should not change the general conclusions of the model (see Salathé *et al.* 2009). In particular, we ask how the migration between  $E_x$  and  $E_y$ , which is assumed to occur at rate  $m$  per generation, affects this stable switching rate. We shall see that much of the analytical structure of the mathematical problem recapitulates that seen in Balkau and Feldman

(1973), where there was no switching from  $A$  to  $a$  but where locus  $B/b$  controlled the migration rate between  $E_x$  and  $E_y$  with constant fitnesses in the two demes.

We first present the case in which selection is constant in the two demes. We find that the reduction principle (zero is the uninvadable switching rate) holds for all possible migration rates. We then consider the case in which there is periodic fluctuating selection within each of the demes. In this case, we find parameter regimes in which the reduction principle does not hold, and nonzero switching rates can be favored. These evolutionarily stable rates are shown to depend both on the environmental volatility and on the migration rate of the system.

### Constant Selection

Let the frequencies and the fitnesses of the four genotypes in the two demes be as follows:

deme	$E_x$				$E_y$			
	$AB$	$Ab$	$aB$	$ab$	$AB$	$Ab$	$aB$	$ab$
fitness	$w_{x_1}$	$w_{x_2}$	$w_{x_3}$	$w_{x_4}$	$w_{y_1}$	$w_{y_2}$	$w_{y_3}$	$w_{y_4}$
frequency	$x_1$	$x_2$	$x_3$	$x_4$	$y_1$	$y_2$	$y_3$	$y_4$

(1)

As the modifier locus is selectively neutral, we have

$$w_{x_1} = w_{x_2}, w_{y_1} = w_{y_2}, w_{x_3} = w_{x_4}, w_{y_3} = w_{y_4}. \quad (2)$$

With allele  $B$  at the modifier locus, the switching rate in both directions between alleles  $A$  and  $a$  in each deme is  $\mu_B$ , and it is  $\mu_b$  with the modifier allele  $b$ . We denote by  $r$  the recombination rate between the two loci in both demes and by  $m$  the rate of migration between the demes.

Thus, in the next generation, after recombination, selection, phenotypic switching, and migration, the new frequencies  $x'_1, x'_2, x'_3, x'_4$  and  $y'_1, y'_2, y'_3, y'_4$  of the genotypes in the two demes are

$$\begin{aligned} x'_1 &= (1 - m)\tilde{x}_1 + m\tilde{y}_1 & y'_1 &= (1 - m)\tilde{y}_1 + m\tilde{x}_1 \\ x'_2 &= (1 - m)\tilde{x}_2 + m\tilde{y}_2 & y'_2 &= (1 - m)\tilde{y}_2 + m\tilde{x}_2 \\ x'_3 &= (1 - m)\tilde{x}_3 + m\tilde{y}_3 & y'_3 &= (1 - m)\tilde{y}_3 + m\tilde{x}_3 \\ x'_4 &= (1 - m)\tilde{x}_4 + m\tilde{y}_4 & y'_4 &= (1 - m)\tilde{y}_4 + m\tilde{x}_4, \end{aligned} \quad (3)$$

where

$$\begin{aligned} w_x \tilde{x}_1 &= (1 - \mu_B)w_{x_1}(x_1 - rD_x) + \mu_B w_{x_3}(x_3 + rD_x) \\ w_x \tilde{x}_2 &= (1 - \mu_b)w_{x_2}(x_2 + rD_x) + \mu_b w_{x_4}(x_4 - rD_x) \\ w_x \tilde{x}_3 &= (1 - \mu_B)w_{x_3}(x_3 + rD_x) + \mu_B w_{x_1}(x_1 - rD_x) \\ w_x \tilde{x}_4 &= (1 - \mu_b)w_{x_4}(x_4 - rD_x) + \mu_b w_{x_2}(x_2 + rD_x) \\ w_y \tilde{y}_1 &= (1 - \mu_B)w_{y_1}(y_1 - rD_y) + \mu_B w_{y_3}(y_3 + rD_y) \\ w_y \tilde{y}_2 &= (1 - \mu_b)w_{y_2}(y_2 + rD_y) + \mu_b w_{y_4}(y_4 - rD_y) \\ w_y \tilde{y}_3 &= (1 - \mu_B)w_{y_3}(y_3 + rD_y) + \mu_B w_{y_1}(y_1 - rD_y) \\ w_y \tilde{y}_4 &= (1 - \mu_b)w_{y_4}(y_4 - rD_y) + \mu_b w_{y_2}(y_2 + rD_y), \end{aligned} \quad (4)$$

with

$$D_x = x_1x_4 - x_2x_3, \quad D_y = y_1y_4 - y_2y_3, \quad (5)$$

$$w_x = \sum_{i=1}^4 w_{x_i}x_i, \quad w_y = \sum_{i=1}^4 w_{y_i}y_i. \quad (6)$$

This general model is very complicated and, to gain some mathematical insight, we assume the following symmetry relations among the fitness parameters:

$$\begin{aligned} w_{x_1} = w_{x_2} = w_{y_3} = w_{y_4} &= (1 + s), \\ w_{y_1} = w_{y_2} = w_{x_3} = w_{x_4} &= 1. \end{aligned} \quad (7)$$

Thus, we assume that with  $s > 0$ , phenotype  $A$  is preferred in deme  $E_x$ , while phenotype  $a$  is preferred in deme  $E_y$  and there is symmetry in these fitness differences between the demes (Ishii *et al.* 1989; Lachmann and Jablonka 1996). Symmetry assumptions in selection pressures are made throughout this article and are important for the mathematical tractability of the model. Future work will focus on the effect of relaxing these symmetry assumptions on the general conclusions of the model.

### Boundary equilibria with symmetric selection

We start our analysis by assuming that initially only allele  $B$  is present at the modifier locus; *i.e.*,  $x_2 = x_4 = y_2 = y_4 = 0$ . To simplify the notation, since  $x_3 = 1 - x_1$  and  $y_1 = 1 - y_3$ , let  $x_1 = x$  and  $y_3 = y$ . With this notation, the change in the population state following the transformations given in (3) and (4) becomes

$$\begin{aligned} x' &= (1 - m) \frac{(1 - \mu_B)(1 + s)x + \mu_B(1 - x)}{1 + sx} \\ &\quad + m \frac{(1 - \mu_B)(1 - y) + \mu_B(1 + s)y}{1 + sy} \\ y' &= (1 - m) \frac{(1 - \mu_B)(1 + s)y + \mu_B(1 - y)}{1 + sy} \\ &\quad + m \frac{(1 - \mu_B)(1 - x) + \mu_B(1 + s)x}{1 + sx}. \end{aligned} \quad (8)$$

Here,

$$\begin{aligned} w_x &= (1 + s)x + (1 - x) = 1 + sx \\ w_y &= (1 + s)y + (1 - y) = 1 + sy. \end{aligned} \quad (9)$$

We first study the equilibrium points of (8). With  $0 < m, \mu_B < 1$ , (8) has no equilibria where one of the two alleles  $A$  or  $a$  is fixed in either deme (*i.e.*,  $x = 0, x = 1, y = 0, y = 1$  are impossible). To find the polymorphic equilibria, we examine the change in  $(y - x)$ . A simple computation shows that

$$y' - x' = \frac{(1 - 2\mu_B)(1 + s)}{(1 + sx)(1 + sy)}(y - x). \quad (10)$$

Hence, at equilibrium, either  $y = x$  or

$$\frac{(1 - 2\mu_B)(1 + s)}{(1 + sx)(1 + sy)} = 1. \quad (11)$$

We call an equilibrium with  $y = x$  a “symmetric equilibrium” and any equilibrium with  $y \neq x$ , such that (11) holds, an “asymmetric equilibrium.”

Let us write

$$\mathbf{x} = (x_1, x_2, x_3, x_4), \quad \mathbf{y} = (y_1, y_2, y_3, y_4). \quad (12)$$

The following result holds when  $0 < m, \mu_B < 1$ :

#### Result 1.

1. On the boundary, where only allele  $B$  is present, there is a unique symmetric equilibrium  $(\mathbf{x}^*, \mathbf{y}^*)$  with

$$\mathbf{x}^* = (x^*, 0, 1 - x^*, 0), \quad \mathbf{y}^* = (1 - x^*, 0, x^*, 0), \quad (13)$$

where  $x^*$  is the unique positive root of the quadratic equation

$$Q(x) = sx^2 + [(s + 2)(m + \mu_B - 2m\mu_B) - s]x - (m + \mu_B - 2m\mu_B) = 0. \quad (14)$$

2.  $(\mathbf{x}^*, \mathbf{y}^*)$  is internally stable on the boundary with only  $B$  present.
3. If  $0 < m, \mu_B < \frac{1}{2}$ , then  $x^* > \frac{1}{2}$ .

The proof of this result is in Supporting Information, File S1.

In addition to the symmetric equilibria, asymmetric equilibria may exist when (11) is satisfied. In fact, we have the following result:

**Result 2.** If  $0 < m < m_0$  and  $0 < \mu_B < \mu_0$ , where

$$m_0 = \frac{s(s + 1)}{2(s(s + 2) + 1)}, \quad \mu_0 = \frac{s}{2(s + 1)}, \quad (15)$$

then there is a unique asymmetric equilibrium  $(\hat{x}, \hat{y})$ . Here,

$$\hat{\mathbf{x}} = (\hat{x}, 0, 1 - \hat{x}, 0), \quad \hat{\mathbf{y}} = (\hat{y}, 0, 1 - \hat{y}, 0), \quad (16)$$

where  $\hat{x}$  is the unique positive root of the quadratic equation

$$T(x) = (1 - m)s^2x^2 - sx[s(1 - m) - \mu_B(s + 2)(1 - 2m)] - \mu_B(2m + s) = 0, \quad (17)$$

and

$$\hat{y} = \frac{s(1 + s)\hat{x} + 2\mu_B(1 + s)}{s(1 + s\hat{x})}. \quad (18)$$

A proof of Result 2 can be found in File S2.

An interesting question is whether the asymmetric equilibrium can be internally stable. The mathematical

analysis of this issue is complicated, but we conjecture that the asymmetric equilibrium is never internally stable in the boundary with only the  $B$  allele present. Our conjecture is based on extensive simulations of the model and the following result, which holds when the switching rate  $\mu_B$  is small.

**Result 3.** *When  $\mu_B$  is “small”, the asymmetric equilibrium  $(\hat{x}, \hat{y})$  is “near” the fixation of  $B$  in the two demes and is internally unstable.*

See File S3 for the proof of this result.

### External stability of the symmetric equilibrium

We now check the local stability of the symmetric equilibrium  $(x^*, y^*)$  to the introduction of the modifier allele  $b$ , which changes the switching rate from  $\mu_B$  to  $\mu_b$ . To this end, we investigate the linear approximation  $L^*$  to the transformation of the population state near  $(x^*, y^*)$ , such that, up to nonlinear terms,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x^* \\ y^* \end{pmatrix} + \begin{pmatrix} \epsilon' \\ \delta' \end{pmatrix} = \begin{pmatrix} x^* \\ y^* \end{pmatrix} + L^* \begin{pmatrix} \epsilon \\ \delta \end{pmatrix}. \quad (19)$$

Here,  $\epsilon = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)$ ,  $\delta = (\delta_1, \delta_2, \delta_3, \delta_4)$  with  $\epsilon_i$  and  $\delta_i$  small, and  $\sum_i \epsilon_i = \sum_i \delta_i = 0$  and similarly for  $\epsilon'$  and  $\delta'$ .

The stability of  $(x^*, y^*)$  is determined by the eigenvalues of  $L^*$ . This approach has been used extensively in *modifier theory* (see, for example, Feldman and Liberman 1986). The matrix  $L^*$  has the structure

$$L^* = \begin{pmatrix} \epsilon_1 & \epsilon_3 & \delta_1 & \delta_3 & \epsilon_2 & \epsilon_4 & \delta_2 & \delta_4 & \epsilon_1 \\ & & & & & & & & \epsilon_3 \\ & & & & & & & & \delta_1 \\ & & & & & & & & \delta_3 \\ & & & & & & & & \epsilon_2 \\ & & & & & & & & \epsilon_4 \\ & & & & & & & & \delta_2 \\ & & & & & & & & \delta_4 \\ & & & & & & & & \epsilon_1 \end{pmatrix} \begin{matrix} \epsilon_1 \\ \epsilon_3 \\ \delta_1 \\ \delta_3 \\ \epsilon_2 \\ \epsilon_4 \\ \delta_2 \\ \delta_4 \\ \epsilon_1 \end{matrix}, \quad (20)$$

where we have swapped the columns and the rows. The matrix  $*$  does not affect the eigenvalues of  $L^*$ , which are those of the two submatrices  $L_{in}$  and  $L_{ex}$ .  $L_{in}$  determines the internal stability of  $(x^*, y^*)$  in the boundary where only  $B$  is present. Since  $(x^*, y^*)$  is assumed to be internally stable, the eigenvalues of  $L_{in}$  are less than 1 in magnitude.  $L_{ex}$  corresponds to the linear approximation of the evolution near  $(x^*, y^*)$  of the genotypes  $Ab$  and  $ab$  in the two demes.

A straightforward computation shows that the  $4 \times 4$  matrix  $L_{ex}$  can be written as

$$L_{ex} = \begin{pmatrix} (1-m)A & (1-m)B & mC & mD \\ (1-m)D & (1-m)C & mB & mA \\ mA & mB & (1-m)C & (1-m)D \\ mD & mC & (1-m)B & (1-m)A \end{pmatrix}, \quad (21)$$

where  $A, B, C, D$  are positive and depend on  $(x^*, y^*)$  and on  $s, r$ , and  $\mu_b$ .  $L_{ex}$  has exactly the same form as the matrix that appears in the external stability analysis of Balkau and Feldman

(1973). What enables us to analyze the positive largest eigenvalue of  $L_{ex}$ , which determines the external stability of  $(x^*, y^*)$ , are the following properties of matrices  $L$  of the form (21):

1. If  $P$  is the  $4 \times 4$  matrix given by

$$P = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (22)$$

then

$$PLP^{-1} = \begin{pmatrix} M & \mathbf{0} \\ * & N \end{pmatrix}, \quad (23)$$

where  $\mathbf{0}$  is the  $2 \times 2$  zero matrix and

$$M = \begin{pmatrix} (1-m)A + mD & (1-m)B + mC \\ mA + (1-m)D & mB + (1-m)C \end{pmatrix}, \quad (24)$$

$$N = \begin{pmatrix} (1-m)C - mB & (1-m)D - mA \\ (1-m)B - mC & (1-m)A - mD \end{pmatrix}.$$

Thus  $L$  is “similar”, via  $P$ , to the block matrix (23).

2. The fourth degree characteristic polynomial of  $L$  factors into the product of the two quadratic characteristic polynomials of  $M$  and  $N$ . These two polynomials share the same constant term

$$\det(M) = \det(N) = (1-2m)(AC - BD). \quad (25)$$

3. In general if, for fixed  $m, L_1, L_2, \dots, L_n$  are matrices of the form (21) with corresponding positive entries  $A_i, B_i, C_i, D_i$  for  $i = 1, 2, \dots, n$ , then

$$P \left( \prod_{i=1}^n L_i \right) P^{-1} = \begin{pmatrix} \prod_{i=1}^n M_i & \mathbf{0} \\ * & \prod_{i=1}^n N_i \end{pmatrix}. \quad (26)$$

Here  $M_i$  and  $N_i$  are the  $M$  and  $N$  matrices of (23) associated with  $L_i$ . Moreover, the fourth-degree characteristic polynomial of  $\prod_{i=1}^n L_i$  factors into the two second-degree characteristic polynomials of  $\prod_{i=1}^n M_i$  and  $\prod_{i=1}^n N_i$ . These two polynomials share the same constant term

$$\prod_{i=1}^n \det(M_i) = \prod_{i=1}^n \det(N_i) = (1-2m)^n \prod_{i=1}^n (A_i C_i - B_i D_i). \quad (27)$$

In File S4 we use properties 1 and 2 to obtain the following external stability result.

**Result 4.** *When  $0 < m, \mu_B, \mu_b < \frac{1}{2}$ , the symmetric equilibrium  $(x^*, y^*)$  is externally stable to the introduction of the new modifier allele  $b$  if  $\mu_b > \mu_B$  and is unstable if  $\mu_b < \mu_B$ .*

Therefore, the reduction principle holds and zero is the only uninhabitable switching rate.

### Overview of results when selection pressures are constant

Intuitively, in this model, both migration and phenotypic switching conspire to reintroduce the deleterious phenotype in each of the two spaces, which selection is trying to remove. Therefore, it is expected that both the migration and the switching rates evolve toward zero (Balkau and Feldman 1973; Liberman and Feldman 1986). For a fixed migration rate, as presented here, phenotypic switching will not evolve, and the population will approach a migration–selection balance at the major loci.

### Periodically Fluctuating Selection

Suppose that within each deme the selection regime varies temporally. We assume periodically fluctuating selection and explore the evolution of the switching rate as a function of both the migration rate and the length of the period of the fluctuating selection, which we also refer to as the environmental period. Specifically, we assume two types of selection regimes, *type 1* with the fitness parameters

deme	$E_x$		$E_y$		(28)
allele	A	a	A	a	
fitness	$1 + s$	1	1	$1 + s$	

and *type 2* with

deme	$E_x$		$E_y$		(29)
allele	A	a	A	a	
fitness	1	$1 + s$	$1 + s$	1.	

In the general case, an environmental cycle consists of  $(k + \ell)$  phases of selection, the first  $k$  of *type 1* followed by  $\ell$  phases of *type 2* selection. We first mathematically explore the dynamics of the system in the cases  $k = \ell = 1$  (cycle of period 2) and  $k = \ell = 2$  (cycle of period 4). These are the cases in which we are able to derive analytical results. We then study the system in the case of larger environmental periods, using numerical approximations and simulation.

### Assumptions and limitations of the fluctuating selection model

For analytical tractability, we make stringent symmetry assumptions on the fitnesses of the two phenotypes A and a, both in time and in space. We assume that, within each environment, each deme has an optimal phenotype and there is symmetry in these fitness differences between the demes. We also assume that, within each deme, each temporal environment produces a different optimal phenotype and, again, symmetry in these fitness differences is assumed. Another important assumption of the model is that there are only two different environments that fluctuate periodically.

We do not expect the results to change qualitatively if there are more than two environments; we also hypothesize that introducing random waiting times for environmental change will affect the evolutionarily stable switching rate, but not whether nonzero switching rates can evolve (Salathé *et al.* 2009). Moreover, it is known that asymmetry in selection pressures can substantially affect evolved switching rates in a model with just temporal heterogeneity (Salathé *et al.* 2009). Therefore, future work will need to focus on the effect of relaxing these symmetry assumptions on the conclusions of our model.

### Cycle of period 2 ( $k = \ell = 1$ )

Here, selection goes through two phases. The first phase is *type 1* as in (28) and the second is *type 2* as in (29).

Again, we first study the boundary equilibria where only the modifier allele B is present and the switching rate is  $\mu_B$ . The initial frequencies of AB in the two demes are denoted by  $x$  and  $y$ , respectively. After phase 1, they change to  $\tilde{x}$  and  $\tilde{y}$ , and after phase 2 they become  $x'$  and  $y'$ , the frequencies at the start of the next cycle. We have

$$\begin{aligned}\tilde{x} &= (1 - m) \frac{(1 - \mu_B)(1 + s)x + \mu_B(1 - x)}{(1 + s)x + (1 - x)} \\ &\quad + m \frac{(1 - \mu_B)(1 - y) + \mu_B(1 + s)y}{(1 - y) + (1 + s)y}, \\ \tilde{y} &= (1 - m) \frac{(1 - \mu_B)(1 + s)y + \mu_B(1 - y)}{(1 + s)y + (1 - y)} \\ &\quad + m \frac{(1 - \mu_B)(1 - x) + \mu_B(1 + s)x}{(1 + s)x + (1 - x)},\end{aligned}\tag{30}$$

and

$$\begin{aligned}x' &= (1 - m) \frac{(1 - \mu_B)\tilde{x} + \mu_B(1 + s)(1 - \tilde{x})}{\tilde{x} + (1 + s)(1 - \tilde{x})} \\ &\quad + m \frac{(1 - \mu_B)(1 + s)(1 - \tilde{y}) + \mu_B\tilde{y}}{\tilde{y} + (1 + s)(1 - \tilde{y})}, \\ y' &= (1 - m) \frac{(1 - \mu_B)\tilde{y} + \mu_B(1 + s)(1 - \tilde{y})}{\tilde{y} + (1 + s)(1 - \tilde{y})} \\ &\quad + m \frac{(1 - \mu_B)(1 + s)(1 - \tilde{x}) + \mu_B\tilde{x}}{\tilde{x} + (1 + s)(1 - \tilde{x})}.\end{aligned}\tag{31}$$

Due to the symmetry in the system, we have

$$\begin{aligned}(\tilde{y} - \tilde{x}) &= \frac{(1 - 2\mu_B)(1 + s)}{(1 + sx)(1 + sy)}(y - x), \\ (y' - x') &= \frac{(1 - 2\mu_B)(1 + s)}{[1 + s(1 - \tilde{x})][1 + s(1 - \tilde{y})]}(\tilde{y} - \tilde{x}).\end{aligned}\tag{32}$$

Hence,

$$(y' - x') = \frac{(1 - 2\mu_B)^2(1 + s)^2}{(1 + sx)(1 + sy)[1 + s(1 - \tilde{x})][1 + s(1 - \tilde{y})]}(y - x).\tag{33}$$

Thus, at equilibrium where  $(y' - x') = (y - x)$ , we have either  $y = x$  or

$$\frac{(1-2\mu_B)^2(1+s)^2}{(1+sx)(1+sy)[1+s(1-\bar{x})][1+s(1-\bar{y})]} = 1. \quad (34)$$

Equilibria with  $y = x$  are the *symmetric equilibria*, while if  $y \neq x$  and (34) holds, we have *asymmetric equilibria*. For the symmetric equilibria we have the following:

**Result 5.** *When  $0 < m$ ,  $\mu_B < 1$ , and  $s > 0$ , a unique symmetric equilibrium  $(\bar{x}, \bar{y})$  exists such that*

$$\bar{x} = (\bar{x}, 0, 1 - \bar{x}, 0), \quad \bar{y} = (1 - \bar{x}, 0, \bar{x}, 0). \quad (35)$$

$\bar{x}$  is the unique positive root of  $R(x) = 0$ , where

$$R(x) = sx^2 + [2 - m_B(s + 2)]x - (1 - m_B), \quad (36)$$

with  $m_B = m + \mu_B - 2m\mu_B$ . If, in addition,  $0 < m, \mu_B < \frac{1}{2}$ , then  $0 < \bar{x} < \frac{1}{2}$ .

The proof of this result is presented in File S5. In File S6 we also show the following:

**Result 6.** *The unique symmetric equilibrium  $(\bar{x}, \bar{y})$  is internally stable on the boundary where only the modifier allele B is present.*

For the local stability of this symmetric equilibrium to the introduction of the “new” modifier allele  $b$ , the external stability matrix  $\mathbf{L}_{\text{ex}}$  is given by

$$\mathbf{L}_{\text{ex}} = \mathbf{L}_{\text{ex}}^2 \mathbf{L}_{\text{ex}}^1, \quad (37)$$

where both  $\mathbf{L}_{\text{ex}}^1$  and  $\mathbf{L}_{\text{ex}}^2$  have the form (21). Thus, using the properties of these matrices, the characteristic polynomial of  $\mathbf{L}_{\text{ex}}$  factors into two quadratic polynomials, which allows us to find the conditions needed for external stability. The complete analysis is given in File S7, where the following result is proved:

**Result 7.** *In the case of selection fluctuating with period 2, the internally stable unique symmetric equilibrium  $(\bar{x}, \bar{y})$  is externally stable when  $\mu_B > \mu_b$  and is unstable when  $\mu_B < \mu_b$ , for all  $0 \leq r \leq 1$  and  $0 < m < \frac{1}{2}$ . Thus, in this case, higher switching rates are favored, and the evolutionarily stable phenotypic switching rate is 1.*

### The mean fitness and external stability

We saw that in a constant environment the symmetric equilibrium is externally stable if  $\mu_b > \mu_B$  and unstable if  $\mu_b < \mu_B$ , so that smaller switching rates are favored. On the other hand, with fluctuating selection of period 2, non-zero switching rates evolve; the symmetric equilibrium is stable if  $\mu_B > \mu_b$  and unstable if  $\mu_B < \mu_b$ . This phenomenon is related to the behavior of the mean fitness at equilibrium in the following way:

**Result 8.** *The mean fitness at the symmetric equilibrium is (i) a decreasing function of  $\mu_B$  in a constant environment or (ii) an increasing function of  $\mu_B$  in a period 2 cycling environment.*

The proof is given in File S8.

### Cycle of period 4 ( $k = \ell = 2$ )

In this case, in each generation, the fitnesses of the genotypes in each population go through four phases (“1”, “2”, “3”, “4”). The fitnesses at these phases are

deme allele	$E_x$		$E_y$	
	A	a	A	a
fitness “1”	$1 + s$	1	1	$1 + s$
fitness “2”	$1 + s$	1	1	$1 + s$
fitness “3”	1	$1 + s$	$1 + s$	1
fitness “4”	1	$1 + s$	$1 + s$	1

The analysis of this case for  $r = 0$  parallels that of Liberman et al. (2011) if we replace  $\mu_M$  by  $m_B = m + \mu_B - 2m\mu_B$  and  $\mu_m$  by  $m_b = m + \mu_b - 2m\mu_b$  (see Result 5). Our results for this case can be summarized as follows.

On the boundary where B is fixed, if we assume that  $x_1 = y_3, x_3 = y_1$  through the four phases of a cycle, then the four-step recursion is

$$\bar{\mathbf{w}}\mathbf{x}' = T_2 \circ T_2 \circ T_1 \circ T_1(\mathbf{x}), \quad (39)$$

where  $\mathbf{x} = (x_1, x_3) = (y_3, y_1)$ ,  $\mathbf{x}' = (x'_1, x'_3) = (y'_3, y'_1)$  and  $T_1$  and  $T_2$  are the transformations associated with fitness “1” and fitness “3”, respectively. The mean fitness  $\bar{w}$ , which is the normalizing factor, is

$$\bar{w} = (1 + s)^2 + m_B(1 - m_B)s^2(1 + s) + m_B(1 - m_B)s(s + 2)^2[(s + 2)x_1 - 1]. \quad (40)$$

If  $r = 0$ , we can show that, at the symmetric equilibrium  $(\bar{x}, \bar{y})$ , we have  $[(s + 2)\bar{x} - 1] > 0$ . Therefore, at equilibrium,  $\bar{w} > (1 + s)^2$ .

For the external stability of  $(\bar{x}, \bar{y})$  note that if  $m_b = 0$ , the matrices governing the evolutionary dynamics of the system in steps  $T_1$  and  $T_2$  become

$$\begin{pmatrix} 1 + s & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 + s \end{pmatrix}, \quad (41)$$

respectively. If  $m_b = 1$ , the matrices in steps  $T_1$  and  $T_2$  become

$$\begin{pmatrix} 0 & 1 \\ 1 + s & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 + s \\ 1 & 0 \end{pmatrix}, \quad (42)$$

respectively. As a result, when  $m_b = 0$ , the local stability is governed by the matrix  $\mathbf{M}^*$ :

$$\begin{aligned} \bar{\mathbf{w}}\mathbf{M}^* &= \begin{pmatrix} 1 & 0 \\ 0 & 1 + s \end{pmatrix}^2 \begin{pmatrix} 1 + s & 0 \\ 0 & 1 \end{pmatrix}^2 \\ &= \begin{pmatrix} (1 + s)^2 & 0 \\ 0 & (1 + s)^2 \end{pmatrix}, \end{aligned} \quad (43)$$

and when  $m_b = 1$ ,

$$\begin{aligned}\bar{w}\mathbf{M}^* &= \begin{pmatrix} 0 & 1+s \\ 1 & 0 \end{pmatrix}^2 \begin{pmatrix} 0 & 1 \\ 1+s & 0 \end{pmatrix}^2 \\ &= \begin{pmatrix} (1+s)^2 & 0 \\ 0 & (1+s)^2 \end{pmatrix}. \end{aligned} \quad (44)$$

Since  $\bar{w} > (1+s)^2$ , in the case  $r = 0$  ( $\bar{x}, \bar{y}$ ) is externally stable when either  $m_b = 0$  or  $m_b = 1$ . Since  $m_b = 0$  if  $m = 0, \mu_b = 0$ , or  $m = 1, \mu_b = 1$  and  $m_b = 1$  if  $m = 0, \mu_b = 1$  or  $m = 1, \mu_b = 0$  (which is not biologically relevant), we conclude that when both  $r$  and  $m$  are small for any  $0 < \mu_b < 1$ , a new mutation-modifying allele cannot invade ( $\bar{x}, \bar{y}$ ) if it produces switching rates that are too high or too low. This suggests that there might be an intermediate switching rate  $\mu_B$  that gives rise to a symmetric equilibrium ( $\bar{x}, \bar{y}$ ), which cannot be invaded by allele  $b$ . Actually, we have the following result:

**Result 9.** *In the period 4 symmetric case, if  $r = 0$ , the mean fitness  $\bar{w} = \bar{w}(m_B)$  achieves a maximum at  $m_B = \frac{1}{2}$ .*

The proof is similar to that in Liberman *et al.* (2011). Note that since  $m_B = m + \mu_B(1 - 2m)$ ,  $m_B = \frac{1}{2}$  implies that  $(1 - 2m)(1 - 2\mu_B) = 0$ , and as  $0 \leq m < \frac{1}{2}$ ,  $m_B = \frac{1}{2}$  coincides with  $\mu_B = \frac{1}{2}$ . Hence, for fixed  $0 \leq m < \frac{1}{2}$ ,  $\bar{w} = \bar{w}(\mu_B)$  has its maximum when  $\mu_B = \frac{1}{2}$ . The next result is based on extensive simulations.

**Result 10.** *In the period 4 symmetric case with  $0 \leq r \leq 1$  and for all migration rates, the phenotypic switching rate  $\mu_B = \frac{1}{2}$  cannot be invaded by rates either smaller or larger than  $\frac{1}{2}$ . Thus,  $\mu_B = \frac{1}{2}$  is the stable switching rate.*

Here, again, the migration rate does not affect the uninvadable switching rate. Moreover, with a switching rate of  $1/2$ , both alleles  $A$  and  $a$  reach a frequency of  $0.5$  in each deme, and thus migration acts to neutrally reshuffle these alleles only between the two demes.

### Cycle of period $k + \ell$

Following our previous analyses where  $k = \ell = 1$  and  $k = \ell = 2$ , we conjecture that, in the general case, on the boundary, where the modifier allele  $B$  is fixed, we have the following result:

**Result 11.** *If  $0 < m, \mu_B < 1$ , and  $s > 0$ , a unique symmetric equilibrium ( $\bar{x}, \bar{y}$ ) exists such that*

$$\bar{x} = (\bar{x}, 0, 1 - \bar{x}, 0), \quad \bar{y} = (1 - \bar{x}, 0, \bar{x}, 0) \quad (45)$$

and  $\bar{x}$  is the unique positive root of a quadratic equation

$$G(x) = x^2 + px + q = 0 \quad (46)$$

with  $G(0) < 0$  and  $G(1) > 0$ , where  $p$  and  $q$  are functions of  $s, \mu_B, k$ , and  $l$ . Moreover, this symmetric equilibrium is internally stable.

Our conjecture is based on analytical examination of several choices of the  $(k + \ell)$  phases, using Mathematica, and on extensive simulations.

The external stability analysis of  $(\bar{x}, \bar{y})$  is very complicated in the general case with arbitrary  $k$  and  $\ell$ , as the associated external stability matrix  $\mathbf{L}_{\text{ex}}$  is then

$$\mathbf{L}_{\text{ex}} = \tilde{\mathbf{L}}_{\text{ex}}^\ell \cdot \tilde{\mathbf{L}}_{\text{ex}}^{\ell-1} \cdots \tilde{\mathbf{L}}_{\text{ex}}^1 \cdot \bar{\mathbf{L}}_{\text{ex}}^k \cdot \bar{\mathbf{L}}_{\text{ex}}^{k-1} \cdots \bar{\mathbf{L}}_{\text{ex}}^1. \quad (47)$$

$\bar{\mathbf{L}}_{\text{ex}}^i$  corresponds to selection phase  $i$  in the  $k$  phases where selection is of type 1, and  $\tilde{\mathbf{L}}_{\text{ex}}^j$  is associated with phase  $j$  of the  $\ell$  phases where selection is of type 2.

Each of these  $(k + \ell)$  matrices is of the form

$$\begin{pmatrix} (1-m)A & (1-m)B & mC & mD \\ (1-m)D & (1-m)C & mB & mA \\ mA & mB & (1-m)C & (1-m)D \\ mD & mC & (1-m)B & (1-m)A \end{pmatrix} \quad (48)$$

with corresponding parameters  $\bar{A}^i, \bar{B}^i, \bar{C}^i, \bar{D}^i$  for  $\bar{\mathbf{L}}_{\text{ex}}^i$  and  $\tilde{A}^j, \tilde{B}^j, \tilde{C}^j, \tilde{D}^j$  for  $\tilde{\mathbf{L}}_{\text{ex}}^j$  (similar to the matrices presented in File S7). All the matrices  $\bar{\mathbf{L}}_{\text{ex}}^i$  and  $\tilde{\mathbf{L}}_{\text{ex}}^j$  for  $i = 1, 2, \dots, k$  and  $j = 1, 2, \dots, \ell$  are positive.

Applying the properties of matrices of the form (48), we have the following result:

**Result 12.** *The characteristic polynomial of  $\mathbf{L}_{\text{ex}}$  factors into two quadratic polynomials. The constant terms of both polynomials are*

$$(1-2m)^{k+\ell} \prod_{i=1}^k (\bar{A}^i \bar{C}^i - \bar{B}^i \bar{D}^i) \cdot \prod_{j=1}^{\ell} (\tilde{A}^j \tilde{C}^j - \tilde{B}^j \tilde{D}^j). \quad (49)$$

The constant term (49) is always positive, and based on extensive mathematical checks using Mathematica it is always  $< 1$ . Therefore, using the same argument as in the case  $k = \ell = 1$ , the external stability of the symmetric equilibrium is determined by a sole condition  $C_1(1) > 0$  where  $C_1(\lambda)$  is the “first” quadratic polynomial in the factorization of the characteristic polynomial of  $\mathbf{L}_{\text{ex}}$ .

**Overview of analytic results when selection pressures are periodic, periods 2 and 4:** When environments change periodically, the reduction principle no longer holds and nonzero switching rates are evolutionarily stable and uninvadable. In fact, if the environment changes every generation or every two generations, this evolutionarily stable switching rate does not depend on the migration rate in the system and is  $1/n$  as reported by previous studies (where  $n$ , the number of generations in each environmental phase, is 1 or 2, respectively). An important result is that this evolutionarily stable switching rate maximizes mean fitness at equilibrium. This result is deeply rooted in classic population genetics theory and is connected to early work on neutral modifiers and the mean fitness principle of Karlin and McGregor (1972).

We were not able to derive mathematical expressions for the external stability conditions for cycles  $k = l > 2$ . For these longer cycles of equal period (which we from now on denote by  $2n$ , where  $n$  is the number of generations in each



phase), we use simulation and numerical analyses and find that the leading eigenvalue depends not only on the difference between  $\mu_B$  and  $\mu_b$ , but also on the initial switching rate  $\mu_B$  and the migration rate. We present these results in the next section.

## Simulations and Numerical Analyses

### Description of the simulation

We start with a population fixed on the  $B$  allele at the modifier locus, which determines a switching rate  $\mu_B$  sampled randomly between 0 and 1 and then held constant. The population evolves for 1000 generations until it reaches an internally stable symmetric equilibrium at the  $A/a$  locus. We then introduce the  $b$  allele at a small frequency ( $10^{-4}$ ) near this symmetric equilibrium. The switching rate  $\mu_b$  set by allele  $b$  is chosen as the product of the resident switching rate and a number generated from an exponential distribution with mean 1. This allows us to test invading mutation rates that are more often close to the resident rate, but also rates that are far from it. After 5000 generations with these fixed values of  $\mu_B$  and  $\mu_b$ , we determine whether the newly introduced allele was able to invade in the population; if the frequency of  $b$  is larger than its initial frequency of  $10^{-4}$ , we declare invasion. This is the initial invasion trial. If the  $b$  allele was able to invade, we start the next invasion trial with this invading switching rate chosen as the new resident switching rate of the  $B$  allele. If there was no invasion, the  $B$  allele determines the same resident switching rate as the previous trial. We then repeat the invasion steps described above. After at least 500 invasion trials and after the resident switching rate cannot be invaded in 50 consecutive trials, the final switching rate in the population is declared to be the stable switching rate.

### Results

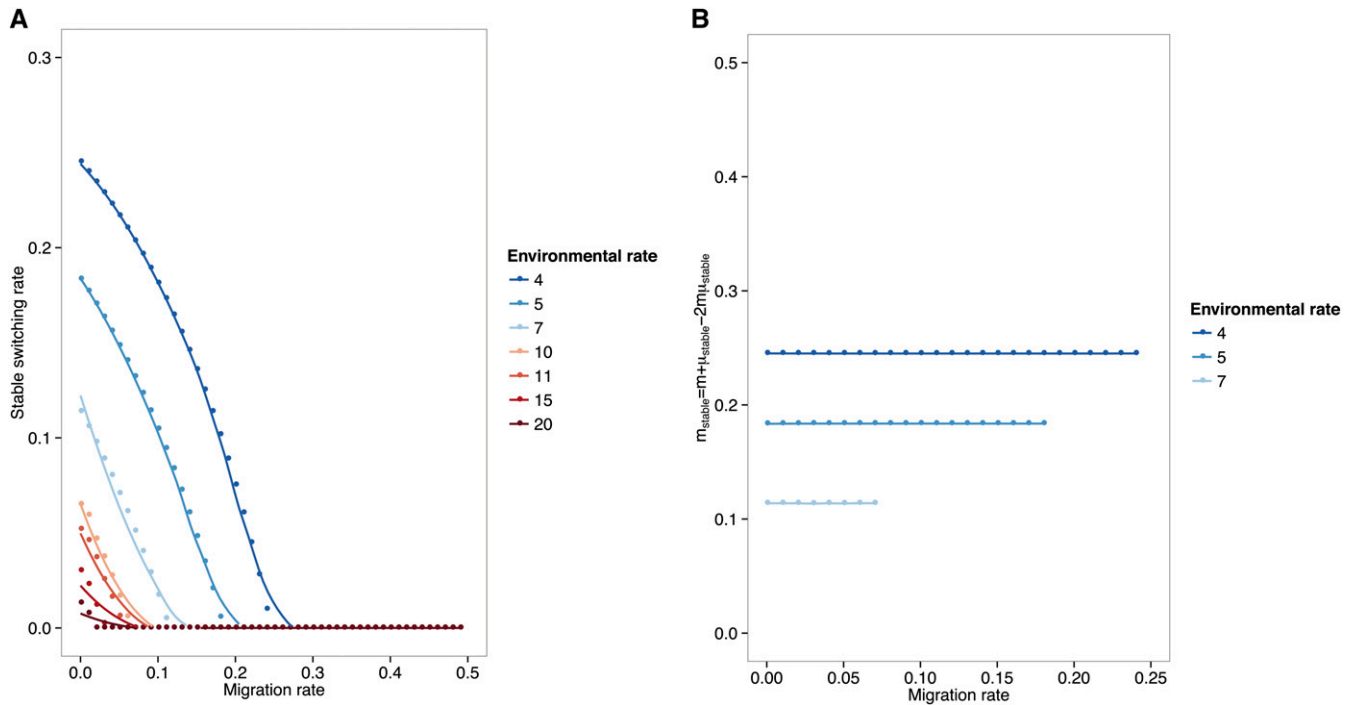
The case in which each environment persists for three generations before a change shows some anomalies is presented in Figure S1 as a function of the migration and recombination rates. In the case of no migration, Ishii *et al.* (1989) first showed that there are two equilibria for the mutation rate, one stable and one unstable. It was also shown that increasing the recombination rate  $r$  can change the stability and the domains of attraction of the two equilibria. With positive migration, we show that the uninvadable switching rate is sensitive to the interplay of recombination and migration, with possible sharp discontinuities in the stable switching rate. For no recombination and small migration rate, any switching rate can be invaded by a larger one. As the recombination rate increases, for small migration, the stable mutation rate is on the order of  $1/3$ . As the migration rate increases, the stable switching rate decreases in a manner consistent with larger values of  $n$ .

The results for the evolutionarily stable switching rate as a function of migration rate for periods  $2n > 6$  are presented

in Figure 1A. The stable switching rate is seen to be a decreasing function of both the migration rate and the number of generations  $n$  before an environmental change. Indeed, we confirm that, in the absence of migration, the stable switching rate is of the order of  $1/n$  as reported by previous studies. As the migration rate increases, the stable switching rate decreases until, above a threshold migration rate, it becomes zero. Note that this threshold in migration also is on the order of  $1/n$ . This is intuitive, since in this case where the environments both change at the same time, a migration rate inversely proportional to the rate of environmental change would place the migrants in a deme and environment with the same selection pressure as the one to which they were adapted. This effectively eliminates the need for switching. For migration rates below this threshold, the uninvadable switching rate is such that, together with migration, it generates the required variability that allows the system to cope with the environmental change it experiences. In fact, Figure 1B shows that the stable switching rate is such that the parameter  $m_{\text{stable}} = m + \mu_{\text{stable}} - 2m\mu_{\text{stable}}$  (the equivalent of  $m_b$  in Result 5) is of the order  $1/n$ .

In Figure 2 we show that, for recombination rate  $r = 0$ , there is an alternative way to obtain the evolutionarily stable switching rate. Ishii *et al.* (1989) found that, when  $r = 0$ , the evolutionarily stable switching rate is the one “that maximizes the long-term geometric average of population fitness.” Therefore, we expect that the switching rate that maximizes mean fitness at equilibrium is also the uninvadable switching rate (also see Result 8). The invasibility condition is again given by the properties of the leading eigenvalue  $\lambda_0$  of the local stability matrix of the system. To determine invasibility numerically, we can differentiate  $\lambda_0$  with respect to  $\mu_b$  and then evaluate this derivative at  $\mu_b = \mu_B$ , where we know that  $\lambda_0 = 1$ . The sign of this derivative then tells us whether a larger or a smaller value of  $\mu_b$  will produce a leading eigenvalue  $\lambda_0$  larger or smaller than 1 and thus whether  $b$  will invade or not. Therefore, the values of  $\mu_b$  where this derivative crosses zero with a negative slope correspond to values of  $\mu_B$  that are uninvadable. We see that the switching rates that maximize the mean fitness at equilibrium (and are zeros for the derivative of mean fitness with respect to  $\mu_B$ ) are also the zeros of the derivative of  $\lambda_0$  and therefore can be regarded as uninvadable. Figure 2B shows this equality of the critical points for the two derivatives, using an example where the environment changes every  $n = 4$  generations, the symmetric selection coefficient is  $s = 0.4$ , and the migration rate  $m = 0.1$ . Figure 2C demonstrates the match between the stable, uninvadable switching rates found by simulation and the switching rate that maximizes mean fitness at equilibrium.

Figure 3 shows the stable switching rate when the environment alternates every four generations as a function of the migration and recombination parameters. We find that the stable switching rate is a decreasing function of the recombination rate, for any migration rate (see also Liberman *et al.* 2011). As recombination increases, the force of



**Figure 1** The evolutionarily stable switching rate as function of migration and number of generations before an environmental change,  $n > 3$ . (A) The symmetric selection coefficient  $s = 0.4$ . Recombination rate is  $r = 0.2$ . We vary the rate of environmental change  $n$  as in the legend. The dots represent the average over 10 runs of the simulation. The plotted curves represent a fit to the data using a generalized additive model with penalized cubic regression splines. The stable switching rate is a decreasing function of both the migration rate and the rate of environmental change. (B) The parameter  $m_{\text{stable}} = m + \mu_{\text{stable}} - 2m\mu_{\text{stable}}$ . The symmetric selection coefficient  $s = 0.4$ . Recombination rate is  $r = 0.2$ . The rate of environmental change  $n$  varies as in the key. The dots represent the average over 10 runs of the simulation.

secondary selection on the modifier decreases as the linkage disequilibrium between the two loci decreases. This is expected to be a relatively weak force, which explains the relatively small difference between the curves in Figure 3 compared to the effect of migration (see Figure 1) for a given recombination rate. In Figure S2, the effect of increasing period is also seen to be more pronounced than that of increasing the recombination rate. In earlier studies of neutral modifiers of mutation and recombination, the induced selection on the modifier locus was of the order of the square of the disequilibrium between the major and the modifier loci. Even if this effect is exaggerated somewhat by the environmental fluctuations, it remains weaker than the effect of either the period of fluctuation or the migration rate.

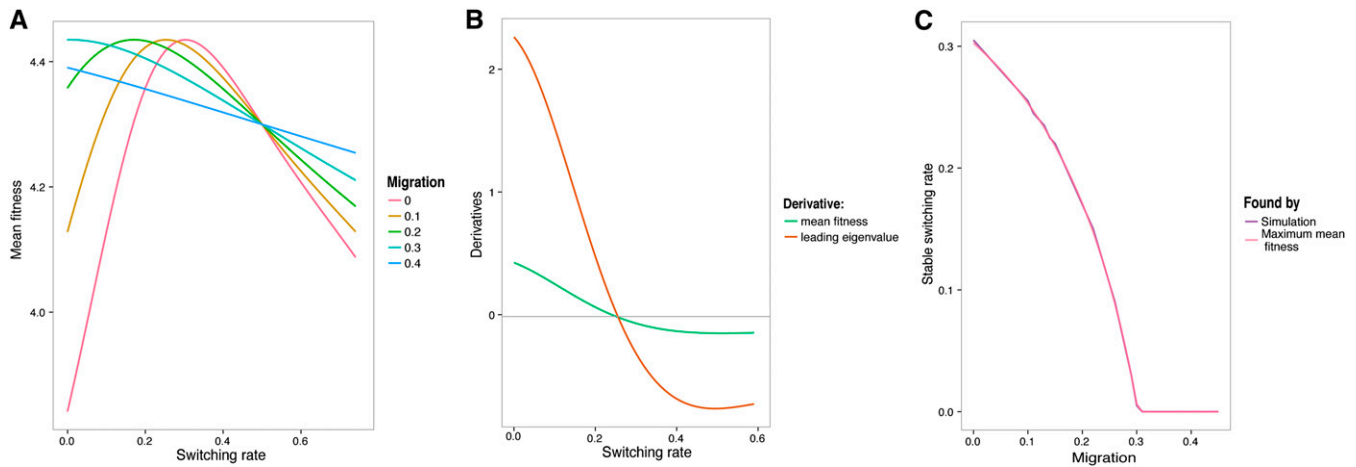
In Figure S3, we show that the strength of the symmetric selection coefficient  $s$  does not qualitatively influence the evolutionary dynamics of the system. This result is in accordance with previous analyses in temporally varying environments showing that changing the symmetric selection coefficient does not affect the evolutionarily stable switching rate (Ishii *et al.* 1989; Salathé *et al.* 2009).

## Discussion

In recent years, the evolution of microbial populations has become a question of significant biological importance. The

medical threat posed by the ability of some bacterial species to resist antibiotics by using phenotypic heterogeneity (Dubnau and Losick 2006; Smits *et al.* 2006) is significant (Centers for Disease Control 2013). Stochasticity and growth bistability allow bacteria to create diverse phenotypes and specialized cells in anticipation of possible adverse changes in the selection pressures (Barrett Deris *et al.* 2013). For example, bacterial persistence is a well-known mechanism that depends on stochastic switching between normal and persister phenotypes, which are able to survive better in the presence of antibiotics but are virtually growth arrested in standard media (Balaban *et al.* 2004). Given the role that stochastic switching has in bacterial populations, there is a need to consider the kinds of selective pressures that promote the evolution of switching behaviors and may be associated with the efficacy and evolvability of drug resistance.

We have investigated the evolution of stochastic switching under both temporal and spatial volatility, using a population genetic model with a modifier gene that controls the rate of switching (for examples see Smits *et al.* 2006; Veening *et al.* 2008; Beaumont *et al.* 2009; Rainey *et al.* 2011). We study conditions under which switching is adaptive and determine the evolutionarily stable (*i.e.*, uninvadable) switching rates for a wide range of migration rates, recombination rates, symmetric selection, and periodic rates of environmental fluctuation. Closed-form analytical results such as those



**Figure 2** The link between the evolutionarily stable switching rate and the switching rate that maximizes mean fitness at equilibrium. The symmetric selection coefficient  $s = 0.4$ . Recombination rate is  $r = 0$ . The environment changes every  $n = 4$  generations. (A) Mean fitness at equilibrium, for the range of migration rates in the key. In B, using the migration rate  $m = 0.1$  as an example, we show that the switching rate that maximizes mean fitness at equilibrium is also the switching rate that is uninvadable by other rates; in other words, the zeros of the derivative of the mean fitness at equilibrium are also the zeros of the derivative of the leading eigenvalue of the local stability matrix. (C) The match between the stable switching rates found by simulation and the switching rates that maximize mean fitness at equilibrium, as a function of the migration rate.

presented here help in gaining insight into the driving evolutionary process that we can then use to inform experimental and possibly medically motivated studies.

Our model reproduces the results found in previous studies with no spatial heterogeneity, but the migration rate between demes with different selection pressures can dramatically alter the evolutionarily stable switching rate. Our results demonstrate that the stochastic phenotypic switching rate is a decreasing function of the migration rate between demes. Moreover, with both spatial and temporal heterogeneity in the system, for environmental periods  $2n > 6$ , there exists a threshold migration rate above which stochastic switching cannot evolve. This threshold is also of the order of  $1/n$ , where  $n$  is the number of generations spent in each environment. Below this threshold, the switching rate evolves to the order of  $1/n$  minus the migration rate in the system. We also find that the switching rate is a decreasing function of the recombination rate in the system, although this effect is small compared to that of the migration rate.

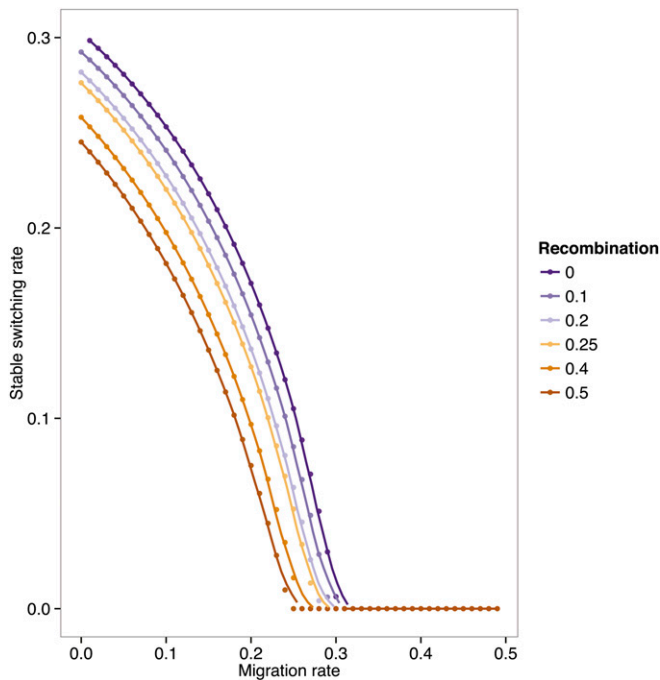
Our work highlights the importance of incorporating spatial heterogeneity when studying stochastic switching in the laboratory. Thus space-dependent fitness conditions ought to play a central role in the design and interpretation of experiments that seek to understand the behavior of natural populations and should try to recreate the natural time and space variations to which these populations would be subject. Populations of cells or strains of bacteria in the wild can experience both temporally and spatially heterogeneous selection pressures, and understanding the different contributions of these two inherently different selection regimes may help to interpret the differences observed between wild and laboratory populations (Dubnau and Losick 2006). Moreover, current methods of studying drug

efficacies are often performed in bulk growth conditions, which might fluctuate spatially in unforeseen ways. A better understanding of the role of phenotypic variability in these populations could enable the development of new techniques and treatment strategies against drug-resistant bacteria.

Although this work is framed in terms of a gene controlling phenotypic switching, the model is general enough to be applicable to a wider range of dynamical systems. The evolutionary dynamics of the modifier locus  $B/b$  can provide insight into the evolution of an epigenetic locus that is able to influence the transition between two different phenotypes, a locus controlling the mutation rate between the two alleles at the major locus  $A/a$ , or the rate of switching between two expression levels.

Here we have focused exclusively on the case in which the environment in the two demes varies periodically, with equal numbers of generations before environmental changes in each deme. We have also studied only the case of symmetric selection, where the selection coefficients in the two demes are equal. Previous work in temporally fluctuating environments (Salathé *et al.* 2009; Liberman *et al.* 2011) has shown that asymmetry in selection pressures or variance in times between environmental changes can substantially change the optimal switching rate. Incorporating selection asymmetries and variability in the environmental fluctuations is an important follow-up to this work.

The question of optimal phenotypic variability and the evolution of stochastic switching needs to be further explored under more general ecological scenarios driven by both temporal and spatial change. Using both analytic and numerical arguments we have made a first step in this direction: we find that under spatial and temporal fluctuations in selection pressures, phenotypic switching can evolve



**Figure 3** The evolutionarily stable switching rate as a function of migration and recombination. The symmetric selection coefficient  $s = 0.4$ . The environment changes every  $n = 4$  generations. The plotted curves represent a fit to the data using a generalized additive model with penalized cubic regression splines.

only when migration is not too high, and its evolution is heavily controlled by the forces driving variability in the population: migration, recombination, and the rate of change in selection pressure.

### Acknowledgments

Research supported in part by the Center for Computational Evolution and Human Genomics at Stanford. The authors thank Rob Furrow and two anonymous reviewers for helpful comments on the manuscript.

### Literature Cited

Acar, M., A. Becskei, and A. van Oudenaarden, 2005 Enhancement of cellular memory by reducing stochastic transitions. *Nature* 435: 228–232.

Acar, M., J. T. Mettetal, and A. van Oudenaarden, 2008 Stochastic switching as a survival strategy in fluctuating environments. *Nat. Genet.* 40: 471–475.

Arnoldini, M., R. Mostowy, S. Bonhoeffer, and M. Ackermann, 2012 Evolution of stress response in the face of unreliable environmental signals. *PLoS Comput. Biol.* 8(8): e1002627.

Balaban, N. Q., J. Merrin, R. Chait, L. Kowalik, and S. Leibler, 2004 Bacterial persistence as a phenotypic switch. *Science* 305: 1622–1625.

Balkau, B. J., and M. W. Feldman, 1973 Selection for migration modification. *Genetics* 74: 171–174.

Beaumont, H. J. E., J. Gallie, C. Kost, G. C. Ferguson, and P. B. Rainey, 2009 Experimental evolution of bet hedging. *Nature* 462: 90–94.

Bird, A., 2002 DNA methylation patterns and epigenetic memory. *Genes Dev.* 16: 6–21.

Bonduriansky, R., and T. Day, 2009 Nongenetic inheritance and its evolutionary implications. *Annu. Rev. Ecol. Evol. Syst.* 40: 103–125.

Carja, O., U. Liberman, and M. W. Feldman, 2013 Evolution with stochastic fitnesses: a role for recombination. *Theor. Popul. Biol.* 86: 29–42.

Centers for Disease Control, 2013 Antibiotic resistance threats in the United States. Available at: <http://www.cdc.gov/drugresistance/threat-report-2013/pdf/ar-threats-2013-508.pdf>.

Danchin, E., and R. H. Wagner, 2010 Inclusive heritability: combining genetic and nongenetic information to study animal culture. *Oikos* 119: 210–218.

Deris, J. B., M. Kim, Z. Zhang, H. Okano, R. Hermsen, A. Groisman, and T. Hwa, 2013 The innate growth bistability and fitness landscapes of antibiotic-resistant bacteria. *Science* 342: 1237435.

Dubnau, D., and R. Losick, 2006 Bistability in bacteria. *Mol. Microbiol.* 61(3): 564572.

Feldman, M. W., and U. Liberman, 1986 An evolutionary reduction principle for genetic modifiers. *Proc. Natl. Acad. Sci. USA* 83: 4824–4827.

Furrow, R. E., F. B. Christiansen, and M. W. Feldman, 2011 Environment-sensitive epigenetics and the heritability of complex diseases. *Genetics* 189: 1377–1387.

Gaal, B., J. W. Pitchford, and A. J. Wood, 2010 Exact results for the evolution of stochastic switching in variable asymmetric environments. *Genetics* 184: 1113–1119.

Ishii K., H. Matsuda, Y. Iwasa, and A. Sasaki, 1989 Evolutionarily stable mutation rate in a periodically changing environment. *Genetics* 121: 163–174.

Karlin, S., and J. L. McGregor, 1972 The evolutionary development of modifier genes. *Proc. Natl. Acad. Sci. USA* 69: 3611–3614.

Kaufmann, B. B., Q. Yang, J. T. Mettetal, and A. van Oudenaarden, 2007 Heritable stochastic switching revealed by single-cell genealogy. *PLoS Biol.* 5: e239.

Kussell, E., and S. Leibler, 2005 Phenotypic diversity, population growth, and information in fluctuating environments. *Science* 309: 2075–2078.

Lachmann, M., and E. Jablonka, 1996 The inheritance of phenotypes: an adaptation to fluctuating environments. *J. Theor. Biol.* 181: 1–9.

Leibler, S., and E. Kussell, 2010 Individual histories and selection in heterogeneous populations. *Proc. Natl. Acad. Sci. USA* 107(29): 13183–13188.

Leigh, E. G., 1970 Natural selection and mutability. *Am. Nat.* 104(937): 301–305.

Liberman, U., and M. W. Feldman, 1986 Modifiers of mutation rate: a general reduction principle. *Theor. Popul. Biol.* 30(1): 125–142.

Liberman, U., and M. W. Feldman, 1989 The reduction principle for genetic modifiers of the migration rate, pp. 111–137 in *Mathematical Evolutionary Theory*, edited by M. W. Feldman. Princeton University Press, Princeton, NJ.

Liberman, U., J. van Cleve, and M. W. Feldman, 2011 On the evolution of mutation in changing environments: recombination and phenotypic switching. *Genetics* 187: 837–851.

Lim, H. N., and A. van Oudenaarden, 2007 A multistep epigenetic switch enables the stable inheritance of DNA methylation states. *Nat. Genet.* 39(2): 269–275.

Mettetal, J. T., D. Muzzey, and J. M. Pedraza, E. M. Ozbudak, and A. van Oudenaarden, 2006 Predicting stochastic gene expression dynamics in single cells. *Proc. Natl. Acad. Sci. USA* 103: 7304–7309.

Moran, N. A., 1992 The evolutionary maintenance of alternative phenotypes. *Am. Nat.* 139(5): 971–989.

- Rainey, P. B., H. J. E. Beaumont, G. C. Ferguson, J. Gallie, C. Kost *et al.*, 2011 The evolutionary emergence of stochastic phenotype switching in bacteria. *Microb. Cell Fact.* 10: S14.
- Salathe, M., J. Van Cleve, and M. W. Feldman, 2009 Evolution of stochastic switching rates in asymmetric fitness landscapes. *Genetics* 182: 1159–1164.
- Smits, W. K., O. P. Kuipers, and J. -W. Veening, 2006 Phenotypic variation in bacteria: the role of feedback regulation. *Nat. Rev. Microbiol.* 4(4): 259–271.
- Starrfelt, J., and H. Kokko, 2012 Bet-hedging—a triple trade-off between means, variances and correlations. *Biol. Rev. Camb. Philos. Soc.* 87: 742–755.
- Thattai, M., and A. van Oudenaarden, 2004 Stochastic gene expression in fluctuating environments. *Genetics* 167: 523–530.
- Veening, J.-W., W. K. Smits, and O. P. Kuipers, 2008 Bistability, epigenetics, and bet-hedging in bacteria. *Annu. Rev. Microbiol.* 62: 193–210.

*Communicating editor: L. M. Wahl*

# GENETICS

**Supporting Information**

<http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.161364/-/DC1>

## **The Evolution of Phenotypic Switching in Subdivided Populations**

**Oana Carja, Uri Liberman, and Marcus W. Feldman**

## File S1

### Proof of Result 1

1. When  $y = x$ , the mean fitnesses in the two demes  $E_x$  and  $E_y$  are equal:

$$w_x = 1 + sx = 1 + sy = w_y, \quad (S1.1)$$

and, from (8), the equilibrium equation is, (with  $\mu = \mu_B$ ),

$$(1 + sx)x = (1 - m)[(1 - \mu)(1 + s)x + \mu(1 - x)] + m[(1 - \mu)(1 - x) + \mu(1 + s)x]. \quad (S1.2)$$

Thus

$$(1 + sx)x = (1 + s)x[(1 - m)(1 - \mu) + m\mu] + (1 - x)[\mu(1 - m) + m(1 - \mu)], \quad (S1.3)$$

or

$$x + sx^2 = (1 + s)x[1 - m - \mu + 2m\mu] + (1 - x)[m + \mu - 2m\mu]. \quad (S1.4)$$

This is equivalent to

$$Q(x) = sx^2 + [(s + 2)(m + \mu - 2m\mu) - s]x - (m + \mu - 2m\mu) = 0. \quad (S1.5)$$

Now, as  $0 < m, \mu < 1$ , we have

$$(m + \mu - 2m\mu) = m(1 - \mu) + \mu(1 - m) > 0. \quad (S1.6)$$

Therefore

$$Q(0) = -(m + \mu - 2m\mu) < 0 \quad (S1.7)$$

and

$$Q(1) = (s + 1)(m + \mu - 2m\mu) > 0. \quad (S1.8)$$

As  $Q(\pm\infty) > 0$ , we conclude that the equation (S1.5) has a unique root  $x^*$  with  $0 < x^* < 1$ . Thus there is a unique symmetric polymorphism  $(\mathbf{x}^*, \mathbf{y}^*)$ , given by (13).

2. Near the equilibrium  $(\mathbf{x}^*, \mathbf{y}^*)$ , on the boundary where only  $B$  is present,  $(z - x)$  is small, and from (10), the internal local stability of  $(\mathbf{x}^*, \mathbf{y}^*)$  in the boundary is determined by the factor

$$C^* = \frac{(1 - 2\mu)(1 + s)}{(1 + sx^*)(1 + sz^*)}. \quad (S1.9)$$

As  $x^* = z^*$ ,  $C^* < 1$  if  $(1+s) < (1+sx^*)^2$ , and as  $s > 0$  this is true if  $s(x^*)^2 + 2x^* > 1$ .

From the equilibrium equation (S1.5), as  $Q(x^*) = 0$  we have

$$\begin{aligned} s(x^*)^2 + 2x^* &= -[(s+2)(m+\mu-2m\mu) - s]x^* + (m+\mu-2m\mu) + 2x^* \\ &= -(s+2)(m+\mu-2m\mu-1)x^* + (m+\mu-2m\mu). \end{aligned} \quad (S1.10)$$

Thus  $s(x^*)^2 + 2x^* > 1$  if and only if

$$(s+2)(1-m-\mu+2m\mu)x^* > (1-m-\mu+2m\mu). \quad (S1.11)$$

But  $(1-m-\mu+2m\mu) = (1-m)(1-\mu) + m\mu > 0$  as  $0 < m, \mu < 1$ , and so  $C^* < 1$  provided  $x^* > \frac{1}{s+2}$ . As  $Q(1) > 0$  and  $Q(x^*) = 0$ , it is sufficient to show that  $Q(\frac{1}{s+2}) < 0$ . Indeed

$$\begin{aligned} Q\left(\frac{1}{s+2}\right) &= \frac{s}{(s+2)^2} + [(s+2)(m+\mu-2m\mu) - s]\frac{1}{s+2} - (m+\mu-2m\mu) \\ &= \frac{s}{(s+2)^2} - \frac{s}{s+2} = -\frac{s(s+1)}{(s+2)^2} < 0. \end{aligned} \quad (S1.12)$$

3. We compute  $Q(\frac{1}{2})$  using (14),

$$Q\left(\frac{1}{2}\right) = \frac{s}{4} + \frac{1}{2}[(s+2)(m+\mu-2m\mu) - s] - (m+\mu-2m\mu). \quad (S1.13)$$

In fact,

$$Q\left(\frac{1}{2}\right) = -\frac{s}{4}[1 - 2(m+\mu-2m\mu)]. \quad (S1.14)$$

But  $1 - 2(m+\mu-2m\mu) = (1-2m)(1-2\mu) > 0$  when  $0 < m, \mu < \frac{1}{2}$ , in which case  $Q(\frac{1}{2}) < 0$  and  $x^* > \frac{1}{2}$  as  $Q(1) > 0$ .



**File S2**

**Proof of Result 2**

If an asymmetric polymorphism exists, then (11) holds, namely, (with  $\mu = \mu_B$ ),

$$1 + sy = \frac{(1 - 2\mu)(1 + s)}{1 + sx}. \quad (\text{S2.1})$$

That is,

$$y = \frac{s(1 - x) - 2\mu(1 + s)}{s(1 + sx)}, \quad 1 - y = \frac{s(1 + s)x + 2\mu(1 + s)}{s(1 + sx)}. \quad (\text{S2.2})$$

Substituting these relations into the equilibrium equation for  $x$  from (8), we find, after some simplification, that

$$x = \frac{1 - m}{1 + sx} [(1 - \mu)(1 + s)x + \mu(1 - x)] + \frac{m}{s}(sx + 2\mu + \mu s). \quad (\text{S2.3})$$

Equation (S2.3) is equivalent to the quadratic equation

$$T(x) = (1 - m)s^2x^2 - sx[s(1 - m) - \mu(s + 2)(1 - 2m)] - \mu(2m + s) = 0. \quad (\text{S2.4})$$

As  $\mu, m, s$  are positive and  $m < 1$ , we have  $T(0) < 0$  and  $T(\pm\infty) > 0$ , implying that  $T(x)$  has two real roots, one positive and one negative. Now

$$\begin{aligned} T(1) &= (1 - m)s^2 - s[s(1 - m) - \mu(s + 2)(1 - 2\mu)] - \mu(2m + s) \\ &= \mu[s(s + 2)(1 - 2m) - (2m + s)]. \end{aligned} \quad (\text{S2.5})$$

$T(1; m)$  is a linear function of  $m$  and

$$\begin{aligned} T(1; 0) &= \mu s(s + 1) > 0 \\ T(1; \frac{1}{2}) &= -\mu(2m + s) < 0 \\ T(1; m_0) &= 0. \end{aligned} \quad (\text{S2.6})$$

Hence if  $0 < m < m_0$ ,  $T(1; m) > 0$  and a unique  $0 < \hat{x} < 1$  exists such that  $T(\hat{x}) = 0$ . In order for  $\hat{x}$  to be an equilibrium, its corresponding  $\hat{y}$  should satisfy  $0 < \hat{y} < 1$ , where

$$1 - \hat{y} = \frac{1 + s}{1 + s\hat{x}} \frac{s\hat{x} + 2\mu}{s} \quad (\text{S2.7})$$

and  $0 < \hat{y} < 1$  if and only if

$$(1 + s)(s\hat{x} + 2\mu) < s(1 + s\hat{x}) \quad (\text{S2.8})$$

or

$$\hat{x} < \frac{s - 2\mu(1 + s)}{s}. \quad (S2.9)$$

So  $0 < \hat{x} < 1$  if  $0 < \mu < \mu_0 = \frac{1}{2} \frac{s}{s+1}$ , and  $[s - 2\mu(1 + s)] > 0$ . We compute  $T\left(\frac{s-2\mu(1+s)}{s}\right)$ , which equals

$$(1 - m)[s - 2\mu(1 + s)]^2 - [s - 2\mu(1 + s)][s(1 - m) - \mu(s + 2)(1 - 2m)] - \mu(2m + s). \quad (S2.10)$$

So

$$T\left(\frac{s-2\mu(1+s)}{s}\right) = 2\mu^2(1 + s)(s + 2m) + s\mu(s + 2)(1 - 2m) - \mu(2m + s). \quad (S2.11)$$

But when  $0 < m < m_0$ ,

$$T(1) = s\mu(s + 2)(1 - 2m) - \mu(2m + s) > 0, \quad (S2.12)$$

therefore  $T\left(\frac{s-2\mu(1+s)}{s}\right) > 0$ , and (S2.9) holds.

**File S3**

**Proof of Result 3**

The asymmetric equilibrium  $(\hat{x}, \hat{y})$  is determined by the solutions of the quadratic equation

$$T(x; \mu_B) = (1 - m)s^2x^2 - sx[s(1 - m) - \mu_B(s + 2)(1 - 2m)] - \mu_B(2m + s) = 0. \quad (\text{S3.1})$$

When  $\mu_B = 0$ , equation (S3.1) reduces to

$$T(x; 0) = -(1 - m)s^2x(1 - x) = 0, \quad (\text{S3.2})$$

giving the two solutions  $\hat{x} = 0$  and  $\hat{x} = 1$ . As  $T(0; \mu_B) < 0$ , when  $\mu_B > 0$  the solution  $x = 0$  shifts to a negative solution of (S3.1). Hence, when  $\mu_B$  is positive and small, the positive root  $\hat{x}(\mu_B)$  of  $T(x; \mu_B) = 0$  is close to  $x = 1$ . That is, when  $\mu_B$  is small the corresponding asymmetric equilibrium is close to the fixation of  $AB$  where  $\hat{x} = \hat{y} = 1$ . Moreover, by continuity, if  $\mu_B$  is small, their stability is the same. Near fixation of  $AB$ ,  $w = 1 - x$  and  $z = 1 - y$  are small, and up to non-linear terms, when  $\mu_B = 0$ , we have

$$\begin{aligned} w' &= \frac{1 - m}{1 + s}w + m(s + 1)z \\ z' &= \frac{m}{1 + s}w + (1 - m)(s + 1)z. \end{aligned} \quad (\text{S3.3})$$

The characteristic polynomial  $P(\lambda)$  of (S3.3) is

$$P(\lambda) = \lambda^2 - (1 - m) \left[ (1 + s) + \frac{1}{1 + s} \right] \lambda + (1 - 2m) \quad (\text{S3.4})$$

and

$$P(1) = 1 - (1 - m) \frac{(1 + s)^2 + 1}{1 + s} + 1 - 2m. \quad (\text{S3.5})$$

In fact, it can be easily seen that

$$(1 + s)P(1) = -s^2(1 - m). \quad (\text{S3.6})$$

As  $P(+\infty) > 0$  and  $P(1) < 0$ , since  $s > 0$ ,  $0 < m < 1$ ,  $P(\lambda)$  has a root larger than 1. Thus, when  $\mu_B$  is small, fixation in  $AB$  is internally locally unstable and so is the asymmetric equilibrium when  $\mu_B$  is small.

**File S4**

**Proof of Result 4**

A straightforward computation shows that the  $4 \times 4$  matrix  $\mathbf{L}_{\text{ex}}$  can be written as

$$\mathbf{L}_{\text{ex}} = \begin{pmatrix} (1-m)A & (1-m)B & mC & mD \\ (1-m)D & (1-m)C & mB & mA \\ mA & mB & (1-m)C & (1-m)D \\ mD & mC & (1-m)B & (1-m)A \end{pmatrix}, \quad (\text{S4.1})$$

where

$$\begin{aligned} (1+sx^*)A &= (1+s)(1-\mu_b) + r(1-x^*)[(s+2)\mu_b - (s+1)] \\ (1+sx^*)B &= (1+s)rx^* + \mu_b[1 - (s+2)rx^*] \\ (1+sx^*)C &= (1-\mu_b) + rx^*[(s+2)\mu_b - 1] \\ (1+sx^*)D &= (1+s)\mu_b + r(1-x^*)[1 - (s+2)\mu_b]. \end{aligned} \quad (\text{S4.2})$$

Observe that “formally”  $A, B, C, D$  are linear in  $\mu_b$ . Let  $A_0$  be the value of  $A$  when  $\mu_b = 0$  and  $A_1$  be its value when  $\mu_b = 1$ . Similarly we have  $B_0, B_1, C_0, C_1, D_0, D_1$ . In fact,

$$\begin{aligned} (1+sx^*)A_0 &= (1+s)[1 - r(1-x^*)] \\ (1+sx^*)A_1 &= r(1-x^*) \\ (1+sx^*)B_0 &= (1+s)rx^* \\ (1+sx^*)B_1 &= 1 - rx^* \\ (1+sx^*)C_0 &= 1 - rx^* \\ (1+sx^*)C_1 &= (1+s)rx^* \\ (1+sx^*)D_0 &= r(1-x^*) \\ (1+sx^*)D_1 &= (1+s)[1 - r(1-x^*)]. \end{aligned} \quad (\text{S4.3})$$

As  $0 < r < 1$ ,  $0 < x^* < 1$  we have  $A_i, B_i, C_i, D_i$  positive for  $i = 0, 1$ . Hence, as  $A, B, C, D$  are linear in  $\mu_b$ ,  $A, B, C, D$  are all positive for  $0 < \mu_b < 1$ . Moreover we have

$$C_0 = B_1, \quad C_1 = B_0, \quad D_0 = A_1, \quad D_1 = A_0. \quad (\text{S4.4})$$

Let  $S(\lambda) = \det(\mathbf{L}_{\text{ex}} - \lambda\mathbf{I})$  be the characteristic polynomial of  $\mathbf{L}_{\text{ex}}$ . The structure of  $\mathbf{L}_{\text{ex}}$  given in (S4.1) entails that  $S(\lambda)$  factors into the product of two quadratic polynomials  $S_1(\lambda)$  and  $S_2(\lambda)$ :

$$S(\lambda) = S_1(\lambda)S_2(\lambda), \quad (\text{S4.5})$$

where

$$\begin{aligned} S_1(\lambda) &= \lambda^2 - \lambda[(1-m)(A+C) + m(B+D)] + (1-2m)(AC - BD) \\ S_2(\lambda) &= \lambda^2 - \lambda[(1-m)(A+C) - m(B+D)] + (1-2m)(AC - BD). \end{aligned} \quad (S4.6)$$

See Balkau and Feldman (1973) for analogous calculations with migration modification. Consider first the roots of  $S_1(\lambda) = 0$ . These are real since the discriminant of  $S_1(\lambda) = 0$  is

$$\begin{aligned} & [(1-m)(A+C) + m(B+D)]^2 - 4(1-2m)(AC - BD) = \\ & = [(1-m)(A-C) + m(B-D)]^2 + 4m(1-m)(AD + BC) + 4(1-m)^2 BD + 4m^2 AC, \end{aligned} \quad (S4.7)$$

which is positive since  $A, B, C, D$  are positive and  $0 < m < 1$ .

In addition,

$$\begin{aligned} AC - BD &= [(1-\mu_b)A_0 + \mu_b A_1] [(1-\mu_b)(C_0 + \mu_b C_1)] \\ &\quad - [(1-\mu_b)B_0 + \mu_b B_1] [(1-\mu_b)D_0 + \mu_b D_1]. \end{aligned} \quad (S4.8)$$

Since  $C_0 = B_1, C_1 = B_0, D_0 = A_1, D_1 = A_0$ , (S4.8) reduces to

$$AC - BD = (1 - 2\mu_b)(A_0 B_1 - A_1 B_0). \quad (S4.9)$$

Substituting  $A_0, A_1, B_0, B_1$  from (S4.3) we have

$$(1 + sx^*)^2 (AC - BD) = (1 - 2\mu_b)(1 + s)(1 - r). \quad (S4.10)$$

Since we assume  $0 < m, \mu_b < \frac{1}{2}$ , the two roots of  $S_1(\lambda) = 0$  are positive. Both of these roots are less than 1 if and only if  $S_1(1) > 0$  and  $S'_1(1) > 0$ .

$$S'_1(1) = 2 - [(1-m)(A+C) + m(B+D)]. \quad (S4.11)$$

As  $C_0 = B_1, C_1 = B_0, D_0 = A_1, D_1 = A_0$ , we have

$$\begin{aligned} (A+C) &= (1-\mu_b)(A_0 + C_0) + \mu_b(A_1 + C_1) \\ &= (1-\mu_b)(A_0 + C_0) + \mu_b(B_0 + D_0), \end{aligned} \quad (S4.12)$$

$$\begin{aligned} (B+D) &= (1-\mu_b)(B_0 + D_0) + \mu_b(B_1 + D_1) \\ &= (1-\mu_b)(B_0 + D_0) + \mu_b(A_0 + C_0). \end{aligned} \quad (S4.13)$$

Hence,

$$(1 - m)(A + C) + m(B + D) = (1 - m_b)(A_0 + C_0) + m_b(B_0 + D_0), \quad (S4.14)$$

where

$$m_b = m + \mu_b - 2m\mu_b. \quad (S4.15)$$

Substituting for  $A_0, B_0, C_0, D_0$ , gives

$$S'_1(1) = (1 + sx^*)^{-1} [r + rsx^* + s(1 - r)(2x^* - 1) + m_b(1 - r)(s + 2)]. \quad (S4.16)$$

Now  $s > 0$ ,  $0 < r < 1$ ,  $m_b = m(1 - \mu_b) + \mu_b(1 - m) > 0$ , and  $x^* > \frac{1}{2}$  if  $0 < m, \mu_B < \frac{1}{2}$ . Therefore  $S'(1) > 0$  provided  $0 < m, \mu_B < \frac{1}{2}$ . Using (S4.11) and (S4.16) it is easily seen that  $S(1) > 0$  if

$$(1 + sx^*)^{-2}(1 - r) \left\{ (x^*)^2 s^2 + sx^* [-s + m_b(s + 2)] - sm_b \right\} > 0. \quad (S4.17)$$

Using the equation  $Q(x^*) = 0$  from (14), we have

$$s(x^*)^2 + [(s + 2)m_B - s]x^* - m_B = 0, \quad (S4.18)$$

where

$$m_B = m + \mu_B - 2m\mu_B. \quad (S4.19)$$

Therefore (S4.17) is satisfied if and only if

$$(m_b - m_B)(1 + sx^*)^{-2}(1 - r)s[x^*(2 + s) - 1] > 0. \quad (S4.20)$$

As  $x^* > \frac{1}{2}$ , by Result 1, and  $0 < m < \frac{1}{2}$ , (S4.20) holds if and only if  $m_b > m_B$ , which is true if and only if  $\mu_b > \mu_B$ .

It is not obvious that the roots of  $S_2(\lambda) = 0$  are real. However, as the matrix  $\mathbf{L}_{\text{ex}}^*$  is positive, the Perron-Frobenius theory ensures that its largest eigenvalue in magnitude is positive. Therefore we just have to ensure that when both eigenvalues are real and positive they are less than 1; when they are real, both are positive or both are negative

since  $(1 - 2m)(1 - 2\mu_b)(AC - BD)$  is positive for  $0 < m, \mu_b < \frac{1}{2}$ . The conditions for this are that both  $S_2(1)$  and  $S'_2(1)$  are positive. But

$$\begin{aligned} S_2(1) &= 1 - [(1 - m)(A + C) - m(B + D)] + (1 - 2m)(AC - BD) \\ &> 1 - [(1 - m)(A + C) + m(B + D)] + (1 - 2m)(AC - BD) = S_1(1), \end{aligned} \tag{S4.21}$$

and  $S_1(1) > 0$  when  $0 < m, \mu_b < \frac{1}{2}$  and  $\mu_b > \mu_B$ , so also  $S_2(1) > 0$ . Similarly

$$\begin{aligned} S'_2(1) &= 2 - [(1 - m)(A + C) - m(B + D)] \\ &> 2 - [(1 - m)(A + C) + m(B + D)] = S'_1(1). \end{aligned} \tag{S4.22}$$

Thus, when  $S'_1(1) > 0$  also  $S'_2(1) > 0$ .

**File S5**

**Proof of Result 5**

At a symmetric equilibrium  $y = x$ , and also, by (32),  $\tilde{y} = \tilde{x}$ . Thus (30) and (31) imply that

$$\tilde{x} = \frac{[(s+1)(1-m_B) - m_B]x + m_B}{sx + 1} \quad (S5.1)$$

and

$$x = \frac{[(1-m_B) - m_B(s+1)]\tilde{x} + m_B(1+s)}{(1+s) - s\tilde{x}}. \quad (S5.2)$$

Substituting (S5.1) into (S5.2) gives the quadratic equation

$$(s+2)m_B \{sx^2 + [2 - m_B(s+2)]x - (1 - m_B)\} = 0. \quad (S5.3)$$

As  $0 < m, \mu_B < 1$ ,  $s > 0$  and  $m_B = m(1 - \mu_B) + \mu_B(1 - m) > 0$ ,  $x$  satisfies the equation  $R(x) = 0$  with  $R(x)$  given in (36). As  $0 < m_B < 1$  we have  $R(0) < 0$ , and as  $R(\pm\infty) > 0$ ,  $R(x) = 0$  has two real roots, one positive and one negative. Observe that

$$R(1) = s + [2 - m_B(s+2)] - (1 - m_B) = (1 - m_B)(s+1) > 0 \quad (S5.4)$$

and

$$R\left(\frac{1}{2}\right) = \frac{s}{4} + \frac{1}{2} \cdot [2 - m_B(s+2)] - (1 - m_B) = \frac{s}{4}(1 - 2m)(1 - 2\mu_B) \quad (S5.5)$$

as  $1 - 2m_B = 1 - 2m - 2\mu_B + 4m\mu_B = (1 - 2m)(1 - 2\mu_B)$ . Therefore when  $0 < m, \mu_B < \frac{1}{2}$  we have  $R(\frac{1}{2}) > 0$  and  $0 < \bar{x} < \frac{1}{2}$ .



**File S6**

**Proof of Result 6**

In view of (33), the symmetric equilibrium  $(\bar{x}, \bar{y})$  is internally stable if

$$\frac{(1 - 2\mu)^2(1 + s)^2}{(1 + s\bar{x})^2[1 + s(1 - \tilde{x})]^2} < 1, \quad (S6.1)$$

as  $\bar{x} = \bar{y}$  and  $\tilde{x} = \tilde{y}$ , where, by (S5.1)

$$\tilde{x} = \frac{[(s + 1)(1 - m_B) - m_B]\bar{x} + m_B}{s\bar{x} + 1}. \quad (S6.2)$$

Thus

$$1 + s(1 - \tilde{x}) = (1 + s) - s \cdot \frac{[(1 + s) - m_B(2 + s)]\bar{x} + m_B}{s\bar{x} + 1}. \quad (S6.3)$$

Hence

$$(1 + s\bar{x})[1 + s(1 - \tilde{x})] = (1 + s)(1 + s\bar{x}) - s[(1 + s) - m_B(s + 2)]\bar{x} - sm_B, \quad (S6.4)$$

or

$$(1 + s\bar{x})[1 + s(1 - \tilde{x})] = (1 + s) + sm_B[(s + 2)\bar{x} - 1]. \quad (S6.5)$$

For condition (S6.1) to be satisfied, since  $(1 + s\bar{x})[1 + s(1 - \tilde{x})] > 0$ , it is sufficient that

$$(1 + s) + sm_B[(s + 2)\bar{x} - 1] > (1 + s), \quad (S6.6)$$

or that  $\bar{x} > \frac{1}{s+2}$ . But

$$R\left(\frac{1}{s+2}\right) = \frac{s}{(s+2)^2} + [2 - m_B(s+2)]\frac{1}{s+2} - (1 - m_B), \quad (S6.7)$$

or

$$R\left(\frac{1}{s+2}\right) = \frac{s}{(s+2)^2} + \frac{2}{s+2} - 1 = -\frac{s(1+s)}{(s+2)^2} < 0. \quad (S6.8)$$

Thus  $R(\frac{1}{s+2}) < 0$  and  $R(1) > 0$ , and so  $\bar{x} > \frac{1}{s+2}$  as desired.

**File S7**

**Proof of Result 7**

As the transformation  $T$  of the population state is  $T = T_2 \circ T_1$ , where  $T_i$  corresponds to phase  $i$ , with selection of type  $i$ , for  $i = 1, 2$ , and as  $\tilde{\mathbf{x}} = T_1 \bar{\mathbf{x}}$ ,  $\bar{\mathbf{x}} = T_2 \tilde{\mathbf{x}}$ , following the analysis for the case without cycles, the linear approximation matrix  $\mathbf{L}_{\text{ex}}$  becomes

$$\mathbf{L}_{\text{ex}} = \mathbf{L}_{\text{ex}}^2 \cdot \mathbf{L}_{\text{ex}}^1, \quad (\text{S7.1})$$

where, as in (S4.1) and (S4.2), we have

$$\mathbf{L}_{\text{ex}}^1 = \begin{pmatrix} (1-m)\bar{A} & (1-m)\bar{B} & m\bar{C} & m\bar{D} \\ (1-m)\bar{D} & (1-m)\bar{C} & m\bar{B} & m\bar{A} \\ m\bar{A} & m\bar{B} & (1-m)\bar{C} & (1-m)\bar{D} \\ m\bar{D} & m\bar{C} & (1-m)\bar{B} & (1-m)\bar{A} \end{pmatrix}, \quad (\text{S7.2})$$

$$\mathbf{L}_{\text{ex}}^2 = \begin{pmatrix} (1-m)\tilde{A} & (1-m)\tilde{B} & m\tilde{C} & m\tilde{D} \\ (1-m)\tilde{D} & (1-m)\tilde{C} & m\tilde{B} & m\tilde{A} \\ m\tilde{A} & m\tilde{B} & (1-m)\tilde{C} & (1-m)\tilde{D} \\ m\tilde{D} & m\tilde{C} & (1-m)\tilde{B} & (1-m)\tilde{A} \end{pmatrix}, \quad (\text{S7.3})$$

and

$$\begin{aligned} (1+s\bar{x})\bar{A} &= (1+s)(1-\mu_b) + r(1-\bar{x})[(s+2)\mu_b - (s+1)] \\ (1+s\bar{x})\bar{B} &= (1+s)r\bar{x} + \mu_b[1 - (s+2)r\bar{x}] \\ (1+s\bar{x})\bar{C} &= (1-\mu_b) + r\bar{x}[(s+2)\mu_b - 1] \\ (1+s\bar{x})\bar{D} &= (1+s)\mu_b + r(1-\bar{x})[1 - (s+2)\mu_b], \end{aligned} \quad (\text{S7.4})$$

$$\begin{aligned} [1+s(1-\tilde{x})]\tilde{A} &= (1-\mu_b) + r(1-\tilde{x})[(2+s)\mu_b - 1] \\ [1+s(1-\tilde{x})]\tilde{B} &= (1+s)\mu_b + r\tilde{x}[1 - (s+2)\mu_b] \\ [1+s(1-\tilde{x})]\tilde{C} &= (1+s)(1-\mu_b) + r\tilde{x}[(s+2)\mu_b - (s+1)] \\ [1+s(1-\tilde{x})]\tilde{D} &= \mu_b + r(1-\tilde{x})[(s+1) - (s+2)\mu_b]. \end{aligned} \quad (\text{S7.5})$$

When we multiply  $\mathbf{L}_{\text{ex}}^2$  by  $\mathbf{L}_{\text{ex}}^1$  we find that the product  $\mathbf{L}_{\text{ex}}$  has the following structure:

$$\mathbf{L}_{\text{ex}} = \begin{pmatrix} a & e & h & d \\ b & f & g & c \\ c & g & f & b \\ d & h & e & a \end{pmatrix}, \quad (\text{S7.6})$$

where

$$\begin{aligned}
a &= (1-m)^2 \tilde{A}\bar{A} + (1-m)^2 \tilde{B}\bar{D} + m^2 \tilde{C}\bar{A} + m^2 \tilde{D}\bar{D} \\
b &= (1-m)^2 \tilde{D}\bar{A} + (1-m)^2 \tilde{C}\bar{D} + m^2 \tilde{B}\bar{A} + m^2 \tilde{A}\bar{D} \\
c &= m(1-m) \left[ \tilde{A}\bar{A} + \tilde{B}\bar{D} + \tilde{C}\bar{A} + \tilde{D}\bar{D} \right] \\
d &= m(1-m) \left[ \tilde{D}\bar{A} + \tilde{C}\bar{D} + \tilde{B}\bar{A} + \tilde{A}\bar{D} \right] \\
e &= (1-m)^2 \tilde{A}\bar{B} + (1-m)^2 \tilde{B}\bar{C} + m^2 \tilde{C}\bar{B} + m^2 \tilde{D}\bar{C} \\
f &= (1-m)^2 \tilde{D}\bar{B} + (1-m)^2 \tilde{C}\bar{C} + m^2 \tilde{B}\bar{B} + m^2 \tilde{A}\bar{C} \\
g &= m(1-m) \left[ \tilde{A}\bar{B} + \tilde{B}\bar{C} + \tilde{C}\bar{B} + \tilde{D}\bar{C} \right] \\
h &= m(1-m) \left[ \tilde{D}\bar{B} + \tilde{C}\bar{C} + \tilde{B}\bar{B} + \tilde{A}\bar{C} \right].
\end{aligned} \tag{S7.7}$$

Let  $D(\lambda) = \det(\mathbf{L}_{\text{ex}} - \lambda \mathbf{I})$  be the characteristic polynomial of  $\mathbf{L}_{\text{ex}}$ . From (S7.6),  $D(\lambda)$  factors into  $2 \times 2$  determinants:

$$D(\lambda) = \begin{vmatrix} a+d-\lambda & e+h & 0 & 0 \\ b+c & f+g-\lambda & 0 & 0 \\ c & g & f-g-\lambda & b-c \\ d & h & e-h & a-d-\lambda \end{vmatrix}. \tag{S7.8}$$

Therefore  $D(\lambda)$  can be written

$$D(\lambda) = D_1(\lambda)D_2(\lambda), \tag{S7.9}$$

where

$$\begin{aligned}
D_1(\lambda) &= \lambda^2 - (a+d+f+g)\lambda + (a+d)(f+g) - (b+c)(e+h) \\
D_2(\lambda) &= \lambda^2 - (a-d+f-g)\lambda + (a-d)(f-g) - (b-c)(e-h).
\end{aligned} \tag{S7.10}$$

As  $0 < m < 1$  and  $\bar{A}, \bar{B}, \bar{C}, \bar{D}$  and  $\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}$  are all positive, the matrix  $\mathbf{L}_{\text{ex}}$  is a positive matrix and its largest eigenvalue in magnitude is positive. Observe that the discriminant of  $D_1(\lambda)$  is

$$(a+d+f+g)^2 - 4[(a+d)(f+g) - (b+c)(e+h)], \tag{S7.11}$$

which is positive and equal to

$$[(a+d) - (f+g)]^2 + 4(b+c)(e+h). \tag{S7.12}$$

In addition,  $(a+d+f+g)$  is positive. Therefore  $D_1(\lambda)$  has real roots, and its largest root in magnitude is positive. Thus this positive root is less than 1 if  $D_1(1) > 0$  and  $D_1'(1) > 0$ .

As the largest eigenvalue of  $\mathbf{L}_{\text{ex}}$  is positive, for stability of  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$  we require that if the eigenvalues associated with  $D_2(\lambda)$  are real and at least one is positive, they are both less than 1. Again the conditions for this are  $D_2(\lambda) > 0$  and  $D_2'(1) > 0$ . Observe that

$$\begin{aligned} D_1'(1) &= 2 - (a + d + f + g) \\ D_2'(1) &= 2 - (a - d + f - g) = D_1'(1) + 2(d + g) > D_1'(1). \end{aligned} \tag{S7.13}$$

In view of (S7.13), for the largest eigenvalue of  $\mathbf{L}_{\text{ex}}$  to be less than one, we require

$$D_1(1) > 0, \quad D_1'(1) > 0, \quad D_2(1) > 0. \tag{S7.14}$$

We now compute the constant terms of  $D_1(\lambda)$  and  $D_2(\lambda)$ . We already know, based on the properties of the matrices  $\mathbf{L}_{\text{ex}}^1$  and  $\mathbf{L}_{\text{ex}}^2$  that the constant terms of both  $D_1(\lambda)$  and  $D_2(\lambda)$  are the same and are equal to

$$(1 - 2m)^2 (\overline{A\overline{C}} - \overline{B\overline{D}}) (\tilde{A}\tilde{C} - \tilde{B}\tilde{D}). \tag{S7.15}$$

With the same technique used to compute (S4.10), we deduce that

$$(1 + s\bar{x})^2 (\overline{A\overline{C}} - \overline{B\overline{D}}) = (1 - 2\mu_b)(1 + s)(1 - r), \tag{S7.16}$$

and similarly

$$[1 + s(1 - \tilde{x})]^2 (\tilde{A}\tilde{C} - \tilde{B}\tilde{D}) = (1 - 2\mu_b)(1 + s)(1 - r). \tag{S7.17}$$

Therefore the constant terms of both  $D_1(\lambda)$  and  $D_2(\lambda)$  are the same and are equal to

$$(1 - 2m)^2 (1 - 2\mu_b)^2 (1 + s\bar{x})^{-2} [1 + s(1 - \tilde{x})]^{-2} (1 + s)^2 (1 - r)^2, \tag{S7.18}$$

which is positive, and so  $D_1(\lambda)$  has two positive roots. Also, as  $a, b, c, d$  are all positive,

$$\begin{aligned} D_2(1) &= 1 - (a - d + f - g) + (a - d)(f - g) - (b - c)(e - h) \\ &> 1 - (a + d + f + g) + (a + d)(f + g) - (b + c)(e + h) = D_1(1). \end{aligned} \tag{S7.19}$$

Hence for the symmetric equilibrium to be externally stable, we require that  $D_1(1)$  and  $D_1'(1)$  are both positive.

Now from (S6.5) we know that

$$(1 + s\bar{x})[1 + s(1 - \tilde{x})] = (1 + s) + sm_B[(s + 2)\bar{x} - 1]. \quad (S7.20)$$

As  $\bar{x} > \frac{1}{s+2}$  we have

$$(1 + s\bar{x})[1 + s(1 - \tilde{x})] > (1 + s). \quad (S7.21)$$

Thus the equal constant terms of  $D_1(\lambda)$  and  $D_2(\lambda)$  given in (S7.18) are positive and less than 1. As a result it is impossible for the two positive roots of  $D_1(\lambda)$  to both be larger than 1, and they are both less than 1 provided  $D_1(1) > 0$ . Hence the external stability of the symmetric equilibrium requires that

$$D_1(1) = 1 - (a + d + f + g) + (a + d)(f + g) - (b + c)(e + h) \quad (S7.22)$$

is positive (the last summand in (S7.22) is given in (S7.18)). We now compute  $a + d + f + g$ .

*Computation of  $(a + d + f + g)$*

We have

$$\begin{aligned} (a + d + f + g) &= \left[ (1 - m)^2 \tilde{A} + m^2 \tilde{C} + m(1 - m) (\tilde{B} + \tilde{D}) \right] \bar{A} \\ &\quad + \left[ (1 - m)^2 \tilde{C} + m^2 \tilde{A} + m(1 - m) (\tilde{B} + \tilde{D}) \right] \bar{C} \\ &\quad + \left[ (1 - m)^2 \tilde{B} + m^2 \tilde{D} + m(1 - m) (\tilde{A} + \tilde{C}) \right] \bar{D} \\ &\quad + \left[ (1 - m)^2 \tilde{D} + m^2 \tilde{B} + m(1 - m) (\tilde{A} + \tilde{C}) \right] \bar{B}. \end{aligned} \quad (S7.23)$$

As  $\bar{A}, \bar{B}, \bar{C}, \bar{D}$  and also  $\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}$  given in (S7.4) and (S7.5), respectively, are all linear functions of  $\mu_b$ , where  $0 \leq \mu_b \leq 1$ , we can represent them as  $\bar{A} = (1 - \mu_b)\bar{A}_0 + \mu_b\bar{A}_1$ , etc.

Hence

$$\begin{aligned} (1 + s\bar{x})\bar{A}_0 &= (1 + s)[1 - r(1 - \bar{x})] = (1 + s\bar{x})\bar{D}_1 \\ (1 + s\bar{x})\bar{B}_0 &= (1 + s)r\bar{x} = (1 + s\bar{x})\bar{C}_1 \\ (1 + s\bar{x})\bar{C}_0 &= 1 - r\bar{x} = (1 + s\bar{x})\bar{B}_1 \end{aligned} \quad (S7.24)$$

$$\begin{aligned} (1 + s\bar{x})\bar{D}_0 &= r(1 - \bar{x}) = (1 + s\bar{x})\bar{A}_1, \\ [1 + s(1 - \tilde{x})]\tilde{A}_0 &= 1 - r(1 - \tilde{x}) = [1 + s(1 - \tilde{x})]\tilde{D}_1 \\ [1 + s(1 - \tilde{x})]\tilde{B}_0 &= r\tilde{x} = [1 + s(1 - \tilde{x})]\tilde{C}_1 \\ [1 + s(1 - \tilde{x})]\tilde{C}_0 &= (1 + s)[1 - r\tilde{x}] = [1 + s(1 - \tilde{x})]\tilde{B}_1 \\ [1 + s(1 - \tilde{x})]\tilde{D}_0 &= (1 + s)r(1 - \tilde{x}) = [1 + s(1 - \tilde{x})]\tilde{A}_1. \end{aligned} \quad (S7.25)$$

Since

$$\begin{aligned} m_b &= m + \mu_b - 2m\mu_b = m(1 - \mu_b) + \mu_b(1 - m) \\ 1 - m_b &= 1 - m - \mu_b + 2m\mu_b = (1 - m)(1 - \mu_b), \end{aligned} \quad (S7.26)$$

we can write

$$\begin{aligned} (a + d + f + g) &= \left[ (1 - m)(1 - m_b)\tilde{A}_0 + (1 - m)m_b\tilde{D}_0 + m(1 - m_b)\tilde{B}_0 + m \cdot m_b\tilde{C}_0 \right] \bar{A} \\ &+ \left[ (1 - m)(1 - m_b)\tilde{C}_0 + (1 - m)m_b\tilde{B}_0 + m(1 - m_b)\tilde{D}_0 + m \cdot m_b\tilde{A}_0 \right] \bar{C} \\ &+ \left[ (1 - m)(1 - m_b)\tilde{B}_0 + (1 - m)m_b\tilde{C}_0 + m(1 - m_b)\tilde{A}_0 + m \cdot m_b\tilde{D}_0 \right] \bar{D} \\ &+ \left[ (1 - m)(1 - m_b)\tilde{D}_0 + (1 - m)m_b\tilde{A}_0 + m(1 - m_b)\tilde{C}_0 + m \cdot m_b\tilde{B}_0 \right] \bar{B}. \end{aligned} \quad (S7.27)$$

Substitute into (S7.27)

$$\begin{aligned} \bar{A} &= (1 - \mu_b)\bar{A}_0 + \mu_b\bar{D}_0, & \bar{B} &= (1 - \mu_b)\bar{B}_0 + \mu_b\bar{C}_0, \\ \bar{C} &= (1 - \mu_b)\bar{C}_0 + \mu_b\bar{B}_0, & \bar{D} &= (1 - \mu_b)\bar{D}_0 + \mu_b\bar{A}_0, \end{aligned} \quad (S7.38)$$

to obtain

$$\begin{aligned} (a + d + f + g) &= (1 - m_b)^2 \left[ \tilde{A}_0\bar{A}_0 + \tilde{B}_0\bar{D}_0 + \tilde{C}_0\bar{C}_0 + \tilde{D}_0\bar{B}_0 \right] \\ &+ m_b(1 - m_b) \left[ \left( \tilde{B}_0 + \tilde{D}_0 \right) (\bar{A}_0 + \bar{C}_0) + \left( \tilde{A}_0 + \tilde{C}_0 \right) (\bar{B}_0 + \bar{D}_0) \right] \\ &+ m_b^2 \left[ \tilde{A}_0\bar{C}_0 + \tilde{B}_0\bar{B}_0 + \tilde{C}_0\bar{A}_0 + \tilde{D}_0\bar{D}_0 \right]. \end{aligned} \quad (S7.29)$$

Equation (S7.29) can also be written as

$$\begin{aligned} (a + d + f + g) &= (1 - 2m_b) \left[ \tilde{A}_0\bar{A}_0 + \tilde{B}_0\bar{D}_0 + \tilde{C}_0\bar{C}_0 + \tilde{D}_0\bar{B}_0 \right] \\ &+ m_b \left[ \left( \tilde{B}_0 + \tilde{D}_0 \right) (\bar{A}_0 + \bar{C}_0) + \left( \tilde{A}_0 + \tilde{C}_0 \right) (\bar{B}_0 + \bar{D}_0) \right] \\ &+ m_b^2 \left[ \left( \tilde{A}_0 + \tilde{C}_0 \right) (\bar{A}_0 + \bar{C}_0) + \left( \tilde{B}_0 + \tilde{D}_0 \right) (\bar{B}_0 + \bar{D}_0) \right. \\ &\quad \left. - \left( \tilde{A}_0 + \tilde{C}_0 \right) (\bar{B}_0 + \bar{D}_0) + \left( \tilde{B}_0 + \tilde{D}_0 \right) (\bar{A}_0 + \bar{C}_0) \right], \end{aligned} \quad (S7.30)$$

or as

$$\begin{aligned} (a + d + f + g) &= (1 - 2m_b) \left[ \tilde{A}_0\bar{A}_0 + \tilde{B}_0\bar{D}_0 + \tilde{C}_0\bar{C}_0 + \tilde{D}_0\bar{B}_0 \right] \\ &+ m_b \left[ \left( \tilde{B}_0 + \tilde{D}_0 \right) (\bar{A}_0 + \bar{C}_0) + \left( \tilde{A}_0 + \tilde{C}_0 \right) (\bar{B}_0 + \bar{D}_0) \right] \\ &+ m_b^2 \left( \tilde{A}_0 + \tilde{C}_0 - \tilde{B}_0 - \tilde{D}_0 \right) (\bar{A}_0 + \bar{C}_0 - \bar{B}_0 - \bar{D}_0). \end{aligned} \quad (S7.31)$$

From (S7.24) and (S7.25),

$$\begin{aligned} (1 + s\bar{x})[1 + s(1 - \tilde{x})] \left[ \tilde{A}_0\bar{A}_0 + \tilde{B}_0\bar{D}_0 + \tilde{C}_0\bar{C}_0 + \tilde{D}_0\bar{D}_0 \right] &= \\ &= (1 + s) \left[ 2(1 - r) + r^2 \right] + r^2s(s + 1)\bar{x} - r^2s\tilde{x} - r^2s^2\bar{x}\tilde{x}, \end{aligned} \quad (S7.32)$$

$$\begin{aligned} (1 + s\bar{x})[1 + s(1 - \tilde{x})] \left[ \left( \tilde{B}_0 + \tilde{D}_0 \right) (\bar{A}_0 + \bar{C}_0) + \left( \tilde{A}_0 + \tilde{C}_0 \right) (\bar{B}_0 + \bar{D}_0) \right] &= \\ &= 2r^2(1 + s\bar{x})[1 + s(1 - \tilde{x})] + r(1 - r)(s + 2) [(s + 2) + s(\bar{x} - \tilde{x})], \end{aligned} \quad (S7.33)$$

$$\begin{aligned} (1 + s\bar{x})[1 + s(1 - \tilde{x})] \left( \tilde{A}_0 + \tilde{C}_0 - \tilde{B}_0 - \tilde{D}_0 \right) (\bar{A}_0 + \bar{C}_0 - \bar{B}_0 - \bar{D}_0) &= \\ &= (s + 2)^2 (1 - r)^2. \end{aligned} \quad (S7.34)$$

Remember that by (S7.18)

$$\begin{aligned} (1 + s\bar{x})^2 [1 + (1 - \tilde{x})]^2 [(a + d)(f + g) - (b + c)(e + h)] &= \\ &= (1 - 2m)^2 (1 - 2\mu_b)^2 (s + 1)^2 (1 - r)^2. \end{aligned} \quad (S7.35)$$

But

$$(1 - 2m)(1 - 2\mu_b) = 1 - 2(m + \mu_b - 2m\mu_b) = 1 - 2m_b. \quad (S7.36)$$

Therefore

$$\begin{aligned} (1 + s\bar{x})^2 [1 + (1 - \tilde{x})]^2 [(a + d)(f + g) - (b + c)(e + h)] &= \\ &= (1 - 2m_b)^2 (s + 1)^2 (1 - r)^2. \end{aligned} \quad (S7.37)$$

Combining all of this, we get that  $D_1(1) = 1 - (a + d + f + g) + (a + d)(f + g) - (b + c)(e + h)$ , which we compute as

$$\begin{aligned} 1 - (1 - 2m_b) \frac{(1 + s) \left[ 2(1 - r) + r^2 \right] + r^2s(s + 1)\bar{x} - r^2s\tilde{x} - r^2s^2\bar{x}\tilde{x}}{(1 + s\bar{x})[1 + s(1 - \tilde{x})]} & \\ - m_b \frac{2r^2(1 + s\bar{x})[1 + s(1 - \tilde{x})] + r(1 - r)(s + 2)[(s + 2) + s(\bar{x} - \tilde{x})]}{(1 + s\bar{x})[1 + s(1 - \tilde{x})]} & \\ - m_b^2 \frac{(s + 2)^2 (1 - r)^2}{(1 + s\bar{x})[1 + s(1 - \tilde{x})]} & \\ + \frac{(1 - 2m_b)^2 (s + 1)^2 (1 - r)^2}{(1 + s\bar{x})^2 [1 + s(1 - \tilde{x})]^2}. & \end{aligned} \quad (S7.38)$$

Observe that

$$r^2[(1 + s) + s(s + 1)\bar{x} - s\tilde{x} - s^2\bar{x}\tilde{x}] = r^2(1 + s\bar{x})[1 + s(1 - \tilde{x})], \quad (S7.39)$$

so (S7.38) simplifies to

$$\begin{aligned}
1 - r^2 - (1 - 2m_b) \frac{2(1+s)(1-r)}{(1+s\bar{x})[1+s(1-\tilde{x})]} - m_b \frac{r(1-r)(s+2)[(s+2)+s(\bar{x}-\tilde{x})]}{(1+s\bar{x})[1+s(1-\tilde{x})]} \\
- m_b^2 \frac{(s+2)^2(1-r)^2}{(1+s\bar{x})[1+s(1-\tilde{x})]} + \frac{(1-2m_b)^2(s+1)^2(1-r)^2}{(1+s\bar{x})^2[1+s(1-\tilde{x})]^2}.
\end{aligned} \tag{S7.40}$$

Clearly  $D_1(1)$  of (S7.40) has a factor of  $(1-r)$ , and in fact

$$D_1(1) = (1-r)f(r), \tag{S7.41}$$

where  $f(r)$  is a linear function of  $r$ , for  $0 \leq r \leq 1$ . Now

$$f(1) = 2 - (1 - 2m_b) \frac{2(1+s)}{(1+s\bar{x})[1+s(1-x)]} - m_b \frac{(s+2)[(s+2)+s(\bar{x}-\tilde{x})]}{(1+s\bar{x})[1+s(1-\tilde{x})]}. \tag{S7.42}$$

Following (S6.7) we have

$$(1+s\bar{x})[1+s(1-\tilde{x})] = (1+s) + sm_B[(s+2)\bar{x} - 1]. \tag{S7.43}$$

We also have an equivalent expression for (S7.43) in terms of  $\tilde{x}$ , namely

$$(1+s\bar{x})[1+s(1-\tilde{x})] = (1+s) + sm_B[(s+1) - (s+2)\tilde{x}]. \tag{S7.44}$$

Also, whereas  $\bar{x} > \frac{1}{s+2}$ , we have  $\tilde{x} < \frac{s+1}{s+2}$ . Applying all of this to (S7.42) and using the fact that

$$(1+s\bar{x})[1+s(1-\tilde{x})] = (1+s) + \frac{1}{2}sm_B[s + (s+2)(\bar{x}-\tilde{x})], \tag{S7.45}$$

we get that

$$\begin{aligned}
(1+s\bar{x})[1+s(1-\tilde{x})]f(1) &= 2(s+1) + sm_B[s + (s+2)(\bar{x}-\tilde{x})] \\
&\quad - 2(1-2m_b)(s+1) - m_b(s+2)[(s+2)+s(\bar{x}-\tilde{x})] \\
&= s^2m_B - m_b \left[ (s+2)^2 - 4(s+1) \right] + s(s+2)(m_B - m_b)(\bar{x}-\tilde{x}) \\
&= s^2(m_B - m_b) + s(s+2)(m_B - m_b)(\bar{x}-\tilde{x}).
\end{aligned} \tag{S7.46}$$

Thus

$$(1+s\bar{x})[1+s(1-\tilde{x})]f(1) = s(m_B - m_b)[s + (s+2)(\bar{x}-\tilde{x})]. \tag{S7.47}$$



But as  $(s+2)\bar{x} > 1$ ,  $(s+2)\tilde{x} < (s+1)$ ,

$$s + (s+2)(\bar{x} - \tilde{x}) = [(s+2)\bar{x} - 1] + [(s+1) - (s+2)\tilde{x}] > 0. \quad (S7.48)$$

It follows that the sign of  $f(1)$  is the same as the sign of  $(m_B - m_b)$ .

We now compute  $f(0)$ :

$$\begin{aligned} f(0) = 1 - (1 - 2m_b) \frac{2(1+s)}{(1+s\bar{x})[1+s(1-\tilde{x})]} - m_b^2 \frac{(s+2)^2}{(1+s\bar{x})[1+s(1-\tilde{x})]} \\ + \frac{(1-2m_b)^2 (s+1)^2}{(1+s\bar{x})^2 [1+s(1-\tilde{x})]^2}. \end{aligned} \quad (S7.49)$$

Using the expression (S7.43) for the product of the two mean fitnesses, we get

$$\begin{aligned} (1+s\bar{x})^2 [1+s(1-\tilde{x})]^2 f(0) = \{(1+s) + sm_B [(s+2)\bar{x} - 1]\}^2 \\ - 2(1-2m_b)(s+1) \{(1+s) + sm_B [(s+2)\bar{x} - 1]\} \\ - m_b^2 (s+2) \{(1+s) + sm_B [(s+2)\bar{x} - 1]\} \\ + (1-2m_b)^2 (s+1)^2. \end{aligned} \quad (S7.50)$$

In (S7.50) we replace the  $\bar{x}^2$  term using the equilibrium equation (36) to give

$$\begin{aligned} (1+s\bar{x})^2 [1+s(1-\tilde{x})]^2 f(0) = (m_B - m_b)s \{ m_b s(s+1) \\ - m_B^2 (s+2)^2 + m_B [(s+1)(s+4) - m_b (s+2)^2] \\ + m_B (s+2)\bar{x} [m_B (s+2)^2 + m_b (s+2)^2 - 4(s+1)] \}. \end{aligned} \quad (S7.51)$$

The right-hand side of (S7.51) is  $(m_B - m_b)s$  multiplied by

$$m_b s(s+1) + m_B (s+2)^2 (m_B + m_b) [\bar{x}(s+2) - 1] + m_B (s+1) [(s+4) - 4\bar{x}(s+2)]. \quad (S7.52)$$

We will show that (S7.52) is always positive. In fact, (S7.52) is equal to

$$\begin{aligned} m_b s(s+1) + m_B \cdot m_b (s+2)^2 [\bar{x}(s+2) - 1] + m_B^2 (s+2)^2 [\bar{x}(s+2) - 1] \\ + m_B (s+1) [(s+4) - 4\bar{x}(s+2)]. \end{aligned} \quad (S7.53)$$

From the equilibrium equation (36) we get that

$$m_B [(s+2)\bar{x} - 1] = s\bar{x}^2 + 2\bar{x} - 1. \quad (S7.54)$$

Hence (S7.53) is equal to

$$\begin{aligned}
& m_b s(s+1) + m_B \cdot m_b (s+2)^2 [\bar{x}(s+2) - 1] + \\
& + m_B (s+2)^2 [s\bar{x}^2 + 2\bar{x} - 1] + m_B(s+1)[(s+4) - 4\bar{x}(s+2)].
\end{aligned} \tag{S7.55}$$

The last two terms have a factor  $m_B$  that multiplies

$$\begin{aligned}
& (s+1)(s+4) - (s+2)^2 + (s+2)^2 \bar{x}(2+s\bar{x}) - 4\bar{x}(s+1)(s+2) = \\
& = s + (s+2)\bar{x}[(s+2)(2+s\bar{x}) - 4(s+1)] \\
& = s + (s+2)\bar{x}[(s+2)s\bar{x} - 2s] \\
& = s \left[ (s+2)^2 \bar{x}^2 - 2(s+2)\bar{x} + 1 \right] = s[(s+2)\bar{x} - 1]^2,
\end{aligned} \tag{S7.56}$$

which is positive. To sum up,  $f(0)$  also has the same sign of  $(m_B - m_b)$ , and so

$$D_1(1) = (1-r)s(m_B - m_b)\Delta(r), \tag{S7.57}$$

where  $\Delta(r)$  is a linear function of  $r$  that is positive for all  $0 \leq r \leq 1$ . As  $(m_B - m_b) = (1-2m)(\mu_B - \mu_b)$ , this proves the following result.

**File S8**

**Proof of Result 8**

*i.* In a constant environment, the mean fitness  $w^*$  at the symmetric equilibrium  $(\bar{x}^*, \bar{y}^*)$  is  $w^* = 1 + sx^*$ , and it is a decreasing function of  $\mu_B$  if  $\frac{\partial x^*}{\partial \mu_B}$  is negative, or equivalently if  $\frac{\partial x^*}{\partial m_B}$  is negative (since  $m_B = m + \mu_B(1 - 2m)$  and  $0 \leq m < \frac{1}{2}$ ). Using the equilibrium equation (14),

$$\frac{\partial x^*}{\partial m_B} = \frac{1 - (s + 2)x^*}{2sx^* + [(s + 2)m_B - s]}. \quad (\text{S8.1})$$

As  $x^* > \frac{1}{s+2}$ , in order for  $\frac{\partial x^*}{\partial m_B}$  to be negative, it is sufficient that

$$x^* > \frac{s - m_B(s + 2)}{2s}. \quad (\text{S8.2})$$

This follows easily from the fact that  $Q(x)$  of (14) satisfies  $Q(0) < 0$ ,  $Q(x^*) = 0$ , and  $Q\left(\frac{s - m_B(s + 2)}{2s}\right) < 0$ .

*ii.* With a fitness cycle of period 2, the mean fitness  $\bar{w}$  at the symmetric equilibrium  $(\bar{x}, \bar{y})$  is

$$\bar{w} = (1 + s) + sm_B[(s + 2)\bar{x} - 1]. \quad (\text{S8.3})$$

$\bar{w}$  is an increasing function of  $\mu_B$  if  $\frac{\partial \bar{w}}{\partial \mu_B} > 0$  or equivalently if  $\frac{\partial \bar{w}}{\partial m_B} > 0$ . Now

$$\frac{\partial \bar{w}}{\partial m_B} = s[(s + 2)\bar{x} - 1] + s(s + 2)m_B \frac{\partial \bar{x}}{\partial m_B}. \quad (\text{S8.4})$$

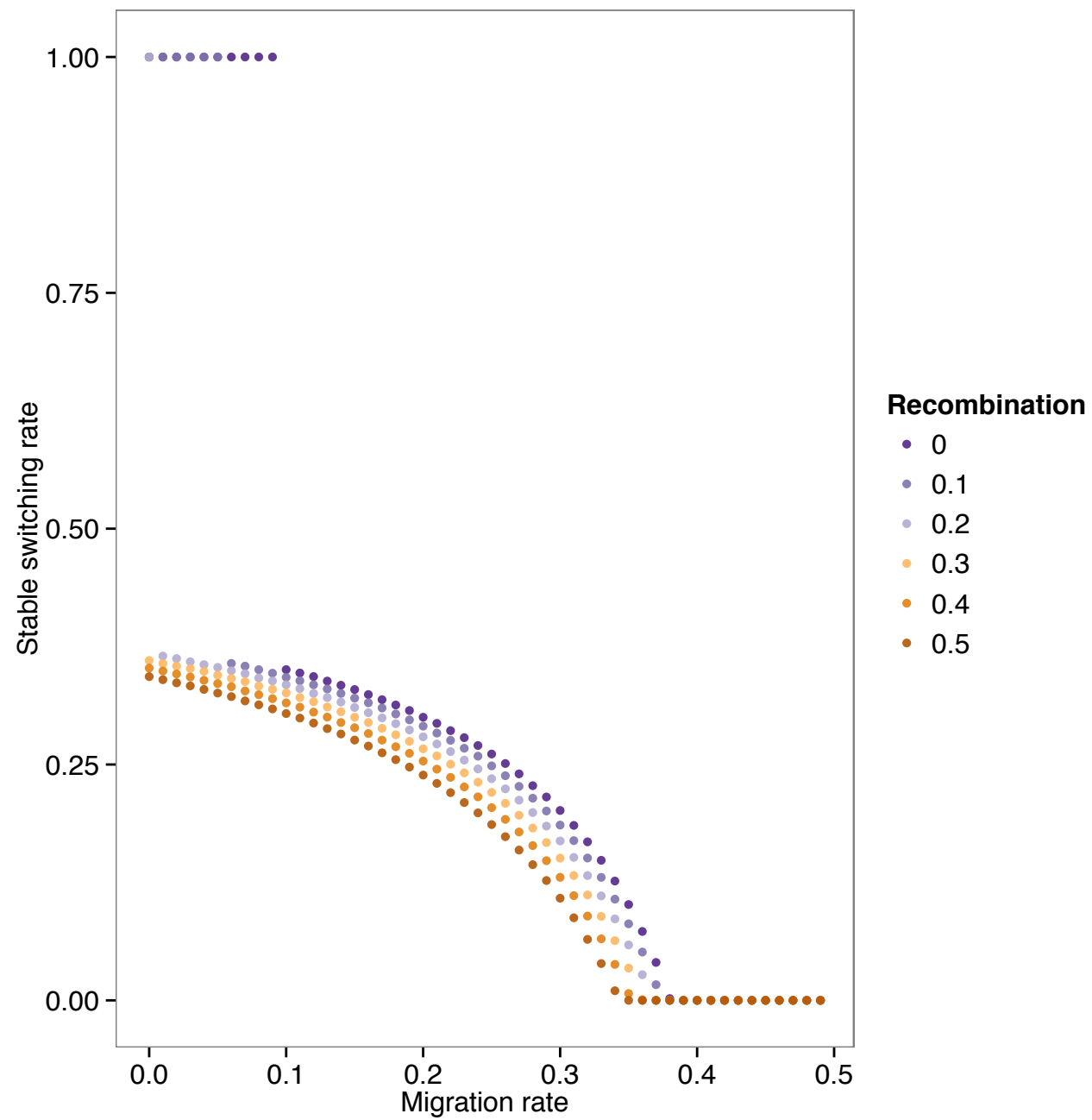
Thus  $\frac{\partial \bar{w}}{\partial m_B} > 0$  provided  $\frac{\partial \bar{x}}{\partial m_B} > 0$ . Using the equilibrium equation  $R(x) = 0$  for  $\bar{x}$ , we have

$$\frac{\partial \bar{x}}{\partial m_B} = \frac{\bar{x}(s + 1) - 1}{2s\bar{x} + [2 - m_B(s + 2)]}. \quad (\text{S8.5})$$

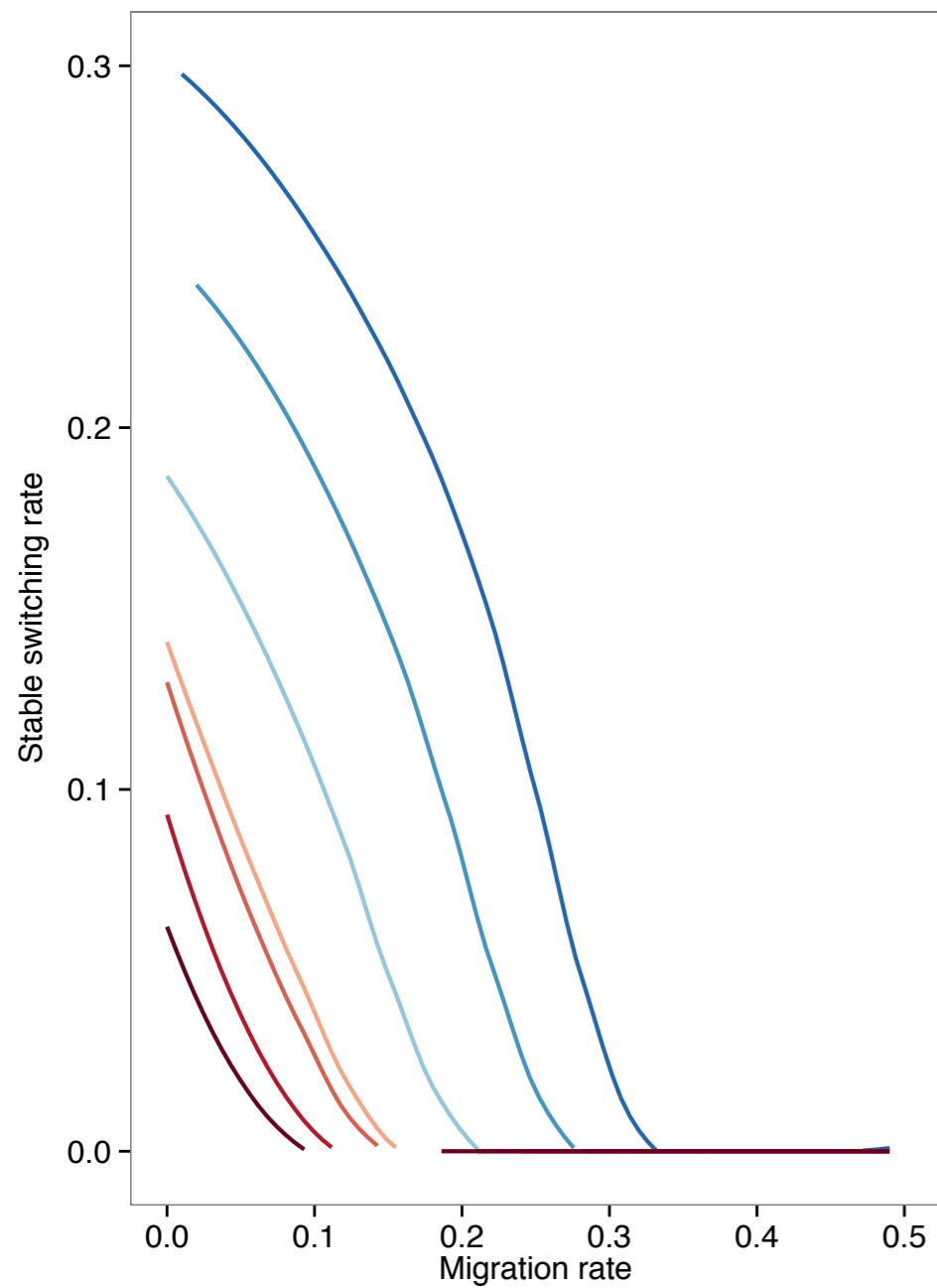
Since  $[\bar{x}(s + 1) - 1] > 0$ , we conclude that  $\frac{\partial \bar{x}}{\partial m_B} > 0$  if

$$\bar{x} > \frac{m_B(s + 2) - 2}{2s}, \quad (\text{S8.6})$$

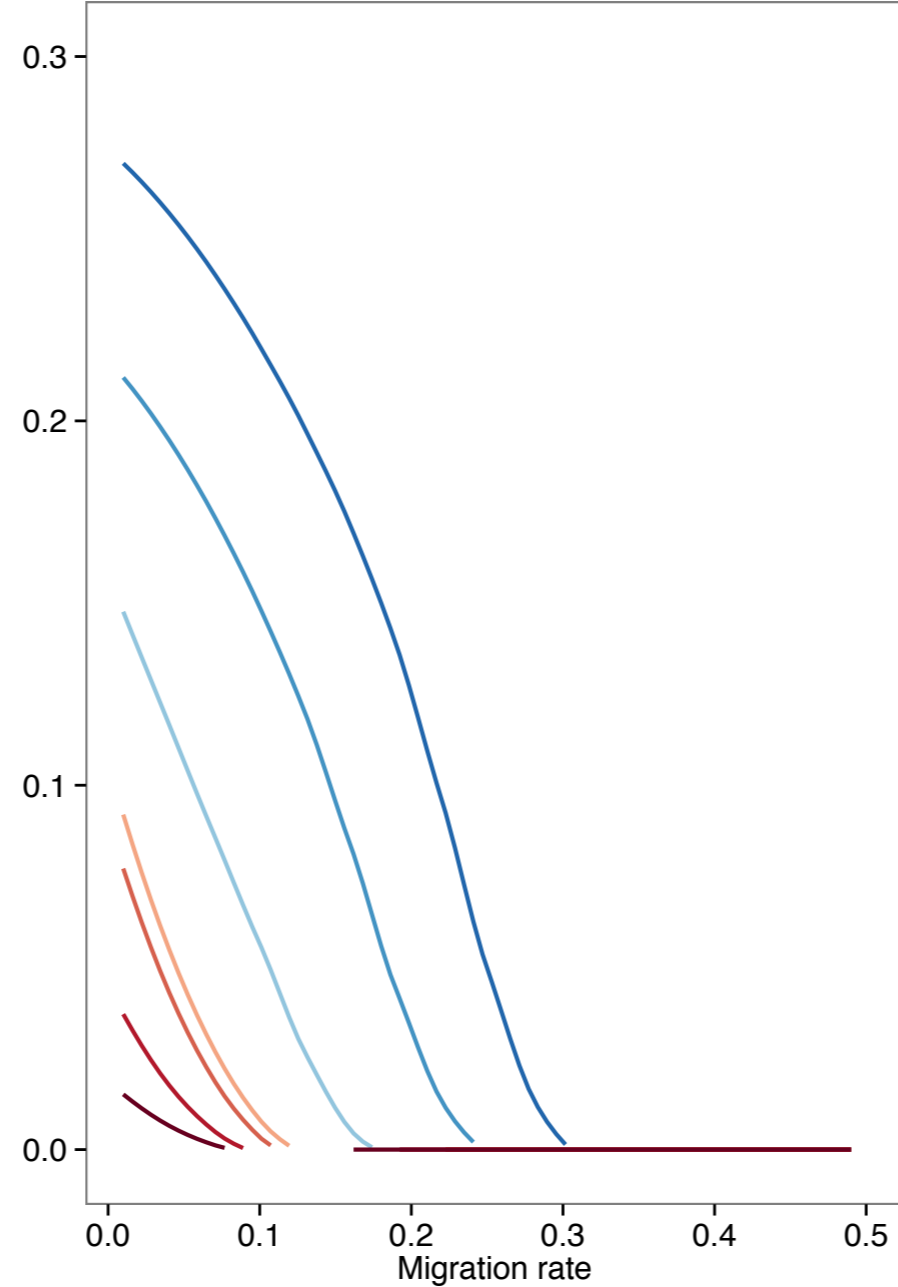
which follows from  $R(0) < 0$ ,  $R(\bar{x}) = 0$ , and  $R\left(\frac{m_B(s + 2) - 2}{2s}\right) < 0$ .



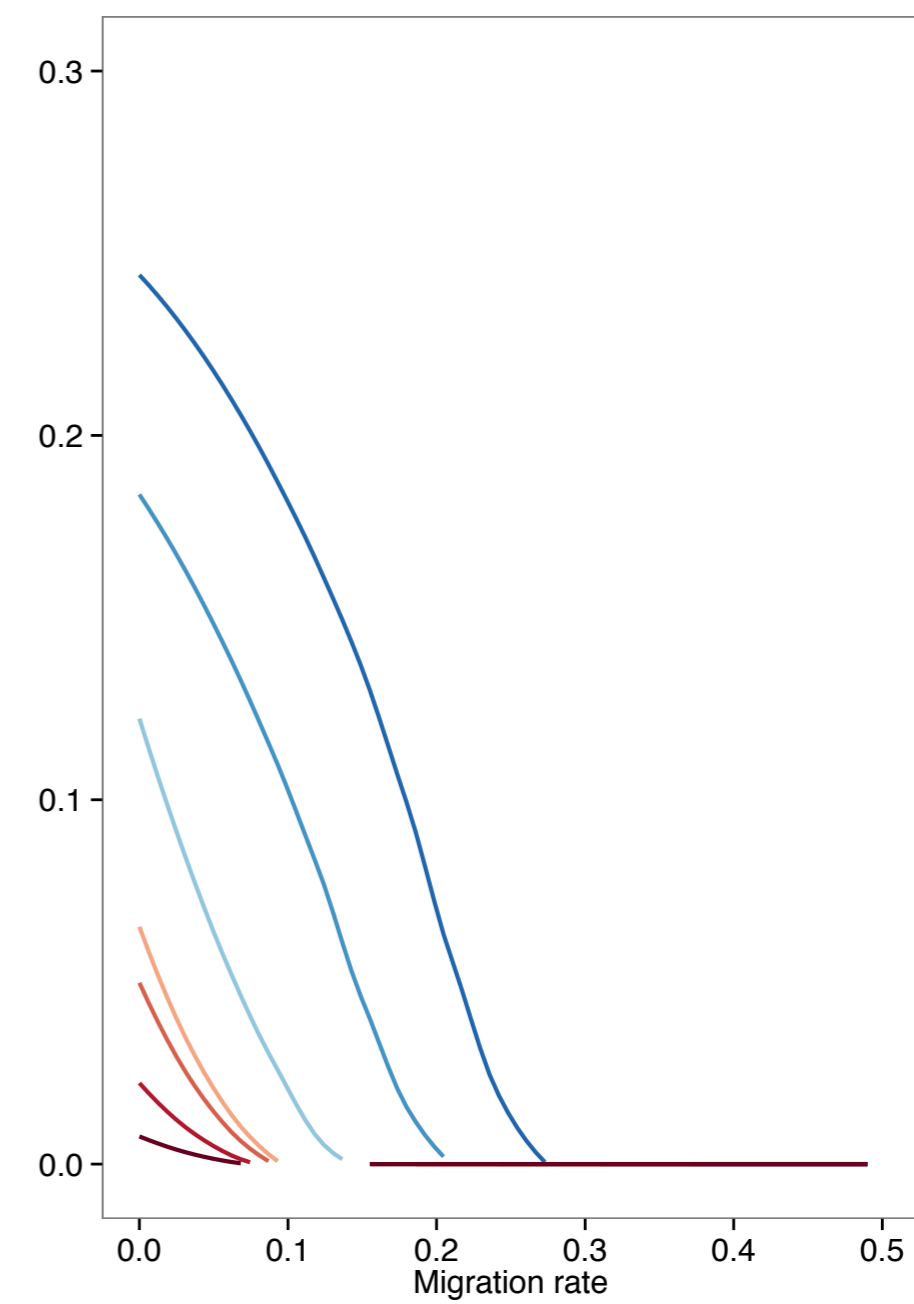
**Figure S1. The evolutionarily stable switching rate as function of migration and recombination rate,  $n = 3$ .** The symmetric selection coefficient  $s = 0.4$ . Recombination rates shown in the legend. The stable switching rate for  $n = 3$  is sensitive to the interplay of recombination and migration rates, with sudden possible discontinuities in the stable switching rate.



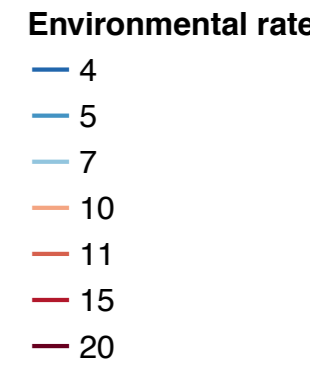
(A)



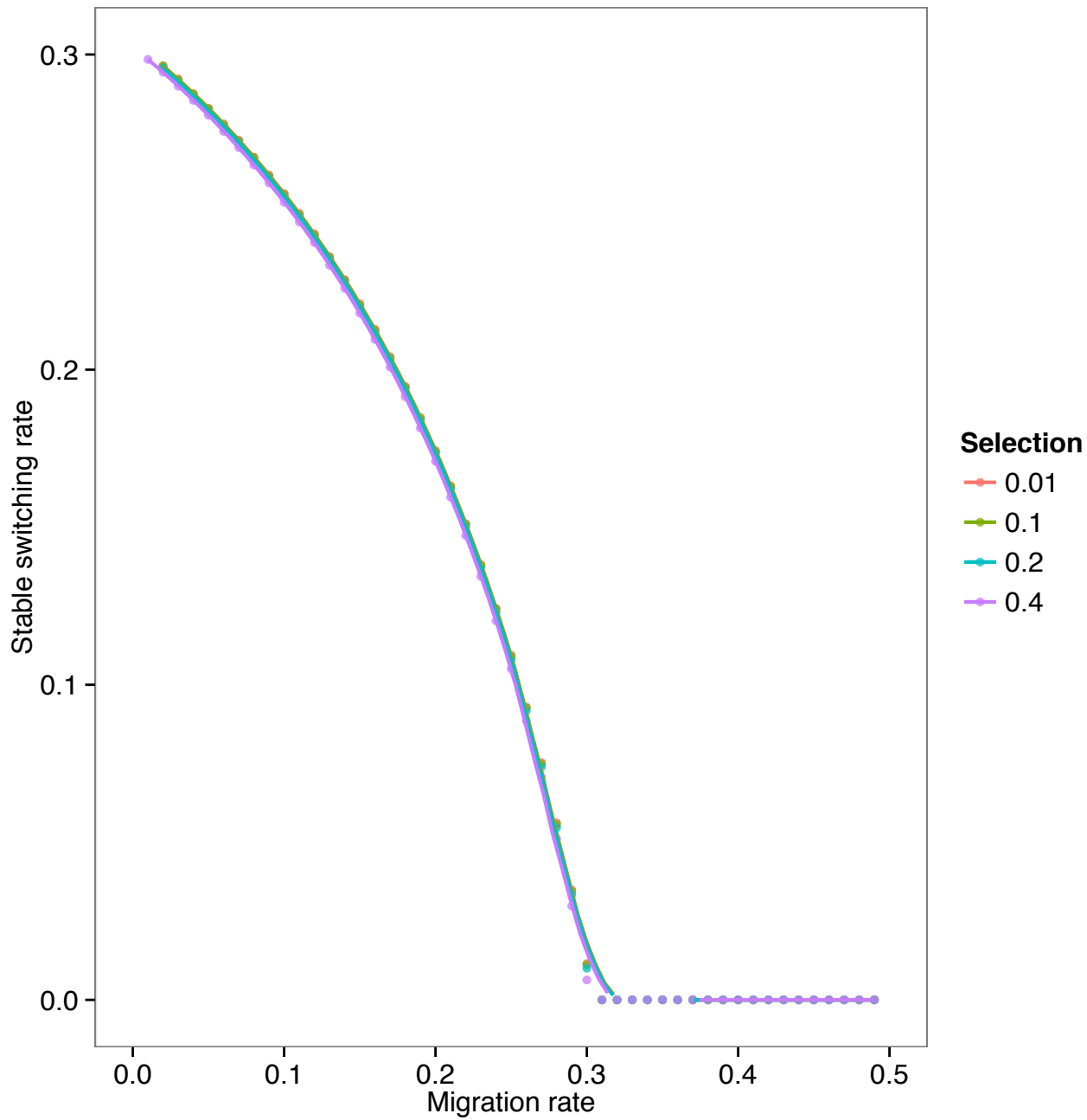
(B)



(C)



**Figure S2. The evolutionarily stable switching rate as function of migration and environmental rate of change,  $n > 3$  for different recombination rates.** The symmetric selection coefficient  $s = 0.4$ . The rate of environmental change  $n$  shown in the legend. The plotted curves represent a fit to the data using a generalized additive model with penalized cubic regression splines. In panel **A**,  $r = 0$ . In panel **B**,  $r = 0.25$ . In panel **C**,  $r = 0.5$ .





**Figure S3. The evolutionarily stable switching rate as function of migration and symmetric selection coefficient  $s$ .** Recombination rate is  $r = 0$ . The environment changes every  $n = 4$  generations. The symmetric selection coefficient  $s$  shown in the legend. The plotted curves represent a fit to the data using a generalized additive model with penalized cubic regression splines. The stable switching rate is invariant to the strength of symmetric selection between the two demes.