

# The consequences of dominance and gene flow for local adaptation and differentiation at two linked loci



Ada Akerman\*, Reinhard Bürger

Department of Mathematics, University of Vienna, Austria

## ARTICLE INFO

### Article history:

Received 9 September 2013  
Available online 30 April 2014

### Keywords:

Effective migration rate  
Selection  
Migration  
Recombination  
Dispersal  
Linkage disequilibrium  
Fixation index

## ABSTRACT

For a subdivided population the consequences of dominance and gene flow for the maintenance of multilocus polymorphism, local adaptation, and differentiation are investigated. The dispersing population inhabits two demes in which selection acts in opposite direction. Fitness is determined additively by two linked diallelic loci with arbitrary intermediate dominance (no over- or underdominance). For weak as well as strong migration, the equilibrium structure is derived. As a special case, a continuous-time continent–island model (CI model) is analyzed, with one-way migration from the continent to the island. For this CI model, the equilibrium and stability configuration is obtained explicitly for weak migration, for strong migration, for independent loci, and for complete linkage. For independent loci, the possible bifurcation patterns are derived as functions of the migration rate. These patterns depend strongly on the degree of dominance. The effects of dominance, linkage, and migration on the amount of linkage disequilibrium (LD) and the degree of local adaptation are explored. Explicit formulas are obtained for  $D (=x_1x_4 - x_2x_3)$  and  $r^2$  (the squared correlation in allelic state). They demonstrate that dominant island alleles increase  $D$  and decrease  $r^2$ . Local adaptation is elevated by dominance of the locally adaptive alleles. The effective migration rate at a linked neutral locus is calculated. If advantageous alleles are dominant, it is decreased only slightly below the actual migration rate. For a quantitative trait that is determined by two additive loci, the influence of dominance on measures of differentiation is studied. Explicit expressions for  $Q_{ST}$  and two types of  $F_{ST}$  at equilibrium are deduced and their relation is discussed.

© 2014 The Authors. Published by Elsevier Inc.  
This is an open access article under the CC BY-NC-ND license  
(<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

## 1. Introduction

Despite extensive treatment in the literature, the interplay of selection, recombination, and migration is far from being well understood. When it comes to local adaptation and its conceptual issues, the examination of these interactions is inevitable (Lenormand, 2002; Kawecki and Ebert, 2004). Also, the exploding amount of data obtained by genome-wide sequencing of individuals from different geographical regions cannot be assessed properly without comprehending the effects of selection and migration on multilocus variation and the effects of genetic architecture (Charlesworth et al., 1997; Nordborg and Tavaré, 2002; Slatkin, 2008).

For one-locus models, both general theory and the study of numerous particular models have provided considerable insight into

the evolutionary consequences of the interaction between migration and selection (reviewed by Karlin, 1982; Lenormand, 2002; Nagylaki and Lou, 2008; Bürger, 2014). Several one-locus models incorporating selection and migration investigate the effect of dominance on the maintenance of polymorphisms (Prout, 1968; Karlin and Campbell, 1980; Nagylaki and Lou, 2001, 2007; Nagylaki, 2009a; Peischl, 2010). In most of these models dominance is intermediate, i.e., overdominance and underdominance are excluded. For studying maintenance of polymorphism in structured populations, intermediate dominance is of particular interest not only because it may be the most common form of dominance but also because, in contrast to panmictic populations, stable polymorphisms can occur in subdivided populations in the absence of overdominance.

Within the past years, progress has been made in the study of the joint evolutionary effects of migration, selection, and recombination. Due to the increased complexity by considering several evolutionary forces simultaneously, mathematical analyses of multilocus migration–selection models focused on limiting or special cases. These include weak or strong migration (Barton, 1983;

\* Correspondence to: Fakultät für Mathematik, Universität Wien, Oskar-Morgenstern-Platz 1, A-1090 Wien, Austria.  
E-mail address: [ada.akerman@univie.ac.at](mailto:ada.akerman@univie.ac.at) (A. Akerman).

Bürger, 2009a,b) or the Levene model (Nagylaki, 2009b; Bürger, 2009c, 2010; Barton, 2010). In these cases, linkage disequilibrium (LD) is weak or absent. For a review, see Bürger (2014).

Recently, simple but informative migration models were studied in which high levels of LD may be maintained. These include the continent–island model (Bürger and Akerman, 2011; henceforth cited as BA11) and the two-deme model with forward and backward migration (Akerman and Bürger, 2014) where genic selection is assumed (i.e., absence of dominance and epistasis). Ignoring dominance, Bank et al. (2012) investigated the continent–island model with epistasis. In these works, the effects of gene flow on local adaptation, on differentiation, and on the maintenance of polymorphism at two linked diallelic loci in the absence of dominance were investigated and analytical characterizations of the possible equilibrium configurations and bifurcation patterns were obtained. In particular, explicit formulas were derived for the maximum migration rate below which a fully polymorphic equilibrium can be maintained.

Only few multilocus migration–selection models investigate the evolutionary consequences of dominance, as it leads to substantial mathematical complications (Bürger, 2009a,b,c, 2010; Nagylaki, 2009b; Chasnov, 2012). With this work, we want to shape our intuition about the interplay of migration and linkage on multilocus variation exhibiting dominance. The main goal of this paper is to explore and quantify the role of intermediate dominance on linked loci for local adaptation in the presence of maladaptive gene flow. For simplicity, we study a dispersing population inhabiting two demes. We consider two diallelic loci A and B, where alleles  $A_1$  and  $B_1$  are favored in deme 1 and alleles  $A_2$  and  $B_2$  are favored in deme 2. Recombination between the loci is arbitrary. We assume that (additive) epistasis absent and neglect mutation and random drift.

The model is introduced in Section 2. In Section 3, we derive the equilibrium and stability structure for two important special cases: weak migration and strong migration. In Section 4, we study the limiting case where selection, recombination, and migration are weak, and where migration is in one direction only. Under these assumptions, the continuous-time continent–island model (CI model) is obtained from the discrete-time two-deme model. For the CI model, we investigate several special cases: weak migration, independent loci (linkage equilibrium), and complete linkage. For these cases, we derive the possible equilibrium structures and the bifurcation patterns. We show that the equilibrium and stability structure maintained in the population depends crucially on the degree of dominance, in particular if linkage is loose. Several technical results and derivations are relegated to the Appendix.

Section 5 is devoted to applications of the results obtained in Sections 3 and 4. In Section 5.1 we investigate how two widely used measures of LD,  $D$  and  $r^2$ , depend on the amount of dominance and on linkage. We show that  $D$  is elevated and  $r^2$  is reduced by increasing dominance of the locally adaptive alleles. In Section 5.2, we use the migration load to quantify the dependence of local adaptation on the recombination rate and the degree of dominance. Our analysis shows that elevated levels of local adaptation are to be expected if the locally adaptive alleles are dominant.

In Section 5.3, we study the strength of barriers to gene flow at neutral sites linked to the selected loci by deriving an approximation for the effective migration rate at a linked neutral site. We show that gene flow at the neutral locus experiences the strongest barrier if the locally adaptive alleles are recessive. If the locally adaptive alleles are completely dominant, linkage has (almost) no effect on the effective migration rate and gene flow at the neutral locus is not significantly reduced.

In Section 5.4, we investigate a quantitative trait subject to diversifying selection. Our main focus is on determining the amount of differentiation under maladaptive gene flow. For this

end, we analyze three different measures of differentiation known from the literature: (i)  $\bar{F}_{ST}$ , which is  $F_{ST}$  averaged over the two loci, (ii)  $F_{ST}^X$ , a multilocus  $F_{ST}$  introduced in Akerman and Bürger (2014), and (iii)  $Q_{ST}$ , which measures differentiation of the trait. We derive informative approximations for these measures and study their relations. Finally, we determine the effects of dominance and linkage on the extent of differentiation achieved at the genotypic and the trait level.

## 2. The model

In the following we introduce the discrete-time two-deme migration–selection model for two diallelic loci. Throughout the paper we ignore mutation and stochastic effects. The population is assumed to be diploid, monoecious, and infinitely large. Generations are discrete and nonoverlapping. The population is subdivided into two discrete demes (denoted by  $k = 1, 2$ ) in which random mating occurs after selection and migration. Therefore, offspring are in Hardy–Weinberg proportions.

Alleles are denoted by  $A_1, A_2$  at locus A and by  $B_1, B_2$  at locus B. The frequencies of the four gametes  $A_1B_1, A_1B_2, A_2B_1,$  and  $A_2B_2$  in deme  $k$  are designated by  $x_{1,k}, x_{2,k}, x_{3,k},$  and  $x_{4,k}$ , respectively. For every deme the state space is given by the probability simplex  $S_4$ , where  $S_4 = \{(x_1, x_2, x_3, x_4) : x_i \geq 0 \text{ and } \sum_{i=1}^4 x_i = 1\}$ . The state space of the full model is  $S_4 \times S_4$ .

We neglect (additive) epistasis. Therefore, the fitness of a two-locus genotype is the sum of the fitness contributions of the constituent one-locus genotypes. Assuming absence of position effects, the fitness matrix in deme  $k$  is given by

$$\begin{matrix} & & B_1B_1 & B_1B_2 & B_2B_2 \\ \begin{matrix} A_1A_1 \\ A_1A_2 \\ A_2A_2 \end{matrix} & \left( \begin{array}{ccc} 1 + a_k + b_k & 1 + a_k + \sigma_k b_k & 1 + a_k - b_k \\ 1 + \vartheta_k a_k + b_k & 1 + \vartheta_k a_k + \sigma_k b_k & 1 + \vartheta_k a_k - b_k \\ 1 - a_k + b_k & 1 - a_k + \sigma_k b_k & 1 - a_k - b_k \end{array} \right), \end{matrix} \quad (2.1)$$

where  $a_k \neq 0$  and  $b_k \neq 0$  are the selection coefficients at locus A and locus B in deme  $k$ , respectively. Dominance coefficients are denoted by  $\vartheta_k$  and  $\sigma_k$  and fulfill

$$|\vartheta_k| \leq 1 \quad \text{and} \quad |\sigma_k| \leq 1 \quad (k = 1, 2). \quad (2.2)$$

We call this intermediate dominance. There is no dominance if

$$\vartheta_k = \sigma_k = 0 \quad \text{for } k = 1, 2, \quad (2.3)$$

and there is deme-independent degree of dominance (DIDID) if

$$\vartheta_1 = \vartheta_2 = \vartheta \quad \text{and} \quad \sigma_1 = \sigma_2 = \sigma; \quad (2.4)$$

see Nagylaki (2009a). We assume

$$|a_k| + |b_k| < 1, \quad (2.5)$$

whence fitnesses are positive, and, without loss of generality,

$$a_1 > 0 > a_2 \quad \text{and} \quad b_1 > 0 > b_2. \quad (2.6)$$

Let  $r \in [0, \frac{1}{2}]$  denote the recombination probability between the two loci, where  $r = 0$  corresponds to complete linkage and  $r = \frac{1}{2}$  to free recombination. The probability that an individual in deme 1 (deme 2) immigrated from deme 2 (deme 1) is denoted by  $m_1$  ( $m_2$ ). We assume  $0 \leq m_k < \frac{1}{2}$  for  $k = 1, 2$ .

We denote the marginal fitness of gamete  $i = 1, 2, 3, 4$  and the mean fitness in deme  $k$  by  $W_{i,k}$  and  $\bar{W}_k$ , respectively. Let

$$W_{14,k} = 1 + \vartheta_k \alpha_k + \sigma_k \beta_k \quad (2.7)$$

be the fitness of the double heterozygote in deme  $k$  and  $\eta_1 = \eta_4 = -\eta_2 = -\eta_3 = 1$ . Then the frequency of gamete  $i = 1, 2, 3, 4$  in deme  $k$  after selection and recombination is given by

$$x_{i,k}^* = \frac{x_{i,k}W_{i,k} - \eta_i r D_k W_{14,k}}{\bar{W}_k}, \quad (2.8a)$$

where  $D_k = x_{1,k}x_{4,k} - x_{2,k}x_{3,k}$  is the usual measure of LD in deme  $k$ . If  $D_k = 0$ , there is linkage equilibrium (LE) in deme  $k$ . After migration, hence in the next generation, the gamete frequencies are

$$\begin{aligned} x'_{i,1} &= (1 - m_1)x_{i,1}^* + m_1x_{i,2}^* \quad \text{and} \\ x'_{i,2} &= m_2x_{i,1}^* + (1 - m_2)x_{i,2}^*. \end{aligned} \quad (2.8b)$$

Instead of gamete frequencies it is often more convenient to use the allele frequencies of  $A_1$  and  $B_1$ ,

$$p_k = x_{1,k} + x_{2,k} \quad \text{and} \quad q_k = x_{1,k} + x_{3,k}, \quad (2.9)$$

respectively, and the measure  $D_k$  of LD in deme  $k = 1, 2$ . Then the gamete frequencies  $x_{i,k}$  are obtained from  $p_k, q_k$ , and  $D_k$  by

$$x_{1,k} = p_k q_k + D_k, \quad x_{2,k} = p_k(1 - q_k) - D_k, \quad (2.10a)$$

$$x_{3,k} = (1 - p_k)q_k - D_k, \quad x_{4,k} = (1 - p_k)(1 - q_k) + D_k. \quad (2.10b)$$

The constraints  $x_{i,k} \geq 0$  and  $\sum_{i=1}^4 x_{i,k} = 1$  for  $i = 1, 2, 3, 4$ , and  $k = 1, 2$  transform into

$$0 \leq p_k \leq 1 \quad \text{and} \quad 0 \leq q_k \leq 1 \quad (2.11a)$$

and

$$\begin{aligned} -\min\{p_k q_k, (1 - p_k)(1 - q_k)\} &\leq D_k \\ &\leq \min\{p_k(1 - q_k), (1 - p_k)q_k\}. \end{aligned} \quad (2.11b)$$

The mean fitness depends only on allele frequencies and is given by

$$\begin{aligned} \bar{W}_k &= 1 - \alpha_k[1 - 2p_k - 2\vartheta_k p_k(1 - p_k)] \\ &\quad - \beta_k[1 - 2q_k - 2\sigma_k q_k(1 - q_k)]. \end{aligned} \quad (2.12)$$

Simple calculations show that the recursion relations for the allele frequencies and LD in demes 1 and 2 are

$$p'_1 = (1 - m_1)p_1^* + m_1p_2^*, \quad p'_2 = m_2p_1^* + (1 - m_2)p_2^*, \quad (2.13a)$$

$$q'_1 = (1 - m_1)q_1^* + m_1q_2^*, \quad q'_2 = m_2q_1^* + (1 - m_2)q_2^*, \quad (2.13b)$$

$$\begin{aligned} D'_1 &= (1 - m_1)D_1^* + m_1D_2^* \\ &\quad + m_1(1 - m_1)(p_1^* - p_2^*)(q_1^* - q_2^*), \end{aligned} \quad (2.13c)$$

$$\begin{aligned} D'_2 &= m_2D_1^* + (1 - m_2)D_2^* \\ &\quad + m_2(1 - m_2)(p_1^* - p_2^*)(q_1^* - q_2^*), \end{aligned} \quad (2.13d)$$

where an asterisk, \*, indicates values after selection and recombination; cf. Li and Nei (1974).

### 3. Equilibrium structure

Equilibria of (2.13) satisfy  $p'_k = p_k, q'_k = q_k$ , and  $D'_k = D_k$  for  $k = 1, 2$ . Three types of equilibria are distinguished: (i) monomorphisms, (ii) single-locus polymorphisms, and (iii) two-locus (full) polymorphisms. Monomorphisms and single-locus polymorphisms are located on the boundary of the state space. By full polymorphisms we mean states at which every allele is present with positive frequency.

The four monomorphisms are

$$\begin{aligned} M_1(A_1B_1 \text{ fixed}): \\ \hat{p}_k = 1, \quad \hat{q}_k = 1, \quad \hat{D}_k = 0 \quad \text{for } k = 1, 2, \end{aligned} \quad (3.1a)$$

$M_2(A_1B_2 \text{ fixed}):$

$$\hat{p}_k = 1, \quad \hat{q}_k = 0, \quad \hat{D}_k = 0 \quad \text{for } k = 1, 2, \quad (3.1b)$$

$M_3(A_2B_1 \text{ fixed}):$

$$\hat{p}_k = 0, \quad \hat{q}_k = 1, \quad \hat{D}_k = 0 \quad \text{for } k = 1, 2, \quad (3.1c)$$

$M_4(A_2B_2 \text{ fixed}):$

$$\hat{p}_k = 0, \quad \hat{q}_k = 0, \quad \hat{D}_k = 0 \quad \text{for } k = 1, 2, \quad (3.1d)$$

where a  $\hat{\cdot}$  signifies equilibrium.

Within each marginal one-locus systems, the stability of the monomorphisms is easily determined from the well known conditions for protection of an allele (Maynard Smith, 1970; Bulmer, 1972; Nagylaki, 1992). In particular, there is a protected polymorphism (in a marginal one-locus system) if the two neighboring monomorphisms are unstable. However, stability of an equilibrium within a marginal one-locus system does not imply stability with respect to the full two-locus system. The local stability of the monomorphisms with respect to the interior of the state space can be derived, but the eigenvalues are complicated and, in general, uninformative. For the single-locus polymorphisms, these eigenvalues are zeros of complicated quartic polynomials.

It is also well known that the existence of a protected polymorphism (in a one-locus system) does not imply that there is a unique stable polymorphic equilibrium. In fact, there can be at least three internal equilibria, whence one has to be unstable (e.g. Karlin and Campbell, 1980; Novak, 2011). If, however, in each deme the geometric mean fitness of the homozygotes is greater than or equal to the fitness of the heterozygote, the existence of a protected polymorphism implies a unique, globally stable polymorphic equilibrium (Karlin and Campbell, 1980; Bürger, 2014). This applies in particular if dominance is absent (2.3).

Below, we study several informative special cases for which we derive the possible equilibrium and stability structure explicitly. We investigate the limiting case of weak migration in Section 3.1 and that of strong migration in Section 3.2.

The maximum number of admissible full polymorphisms derived by the subsequent analysis and encountered in numerical work supplementing our analysis is nine.

#### 3.1. Weak migration

In the absence of migration,  $m_1 = m_2 = 0$ , the dynamics in the two demes are decoupled. If there is intermediate dominance (2.2), the fittest haplotype becomes fixed in the respective deme. Considering the two demes collectively, the globally asymptotically stable equilibrium in the absence of migration is given by

$$F_0: \quad \hat{p}_1 = \hat{q}_1 = 1, \quad \hat{p}_2 = \hat{q}_2 = 0, \quad \hat{D}_1 = \hat{D}_2 = 0. \quad (3.2)$$

$F_0$  is globally asymptotically stable as mean fitness increases in every deme if fitnesses are additive (Ewens, 1969).

The eigenvalues at  $F_0$  are easily calculated (Appendix A.1.1). They are positive. In particular,  $F_0$  is hyperbolic (i.e., no eigenvalue has modulus 1) if and only if the dominance coefficients satisfy

$$\vartheta_1 \neq 1 \quad \text{and} \quad \sigma_1 \neq 1 \quad \text{and} \quad \vartheta_2 \neq -1 \quad \text{and} \quad \sigma_2 \neq -1. \quad (3.3)$$

This condition excludes (complete) dominance of an advantageous allele.

Assuming weak migration, i.e., the migration rates  $m_1$  and  $m_2$  are sufficiently small relative to the recombination rate and the fitness differences between genotypes, straightforward calculation using series expansion gives the perturbed equilibrium, denoted by  $F_m$ , up to first order in  $m_1$  and  $m_2$ :

$$\begin{aligned} \hat{p}_1 &= 1 - m_1 \frac{1 + a_1 + b_1}{a_1(1 - \vartheta_1)} \\ &\quad \times \frac{a_1(1 - \vartheta_1) + rW_{14,1}}{a_1(1 - \vartheta_1) + b_1(1 - \sigma_1) + rW_{14,1}}, \end{aligned} \quad (3.4a)$$

$$\hat{q}_1 = 1 - m_1 \frac{1 + a_1 + b_1}{b_1(1 - \sigma_1)} \times \frac{b_1(1 - \sigma_1) + rW_{14,1}}{a_1(1 - \vartheta_1) + b_1(1 - \sigma_1) + rW_{14,1}}, \quad (3.4b)$$

$$\hat{D}_1 = m_1 \frac{1 + a_1 + b_1}{a_1(1 - \vartheta_1) + b_1(1 - \sigma_1) + rW_{14,1}}, \quad (3.4c)$$

$$\hat{p}_2 = m_2 \frac{1 - a_2 - b_2}{-a_2(1 + \vartheta_2)} \times \frac{-a_2(1 + \vartheta_2) + rW_{14,2}}{-a_2(1 + \vartheta_2) - b_2(1 + \sigma_2) + rW_{14,2}}, \quad (3.4d)$$

$$\hat{q}_2 = m_2 \frac{1 - a_2 - b_2}{-b_2(1 + \sigma_2)} \times \frac{-b_2(1 + \sigma_2) + rW_{14,2}}{-a_2(1 + \vartheta_2) - b_2(1 + \sigma_2) + rW_{14,2}}, \quad (3.4e)$$

$$\hat{D}_2 = m_2 \frac{1 - a_2 - b_2}{-a_2(1 + \vartheta_2) - b_2(1 + \sigma_2) + rW_{14,2}}. \quad (3.4f)$$

It should be noted that, to this order of approximation, allele frequencies and LD in deme  $k$  depend only on the selection and dominance parameters in deme  $k$ .

From the assumptions (2.5) and (2.6) on the selection coefficients, and assumption (2.2) on the dominance coefficients, it follows easily that  $F_m$  exhibits positive LD and that tighter linkage increases LD and elevates the equilibrium frequency of the locally adaptive alleles within the respective deme; see (A.3). It is well known that in the absence of migration and of additive epistasis, no LD can be maintained. Also migration cannot generate or maintain LD in the presence of recombination (Bürger, 2009a). Therefore, LD is generated by the interaction of (non)epistatic selection and migration.

Increasing dominance of a locally adaptive allele decreases its frequency in the deme where it is advantageous (A.4). The reason is that the deleterious effect of the other allele is masked. Finally, increasing dominance of the locally adaptive alleles elevates the amount of LD in their respective deme and decreases the frequency of the haplotype with highest local fitness; see (A.5).

Because in an isolated population subject to the selection scheme (2.1) there are four equilibria (the monomorphic states), in the full two-deme model with migration absent there are 16 equilibria. These are the four monomorphisms  $M_i$  (3.1), eight single-locus polymorphisms,  $F_0$  and three other (unstable) full polymorphisms; see Appendix A.1.2 for details. The monomorphisms and the single-locus polymorphisms are unstable with respect to the full state space  $S_4 \times S_4$ .

If the dominance coefficients satisfy

$$-1 < \vartheta_k < 1 \quad \text{and} \quad -1 < \sigma_k < 1 \quad \text{for } k = 1 \text{ and } k = 2, \quad (3.5)$$

all equilibria are hyperbolic in the absence of migration (results not shown). Hence, the results of Karlin and McGregor (1972) imply that, for sufficiently weak migration (where ‘sufficiently’ may depend also on the dominance coefficients), there is at most one admissible equilibrium in the neighborhood of each equilibrium of the unperturbed system. Because unstable equilibria may leave the state space if migration is turned on, the number of admissible equilibria with migration is (much) smaller than 16.

If the dominance coefficients satisfy (3.5), four of the eight single-locus polymorphisms and the three unstable full polymorphisms leave the state space when migration is turned on (Appendix A.1). In particular,  $F_m$  is the unique internal equilibrium for sufficiently small  $m_1 > 0$  and  $m_2 > 0$ , and it is globally asymptotically stable by Theorem 5.4 in Bürger (2009a). This theorem

shows that for an arbitrary number of loci and strictly intermediate dominance at every locus, there is a unique, globally asymptotically stable equilibrium if migration is sufficiently weak. However, the calculation of the (approximate) coordinates would be cumbersome because higher-order linkage disequilibria come into play. For symmetric migration and equivalent loci without dominance, see Barton (1983).

If the dominance coefficients satisfy

$$\vartheta_1 = -1 \quad \text{or} \quad \sigma_1 = -1 \quad \text{or} \quad \vartheta_2 = 1 \quad \text{or} \quad \sigma_2 = 1, \quad (3.6)$$

the weak-migration approximation (3.4) of  $F_m$  applies, but  $F_m$  is not necessarily unique. Indeed, if a locally advantageous allele is completely recessive (i.e., (3.6) holds), additional internal equilibria emerge under weak migration. This is already the case in the one-locus CI model (e.g. Nagylaki, 1992, Chap. 6.1).

**Example 3.1.** 1. For a single locus with migration between two demes and one allele (completely) dominant in both demes, two internal equilibria may coexist, one of them simultaneously stable with a monomorphism (Nagylaki, 2009a). If the advantageous alleles are recessive in ‘their’ deme, three internal equilibria may coexist for weak migration. Assuming weak migration and fitnesses at locus A as in (2.1) with  $\vartheta_1 = -1$  and  $\vartheta_2 = 1$ , and using coordinates  $(p_1, p_2)$ , these three internal equilibria are perturbations of the equilibria  $(0, 0)$ ,  $(1, 1)$ , and  $(1, 0)$ . The latter is locally stable, the others are unstable. The monomorphic equilibria through which the unstable equilibria bifurcate,  $(0, 0)$  and  $(1, 1)$ , are locally stable. If the locally advantageous alleles are nearly recessive, then three internal equilibria may coexist for intermediate or high migration rates (e.g., Karlin, 1977).

2. For the present two-locus model, nine internal equilibria can coexist if  $\vartheta_1 = \sigma_1 = -1$  and  $\vartheta_2 = \sigma_2 = 1$ . These are the combinations of the three one-locus polymorphisms mentioned above. They exist if recombination is strong. Only one ( $F_m$ ) is stable.  $F_m$  is simultaneously stable with the four monomorphisms and four single-locus polymorphisms. This example illustrates that with dominance, equilibrium configurations and dynamics can become very complicated.

If  $\vartheta_1 = 1$  or  $\vartheta_2 = 1$  or  $\sigma_1 = -1$  or  $\sigma_2 = -1$ , the approximation (3.4) does not apply. It is straightforward, though tedious, to derive approximations in the resulting cases. Then, to leading order, at least one allele frequency depends on the square root of a migration rate, which is reminiscent of mutation–selection balance of dominant alleles. We refer to Appendix A.1.3 for the special case when the adaptive alleles are dominant.

### 3.2. Strong migration

We assume that migration is much stronger than selection and recombination. To this end, we set

$$a_k = \epsilon \alpha_k, \quad b_k = \epsilon \beta_k \quad (k = 1, 2) \quad (3.7)$$

and

$$r = \epsilon \rho, \quad (3.8)$$

where  $\epsilon > 0$  is sufficiently small and  $\alpha_k, \beta_k$ , and  $\rho$  are defined by these relations whence,  $\alpha_1 > 0 > \alpha_2$  and  $\beta_1 > 0 > \beta_2$ . In the limit  $\epsilon \rightarrow 0$ , the discrete-time dynamics (2.8) can be approximated by the so-called strong-migration limit, in which the population is well mixed, i.e., spatially homogeneous. The strong-migration limit is a system of ordinary differential equations that is formally equivalent to a two-locus selection–recombination system for a panmictic population, in which the fitnesses of genotypes are spatial averages of the fitnesses in the two demes (Section 4.2 in Bürger, 2009a).

To derive the strong-migration limit, we define the ratio

$$\phi = \frac{m_1}{m_1 + m_2}, \quad (3.9)$$

where  $(1 - \phi, \phi)$  is the leading left eigenvector of the backward migration matrix (which has the diagonal elements  $1 - m_1$  and  $1 - m_2$ ). We assume that migration occurs in both directions, i.e.,  $0 < \phi < 1$ , and define the spatially averaged selection coefficients,

$$\bar{\alpha} = (1 - \phi)\alpha_1 + \phi\alpha_2 \quad \text{and} \quad \bar{\beta} = (1 - \phi)\beta_1 + \phi\beta_2, \quad (3.10)$$

and the spatially averaged fitness contributions of the single-locus heterozygotes,

$$\begin{aligned} \bar{\vartheta}\alpha &= (1 - \phi)\vartheta_1\alpha_1 + \phi\vartheta_2\alpha_2 \quad \text{and} \\ \bar{\sigma}\beta &= (1 - \phi)\sigma_1\beta_1 + \phi\sigma_2\beta_2. \end{aligned} \quad (3.11)$$

The averaged genotypic selection coefficients are still additive between loci, i.e., nonepistatic. At locus A, the selection coefficients of the genotypes  $A_1A_1$ ,  $A_1A_2$ , and  $A_2A_2$  are  $\bar{\alpha}$ ,  $\bar{\vartheta}\alpha$ , and  $-\bar{\alpha}$ , and similarly at locus B.

If there is no dominance at locus A, then  $\bar{\vartheta}\alpha = 0$ . If the degree of dominance is deme independent at locus A ( $\vartheta_1 = \vartheta_2 = \vartheta$ ), then  $\bar{\vartheta}\alpha = \vartheta\bar{\alpha}$  (and analogously for locus B).

Straightforward calculations show that the strong-migration limit is given by

$$\frac{d\bar{p}}{dt} = \bar{p}(1 - \bar{p}) [\bar{\alpha} + \bar{\vartheta}\alpha(1 - 2\bar{p})] + \bar{D} [\bar{\beta} + \bar{\sigma}\beta(1 - 2\bar{q})], \quad (3.12a)$$

$$\frac{d\bar{q}}{dt} = \bar{q}(1 - \bar{q}) [\bar{\beta} + \bar{\sigma}\beta(1 - 2\bar{q})] + \bar{D} [\bar{\alpha} + \bar{\vartheta}\alpha(1 - 2\bar{p})], \quad (3.12b)$$

$$\begin{aligned} \frac{d\bar{D}}{dt} &= \bar{D} [(1 - 2\bar{p})(\bar{\alpha} + \bar{\vartheta}\alpha(1 - 2\bar{p})) \\ &\quad + (1 - 2\bar{q})(\bar{\beta} + \bar{\sigma}\beta(1 - 2\bar{q})) - \rho]. \end{aligned} \quad (3.12c)$$

Here,

$$\bar{p} = (1 - \phi)p_1 + \phi p_2 \quad \text{and} \quad \bar{q} = (1 - \phi)q_1 + \phi q_2 \quad (3.13)$$

indicate the spatially averaged allele frequencies and  $\bar{D}$  the LD of spatially averaged gamete frequencies.

The dynamics of (3.12) is well understood (Karlin and Liberman, 1978; Bürger, 2000, Chap. II.1.1). Because, there is no epistasis, mean fitness is a (strict) Lyapunov function (Ewens, 1969), whence every trajectory converges to an equilibrium point. In addition, every equilibrium is in LE and there is at most one internal (fully polymorphic) equilibrium. It exists if and only if at each locus there is overdominance ( $\bar{\vartheta}\alpha > |\bar{\alpha}|$  and  $\bar{\sigma}\beta > |\bar{\beta}|$ ) or underdominance ( $\bar{\vartheta}\alpha < |\bar{\alpha}|$  and  $\bar{\sigma}\beta < |\bar{\beta}|$ ). The internal equilibrium is globally asymptotically stable if there is overdominance at both loci; otherwise, it is unstable. The equilibrium allele frequencies at this internal equilibrium are

$$\hat{p} = \frac{1}{2} (1 + \bar{\alpha}/\bar{\vartheta}\alpha) \quad \text{and} \quad \hat{q} = \frac{1}{2} (1 + \bar{\beta}/\bar{\sigma}\beta). \quad (3.14)$$

For the discrete-time dynamics (2.8), Proposition 4.10 and Theorem 4.3 in Bürger (2009a) imply that there exists a stable fully polymorphic equilibrium for arbitrarily strong migration if and only if there is average overdominance. In this case, it is globally asymptotically stable (cf. Theorem 2.2 and Proposition 2.6 in Bürger, 2009b). If there is average overdominance at one locus, this locus will be maintained polymorphic for arbitrarily strong migration. Otherwise, in particular, if average dominance is intermediate or absent, no polymorphism can be maintained for strong migration.

#### 4. Continent–island model

We analyze the so-called continent–island model (CI model), in which there is one-way migration from one deme (the continent) into the second deme (the island). It is assumed that the continental population is at equilibrium, in our case fixed for the locally advantageous haplotype. We assume that  $m_2 = 0$  in (2.13), hence our island population inhabits deme 1, in which alleles  $A_1$  and  $B_1$  are advantageous. The fitnesses of the genotypes are given by (2.1) for  $k = 1$ . All immigrants are of type  $A_2B_2$ . It is sufficient to keep track of the frequencies on the island, which considerably simplifies the dynamics and analysis.

Because the continuous-time version of the model is much more accessible to analysis than the discrete-time version (compare Nagylaki's treatment of discrete time in 1992 with his analysis of continuous time in 1975), and because their properties do not differ qualitatively, we restrict attention to continuous time. We assume that selection, recombination, and migration are weak, and rescale the selection coefficients, the recombination probability, and the migration rate according to

$$a_1 = \epsilon\alpha, \quad b_1 = \epsilon\beta, \quad r = \epsilon\rho, \quad m_1 = \epsilon\mu. \quad (4.1)$$

Rescaling also time and letting  $\epsilon \rightarrow 0$  (cf. Bürger, 2009a), we obtain the continuous-time dynamics on the island,

$$\begin{aligned} \frac{dp}{dt} &= \alpha p(1 - p)(1 + \vartheta(1 - 2p)) \\ &\quad + \beta D(1 + \sigma(1 - 2q)) - \mu p, \end{aligned} \quad (4.2a)$$

$$\begin{aligned} \frac{dq}{dt} &= \beta q(1 - q)(1 + \sigma(1 - 2q)) \\ &\quad + \alpha D(1 + \vartheta(1 - 2p)) - \mu q, \end{aligned} \quad (4.2b)$$

$$\begin{aligned} \frac{dD}{dt} &= D[\alpha(1 - 2p)(1 + \vartheta(1 - 2p)) \\ &\quad + \beta(1 - 2q)(1 + \sigma(1 - 2q))] - \rho D + \mu(pq - D), \end{aligned} \quad (4.2c)$$

where we omitted the subscripts for  $p$ ,  $q$ , and  $D$ . Because we are treating a continuous-time model, the parameters  $\alpha$ ,  $\beta$ ,  $\rho$ , and  $\mu$  are rates (of growth, recombination, and migration), whence they can be arbitrarily large. Their magnitude is determined by the time scale. By rescaling time, for instance to units of  $\rho$  or  $\mu$ , the number of independent parameters could be reduced by one without changing the equilibrium properties. We refrain from doing so because in our applications it will be illuminating to use either  $\rho$  or  $\mu$  as an independent parameter.

Two loci with genic selection were studied in BA11, and two loci with epistasis but no dominance by Bank et al. (2012). A complete analysis of the full model (4.2), in particular of the internal equilibria, seems impossible. Therefore, we focus on the limiting cases of weak migration (Section 4.2), strong migration (Section 4.3), linkage equilibrium (Section 4.4), and complete linkage (Section 4.5).

The following critical migration rates will be useful:

$$\mu_1^A = \alpha(1 + \vartheta), \quad (4.3a)$$

$$\mu_1^B = \beta(1 + \sigma), \quad (4.3b)$$

$$\mu^M = \alpha(1 + \vartheta) + \beta(1 + \sigma) - \rho. \quad (4.3c)$$

##### 4.1. Boundary equilibria

The only monomorphic state that is an equilibrium if  $\mu > 0$  is fixation of the continental haplotype  $A_2B_2$ :

$$M: \quad \hat{p} = 0, \quad \hat{q} = 0, \quad \hat{D} = 0. \quad (4.4)$$

The eigenvalues at M are

$$\begin{aligned} \alpha(1 + \vartheta) - \mu, \quad \beta(1 + \sigma) - \mu, \\ \alpha(1 + \vartheta) + \beta(1 + \sigma) - \rho - \mu. \end{aligned} \quad (4.5)$$

Therefore,

M is asymptotically stable if and only if

$$\mu > \max\{\mu_1^A, \mu_1^B, \mu^M\}. \quad (4.6)$$

Up to four single-locus polymorphisms of (4.2) are admissible. Their equilibrium frequencies and their stability within the one-locus system is known from one-locus theory (Nagylaki, 1975, Section III). Setting

$$p_{\pm} = \frac{1 + 3\vartheta \pm \sqrt{(1 - \vartheta)^2 + 8\mu\vartheta/\alpha}}{4\vartheta}, \quad (4.7a)$$

$$q_{\pm} = \frac{1 + 3\sigma \pm \sqrt{(1 - \sigma)^2 + 8\mu\sigma/\beta}}{4\sigma}, \quad (4.7b)$$

we can write the equilibrium frequencies of the four single-locus polymorphisms as

$$E_{A,1}: \hat{p} = p_-, \quad \hat{q} = 0, \quad \hat{D} = 0, \quad (4.8a)$$

$$E_{A,2}: \hat{p} = p_+, \quad \hat{q} = 0, \quad \hat{D} = 0, \quad (4.8b)$$

$$E_{B,1}: \hat{p} = 0, \quad \hat{q} = q_-, \quad \hat{D} = 0, \quad (4.8c)$$

$$E_{B,2}: \hat{p} = 0, \quad \hat{q} = q_+, \quad \hat{D} = 0. \quad (4.8d)$$

The admissibility conditions for these equilibria are given in Appendix A.4.

Within their respective one-locus systems,  $E_{A,1}$  and  $E_{B,1}$  are asymptotically stable whenever admissible.  $E_{A,2}$  and  $E_{B,2}$  are always unstable (Nagylaki, 1975). Although the eigenvalues at the single-locus polymorphisms with respect to the complete state space  $S_4$  can be determined explicitly, the stability conditions are too complicated to be informative. We refrain from presenting them and derive stability only in the special cases below.

#### 4.2. Weak migration

As in Section 3.1, we assume that the migration rate is small compared to the recombination rate and to the fitness differences among genotypes. This requires the assumption

$$-1 < \vartheta < 1 \quad \text{and} \quad -1 < \sigma < 1. \quad (4.9)$$

It follows that a unique, globally asymptotically stable full polymorphism  $F_{\mu}$  exists. Up to first order in  $\mu$ , its coordinates are given by

$$\hat{p} = 1 - \frac{\mu}{\alpha(1 - \vartheta)} \frac{\alpha(1 - \vartheta) + \rho}{\alpha(1 - \vartheta) + \beta(1 - \sigma) + \rho}, \quad (4.10a)$$

$$\hat{q} = 1 - \frac{\mu}{\beta(1 - \sigma)} \frac{\beta(1 - \sigma) + \rho}{\alpha(1 - \vartheta) + \beta(1 - \sigma) + \rho}, \quad (4.10b)$$

$$\hat{D} = \frac{\mu}{\alpha(1 - \vartheta) + \beta(1 - \sigma) + \rho}, \quad (4.10c)$$

which is simpler than the corresponding Eqs. (3.4a)–(3.4c) in discrete time. In the absence of dominance ( $\vartheta = \sigma = 0$ ), (4.10) was obtained in BA11.

The approximation (4.10) is not admissible if  $\vartheta = 1$  or  $\sigma = 1$ , i.e., if an island allele is completely dominant. If  $\vartheta = -1$  or  $\sigma = -1$ ,  $F_{\mu}$  is admissible and asymptotically stable, and its approximation is (4.10), but in general it is not unique. Up to two additional full polymorphisms (if either  $\vartheta = -1$  or  $\sigma = -1$ ) or

up to three additional full polymorphisms (if  $\vartheta = \sigma = -1$ ) may enter the state space under weak migration; see Appendix A.5.

The influence of recombination and dominance on the allele frequencies and LD is the same as in the more general case (3.4) because, as already noted, to this order of approximation, allele frequencies and LD at equilibrium are independent of the other deme. Hence, the remarks following (3.4) apply to (4.10) as well.

The special case where the island alleles are completely dominant ( $\vartheta = \sigma = 1$ ) is treated in Appendix A.7.

#### 4.3. Strong migration

If migration is sufficiently strong, the island is swamped by the continental haplotype. Indeed, it can be proved that global convergence to M occurs for arbitrary intermediate dominance if  $\mu > 2(\alpha + \beta)$  (Appendix A.6). For specific parameters lower bounds apply; see, e.g., Theorem 4.2 for  $\rho = 0$  or Theorem 2 in BA11 for  $\vartheta = \sigma = 0$ .

#### 4.4. Linkage equilibrium

If recombination is strong relative to selection and migration,  $\rho \gg \max\{\alpha, \beta, \mu\}$ , the dynamics (4.2) can be approximated by a dynamical system on  $[0, 1]^2$  for the allele frequencies,

$$\frac{dp}{dt} = \alpha p(1 - p)(1 + \vartheta(1 - 2p)) - \mu p, \quad (4.11a)$$

$$\frac{dq}{dt} = \beta q(1 - q)(1 + \sigma(1 - 2q)) - \mu q, \quad (4.11b)$$

by assuming LE ( $D = 0$ ). Because epistasis is absent, the equations are decoupled and the dynamics for locus A and locus B can be studied separately. Therefore, (4.11) is the Cartesian product of the dynamics at locus A and B. Its properties can be derived from Nagylaki's (1975) one-locus analysis. Moreover, (4.11), and its generalization to an arbitrary number of loci, is a gradient system because each equation is equivalent to a one-locus mutation–selection model, hence a gradient system; cf. Remark 4.3.

In addition to M and the four single-locus polymorphism, the following four full polymorphisms may exist:

$$F_1: \hat{p} = p_-, \quad \hat{q} = q_-, \quad (4.12a)$$

$$F_2: \hat{p} = p_-, \quad \hat{q} = q_+, \quad (4.12b)$$

$$F_3: \hat{p} = p_+, \quad \hat{q} = q_-, \quad (4.12c)$$

$$F_4: \hat{p} = p_+, \quad \hat{q} = q_+, \quad (4.12d)$$

with  $p_{\pm}$  and  $q_{\pm}$  as in (4.7). From Section 4.1, we infer that

$$F_1 \text{ is admissible if and only if (A.22) and (A.26) hold.} \quad (4.13)$$

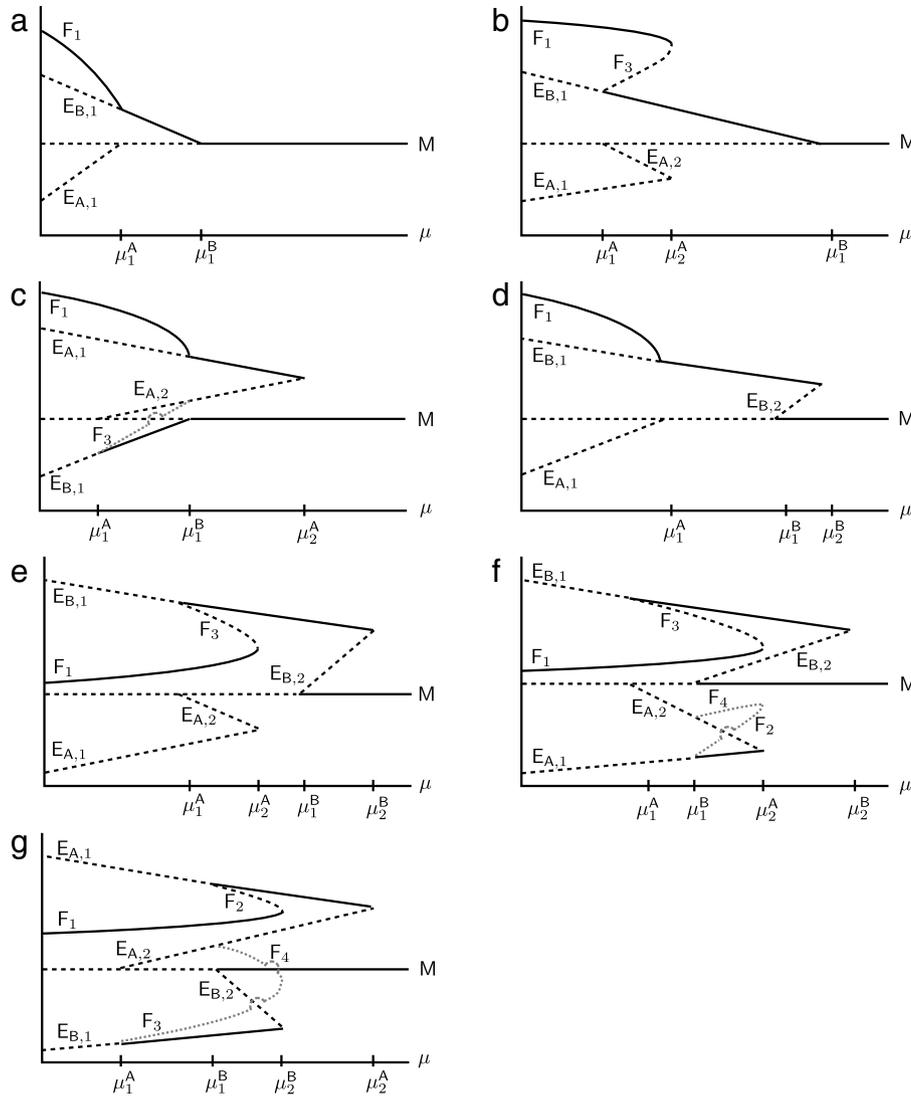
Because  $F_1$  is the Cartesian product of two asymptotically stable equilibria of the respective single-locus dynamics, it is asymptotically stable if admissible. Analogously, the other three internal equilibria are unstable (if admissible). For weak migration,  $F_1$  corresponds to  $F_{\mu}$  (4.10).

Under the assumption (4.9), the following theorem provides all equilibrium configurations of (4.11) as a function of the migration rate. For an efficient presentation, we define

$$\zeta_1 = \frac{1 + \sigma}{1 + \vartheta}, \quad (4.14a)$$

$$\zeta_2 = \frac{\vartheta(1 - \sigma)^2}{\sigma(1 - \vartheta)^2}, \quad (4.14b)$$

$$\zeta_3 = -8\vartheta \frac{1 + \sigma}{(1 - \vartheta)^2}. \quad (4.14c)$$



**Fig. 1.** Bifurcation diagrams for the CI model under linkage equilibrium. Diagrams (a)–(g) represent all possible equilibrium configurations, corresponding to the cases (i)–(vii) in Theorem 4.1. Each diagram displays the possible equilibria as a function of the migration rate  $\mu$ . Each line indicates one equilibrium. The lines are drawn such that intersections of lines occur if and only if the corresponding equilibria bifurcate (note that the gray lines in diagrams (c), (f), and (g) intersect only at their origin and their end). Solid lines represent asymptotically stable equilibria, dashed lines unstable equilibria. Equilibria are shown if and only if they are admissible.

We assume without loss of generality that

$$0 < \alpha < \zeta_1\beta. \tag{4.15}$$

The case  $\alpha > \zeta_1\beta > 0$  can be obtained by exchanging the loci, i.e., by the transformation  $\alpha \leftrightarrow \beta, \vartheta \leftrightarrow \sigma, E_{A,1} \leftrightarrow E_{B,1}$ , etc.

**Theorem 4.1.** Assume (4.11), (4.9) and (4.15). Fig. 1 shows all possible bifurcation diagrams that occur for an open set of parameters  $(\alpha, \beta, \vartheta, \sigma, \rho)$ , where  $\mu$  is the bifurcation parameter.

(i) Diagram (a) in Fig. 1 applies if and only if

$$-\frac{1}{3} < \vartheta < 1 \quad \text{and} \quad -\frac{1}{3} < \sigma < 1 \tag{4.16}$$

hold.

(ii) Diagram (b) in Fig. 1 applies if and only if

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -\frac{1}{3} < \sigma < 1 \quad \text{and} \quad 0 < \alpha < \zeta_3\beta \tag{4.17}$$

hold.

(iii) Diagram (c) in Fig. 1 applies if and only if

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -\frac{1}{3} < \sigma < 1 \quad \text{and} \quad \zeta_3\beta < \alpha < \zeta_1\beta \tag{4.18}$$

hold.

(iv) Diagram (d) in Fig. 1 applies if and only if

$$-\frac{1}{3} < \vartheta < 1 \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \tag{4.19}$$

hold.

(v) Diagram (e) in Fig. 1 applies if and only if

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad 0 < \alpha < \zeta_3\beta \tag{4.20}$$

hold.

(vi) Diagram (f) in Fig. 1 applies if and only if either

$$-1 < \sigma \leq \vartheta < -\frac{1}{3} \quad \text{and} \quad \zeta_3\beta < \alpha < \zeta_1\beta, \tag{4.21}$$

or

$$-1 < \vartheta < \sigma < -\frac{1}{3} \quad \text{and} \quad \zeta_3\beta < \alpha < \zeta_2\beta \quad (4.22)$$

hold.

(vii) Diagram (g) in Fig. 1 applies if and only if

$$-1 < \vartheta < \sigma < -\frac{1}{3} \quad \text{and} \quad \zeta_2\beta < \alpha < \zeta_1\beta \quad (4.23)$$

hold.

The proof is given in Appendix A.8.

Case (i) in Theorem 4.1 is the only case in which there is a unique asymptotically stable equilibrium for every migration rate. Case (i) applies, in particular, if at each locus there is no dominance (see BA11) or the locally advantageous alleles ( $A_1$  and  $A_2$ ) are partially dominant. In cases (ii)–(iv), there are two simultaneously stable equilibria for a range of intermediate migration rates, and in cases (v) and (vi) there are four.

By a perturbation analysis of  $F_1$ , a so-called quasi-linkage-equilibrium approximation can be obtained which yields the stable internal equilibrium for strong recombination. In the absence of dominance, the result is quite simple (Eqs. 4.3 in BA11). The expressions with dominance are much more complicated and not presented.

#### 4.5. No recombination

We assume completely linked loci ( $\rho = 0$ ), set

$$h = \frac{\alpha\vartheta + \beta\sigma}{\alpha + \beta}, \quad (4.24)$$

and define

$$\varphi_{\pm} = \frac{1}{4h} \left[ 1 + 3h \pm \sqrt{(1-h)^2 + 8h\mu/(\alpha + \beta)} \right] \quad \text{if } h \neq 0, \quad (4.25a)$$

$$\varphi_{-} = 1 - \mu/(\alpha + \beta) \quad \text{if } h = 0. \quad (4.25b)$$

We admit  $|\vartheta| = 1$  or  $|\sigma| = 1$ , hence  $|h| = 1$ .

**Theorem 4.2.** Assume (4.2),  $\rho = 0$ , and that initially the locally adaptive haplotype  $A_1B_1$  is present, i.e.,  $x_1 > 0$ .

- (i) Every solution converges to an equilibrium on the edge  $x_1 + x_4 = 1$ .
- (ii) Let  $-\frac{1}{3} \leq h \leq 1$ . If  $\mu < \mu^M$ , i.e.,  $\mu < (\alpha + \beta)(1 + h)$ , then the equilibrium  $F_{\rho,-}$ , given by

$$\hat{p} = \hat{q} = \varphi_{-} \quad \text{and} \quad \hat{D} = \hat{p}(1 - \hat{q}), \quad (4.26)$$

is globally asymptotically stable.

If  $\mu \geq \mu^M$ , the monomorphic equilibrium M is globally asymptotically stable.

- (iii) Let  $-1 \leq h < -\frac{1}{3}$ . If  $\mu < \mu^M$ , then  $F_{\rho,-}$  is globally asymptotically stable.

If

$$\mu^M < \mu \leq \mu^* = \frac{(1-h)^2(\alpha + \beta)}{-8h}, \quad (4.27)$$

then  $F_{\rho,-}$  is simultaneously stable with M, and there is an unstable equilibrium  $F_{\rho,+}$  between  $F_{\rho,-}$  and M. It is given by

$$\hat{p} = \hat{q} = \varphi_{+}^{(1)} \quad \text{and} \quad \hat{D} = \hat{p}(1 - \hat{q}) \quad (4.28)$$

and satisfies  $F_{\rho,+} = M$  if  $\mu = \mu^M$  and  $F_{\rho,+} = F_{\rho,-}$  if  $\mu = \mu^*$ .

If  $\mu > \mu^*$ , then M is globally asymptotically stable.

**Proof.** Since  $\rho = 0$ , every face and every edge of the simplex  $S_4$  is invariant. Our first and main step is to prove that every trajectory that does not start on  $x_1 = 0$  approaches the edge  $x_1 + x_4 = 1$ . If  $x_1 > 0$ , we obtain from (A.16)

$$\frac{d}{dt} \left( \frac{x_2}{x_1} \right) = -\beta x_2 \frac{1 + \sigma(1 - 2(x_1 + x_3))}{x_1} < 0 \quad \text{if } x_2 > 0, \quad (4.29a)$$

$$\frac{d}{dt} \left( \frac{x_3}{x_1} \right) = -\alpha x_3 \frac{1 + \vartheta(1 - 2(x_1 + x_2))}{x_1} < 0 \quad \text{if } x_3 > 0, \quad (4.29b)$$

where the inequalities in (4.29) follow because  $(x_1, x_2, x_3, x_4) \in S_4$ ,  $1 + \vartheta(1 - 2(x_1 + x_2)) > 0$  if  $x_1 > 0$  and  $x_3 > 0$ , and  $1 + \sigma(1 - 2(x_1 + x_3)) > 0$  if  $x_1 > 0$  and  $x_2 > 0$ . This allows us to conclude that the  $\omega$ -limit of every trajectory with  $x_1 > 0$  is contained in the invariant edge  $x_1 + x_4 = 1$  (LaSalle, 1976).

Therefore, it is sufficient to study the dynamics of trajectories on the edge  $x_1 + x_4 = 1$ . There, the dynamics is equivalent to that of a one-locus CI model, in which the haplotypes  $A_1A_1$  and  $A_2A_2$  play the role of the alleles and the fitnesses of the one-locus genotypes are  $\alpha + \beta$ ,  $h(\alpha + \beta)$ , and  $-(\alpha + \beta)$ . Now, the results of the theorem follow immediately from Nagylaki's (1975) analysis in Section 3.  $\square$

- Remark 4.3.**
1. If  $\mu$  is small, then  $F_{\rho,-}$  corresponds to  $F_{\mu}$  (4.10).
  2. On the boundary face  $x_1 = 0$ , in addition to M, up to two polymorphic equilibria may exist. One of them can be locally stable within the face  $x_1 = 0$ . However, they cannot be approached from  $x_1 > 0$ .
  3. If  $\rho = 0$ , the CI model is formally equivalent to a one-locus mutation–selection model with the same number of alleles and so-called house-of-cards mutation (Section 4.4 Bürger, 2014). For four alleles, this follows by setting the mutation rates to alleles 1, 2, and 3 (corresponding to gametes  $A_1B_1, A_1B_2, A_2B_1$ ) to zero, and assuming that each of alleles 1, 2, and 3 mutates to allele 4 ( $A_2B_2$ ) at rate  $\mu$ ; cf. Section 3.4.4 in BA11.
  4. For small  $\rho$  a perturbation analysis can be performed and the coordinates of the perturbations of  $F_{\rho,-}$  and  $F_{\rho,+}$  can be calculated to leading order in  $\rho$ . Since they are complicated, we do not present them (see BA11, (4.2) for the case without dominance). More importantly, the arguments in Sections 3.4.4 and 3.4.5 of BA11 extend immediately to the present model with dominance and show that (except on  $x_1 = 0$ ) the global dynamics for small  $\rho$  is qualitatively the same as for  $\rho = 0$ . Hence, the only possibly stable equilibria are the perturbation of  $F_{\rho,-}$  and M, and solutions converge to one of them.

Theorem 4.2 shows that for completely linked loci the number of (stable) equilibria and of possible equilibrium configurations is much lower than in the case of LE (Theorem 4.1). The likely explanation is that in the absence of recombination the haplotypes  $A_1B_1$  and  $A_2B_2$ , which are maintained in the population because of selection and immigration, respectively, can produce only the heterozygote genotype  $A_1B_1/A_2B_2$ , whereas all others (if present initially) are removed by selection. With (strong) recombination, however, all other genotypes are produced continuously and some of them have higher fitness than  $A_1B_1/A_2B_2$ . Thus, the effects of dominance become more evident with strong recombination.

## 5. Applications

### 5.1. Linkage disequilibrium

We investigate the effects of dominance and linkage on the amount of LD. The approximations for  $\hat{D}_k$  in (3.4) deceptively suggest that LD is weak if migration is weak. This is true only in absolute terms but not if LD is measured relative to the existing

allele-frequency variation. In fact, if the squared correlation in allelic state

$$\hat{r}^2 = \frac{D^2}{p(1-p)q(1-q)} \quad (5.1)$$

(a common measure for statistical purposes; e.g., Nordborg and Tavaré, 2002; Slatkin, 2008) is used, the picture changes. A simple calculation shows that  $\hat{r}^2$  in deme  $k$  tends to

$$\frac{a_k(1 - \vartheta_k)b_k(1 - \sigma_k)}{[a_k(1 - \vartheta_k) + rW_{14,k}][b_k(1 - \sigma_k) + rW_{14,k}]} \quad (5.2)$$

as  $m_k \rightarrow 0$ . For tightly linked loci, this is close to its maximum value one.

In the following, we study the properties of both measures of LD in the continuous-time CI model of Section 4. There are three reasons for this: (i) the allele frequencies and linkage disequilibria in deme  $k$  depend, to leading order in the migration rate, only on the fitness coefficients, the dominance parameters and the migration rate into deme  $k$ ; (ii) no subscripts are required in the CI model; (iii) the continuous-time model yields simpler, but qualitatively similar, formulas because many higher-order interaction terms vanish. We focus on weak migration and evaluate quantities at the stable fully polymorphic equilibrium  $F_\mu$ . Numerical results are based on the numerical solution of the equilibrium equations obtained from (4.2).

#### Intermediate dominance

From the approximation (4.10c) it follows that  $\hat{D}$  increases with increasing migration rate  $\mu$ . This holds for arbitrary degree of dominance and is valid up to an intermediate migration rate.

The calculation of  $\hat{r}^2$  requires an approximation of  $F_\mu$  to second order in  $\mu$ , thus an extension of (4.10). This is complicated and not shown, but available on request as a *Mathematica* (Wolfram, 2010) notebook. Assuming  $-1 \leq \vartheta < 1$  and  $-1 \leq \sigma < 1$ , we obtain

$$\hat{r}^2 = \frac{\alpha(1-\vartheta)\beta(1-\sigma)}{[\alpha(1-\vartheta) + \rho][\beta(1-\sigma) + \rho]} - \frac{\mu\rho}{(1-\vartheta)(1-\sigma)[\alpha(1-\vartheta) + \rho]^2[\beta(1-\sigma) + \rho]^2} \times \left( \frac{\alpha\rho(1-\vartheta)^2(1-3\sigma)[\alpha(1-\vartheta) + \rho] + \beta\rho(1-\sigma)^2(1-3\vartheta)[\beta(1-\sigma) + \rho]}{\alpha(1-\vartheta) + \beta(1-\sigma) + \rho} - 2\alpha(1-\vartheta)\beta(1-\sigma) \frac{\alpha\sigma(1-\vartheta)^2 + \beta\vartheta(1-\sigma)^2 + \rho[\sigma(1-\vartheta) + \vartheta(1-\sigma)]}{\alpha(1-\vartheta) + \beta(1-\sigma) + \rho} \right). \quad (5.3)$$

If  $\vartheta = \sigma = 0$ , this simplifies to Eq. 4.9 in BA11.

With the help of *Mathematica* it can be shown that  $\hat{r}^2$  is monotonically decreasing in  $\mu$  if  $-1 \leq \vartheta \leq 0$  and  $-1 \leq \sigma \leq 0$ . This contrasts the dependence of  $\hat{D}$  on  $\mu$ . However, if  $\frac{1}{3} \leq \vartheta < 1$  and  $\frac{1}{3} \leq \sigma < 1$ , then  $\hat{r}^2$  is monotonically increasing in  $\mu$ . If  $\vartheta$  and  $\sigma$  satisfy other relations than stated above, the effect of gene flow on  $\hat{r}^2$  depends strongly on the precise relation between the parameters.

In accordance with intuition, tighter linkage increases both  $\hat{D}$  and  $\hat{r}^2$ , because differentiation in (4.10c) and (5.3) yields

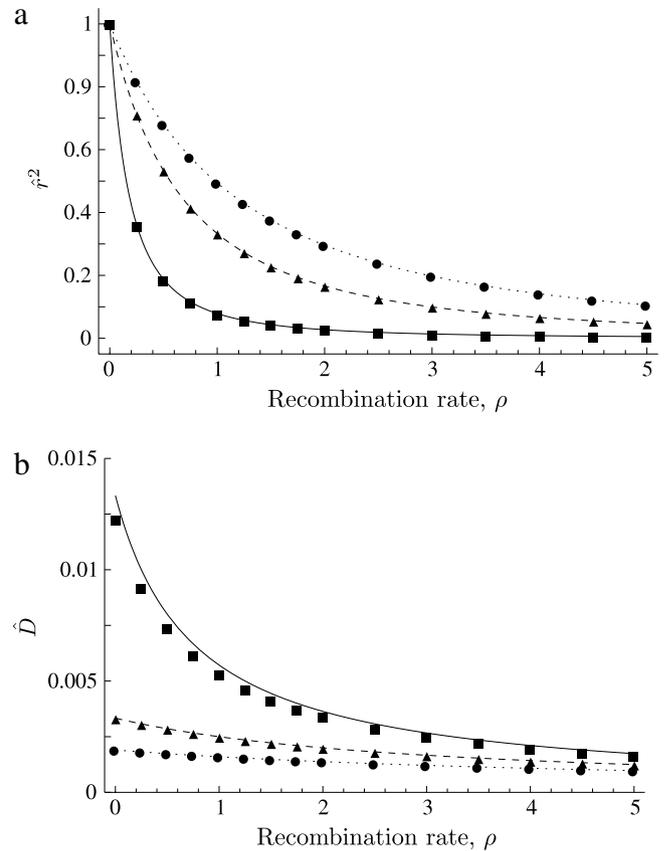
$$\frac{\partial \hat{D}}{\partial \rho} < 0 \quad \text{and} \quad \frac{\partial \hat{r}^2}{\partial \rho} < 0. \quad (5.4)$$

Similarly, we find

$$\frac{\partial \hat{r}^2}{\partial \vartheta} < 0 \quad \text{and} \quad \frac{\partial \hat{r}^2}{\partial \sigma} < 0, \quad (5.5)$$

whereas differentiation in (4.10c) yields

$$\frac{\partial \hat{D}}{\partial \vartheta} > 0 \quad \text{and} \quad \frac{\partial \hat{D}}{\partial \sigma} > 0. \quad (5.6)$$



**Fig. 2.** The amount of LD exhibited at the fully polymorphic equilibrium  $F_\mu$  in the CI model (4.2). Diagrams (a) and (b) present the LD measures  $\hat{r}^2$  and  $\hat{D}$  at  $F_\mu$  as functions of  $\rho$  respectively. The lines in diagram (a) show the approximation (5.3) of  $\hat{r}^2$ . The squares, triangles, and circles correspond to values of  $\hat{r}^2$  which were obtained numerically from (4.2). The lines in diagram (b) show the approximation (4.10c) of  $\hat{D}$ . The squares, triangles, and circles correspond to values of  $\hat{D}$  which were obtained numerically from (4.2). In both diagrams we assume  $\alpha = 2\beta = 1$ ,  $\mu = 0.01$ , and  $\vartheta = \sigma$ . The solid line and the squares correspond to  $\vartheta = \frac{3}{4}$  (dominant island alleles). The dashed line and the triangles correspond to  $\vartheta = 0$  (no dominance). The dotted line and the circles correspond to  $\vartheta = -\frac{3}{4}$  (recessive island alleles).

Thus, dominance has opposite effects on  $\hat{r}^2$  and  $\hat{D}$ : whereas dominant island alleles increase  $\hat{D}$ , they decrease  $\hat{r}^2$ .

The lower  $\hat{r}^2$ , the less statistically associated are the alleles. If the island alleles are partially dominant, i.e.,  $0 < \vartheta < 1$  or  $0 < \sigma < 1$ , the fitness of heterozygotes is similar to that of adapted homozygotes and the frequency of maladapted haplotypes and alleles increases; cf. Eq. (A.5b). Therefore, the allelic variance and the denominator in (5.1) increase and the adaptive alleles are expected to be associated with maladaptive alleles more often than when an island allele is recessive. This may be one explanation why dominant island alleles lead to a lower  $\hat{r}^2$ .

We illustrate the effects of linkage and the distinct effects of dominance on  $\hat{r}^2$  and  $\hat{D}$  in Fig. 2 by comparing  $\hat{r}^2$  and  $\hat{D}$  at  $F_\mu$  for different dominance coefficients as functions of the recombination rate  $\rho$ . Fig. 3 in BA11 displays  $\hat{r}^2$  and  $\hat{D}$  as functions of  $\mu/\rho$  if dominance is absent.

#### Completely dominant island alleles

If the island alleles are completely dominant ( $\vartheta = \sigma = 1$ ), we use (A.37) to obtain the approximation

$$\hat{r}^2 \approx 2\mu \frac{\sqrt{\alpha\beta}}{\rho^2} \quad (5.7)$$

for  $\rho > 0$ . Consequently,  $\hat{r}^2$  increases with increasing  $\mu$ . Reduced recombination has a much more pronounced effect on  $\hat{r}^2$  than increased selection. Eq. (5.7) shows that with completely dominant island alleles,  $\hat{r}^2$  decreases remarkably rapidly as  $\rho$  increases. From (A.37c) we find that, to this order of approximation,  $\hat{D} \approx \mu/\rho$ .

### 5.2. Local adaptation and migration load

In this section, we study the joint effects of migration, recombination, and dominance on local adaptation. Immigration of maladapted individuals may severely reduce the mean fitness of a population. This reduction from the maximum possible fitness is called the migration load,  $L$ . We use it as a measure of local adaptation. For simplicity, we consider only the CI model (4.2). Then the mean fitness on the island is

$$\bar{w} = \alpha[2p - 1 + 2\vartheta p(1 - p)] + \beta[2q - 1 + 2\sigma q(1 - q)]; \quad (5.8)$$

cf. (2.12). It is maximized (and equals  $\alpha + \beta$ ) if the island haplotype ( $\hat{p} = \hat{q} = 1$  and  $\hat{D} = 0$ ) is fixed. Therefore, the migration load is

$$L = \alpha + \beta - \bar{w} \\ = 2\alpha(1 - p)(1 - \vartheta p) + 2\beta(1 - q)(1 - \sigma q). \quad (5.9)$$

#### Intermediate dominance

If migration is sufficiently weak and dominance parameters satisfy  $-1 \leq \vartheta < 1$  and  $-1 \leq \sigma < 1$ , we use (4.10) to approximate the asymptotically stable equilibrium  $F_\mu$  and obtain

$$\hat{L} = 2\mu \left( 1 + \frac{\rho}{\alpha(1 - \vartheta) + \beta(1 - \sigma) + \rho} \right) + \mathcal{O}(\mu^2). \quad (5.10)$$

This shows that stronger recombination and a higher degree of dominance of the advantageous alleles increase the load, hence decrease local adaptation.

Under LE ( $\rho \rightarrow \infty$ ), (5.10) simplifies to

$$\hat{L} = 4\mu + \mathcal{O}(\mu^2), \quad (5.11a)$$

whereas for complete linkage

$$\hat{L} = 2\mu + \mathcal{O}(\mu^2) \quad (5.11b)$$

is obtained. Thus, in these limiting cases intermediate dominance has no effect on the mean fitness to leading order in  $\mu$ .

#### Completely dominant island alleles

For completely dominant island alleles ( $\vartheta = \sigma = 1$ ) and if  $\rho > 0$ , we derive the migration load at  $F_\mu$  using the approximation (A.37) and obtain

$$\hat{L} = 2\mu + \frac{\mu\sqrt{2\mu}}{\rho}(\sqrt{\alpha} + \sqrt{\beta}) + \mathcal{O}(\mu^2). \quad (5.12)$$

It increases with increasing recombination rate  $\rho$ . In the limit of  $\rho \rightarrow \infty$  (LE),

$$\hat{L} = 2\mu \quad (5.13a)$$

is obtained. For complete linkage ( $\rho = 0$ ), the load is

$$\hat{L} = \mu. \quad (5.13b)$$

These equations are exact, because the exact equilibrium frequencies were derived in these cases. In view of Remark 4.3, it is not surprising that the migration loads derived above parallel the well-known formulas for the mutation load if  $\mu$  is interpreted as the mutation rate (e.g., Bürger, 2000, Chap. III.3.1); see also Section 6.

As already suggested by (5.10), Eqs. (5.11)–(5.13) confirm that local adaptation is maximized if linkage is complete ( $\rho = 0$ ) and the island alleles are completely dominant ( $\vartheta = \sigma = 1$ ).

### 5.3. Effective migration rate

Linkage of a neutral to a selected locus may impede gene flow at the neutral locus and generate a barrier to gene flow. The concept of an effective migration rate was introduced to quantify this effect of linkage (Petry, 1983; Bengtsson, 1985; Barton and Bengtsson, 1986). Recently, the effective migration at a neutral locus (N) was studied that is located between two loci (A and B) under genic selection (BA11; Akerman and Bürger, 2014; Aeschbacher and Bürger, 2014). Here, we extend the work in BA11 by deriving an explicit expression for the effective migration rate in the CI model with dominance.

Recombination between locus A (B) and the neutral locus occurs with rate  $\rho_1$  ( $\rho_2$ ) such that  $\rho = \rho_1 + \rho_2$ . Thus, only one crossover event occurs in a sufficiently small time interval. We assume that  $\rho_1$  and  $\rho_2$  are positive, i.e., the neutral locus is not completely linked to a selected site. We consider two variants at the neutral locus,  $N_1$  and  $N_2$ , where we denote the frequency of  $N_1$  on the island by  $n$  and on the continent by  $n_c$ .

Evolution at the three loci is modeled by a system of seven ordinary differential equations for the allele frequencies and linkage disequilibria ( $p, q, D_{AB}, n, D_{AN}, D_{NB}, D_{ANB}$ ). The equations for the change of  $p, q$ , and  $D_{AB}$  are given by (4.2). The equations for the allele frequencies at the neutral locus and the associated linkage disequilibria are given in Eqs. (A.47) in the Appendix. The system has an equilibrium with  $n = n_c$  and  $D_{AN} = D_{NB} = D_{ANB} = 0$ . If migration is weak, the population is maintained at  $F_\mu$ . At this globally stable full polymorphism, the Jacobian has block structure,

$$J = \begin{pmatrix} J_S & 0 \\ 0 & J_N \end{pmatrix}, \quad (5.14)$$

where  $J_S$  is the Jacobian describing convergence of  $(p, q, D_{AB})$  to  $F_\mu$  (4.10), and  $J_N$  is the Jacobian describing convergence of  $(n, D_{AN}, D_{NB}, D_{ANB})$  to  $(n_c, 0, 0, 0)$ .

The rate of convergence to equilibrium at the neutral locus is determined by the leading eigenvalue of  $J_N$ . In the limit of weak migration, we derived an approximation for this leading eigenvalue  $\lambda_N$  (details not shown; a Mathematica notebook is available on request). We define the effective migration rate (into the island) by  $\mu_{\text{eff}} = -\lambda_N$  (Bengtsson, 1985; Kobayashi et al., 2008).

#### Intermediate dominance

If  $-1 \leq \vartheta < 1$  and  $-1 \leq \sigma < 1$ , we obtain the effective migration rate up to first order in  $\mu$  by substituting (4.10) into  $J_N$  and calculating  $\lambda_N$ :

$$\mu_{\text{eff}} = \mu \frac{\rho_1}{\rho_1 + \alpha(1 - \vartheta)} \frac{\rho_2}{\rho_2 + \beta(1 - \sigma)}. \quad (5.15)$$

In the absence of dominance (2.3), Eq. (5.15) was derived in BA11. Eq. (5.15) shows that increasing recessiveness of the island alleles ( $\vartheta \downarrow -1$  and  $\sigma \downarrow -1$ ) reduces the effective migration rate, whereas increasing dominance increases it.

#### Completely dominant island alleles

In the limit  $\vartheta \uparrow 1$  and  $\sigma \uparrow 1$ , (5.15) is not admissible because the weak-migration approximation (4.10) of  $F_\mu$  is not admissible. If we employ the weak-migration approximation (A.37) of  $F_\mu$  for completely dominant island alleles ( $\vartheta = \sigma = 1$ ), we obtain

$$\mu_{\text{eff}} = \mu - \mu\sqrt{2\mu} \left( \frac{\sqrt{\alpha}}{\rho_1} + \frac{\sqrt{\beta}}{\rho_2} \right). \quad (5.16)$$

As we assume weak migration,  $\mu \ll \{\alpha, \beta, \rho\}$ , up to first order in  $\mu$ ,  $\mu_{\text{eff}} \approx \mu$  holds. Thus, if both island alleles are completely

dominant, we do not expect a significant effect of linkage on the effective migration rate of a linked neutral marker allele.

#### 5.4. Differentiation of a quantitative trait

In this section, we consider a quantitative trait that is subject to directional selection in opposite directions in the two demes. We study how differentiation is affected by dominance and linkage if the population exchanges migrants at a sufficiently small rate. For this purpose, we calculate and compare three different measures of differentiation, namely  $Q_{ST}$  and two measures of  $F_{ST}$ .

The trait is determined by two diallelic, nonepistatic loci. The effects of the alleles  $A_1, A_2, B_1,$  and  $B_2$  on the genotypic value are  $-\frac{1}{2}\gamma_1, \frac{1}{2}\gamma_1, -\frac{1}{2}\gamma_2,$  and  $\frac{1}{2}\gamma_2,$  respectively, where  $\gamma_1$  and  $\gamma_2$  are positive, fixed, and (without loss of generality) normalized such that

$$\gamma_1 + \gamma_2 = 1. \tag{5.17}$$

The coefficients of dominance in deme  $k$  are designated  $\vartheta_k$  and  $\sigma_k$ , where (2.2) is assumed. If  $\vartheta_k = \vartheta$  and  $\sigma_k = \sigma$  for  $k = 1, 2,$  there is no genotype–environment interaction (or DIDID). The matrix  $G_k$  of genotypic values in deme  $k$  is given by

$$\begin{matrix} & B_1B_1 & B_1B_2 & B_2B_2 \\ \begin{matrix} A_1A_1 \\ A_1A_2 \\ A_2A_2 \end{matrix} & \begin{pmatrix} -\gamma_1 - \gamma_2 & -\gamma_1 - \sigma_k\gamma_2 & -\gamma_1 + \gamma_2 \\ -\vartheta_k\gamma_1 - \gamma_2 & -\vartheta_k\gamma_1 - \sigma_k\gamma_2 & -\vartheta_k\gamma_1 + \gamma_2 \\ \gamma_1 - \gamma_2 & \gamma_1 - \sigma_k\gamma_2 & \gamma_1 + \gamma_2 \end{pmatrix} \end{matrix}. \tag{5.18}$$

We denote by  $G_{ij,k}$  the genotypic value of a genotype formed by gametes  $i$  and  $j$  ( $i, j \in \{1, 2, 3, 4\}$ ) in deme  $k$ . Let  $s_k$  be the selection coefficient in deme  $k$  acting on the quantitative trait. We assume  $s_1 > 0 > s_2$  and  $|s_k| \leq 1$  for  $k = 1, 2.$  The relationship between fitness and genotypic value is given by

$$W_{ij,k} = 1 - s_k G_{ij,k}, \tag{5.19}$$

whence in (2.1) we obtain  $a_k = s_k\gamma_1$  and  $b_k = s_k\gamma_2.$  Therefore, (2.13) governs the discrete-time dynamics of the alleles and LD in the two demes.

Estimators of multilocus  $F_{ST}$  are usually defined as weighted averages of one-locus  $F_{ST}$  estimators (e.g., Weir and Cockerham, 1984; Levis and Hamilton, 2011). In our two-locus model, we denote this measure by

$$\bar{F}_{ST} = \frac{1}{2} (F_{ST}^A + F_{ST}^B), \tag{5.20}$$

where  $F_{ST}^A = \frac{\text{Var}(p)}{\bar{p}(1-\bar{p})}$  is the classical differentiation measure at the diallelic locus A (analogously for  $F_{ST}^B$  with  $q$  instead of  $p$ ). For simplicity, we assume that the deme proportions  $c_1, c_2$  satisfy  $c_1 = 1 - \phi$  and  $c_2 = \phi,$  where  $\phi$  is the migration ratio (3.9). Then migration is reciprocal, i.e., the number of migrants in both directions is the same (Nagylaki, 1992, p. 136). Therefore, allele frequencies are averaged according to  $\bar{p} = (1 - \phi)p_1 + \phi p_2.$

In Akerman and Bürger (2014), a new multilocus measure of  $F_{ST}$  was introduced as an extension of a fixation index defined by Nagylaki (1998). It is a genuine multilocus version of  $F_{ST}$  that measures the covariance of the frequencies of (multilocus) haplotypes. In our diallelic model, it is given by

$$F_{ST}^X = \frac{1}{\sum_i \bar{x}_i(1 - \bar{x}_i)} \sum_i \bar{x}_i(1 - \bar{x}_i) F_{ST}^{x_i}, \tag{5.21}$$

where

$$F_{ST}^{x_i} = \frac{\text{Var}(x_i)}{\bar{x}_i(1 - \bar{x}_i)} \tag{5.22}$$

for  $i = 1, 2, 3, 4.$  The mean frequency  $\bar{x}_i$  of haplotype  $i$  is  $\bar{x}_i = (1 - \phi)x_{i,1} + \phi x_{i,2}.$

A popular measure for differentiation of a quantitative trait is  $Q_{ST}$  (Lande, 1992; Spitze, 1993). It is given by

$$Q_{ST} = \frac{\text{Var}_T(G)}{\text{Var}_T(G) + 2\text{Var}_S(G)}, \tag{5.23}$$

where  $\text{Var}_S(G)$  is the average genotypic variance within demes, and  $\text{Var}_T(G)$  is the genotypic variance among subpopulations (the subscript  $S$  denotes subpopulation and  $T$  denotes total population). These can be written as

$$\text{Var}_S(G) = (1 - \phi)\text{Var}(G_1) + \phi\text{Var}(G_2), \tag{5.24a}$$

$$\begin{aligned} \text{Var}_T(G) &= (1 - \phi)(\bar{G}_1 - \bar{G})^2 + \phi(\bar{G}_2 - \bar{G})^2 \\ &= (1 - \phi)\phi(\bar{G}_1 - \bar{G}_2)^2, \end{aligned} \tag{5.24b}$$

where

$$\bar{G}_k = \sum_{i,j} G_{ij,k} x_{i,k} x_{j,k}, \tag{5.25a}$$

$$\text{Var}(G_k) = \sum_{i,j} (G_{ij,k} - \bar{G}_k)^2 x_{i,k} x_{j,k}, \tag{5.25b}$$

are the mean genotypic value and the genetic variance in deme  $k,$  respectively, and

$$\bar{G} = (1 - \phi)\bar{G}_1 + \phi\bar{G}_2 \tag{5.25c}$$

is the average genotypic value of the metapopulation. In terms of allele frequencies, the mean genotypic value in deme  $k$  is

$$\begin{aligned} \bar{G}_k &= \gamma_1 [1 - 2p_k(1 + \vartheta_k(1 - p_k))] \\ &\quad + \gamma_2 [1 - 2q_k(1 + \sigma_k(1 - q_k))]. \end{aligned} \tag{5.26}$$

The genetic variance  $\text{Var}(G_k)$  in deme  $k$  can be decomposed into an additive component,  $\text{Var}_{\text{add}}(G_k),$  and a dominance component,  $\text{Var}_{\text{dom}}(G_k).$  The additive variance can be further decomposed into the independent contributions of the two loci and an interaction term due to LD:

$$\text{Var}_{\text{add}}(G_k) = 2\gamma_1^2 p_k(1 - p_k) + 2\gamma_2^2 q_k(1 - q_k) + 4\gamma_1\gamma_2 D_k. \tag{5.27}$$

The dominance variance is of more complicated form:

$$\begin{aligned} \text{Var}_{\text{dom}}(G_k) &= 2\gamma_1^2 \vartheta_k p_k(1 - p_k)[2(1 - 2p_k) \\ &\quad + \vartheta_k(1 - 2p_k(1 - p_k))] + 2\gamma_2^2 \sigma_k q_k(1 - q_k) \\ &\quad \times [2(1 - 2q_k) + \sigma_k(1 - 2q_k(1 - q_k))] \\ &\quad + 4\gamma_1\gamma_2 D_k[\vartheta_k(1 - 2p_k) + \sigma_k(1 - 2q_k) \\ &\quad + \vartheta_k\sigma_k((1 - 2p_k)(1 - 2q_k) + 2D_k)] \end{aligned} \tag{5.28}$$

(Avery and Hill, 1978).

Whenever  $\bar{F}_{ST}, F_{ST}^X,$  or  $Q_{ST}$  attain the value one or zero we speak of complete differentiation or of no differentiation, respectively. In the following, we assume weak migration and present explicit formulas for  $F_{ST}^X, \bar{F}_{ST},$  and  $Q_{ST}$  at the full polymorphism  $F_m$  (a *Mathematica* notebook with detailed derivations is available on request). For ease of presentation, we define the total migration rate by  $m = m_1 + m_2;$  then  $\phi = m_1/m.$

#### Intermediate dominance

If the dominance coefficients satisfy (2.2) and (3.3), the full polymorphism  $F_m$  is admissible for weak migration with approximate equilibrium frequencies given in (3.4). It is unique and globally asymptotically stable if (3.5) holds. We set

$$d_1 = \gamma_1 \vartheta_1 + \gamma_2 \sigma_1 \quad \text{and} \quad d_2 = \gamma_1 \vartheta_2 + \gamma_2 \sigma_2. \tag{5.29}$$

If  $m$  is sufficiently small, we obtain for  $F_{ST}^A$  at  $F_m$ :

$$F_{ST}^A \approx 1 - \frac{m}{\gamma_1} \left( \frac{1 + s_1}{s_1(1 - \vartheta_1)} \frac{s_1\gamma_1(1 - \vartheta_1) + r(1 + s_1d_1)}{s_1(1 - d_1) + r(1 + s_1d_1)} + \frac{1 - s_2}{s_2(1 + \vartheta_2)} \frac{s_2\gamma_1(1 + \vartheta_2) - r(1 + s_2d_2)}{-s_2(1 + d_2) + r(1 + s_2d_2)} \right). \quad (5.30)$$

The formula for  $F_{ST}^B$  is analogous. Therefore, after a brief calculation (5.20) yields the following first-order approximation in  $m$ :

$$\begin{aligned} \bar{F}_{ST} \approx & 1 - \frac{m(1 + s_1)}{s_1(1 - d_1) + r(1 + s_1d_1)} \\ & \times \left[ 1 + \frac{r(1 + s_1d_1)}{2s_1} \left( \frac{1}{\gamma_1(1 - \vartheta_1)} + \frac{1}{\gamma_2(1 - \sigma_1)} \right) \right] \\ & - \frac{m(1 - s_2)}{-s_2(1 + d_2) + r(1 + s_2d_2)} \\ & \times \left[ 1 - \frac{r(1 + s_2d_2)}{2s_2} \left( \frac{1}{\gamma_1(1 + \vartheta_2)} + \frac{1}{\gamma_2(1 + \sigma_2)} \right) \right]. \quad (5.31) \end{aligned}$$

Under sufficiently weak migration,  $F_{ST}^X$  (5.21) at  $F_m$  is approximately given by

$$\begin{aligned} F_{ST}^X \approx & 1 - \frac{m}{\gamma_1\gamma_2} \left( \frac{1 + s_1}{s_1(1 - \vartheta_1)(1 - \sigma_1)} \right. \\ & \times \frac{s_1\gamma_1\gamma_2(1 - \vartheta_1)(1 - \sigma_1) + r(1 - d_1)(1 + s_1d_1)}{s_1(1 - d_1) + r(1 + s_1d_1)} \\ & + \frac{(1 - s_2)}{-s_2(1 + \vartheta_2)(1 + \sigma_2)} \\ & \left. \times \frac{-s_2\gamma_1\gamma_2(1 + \vartheta_2)(1 + \sigma_2) + r(1 + d_2)(1 + s_2d_2)}{-s_2(1 + d_2) + r(1 + s_2d_2)} \right). \quad (5.32) \end{aligned}$$

Both  $\bar{F}_{ST}$  and  $F_{ST}^X$  increase with decreasing recombination probability  $r$ , i.e., for fixed  $s_k, \vartheta_k, \sigma_k, m_k$  ( $k = 1, 2$ ) and fixed allelic effects  $\gamma_1, \gamma_2$ , differentiation is maximized for complete linkage ( $r = 0$ ) and minimized for free recombination ( $r = 1/2$ ). Under sufficiently weak migration, the difference  $F_{ST}^X - \bar{F}_{ST}$  is approximately

$$\begin{aligned} -\frac{mr}{2\gamma_1\gamma_2} \left( \frac{1 + s_1}{s_1(1 - \vartheta_1)(1 - \sigma_1)} \frac{(1 - d_1)(1 + s_1d_1)}{s_1(1 - d_1) + r(1 + s_1d_1)} \right. \\ \left. + \frac{1 - s_2}{-s_2(1 + \sigma_2)(1 + \vartheta_2)} \frac{(1 + d_2)(1 + s_2d_2)}{-s_2(1 + d_2) + r(1 + s_2d_2)} \right). \quad (5.33) \end{aligned}$$

Since the term in brackets is positive,

$$F_{ST}^X < \bar{F}_{ST} \quad (5.34)$$

holds whenever  $r > 0$  (the difference being maximized if  $r = 1/2$ ).

In the absence of dominance ( $\vartheta_k = \sigma_k = 0$  for  $k = 1, 2$ ), the difference is

$$F_{ST}^X - \bar{F}_{ST} \approx -\frac{mr}{2\gamma_1\gamma_2} \left( \frac{1 + s_1}{s_1(r + s_1)} + \frac{1 - s_2}{-s_2(r - s_2)} \right), \quad (5.35)$$

and

$$\begin{aligned} F_{ST}^X \approx & 1 - \frac{m}{\gamma_1\gamma_2} \\ & \times \left( \frac{(r + s_1\gamma_1\gamma_2)(1 + s_1)}{s_1(r + s_1)} + \frac{(r - s_2\gamma_1\gamma_2)(1 - s_2)}{-s_2(r - s_2)} \right). \quad (5.36) \end{aligned}$$

For complete linkage ( $r = 0$ ), (5.33) yields  $F_{ST}^X = \bar{F}_{ST}$  and (5.32) produces

$$F_{ST}^X = \bar{F}_{ST} \approx 1 - m \left( 2 + \frac{1}{s_1} - \frac{1}{s_2} \right). \quad (5.37)$$

In this special case,  $F_{ST}^X (= \bar{F}_{ST})$  is independent of the magnitude of allelic effects.

If  $r = 0$  and dominance coefficients satisfy (3.3), we obtain

$$\begin{aligned} F_{ST}^X = \bar{F}_{ST} \\ \approx & 1 - m \left[ \frac{1}{1 - d_1} \left( 1 + \frac{1}{s_1} \right) + \frac{1}{1 + d_2} \left( 1 - \frac{1}{s_2} \right) \right]. \quad (5.38) \end{aligned}$$

Next, we investigate the effects of selection and of dominance on  $F_{ST}^X$  and  $\bar{F}_{ST}$ . For simplicity, we assume that only one locus exhibits dominance (locus A). Setting  $\sigma_1 = \sigma_2 = 0$  in the respective formulas of  $\bar{F}_{ST}$  (5.31) and  $F_{ST}^X$  (5.32), we obtain that independent of the degree of dominance  $\vartheta_k$  ( $k = 1, 2$ ), both  $\bar{F}_{ST}$  and  $F_{ST}^X$  increase with increasing selection strength,  $|s_k|$ . As expected, stronger diversifying selection leads to increased population differentiation.

Dominance has the same qualitative effects on  $F_{ST}^X$  and  $\bar{F}_{ST}$ : With the help of *Mathematica*, it is easy to show that if  $\vartheta_1$  and  $\vartheta_2$  are varied independently,

$$\frac{\partial F_{ST}^X}{\partial \vartheta_1} < 0, \quad \frac{\partial \bar{F}_{ST}}{\partial \vartheta_1} < 0, \quad \frac{\partial F_{ST}^X}{\partial \vartheta_2} > 0, \quad \frac{\partial \bar{F}_{ST}}{\partial \vartheta_2} > 0 \quad (5.39)$$

hold for fixed but arbitrary selection coefficients and recombination probabilities. This implies that recessiveness of locally adaptive alleles (i.e.,  $A_1$  in deme 1 if  $\vartheta_1 < 0$  and  $A_2$  in deme 2 if  $\vartheta_2 > 0$ ) elevates population differentiation.

Fig. 3(a) shows a representative example of the effect of dominance on  $\bar{F}_{ST}$  and  $F_{ST}^X$  in which we assume  $\vartheta_1 = -\vartheta_2$ . By (5.39), both  $\bar{F}_{ST}$  and  $F_{ST}^X$  decrease with increasing  $\vartheta_1$ . Population differentiation (if measured by  $\bar{F}_{ST}$  or  $F_{ST}^X$ ) is largest if  $\vartheta_1 = -\vartheta_2 = -1$ .

If there is DIDID, i.e.,  $\vartheta_1 = \vartheta_2 = \vartheta$ ,  $F_{ST}^X$  and  $\bar{F}_{ST}$  are not monotone in  $\vartheta$ . Numerical calculations suggest that  $F_{ST}^X$  and  $\bar{F}_{ST}$  are always concave with respect to  $\vartheta$ , with their maximum at intermediate values of  $\vartheta$ . Unfortunately, no general proof is available.

In Fig. 3(b) we assume DIDID ( $\vartheta_1 = \vartheta_2$ ) and illustrate the effect of the degree of dominance on  $F_{ST}^X$  and  $\bar{F}_{ST}$ . Due to the symmetry assumptions on selection,  $F_{ST}^X$  and  $\bar{F}_{ST}$  attain their maximum at  $\vartheta_1 = \vartheta_2 = 0$ .

Finally, we investigate the effects of linkage, selection, and dominance on the third differentiation measure,  $Q_{ST}$  (5.23). For sufficiently weak migration,  $Q_{ST}$  at  $F_m$  is given by

$$Q_{ST} \approx 1 - m \left[ (1 - d_1) \left( 1 + \frac{1}{s_1} \right) + (1 + d_2) \left( 1 - \frac{1}{s_2} \right) \right]. \quad (5.40)$$

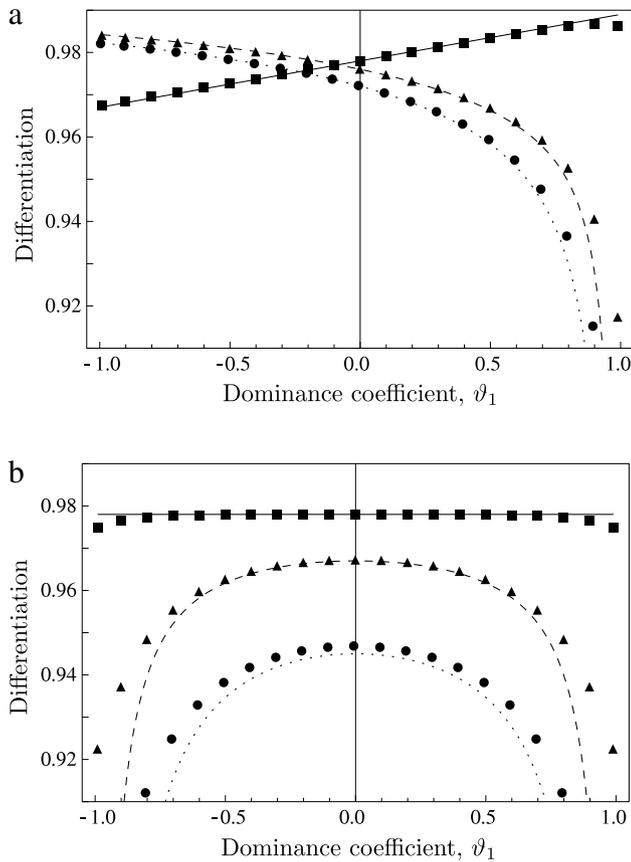
Interestingly,  $Q_{ST}$  is independent of the recombination probability  $r$  to first order in  $m$ . Supplemental numerical investigations (not shown) confirm that  $Q_{ST}$  is rather insensitive to variation in  $r$  if migration is weak. However, in contrast to  $\bar{F}_{ST}$  and  $F_{ST}^X$  (see above),  $Q_{ST}(r = 0) < Q_{ST}(r = 1/2)$  may occur (though the difference is very small).

By our assumption on the selection coefficients and allelic effects, the term in brackets in (5.40) is positive, and it increases with increasing selection strength  $|s_k|$  ( $k = 1, 2$ ). Thus, as for  $F_{ST}^X$  and  $\bar{F}_{ST}$ , stronger diversifying selection leads to increased population differentiation.

If  $\vartheta_1$  and  $\vartheta_2$  vary independently, the effect of dominance on  $Q_{ST}$  is given by

$$\frac{\partial Q_{ST}}{\partial \vartheta_1} > 0 \quad \text{and} \quad \frac{\partial Q_{ST}}{\partial \vartheta_2} < 0 \quad (5.41)$$

(analogous formulas hold for locus B with  $\sigma_1$  and  $\sigma_2$ ). Thus, differentiation  $Q_{ST}$  of the trait is elevated if the locally adaptive alleles are dominant compared with the case where the adaptive alleles are recessive (e.g., trait differentiation will be larger if  $A_1$  is dominant in deme 1,  $\vartheta_1 > 0$ , and if  $A_2$  is dominant in deme 2,  $\vartheta_2 < 0$ ).



**Fig. 3.** Differentiation of a quantitative trait. Diagrams (a) and (b) show  $F_{ST}^X$ ,  $F_{ST}$ , and  $Q_{ST}$  at  $F_m$  in the two-deme model (2.8) as functions of the dominance coefficient  $\vartheta_1$ . In both diagrams we assume  $s_1 = -s_2 = 0.1$ ,  $\gamma_1 = \gamma_2 = 1$ ,  $m = 0.001$ ,  $\phi = \frac{1}{2}$ , and  $\sigma_1 = \sigma_2 = 0$ . Diagram (a) displays the case  $\vartheta_1 = \vartheta_2$ . Diagram (b) displays the case  $\vartheta_1 = -\vartheta_2$  (DIDID). In both diagrams, solid, dashed, and dotted lines correspond to the approximations (5.40) of  $Q_{ST}$ , (5.31) of  $\bar{F}_{ST}$ , and (5.32) of  $F_{ST}^X$  at  $F_m$ , respectively. Squares, triangles, and circles correspond to values of  $Q_{ST}$ ,  $\bar{F}_{ST}$  and  $F_{ST}^X$  at  $F_m$ , respectively. They are obtained by calculating the internal equilibrium numerically from (2.8).

Comparison of (5.39) and (5.41) shows that increasing dominance has opposite effects on the differentiation measures  $F_{ST}^X$  and  $\bar{F}_{ST}$  on the one hand, and  $Q_{ST}$  on the other. This is illustrated by Fig. 3(a) which assumes  $\vartheta_1 = -\vartheta_2$ . There,  $Q_{ST}$  increases in  $\vartheta_1$ , whereas both  $F_{ST}^X$  and  $\bar{F}_{ST}$  decrease.

If there is DIDID, we obtain

$$\frac{\partial Q_{ST}}{\partial \vartheta} < 0 \quad \text{if and only if} \quad s_1 > -s_2. \tag{5.42}$$

In Fig. 3(b), where DIDID is assumed ( $\vartheta_1 = -\vartheta_2$ ),  $Q_{ST}$  is independent of the degree of dominance because  $s_1 = -s_2$  in our example. We note also that under DIDID,  $Q_{ST}$  may show a different dependence on dominance than  $F_{ST}^X$  and  $\bar{F}_{ST}$ .

From (5.38) and (5.40), we obtain

$$F_{ST}^X = \bar{F}_{ST} = Q_{ST} \tag{5.43}$$

if linkage is complete ( $r = 0$ ) and dominance is absent ( $\vartheta_k = \sigma_k = 0$  for  $k = 1, 2$ ). Then (5.37) applies.

In the absence of dominance ( $\vartheta_k = \sigma_k = 0$  for  $k = 1, 2$ ) it follows from the easily established facts  $\partial F_{ST}^X / \partial r < 0$ ,  $\partial \bar{F}_{ST} / \partial r < 0$ , and  $\partial Q_{ST} / \partial r = 0$  that

$$F_{ST}^X \leq \bar{F}_{ST} \leq Q_{ST} \tag{5.44}$$

if migration is sufficiently weak. Here, equality holds if and only if  $r = 0$ .

With dominance, (5.44) may be violated. Numerical analyses showed that  $Q_{ST} < F_{ST}^X$  and  $Q_{ST} < \bar{F}_{ST}$  may hold. An example is displayed in Fig. 3(a), where  $Q_{ST} < F_{ST}^X < \bar{F}_{ST}$  holds if  $\vartheta_1 = -\vartheta_2 < -\frac{1}{4}$ .

In summary, the measures of population differentiation  $\bar{F}_{ST}$  and  $F_{ST}^X$  (based on differences in alleles or haplotypes) show a qualitatively different behavior with respect to dominance than  $Q_{ST}$  (based on differences in trait distributions).

*Completely dominant locally adaptive alleles*

If the dominance parameters do not satisfy (3.6), the above approximations do not hold. As an example, we treat the case when the locally adaptive alleles are completely dominant in their respective deme, i.e., (A.13) holds. Then the fully polymorphic equilibrium  $F_m$  is approximated by (A.14) if  $r > 0$ , and by (A.15) if  $r = 0$ . We use these approximations to derive  $\bar{F}_{ST}$ ,  $F_{ST}^X$ , and  $Q_{ST}$  to leading order in  $m$ .

If  $r > 0$  we obtain

$$F_{ST}^X \approx 1 - \sqrt{\frac{m}{2}} \left( \frac{1}{\sqrt{\gamma_1}} + \frac{1}{\sqrt{\gamma_2}} \right) \times \left( \sqrt{\frac{1+s_1}{s_1\phi}} + \sqrt{\frac{1-s_2}{-s_2(1-\phi)}} \right), \tag{5.45}$$

and if  $r = 0$ , then

$$F_{ST}^X \approx 1 - \sqrt{\frac{m}{2}} \left( \sqrt{\frac{1+s_1}{s_1\phi}} + \sqrt{\frac{1-s_2}{-s_2(1-\phi)}} \right), \tag{5.46}$$

which is independent of the allelic effects  $\gamma_1, \gamma_2$  up to first order in  $\sqrt{m}$ . In both cases,  $F_{ST}^X$  and  $\bar{F}_{ST}$  are independent of  $r$  to first order in  $\sqrt{m}$ .

If the dominance coefficients satisfy (A.13), then

$$\bar{F}_{ST} \approx \frac{1}{2} (1 + F_{ST}^X) \tag{5.47}$$

holds for every  $r \geq 0$ . It follows that  $F_{ST}^X \leq \bar{F}_{ST}$ .  $Q_{ST}$  is given by

$$Q_{ST} \approx 1 - m \left( 2 + \frac{1}{s_1} - \frac{1}{s_2} \right), \tag{5.48}$$

which is independent of  $r$  and of the allelic effects  $\gamma_1, \gamma_2$ . Interestingly, the approximation (5.48) for complete dominance of the form (A.13) equals that for no dominance ( $\vartheta_1 = \sigma_1 = \vartheta_2 = \sigma_2 = 0$ ) in (5.40). We note, however, that if (A.13) is applied to (5.40), then (5.48) is not obtained.

**6. Discussion**

The purpose of our analysis was to shape our understanding of the combined effects of gene flow, dominance, and linkage on the patterns of genetic variation and on the degree of local adaptation and of differentiation in a subdivided population inhabiting a heterogeneous environment. To this end, we studied a deterministic selection–migration model for two linked diallelic loci under selection that exhibit intermediate dominance in each of two demes (Section 2). Alternative alleles are selectively favored in the two demes. Epistasis and genetic drift are ignored.

The first part of our analysis is focused on the existence and the stability of polymorphic equilibria and on the dependence of the equilibrium configurations on migration, selection, dominance, and recombination. From these results, we derived informative approximations for the LD measures  $D$  and  $r^2$ , the migration load, and the effective migration rate at a linked neutral marker locus. In addition, we studied the consequences for differentiation in

a quantitative trait if selection acts in opposite directions in the two demes by deriving approximations for  $Q_{ST}$  and two different variants of multilocus  $F_{ST}$ . We briefly recapitulate the main results before discussing some implications.

The results in Sections 3 and 4 demonstrate that dominance may lead to a much richer equilibrium structure than observed in the absence of dominance. This is already known from one-locus models (Karlin, 1977; Karlin and Campbell, 1980; Nagylaki, 1975, 1992, 2009a) but becomes aggravated with more loci. If dominance at each locus and in each deme is strictly intermediate, i.e., if (3.5) holds, then for sufficiently weak migration there is a unique fully polymorphic equilibrium ( $F_m$ ), which is globally asymptotically stable. At this equilibrium, the locally adaptive alleles occur at high frequency in ‘their’ deme; see (3.4). Increasing dominance of locally adaptive alleles decreases their frequency in ‘their’ deme and increases LD.

If at least one allele is completely dominant in one deme or, more precisely, if (3.6) holds, then more than one stable equilibrium may exist for arbitrarily weak migration. In fact, with two-way migration up to nine internal equilibria may coexist, but only one of them (corresponding to  $F_m$ ) can be stable. In this extreme situation, which occurs if the locally adaptive alleles are recessive, several boundary equilibria are also stable. If an advantageous allele is partially recessive (in at least one deme), then multiple internal equilibria and coexistence of stable internal and boundary equilibria can occur for intermediate migration rates, but not for arbitrarily small ones.

For sufficiently strong migration (Section 3.2), the equilibrium structure is simpler because the population is well mixed. Therefore, and because there is no epistasis, a locus can be maintained stably polymorphic if and only if there is spatially averaged overdominance. This implies that, somewhat counter intuitively, genetic variation at several loci can be maintained under very strong migration. What is required is average overdominance at those loci; cf. Bürger (2009b).

Because of the complexity of the two-deme model with forward and backward migration, we studied a continent–island (CI) model in more detail, in which the continent is fixed for the haplotype  $A_2B_2$  and  $A_1B_1$  is the haplotype with highest fitness on the island (Section 4). To simplify the analysis further we used a continuous-time version (4.2), which amounts to assuming that all evolutionary forces are weak. As for two-way migration, if migration is sufficiently weak and dominance strictly intermediate, there is a unique, globally asymptotically stable equilibrium,  $F_\mu$ , given by (4.10). If an advantageous allele is recessive, then again multiple internal equilibria may coexist and the equilibrium  $F_\mu$  is simultaneously stable with up to three boundary equilibria. Thus, the possible equilibrium configurations are fewer and simpler than with forward and backward migration.

For strong migration, the CI model is much simpler because then the continental haplotype swamps the island and no polymorphism can be maintained.

For arbitrary migration rates and intermediate dominance, we obtained the possible equilibrium configurations for the case that recombination is so strong that LE can be assumed (Section 4.4) and the case of completely linked loci (Section 4.5). Under the assumption of LE, the equilibrium structure and dynamics can be inferred quite straightforwardly from the well known one-locus case (Nagylaki, 1975, 1992). The seven possible bifurcation patterns are summarized in Theorem 4.1 and illustrated in Fig. 1. If  $\rho = 0$ , the number of possible bifurcation patterns is only two (Theorem 4.2) and they are equivalent to the patterns that can occur with one diallelic locus. The reason is that in the absence of recombination, only the immigrating continental haplotype ( $A_2B_2$ ) and the locally beneficial island haplotype ( $A_1B_1$ ) can be maintained. Segregation produces the corresponding double heterozygote. An explanation

why the number of bifurcation scenarios is elevated under strong recombination may be that recombination facilitates the formation of many types of heterozygotes (with different fitnesses), in which the effects of dominance become evident only.

As proved in BA11 under the assumption of no dominance, for intermediate recombination and migration rates more complicated equilibrium structures can occur than for the limiting cases treated above. Because these results are structurally stable, they extend to small deviations from no dominance. We expect that with strong dominance, much more complex equilibrium configurations than discussed above can occur if recombination and migration are about as strong as selection. An analytical study seems hopeless and a comprehensive numerical study may not be worth the effort for the insight to be gained.

Section 5 is devoted to various applications. In Section 5.1, we studied the amount of LD exhibited at the full polymorphism  $F_\mu$  in the CI model for weak migration. We compared two measures of LD: The classical population-genetic measure  $D$  ( $= x_1x_4 - x_2x_3$ ), and  $r^2$  (5.1) which is often used for statistical purposes. Both measures depend quite differently on the underlying parameters. Of course, both decrease with increasing recombination rate  $\rho$  (Fig. 2). However, a comparison of Fig. 2(a) and (b) reveals that increased levels of dominance (higher  $\vartheta$  or  $\sigma$ ) increase  $D$ , but decrease  $r^2$ . This is proved in Eqs. (5.5) and (5.6). We showed that independently of the degree of dominance,  $D$  increases monotonically with increasing migration (until a maximum is reached at an intermediate migration rate). In contrast, the measure  $r^2$  may increase or decrease with increasing  $\mu$ . This depends strongly on the values of  $\vartheta$  and  $\sigma$ . It decreases in the absence of dominance or if the island alleles at both loci are (partially) recessive.

The delicate sensitivity of these measures to the underlying parameters may compromise inference methods which often assume genic selection, thus ignore dominance and epistasis. In practice, neutral markers will rarely be completely linked to selected sites. However, variability at linked neutral sites will be affected by selection until LD decays to average levels. In particular, linked selected sites will act as barriers to gene flow at neutral markers (Petry, 1983; Bengtsson, 1985; Barton and Bengtsson, 1986).

Therefore, we derived the effective migration rate,  $\mu_{\text{eff}}$ , at a neutral site that is located between two selected sites (Section 5.3). The resulting expression (5.15) is simple and provides considerable insight. First, increasing dominance of a locally adaptive allele increases the effective migration rate, although not quite up to  $\mu$  (as shown by (5.16) which holds for complete dominance). Recessiveness of island alleles magnifies the barrier to gene flow considerably, particularly if recombination is weaker than selection. Eq. (5.15) also shows that for two loci the gene-flow factor  $\mu_{\text{eff}}/\mu$  (Bengtsson, 1985) is the product of the single-locus gene-flow factors. It can be shown that this property extends to multiple loci; cf. Eq. (24) in Aeschbacher and Bürger (2014). Hence, linkage to more than one selected site may magnify barriers to gene flow substantially. Extension of (5.15) to two demes with bidirectional migration should parallel the expression (8.3) in Akerman and Bürger (2014).

To explore the effects of dominance and linkage on local adaptation on the island, we calculated the migration load at the globally stable, fully polymorphic equilibrium  $F_\mu$  for weak migration. For intermediate dominance the load is given by (5.10), and for completely dominant island alleles it is given by (5.12). These expressions show that more recombination and a higher degree of dominance of the island alleles increase the load, hence hamper local adaptation. The migration load in the limiting cases of LE, Eqs. (5.11a) and (5.13a), and of complete linkage, Eqs. (5.11b) and (5.13b), coincides with the mutation load if the migration rate is substituted by the average mutation rate of the involved loci.

In fact, this is expected from the formal equivalence of one-locus CI models with two or more alleles to mutation–selection models of house-of-cards type (Section 4.4, Remark 4.3, and Bürger, 2014, Section 4.4).

Simple intuitive explanations of these formulas are obtained by adapting the classical ‘genetic death’ argument of H.J. Muller, as detailed in Crow and Kimura (1970, pp. 301–302). For completely linked loci, only the haplotypes  $A_1B_1$  (which is selectively favored) and  $A_2B_2$  (the maladaptive immigrant) persist in the population. Hence one selective death (of an individual carrying  $A_2B_2$ ) eliminates twice as many deleterious genes as in LE when  $A_2$  and  $B_2$  occur independently. This explains the factor two by which (5.11a) and (5.13a) and (5.11b) and (5.13b) differ. The reduction by an additional factor of two from (5.11a) to (5.11b), and from (5.13a) to (5.13b), occurs because completely recessive deleterious alleles are only eliminated in homozygous individuals, whence only half the number of selective deaths is required compared to alleles with intermediate dominance.

In Section 5.4, we studied differentiation of a quantitative trait that is determined by two linked diallelic loci exhibiting dominance. Linear (nonepistatic) selection acts in opposite direction in the two demes. We used the discrete-time model (2.8) and assumed that migration between the two demes is sufficiently weak, such that the fully polymorphic equilibrium  $F_m$  is globally asymptotically stable with approximate equilibrium frequencies given by (3.4). We calculated and compared three distinct measures of differentiation: (i)  $\bar{F}_{ST}$  (5.20), which measures differentiation of alleles in the population by averaging the single-locus values of  $F_{ST}$ ; (ii)  $F_{ST}^X$  (5.21), which measures differentiation of haplotypes in the population, and (iii)  $Q_{ST}$  (5.23), a measure of differentiation based on trait variances. The corresponding approximations are given by (5.31), (5.32) and (5.40), respectively.

Our results demonstrate that  $F_{ST}^X$  and  $\bar{F}_{ST}$  show qualitatively similar behavior with respect to selection, recombination, and dominance. Increasing dominance of a locally adaptive allele reduces  $F_{ST}^X$  and  $\bar{F}_{ST}$  (5.39). This is in contrast to  $Q_{ST}$ , which increases with increasing dominance of the locally adaptive alleles (5.41). However,  $Q_{ST}$  seems to be much less sensitive to deviations from no dominance than  $\bar{F}_{ST}$  and  $F_{ST}^X$  (Fig. 3).  $\bar{F}_{ST}$ ,  $F_{ST}^X$ , and  $Q_{ST}$  all increase with decreasing recombination or with increasing strength of selection on the trait. In the absence of dominance, we showed that  $F_{ST}^X \leq \bar{F}_{ST} \leq Q_{ST}$  (5.44), holds with equality only if  $r = 0$ . Whereas we proved that  $F_{ST}^X \leq \bar{F}_{ST}$  holds for arbitrary dominance (and weak migration), numerical analyses show that  $Q_{ST} < \bar{F}_{ST}$  or  $Q_{ST} < F_{ST}^X$  may occur (Fig. 3(a)). We also derived approximations of  $F_{ST}^X$ ,  $\bar{F}_{ST}$ , and  $Q_{ST}$  for completely dominant island alleles, (5.45)–(5.48).

Comparisons of  $Q_{ST}$  and  $F_{ST}$  are often used to infer selection. For neutral traits that are genetically controlled by purely additive genes (no dominance and epistasis), the mean  $Q_{ST}$  is equal to the mean  $F_{ST}$  of neutral loci (Lande, 1992; Whitlock, 1999). If the  $Q_{ST}$  of a trait is significantly greater than the  $F_{ST}$  of the neutral loci, this is often taken as evidence that the trait has diversified more than would be expected by genetic drift alone, whereas the opposite relation might be taken as evidence of stabilizing selection on the trait toward the same optimum in each subpopulation. Whitlock (2008) discusses the possible pitfalls of such conclusions lucidly and warns that nonadditive genetic factors can easily cause the  $Q_{ST}$  of a neutral trait to be much lower than  $F_{ST}$  (see also Goudet and Martin, 2007, and the literature cited therein). Our results about  $Q_{ST}$  and  $F_{ST}$  have only an indirect bearing on inference issues because we calculated  $F_{ST}$  for selected loci. However, they clearly show that dominance may induce any relation between  $Q_{ST}$  and  $F_{ST}$ , in particular if markers are linked to selected sites. More detailed conclusions about the utility of  $Q_{ST} - F_{ST}$  contrasts might become available once the properties of our genuine multilocus measure  $F_{ST}^X$  of  $F_{ST}$  has been explored for neutral models.

## Acknowledgments

The authors thank an anonymous referee for constructive comments on the manuscript and Dr. Simon Aeschbacher for pointing out an error in Eqs. (4.25), (4.26) of BA11. The more general equations (A.47) presented here take into account his correction. This work was supported by grants P21305 and P25188 of the Austrian Science Fund FWF. In addition, support by the Vienna Graduate School of Population Genetics funded by the Austrian Science Fund (FWF, W1225) is acknowledged.

## Appendix A

### A.1. Two-deme model: weak migration

#### A.1.1. Properties of $F_0$ and $F_m$

The eigenvalues at  $F_0$  are given by

$$\lambda_1^{F_0} = \frac{1 + \vartheta_1 a_1 + b_1}{1 + a_1 + b_1}, \quad \lambda_2^{F_0} = \frac{1 + a_1 + \sigma_1 b_1}{1 + a_1 + b_1},$$

$$\lambda_3^{F_0} = (1 - r) \frac{1 + \vartheta_1 a_1 + \sigma_1 b_1}{1 + a_1 + b_1}, \quad (A.1)$$

$$\lambda_4^{F_0} = \frac{1 + \vartheta_2 a_2 - b_2}{1 - a_2 - b_2}, \quad \lambda_5^{F_0} = \frac{1 - a_2 + \sigma_2 b_2}{1 - a_2 - b_2},$$

$$\lambda_6^{F_0} = (1 - r) \frac{1 + \vartheta_2 a_2 + \sigma_2 b_2}{1 - a_2 - b_2}. \quad (A.2)$$

They are positive. From the assumptions (2.2) and  $0 \leq r \leq \frac{1}{2}$ , it follows that  $\lambda_3^{F_0} \neq 1$  and  $\lambda_6^{F_0} \neq 1$ . Obviously,  $F_0$  is hyperbolic (i.e., no eigenvalue lies on the unit circle) if and only if the dominance coefficients satisfy (3.3).

The following formulas imply the statements in Section 3.1 about the influence of recombination and dominance on the equilibrium allele frequencies and the linkage disequilibria. Straightforward differentiation of the allele frequencies and linkage disequilibria in (3.4) together with our assumptions on the selection and dominance coefficients, i.e., (2.5), (2.6), and (2.2), yield

$$\frac{\partial \hat{D}_1}{\partial r} < 0, \quad \frac{\partial \hat{D}_2}{\partial r} < 0, \quad (A.3a)$$

$$\frac{\partial \hat{p}_1}{\partial r} < 0, \quad \frac{\partial \hat{q}_1}{\partial r} < 0, \quad \frac{\partial \hat{p}_2}{\partial r} > 0, \quad \frac{\partial \hat{q}_2}{\partial r} > 0 \quad (A.3b)$$

and

$$\frac{\partial \hat{p}_k}{\partial \vartheta_k} < 0, \quad \frac{\partial \hat{p}_k}{\partial \sigma_k} < 0, \quad \frac{\partial \hat{q}_k}{\partial \vartheta_k} < 0, \quad \frac{\partial \hat{q}_k}{\partial \sigma_k} < 0$$

for  $k = 1, 2$ . (A.4)

Moreover,

$$\frac{\partial \hat{D}_1}{\partial \vartheta_1} > 0, \quad \frac{\partial \hat{D}_1}{\partial \sigma_1} > 0, \quad \frac{\partial \hat{D}_2}{\partial \vartheta_2} < 0, \quad \frac{\partial \hat{D}_2}{\partial \sigma_2} < 0, \quad (A.5a)$$

such that in deme 1

$$\frac{\partial \hat{x}_{i,1}}{\partial \vartheta_1} > 0 \quad \text{for } i = 2, 3, 4, \quad \text{and} \quad \frac{\partial \hat{x}_{1,1}}{\partial \vartheta_1} < 0 \quad (A.5b)$$

hold, and in deme 2

$$\frac{\partial \hat{x}_{i,2}}{\partial \vartheta_2} < 0 \quad \text{for } i = 1, 2, 3, \quad \text{and} \quad \frac{\partial \hat{x}_{4,2}}{\partial \vartheta_2} > 0. \quad (A.5c)$$

Analogous relations to (A.5b) or (A.5c) hold for  $\sigma_1$  or  $\sigma_2$ , respectively.

A.1.2. Properties of other equilibria

In the absence of migration ( $m_1 = m_2 = 0$ ), and if considering the complete state space  $S_4 \times S_4$ , there exist four full polymorphisms. In addition to  $F_0$  (3.2), the three remaining full polymorphisms (located on the boundary of  $S_4 \times S_4$ ) are:

$$F_0^{(2)}: \hat{p}_1 = \hat{q}_1 = 0, \quad \hat{p}_2 = \hat{q}_2 = 1, \quad \hat{D}_1 = \hat{D}_2 = 0, \quad (A.6a)$$

$$F_0^{(3)}: \hat{p}_1 = \hat{q}_2 = 1, \quad \hat{q}_1 = \hat{p}_2 = 0, \quad \hat{D}_1 = \hat{D}_2 = 0, \quad (A.6b)$$

$$F_0^{(4)}: \hat{p}_1 = \hat{q}_2 = 0, \quad \hat{q}_1 = \hat{p}_2 = 1, \quad \hat{D}_1 = \hat{D}_2 = 0. \quad (A.6c)$$

All  $F_0^{(i)}$  ( $i = 2, 3, 4$ ) are unstable and generically hyperbolic. We show this for  $F_0^{(2)}$  explicitly, the remaining cases are left to the interested reader.

If  $k = 1, 2$  the eigenvalues at  $F_0^{(2)}$  are given by

$$\begin{aligned} \lambda_{1,k} &= \frac{1 - \alpha_k + \sigma_k \beta_k}{1 - \alpha_k - \beta_k}, & \lambda_{2,k} &= \frac{1 + \vartheta_k \alpha_k - \beta_k}{1 - \alpha_k - \beta_k}, \\ \lambda_{3,k} &= (1 - r) \frac{1 + \vartheta_k \alpha_k + \sigma_k \beta_k}{1 - \alpha_k - \beta_k}. \end{aligned} \quad (A.7)$$

All eigenvalues are positive.  $F_0^{(2)}$  is hyperbolic if and only if  $\vartheta_k \neq -1, \sigma_k \neq 1$  ( $k = 1, 2$ ), or  $r \neq \frac{\alpha_1(1+\vartheta_1)+\beta_1(1+\sigma_1)}{1+\vartheta_1\alpha_1+\sigma_1\beta_1} (\leq \frac{1}{2})$ . As  $\lambda_{1,1} \geq 1$  and  $\lambda_{2,1} \geq 1, F_0^{(2)}$  is unstable.

In the following we assume (2.2). Then  $F_0$  is globally asymptotically stable. If all equilibria are hyperbolic (which is generically the case but is not shown here), Theorem 5.4 in Bürger (2009a) can be applied. According to this theorem, unstable full polymorphisms located on the boundary of  $S_4 \times S_4$  if  $m_1 = m_2 = 0$  may leave the state space under small perturbations by positive migration. This is indeed the case for  $F_0^{(i)}$  ( $i = 2, 3, 4$ ). For  $F_0^{(2)}$  we show this explicitly, the remaining full polymorphisms are left to the interested reader.

If migration is weak, the perturbation of  $F_0^{(2)}$  up to order  $m_1$  and  $m_2$  is given by

$$\hat{p}_1 = -m_1 \frac{(1 - \alpha_1 - \beta_1)}{\alpha_1(1 + \vartheta_1)} \times \frac{r(1 + \vartheta_1\alpha_1 + \sigma_1\beta_1) - \alpha_1(1 + \vartheta_1)}{r(1 + \vartheta_1\alpha_1 + \sigma_1\beta_1) - \alpha_1(1 + \vartheta_1) - \beta_1(1 + \sigma_1)}, \quad (A.8a)$$

$$\hat{q}_1 = -m_1 \frac{(1 - \alpha_1 - \beta_1)}{\beta_1(1 + \sigma_1)} \times \frac{r(1 + \vartheta_1\alpha_1 + \sigma_1\beta_1) - \beta_1(1 + \sigma_1)}{r(1 + \vartheta_1\alpha_1 + \sigma_1\beta_1) - \alpha_1(1 + \vartheta_1) - \beta_1(1 + \sigma_1)}, \quad (A.8b)$$

$$\hat{D}_1 = m_1 \frac{(1 - \alpha_1 - \beta_1)}{r(1 + \vartheta_1\alpha_1 + \sigma_1\beta_1) - \alpha_1(1 + \vartheta_1) - \beta_1(1 + \sigma_1)}, \quad (A.8c)$$

$$\hat{p}_2 = 1 - m_2 \frac{1 + \alpha_2 + \beta_2}{\alpha_2(1 - \vartheta_2)} \times \frac{r(1 + \vartheta_2\alpha_2 + \sigma_2\beta_2) + \alpha_2(1 - \vartheta_2)}{r(1 + \vartheta_2\alpha_2 + \sigma_2\beta_2) + \alpha_2(1 - \vartheta_2) + \beta_2(1 - \sigma_2)}, \quad (A.8d)$$

$$\hat{q}_2 = 1 - m_2 \frac{1 + \alpha_2 + \beta_2}{\beta_2(1 - \sigma_2)} \times \frac{r(1 + \vartheta_2\alpha_2 + \sigma_2\beta_2) + \beta_2(1 - \sigma_2)}{r(1 + \vartheta_2\alpha_2 + \sigma_2\beta_2) + \alpha_2(1 - \vartheta_2) + \beta_2(1 - \sigma_2)}, \quad (A.8e)$$

$$\hat{D}_2 = m_2 \frac{1 + \alpha_2 + \beta_2}{r(1 + \vartheta_2\alpha_2 + \sigma_2\beta_2) + \alpha_2(1 - \vartheta_2) + \beta_2(1 - \sigma_2)}. \quad (A.8f)$$

At this equilibrium, the constraint in (2.11b) in deme 1 is approximately

$$- \min \{0, 1 - \hat{p}_1 - \hat{q}_1\} = 0 < D_1 < \min \{\hat{p}_1, \hat{q}_1\} \quad (A.9)$$

if migration is sufficiently weak, where by (A.8b) the following relation holds:

$$\hat{D}_1 > 0 \iff r > \frac{\alpha_1(1 + \vartheta_1) + \beta_1(1 + \sigma_1)}{1 + \vartheta_1\alpha_1 + \sigma_1\beta_1}. \quad (A.10)$$

From (A.8a) and (A.8b) we note that

$$\hat{p}_1 = \hat{D}_1 \left( 1 - r \frac{1 + \vartheta_1\alpha_1 + \sigma_1\beta_1}{\alpha_1(1 + \vartheta_1)} \right), \quad (A.11)$$

and obtain the following relation if  $\hat{D}_1 > 0$ :

$$\hat{p}_1 > 0 \iff r < \frac{\alpha_1(1 + \vartheta_1)}{1 + \vartheta_1\alpha_1 + \sigma_1\beta_1}. \quad (A.12)$$

This is a contradiction to (A.10). Thus, the perturbation of  $F_0^{(2)}$  is not in the state space.

In the absence of migration, there exist eight single-locus polymorphisms. It can be shown that four single-locus polymorphisms are admissible under weak migration (one single-locus polymorphism at each marginal one-locus system). The remaining four single-locus polymorphisms leave the state space under small perturbations. The details are not shown here and are left to the interested reader (see Akerman and Bürger (2014) for the case of no dominance).

A.1.3. Completely dominant adaptive alleles

We analyze the special case when dominance coefficients satisfy

$$\vartheta_1 = \sigma_1 = -\vartheta_2 = -\sigma_2 = 1. \quad (A.13)$$

Then  $F_0$  (3.2) is not hyperbolic; cf. (3.3). Assuming  $r > 0$ , we obtain the approximation of  $F_m$  by performing a nonregular perturbation of  $F_0$  for weak migration. It is given by

$$\hat{p}_1 = 1 - \sqrt{m_1 \frac{1 + a_1 + b_1}{2a_1}} + \sqrt{m_1 m_2 \frac{(1 + a_1 + b_1)(1 - a_2 - b_2)}{-2a_1 a_2}} + \frac{m_1}{2r} \sqrt{\frac{b_1}{a_1}}, \quad (A.14a)$$

$$\hat{q}_1 = 1 - \sqrt{m_1 \frac{1 + a_1 + b_1}{2b_1}} + \sqrt{m_1 m_2 \frac{(1 + a_1 + b_1)(1 - a_2 - b_2)}{-2b_1 b_2}} + \frac{m_1}{2r} \sqrt{\frac{a_1}{b_1}}, \quad (A.14b)$$

$$\hat{D}_1 = \frac{m_1}{r}, \quad (A.14c)$$

$$\hat{p}_2 = \sqrt{m_2 \frac{1 - a_2 - b_2}{-2a_2}} + \sqrt{m_1 m_2 \frac{(1 + a_1 + b_1)(1 - a_2 - b_2)}{-2a_1 a_2}} - \frac{m_2}{2r} \sqrt{\frac{b_2}{a_2}}, \quad (A.14d)$$

$$\hat{q}_2 = \sqrt{m_2 \frac{1 - a_2 - b_2}{-2b_2}} + \sqrt{m_1 m_2 \frac{(1 + a_1 + b_1)(1 - a_2 - b_2)}{-2b_1 b_2}} - \frac{m_2}{2r} \sqrt{\frac{a_2}{b_2}}, \quad (A.14e)$$

$$\hat{D}_2 = \frac{m_2}{r}. \quad (A.14f)$$

Interestingly, the amount of LD in deme  $k$  is independent of selection to this order of approximation. Note that the expressions

for  $D_k$  in (A.14) coincide with those of  $D_k$  in (3.4) if (A.13) is applied there.

If linkage is complete ( $r = 0$ ),  $F_m$  is approximated by

$$\hat{p}_1 = \hat{q}_1 = 1 - \sqrt{m_1 \frac{1 + a_1 + b_1}{2(a_1 + b_1)}}, \quad \hat{D}_1 = \hat{p}_1(1 - \hat{p}_1), \quad (\text{A.15a})$$

$$\hat{p}_2 = \hat{q}_2 = \sqrt{m_2 \frac{1 - a_2 - b_2}{-2(a_2 + b_2)}}, \quad \hat{D}_2 = \hat{p}_2(1 - \hat{p}_2). \quad (\text{A.15b})$$

At this equilibrium,  $\hat{x}_{2,k} = \hat{x}_{3,k} = 0$  ( $k = 1, 2$ ) and strong LD is maintained.

Motivated by one-locus theory and the CI model (Appendix A.7), we expect, but have not proved, that  $F_m$  is the unique fully polymorphic equilibrium in this case.

### A.2. CI model: gamete dynamics

In Section 4, the dynamics in the continent–island model of the allele frequencies  $p, q$  and the measure of LD  $D$  is given in (4.2). For some derivations (e.g., in the proof of Theorem 4.2) it is more convenient to work with gamete frequencies  $x_i$  ( $i = 1, 2, 3, 4$ ). Using the transformation (2.10) (where we omit the subscript  $k$  for the deme, as we keep track of the island types only), the dynamics (4.2) can be rewritten as

$$\begin{aligned} \dot{x}_1 = \frac{dx_1}{dt} = & \alpha x_1(x_3 + x_4)[1 + \vartheta(1 - 2(x_1 + x_2))] \\ & + \beta x_1(x_2 + x_4)[1 + \sigma(1 - 2(x_1 + x_3))] \\ & - \rho D - \mu x_1, \end{aligned} \quad (\text{A.16a})$$

$$\begin{aligned} \dot{x}_2 = \frac{dx_2}{dt} = & \alpha x_2(x_3 + x_4)[1 + \vartheta(1 - 2(x_1 + x_2))] \\ & - \beta x_2(x_1 + x_3)[1 + \sigma(1 - 2(x_1 + x_3))] \\ & + \rho D - \mu x_2, \end{aligned} \quad (\text{A.16b})$$

$$\begin{aligned} \dot{x}_3 = \frac{dx_3}{dt} = & -\alpha x_3(x_1 + x_2)[1 + \vartheta(1 - 2(x_1 + x_2))] \\ & + \beta x_3(x_2 + x_4)[1 + \sigma(1 - 2(x_1 + x_3))] \\ & + \rho D - \mu x_3, \end{aligned} \quad (\text{A.16c})$$

$$\begin{aligned} \dot{x}_4 = \frac{dx_4}{dt} = & -\alpha x_4(x_1 + x_2)[1 + \vartheta(1 - 2(x_1 + x_2))] \\ & - \beta x_4(x_1 + x_3)[1 + \sigma(1 - 2(x_1 + x_3))] \\ & - \rho D + \mu(1 - x_4). \end{aligned} \quad (\text{A.16d})$$

This system of ordinary differential equations is defined on  $S_4$ . If  $\vartheta = \sigma = 0$ , it simplifies to that used in BA11.

### A.3. CI model: important quantities

In addition to the critical migration rates defined in (4.3), in the proofs of Theorems 4.1 and 4.2 we shall need

$$\mu_2^A = -\frac{\alpha(1 - \vartheta)^2}{8\vartheta}, \quad (\text{A.17a})$$

$$\mu_2^B = -\frac{\beta(1 - \sigma)^2}{8\sigma}. \quad (\text{A.17b})$$

Obviously, the critical migration rates satisfy:

$$\mu_1^A > 0 \quad \text{and} \quad \mu_1^B > 0 \quad \text{always}, \quad (\text{A.18a})$$

$$\mu_2^A > 0 \iff -1 < \vartheta < 0, \quad (\text{A.18b})$$

$$\mu_2^B > 0 \iff -1 < \sigma < 0, \quad (\text{A.18c})$$

$$\mu^M > 0 \iff 0 \leq \rho < \alpha(1 + \vartheta) + \beta(1 + \sigma). \quad (\text{A.18d})$$

Recall the definitions of  $\zeta_i$  ( $i = 1, 2, 3$ ) given in (4.14) and define

$$\zeta_4 = -\frac{(1 - \sigma)^2}{8\sigma(1 + \vartheta)}. \quad (\text{A.19})$$

The critical migration rates  $\mu_1^A, \mu_1^B, \mu_2^A$ , and  $\mu_2^B$  satisfy the following relations (which are independent of the recombination rate  $\rho$ ): For any  $-1 < \vartheta < 1$  and  $-1 < \sigma < 1$ ,

$$0 < \mu_1^A < \mu_1^B \iff 0 < \alpha < \zeta_1\beta, \quad (\text{A.20a})$$

$$0 < \mu_1^B < \mu_1^A \iff 0 < \zeta_1\beta < \alpha. \quad (\text{A.20b})$$

If  $-1 < \vartheta < -\frac{1}{3}$  and  $-\frac{1}{3} \leq \sigma < 1$  hold, the following relations hold:

$$0 < \mu_1^A < \mu_2^A < \mu_1^B \iff \alpha < \zeta_3\beta, \quad (\text{A.20c})$$

$$0 < \mu_1^A < \mu_1^B < \mu_2^A \iff \zeta_3\beta < \alpha < \zeta_1\beta, \quad (\text{A.20d})$$

$$0 < \mu_1^B < \mu_1^A < \mu_2^A \iff \zeta_1\beta < \alpha. \quad (\text{A.20e})$$

If  $-\frac{1}{3} \leq \vartheta < 1$  and  $-1 < \sigma < -\frac{1}{3}$  hold, the following relations hold:

$$0 < \mu_1^A < \mu_1^B < \mu_2^B \iff \alpha < \zeta_1\beta, \quad (\text{A.20f})$$

$$0 < \mu_1^B < \mu_1^A < \mu_2^B \iff \zeta_1\beta < \alpha < \zeta_4\beta, \quad (\text{A.20g})$$

$$0 < \mu_1^B < \mu_2^B < \mu_1^A \iff \zeta_4\beta < \alpha. \quad (\text{A.20h})$$

If  $-1 < \vartheta < -\frac{1}{3}$  and  $-1 < \sigma < -\frac{1}{3}$  hold, the following relations hold:

$$0 < \mu_1^A < \mu_2^A < \mu_1^B < \mu_2^B \iff \alpha < \zeta_3\beta, \quad (\text{A.20i})$$

$$0 < \mu_1^A < \mu_1^B < \mu_2^A < \mu_2^B \iff \text{either } \sigma \leq \vartheta \quad \text{and} \\ \zeta_3\beta < \alpha < \zeta_1\beta, \quad \text{or } \vartheta < \sigma \quad \text{and} \quad \zeta_3\beta < \alpha < \zeta_2\beta, \quad (\text{A.20j})$$

$$0 < \mu_1^A < \mu_1^B < \mu_2^B < \mu_2^A \iff \vartheta < \sigma \quad \text{and} \\ \zeta_2\beta < \alpha < \zeta_1\beta, \quad (\text{A.20k})$$

$$0 < \mu_1^B < \mu_2^B < \mu_1^A < \mu_2^A \iff \zeta_4\beta < \alpha, \quad (\text{A.20l})$$

$$0 < \mu_1^B < \mu_1^A < \mu_2^B < \mu_2^A \iff \text{either } \sigma \leq \vartheta \quad \text{and} \\ \zeta_2\beta < \alpha < \zeta_4\beta, \quad \text{or } \vartheta < \sigma \quad \text{and} \quad \zeta_1\beta < \alpha < \zeta_4\beta. \quad (\text{A.20m})$$

The critical migration rates  $\mu_1^A, \mu_1^B$ , and  $\mu^M$  satisfy the following relations (which depend on the recombination rate  $\rho$ ):

$$0 < \mu^M < \mu_1^A < \mu_1^B \iff 0 < \alpha < \zeta_1\beta \quad \text{and} \\ \beta(1 + \sigma) < \rho < \alpha(1 + \vartheta) + \beta(1 + \sigma), \quad (\text{A.21a})$$

$$0 < \mu_1^A < \mu^M < \mu_1^B \iff 0 < \alpha < \zeta_1\beta \quad \text{and} \\ \alpha(1 + \vartheta) < \rho < \beta(1 + \sigma), \quad (\text{A.21b})$$

$$0 < \mu_1^A < \mu_1^B < \mu^M \iff 0 < \alpha < \zeta_1\beta \quad \text{and} \\ 0 \leq \rho < \alpha(1 + \vartheta), \quad (\text{A.21c})$$

$$0 < \mu^M < \mu_1^B < \mu_1^A \iff 0 < \zeta_1\beta < \alpha \quad \text{and} \\ \alpha(1 + \vartheta) < \rho < \alpha(1 + \vartheta) + \beta(1 + \sigma), \quad (\text{A.21d})$$

$$0 < \mu_1^B < \mu^M < \mu_1^A \iff 0 < \zeta_1\beta < \alpha \quad \text{and} \\ \beta(1 + \sigma) < \rho < \alpha(1 + \vartheta), \quad (\text{A.21e})$$

$$0 < \mu_1^B < \mu_1^A < \mu^M \iff 0 < \zeta_1\beta < \alpha \quad \text{and} \\ 0 \leq \rho < \beta(1 + \sigma). \quad (\text{A.21f})$$

### A.4. CI model: admissibility of the boundary equilibria (4.8)

$E_{A,1}$  is admissible if and only if one of the following two conditions holds

$$-\frac{1}{3} \leq \vartheta < 1 \quad \text{and} \quad 0 \leq \mu \leq \mu_1^A, \quad \text{or} \quad (\text{A.22a})$$

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad 0 \leq \mu \leq \mu_2^A. \quad (\text{A.22b})$$

$E_{A,2}$  is admissible if and only if

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad \mu_1^A \leq \mu \leq \mu_2^A \quad (\text{A.23})$$

holds, where

$$E_{A,2} = M \iff \mu = \mu_1^A. \quad (\text{A.24})$$

If (A.23) holds,  $E_{A,1}$  and  $E_{A,2}$  are simultaneously admissible. For increasing migration  $E_{A,1}$  and  $E_{A,2}$  leave the state space according to the following relations:

$$-\frac{1}{3} \leq \vartheta < 1 \quad \text{and} \quad \mu \uparrow \mu_1^A \iff E_{A,1} \rightarrow M, \quad (\text{A.25a})$$

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad \mu \uparrow \mu_2^A \iff E_{A,1} \rightarrow E_{A,2}. \quad (\text{A.25b})$$

$E_{B,1}$  is admissible if and only if one of the following two conditions holds

$$-\frac{1}{3} \leq \sigma < 1 \quad \text{and} \quad 0 \leq \mu \leq \mu_1^B, \quad \text{or} \quad (\text{A.26a})$$

$$-1 < \sigma < -\frac{1}{3} \quad \text{and} \quad 0 \leq \mu \leq \mu_2^B. \quad (\text{A.26b})$$

$E_{B,2}$  is admissible if and only if

$$-1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu_1^B \leq \mu \leq \mu_2^B \quad (\text{A.27})$$

holds, where

$$E_{B,2} = M \iff \mu = \mu_1^B. \quad (\text{A.28})$$

$E_{B,1}$  and  $E_{B,2}$  leave the state space according to the following relations:

$$-\frac{1}{3} \leq \sigma < 1 \quad \text{and} \quad \mu \uparrow \mu_1^B \iff E_{B,1} \rightarrow M, \quad (\text{A.29a})$$

$$-1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu \uparrow \mu_2^B \iff E_{B,1} \rightarrow E_{B,2}. \quad (\text{A.29b})$$

#### A.5. CI model: weak migration

If  $\mu = 0$  in (4.2), in addition to the globally asymptotically stable monomorphism  $\hat{p} = \hat{q} = 1, \hat{D} = 0$  (in the following denoted  $M^{(1)}$ ) and the unstable monomorphism  $M$  (4.4), there exist two unstable monomorphic equilibria given by

$$M^{(2)}: \hat{p} = 1, \quad \hat{q} = 0, \quad \hat{D} = 0, \quad \text{and} \quad (\text{A.30})$$

$$M^{(3)}: \hat{p} = 0, \quad \hat{q} = 1, \quad \hat{D} = 0. \quad (\text{A.31})$$

In the absence of migration, the eigenvalues at  $M^{(1)}$  are

$$\begin{aligned} &-\alpha(1 - \vartheta), \quad -\beta(1 - \sigma), \\ &-\rho - \alpha(1 - \vartheta) - \beta(1 - \sigma). \end{aligned} \quad (\text{A.32})$$

In this case,  $M^{(1)}$  is not hyperbolic if and only if  $\vartheta = 1$ , or  $\sigma = 1$ , or  $\rho = 0$  and  $\vartheta = \sigma = 1$ . The eigenvalues at  $M^{(2)}$  if  $\mu = 0$  are

$$\begin{aligned} &-\alpha(1 - \vartheta), \quad \beta(1 + \sigma), \\ &-\rho - \alpha(1 - \vartheta) + \beta(1 + \sigma), \end{aligned} \quad (\text{A.33})$$

where  $M^{(2)}$  is not hyperbolic if and only if  $\vartheta = 1$ , or  $\sigma = -1$ , or  $\rho = -\alpha(1 - \vartheta) + \beta(1 + \sigma)$ . The eigenvalues at  $M^{(3)}$  if  $\mu = 0$  are

$$\begin{aligned} &\alpha(1 + \vartheta), \quad -\beta(1 - \sigma), \\ &-\rho + \alpha(1 + \vartheta) - \beta(1 - \sigma), \end{aligned} \quad (\text{A.34})$$

such that  $M^{(3)}$  is not hyperbolic if and only if  $\vartheta = -1$ , or  $\sigma = 1$ , or  $\rho = \alpha(1 + \vartheta) - \beta(1 - \sigma)$ . The eigenvalues at  $M$  if  $\mu = 0$  are

$$\begin{aligned} &\alpha(1 + \vartheta), \quad \beta(1 + \sigma), \\ &-\rho + \alpha(1 + \vartheta) + \beta(1 + \sigma), \end{aligned} \quad (\text{A.35})$$

such that  $M$  is not hyperbolic if and only if  $\vartheta = -1$ , or  $\sigma = -1$ , or  $\rho = \alpha(1 + \vartheta) + \beta(1 + \sigma)$ .

If dominance coefficients satisfy  $-1 < \vartheta < 1$  and  $-1 < \sigma < 1$ , generically,  $F_\mu$  (4.10) is the only full polymorphism entering the state space under weak migration. In the nongeneric cases where recombination is such that a monomorphism is not hyperbolic, an additional full polymorphism may enter the state space under weak migration.

Now we treat the remaining cases where  $|\vartheta| = 1$  or  $|\sigma| = 1$ . If only one locus is completely recessive, in addition to  $F_\mu$  (4.10) up to two full polymorphisms may enter the state space for weak migration: one via  $M$  and either one via  $M^{(3)}$  (if  $\vartheta = -1$ ) or one via  $M^{(2)}$  (if  $\sigma = -1$ ). If both island alleles are recessive ( $\vartheta = \sigma = -1$ ), in addition to  $F_\mu$  up to three additional full polymorphism may enter the state space under weak migration. If one locus is completely dominant, in addition to  $F_\mu$  (approximated by (A.37) if  $\rho > 0$  or given by (4.26) if  $\rho = 0$ , respectively) at most one additional full polymorphisms may enter the state space (if  $\vartheta = 1$  or  $\sigma = 1$  it is a perturbation of  $M^{(2)}$  or  $M^{(3)}$ , respectively). The case  $\vartheta = \sigma = 1$  is treated in Appendix A.7.

#### A.6. CI model: strong migration

We prove global convergence to  $M$  if  $\mu > 2(\alpha + \beta)$ . The proof is similar to that for the continuous-time CI model with genic selection (in Section 3.4.3 of BA11). There it was shown for  $\vartheta = \sigma = 0$  that  $M$  is globally asymptotically stable if  $\mu > \alpha + \beta$ .

Because  $0 \leq 1 + \vartheta(1 - 2p) \leq 2$  or  $0 \leq 1 + \sigma(1 - 2q) \leq 2$  hold by  $-1 \leq \vartheta \leq 1$  or  $-1 \leq \sigma \leq 1$ , respectively, and as  $D \leq p(1 - q)$ , we obtain from (4.2) that

$$\dot{p} \leq p[2(\alpha + \beta) - \mu - 2(\alpha p + \beta q)] \leq 0 \quad (\text{A.36})$$

if  $\mu > 2(\alpha + \beta)$  with equality if and only if  $\hat{p} = 0$ . A similar argument using  $D \leq q(1 - p)$  follows for  $\dot{q}$ . Thus, there exist two global Lyapunov functions, and all trajectories must converge to  $M$  (LaSalle, 1976). If the island alleles are completely dominant ( $\vartheta = \sigma = 1$ ) and recombination is absent ( $\rho = 0$ ), the bound  $2(\alpha + \beta)$  is best possible, because it is equal to  $\mu^M$ ; cf. (4.6), (A.21c) and (A.21f).

#### A.7. CI model: completely dominant island alleles

We briefly analyze the CI model with completely dominant island alleles, i.e.,  $\vartheta = \sigma = 1$ . Because the case  $\rho = 0$  is included in Section 4.5, we assume  $\rho > 0$ .

In Section 4.2, we derived an approximation of the full polymorphism  $F_\mu$  under weak migration under the condition that both  $1 \leq \vartheta < 1$  and  $1 \leq \sigma < 1$  hold. Note that if  $\vartheta = \sigma = 1$ , the denominators in (4.10) are zero, thus the approximation is not admissible in this case. Assuming  $\vartheta = \sigma = 1$ , a nonregular perturbation of the globally asymptotically stable equilibrium under no migration, ( $\hat{p} = \hat{q} = 1$  and  $\hat{D} = 0$ ), gives the following weak-migration approximation of the equilibrium frequencies:

$$\begin{aligned} \hat{p} = &1 - \sqrt{\frac{\mu}{2\alpha}} + \frac{\mu}{2\rho} \sqrt{\frac{\beta}{\alpha}} \\ &- \frac{\mu}{4\rho^2} \sqrt{\frac{\mu}{2\alpha}} (3\beta + 2\rho + 6\sqrt{\alpha\beta}) + \mathcal{O}(\mu^2), \end{aligned} \quad (\text{A.37a})$$

$$\hat{q} = 1 - \sqrt{\frac{\mu}{2\beta}} + \frac{\mu}{2\rho} \sqrt{\frac{\alpha}{\beta}} - \frac{\mu}{4\rho^2} \sqrt{\frac{\mu}{2\beta}} (3\alpha + 2\rho + 6\sqrt{\alpha\beta}) + \mathcal{O}(\mu^2), \quad (\text{A.37b})$$

$$\hat{D} = \frac{\mu}{\rho} - \frac{\mu}{\rho^2} \sqrt{\frac{\mu}{2\alpha\beta}} (\sqrt{\alpha} + \sqrt{\beta})(\rho + 2\sqrt{\alpha\beta}) + \mathcal{O}(\mu^2), \quad (\text{A.37c})$$

(higher order terms are needed in several applications of Section 5.3).

It is not difficult to show that no other internal equilibrium is generated by a bifurcation of a boundary equilibrium if migration is turned on. The (unstable) monomorphic equilibria at which  $x_2 = 1$  or  $x_3 = 1$  become (unstable) single-locus polymorphisms if  $\mu > 0$ , and the equilibrium  $x_4 = 1$  (fixation of the continental type) remains unchanged.

#### A.8. CI model: proof of Theorem 4.1

Because the dynamics at locus A and locus B in (4.11) are decoupled, the results in Theorem 4.1 can be derived straightforwardly from the respective one-locus result of Nagylaki (1975). For sufficiently weak migration, only  $F_1$  (4.12a) is admissible. Up to first order in  $\mu$ , it is given by (4.10).

The admissibility conditions of the other internal equilibria,  $F_i$  (4.12), are

$$F_2 \text{ is admissible if and only if (A.22) and (A.27) hold,} \quad (\text{A.38a})$$

$$F_3 \text{ is admissible if and only if (A.23) and (A.26) hold,} \quad (\text{A.38b})$$

$$F_4 \text{ is admissible if and only if (A.23) and (A.27) hold.} \quad (\text{A.38c})$$

They may enter the state space at higher migration rates according to the following conditions:

$F_2 = E_{A,1}$  if and only if one of the following two conditions holds:

$$-\frac{1}{3} \leq \vartheta < 1 \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu = \mu_1^B < \mu_1^A, \quad \text{or} \quad (\text{A.39a})$$

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu = \mu_1^B < \mu_2^A. \quad (\text{A.39b})$$

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu = \mu_1^B < \mu_2^A. \quad (\text{A.39b})$$

$$\mu = \mu_1^B < \mu_2^A. \quad (\text{A.39b})$$

$F_3 = E_{B,1}$  if and only if one of the following two conditions holds:

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -\frac{1}{3} \leq \sigma < 1 \quad \text{and} \quad \mu = \mu_1^A < \mu_1^B, \quad \text{or} \quad (\text{A.39c})$$

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu = \mu_1^A < \mu_2^B. \quad (\text{A.39d})$$

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu = \mu_1^A < \mu_2^B. \quad (\text{A.39d})$$

$$\mu = \mu_1^A < \mu_2^B. \quad (\text{A.39d})$$

$F_4 = E_{A,2}$  if and only if

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu_1^A < \mu_1^B < \mu_2^A \quad \text{and} \quad \mu = \mu_1^B. \quad (\text{A.39e})$$

$$\mu_1^A < \mu_1^B < \mu_2^A \quad \text{and} \quad \mu = \mu_1^B. \quad (\text{A.39e})$$

$F_4 = E_{B,2}$  if and only if

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu_1^B < \mu_1^A < \mu_2^B \quad \text{and} \quad \mu = \mu_1^A. \quad (\text{A.39f})$$

$$\mu_1^B < \mu_1^A < \mu_2^B \quad \text{and} \quad \mu = \mu_1^A. \quad (\text{A.39f})$$

$F_4 = M$  if and only if

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu = \mu_1^A = \mu_1^B. \quad (\text{A.39g})$$

$$\text{and} \quad \mu = \mu_1^A = \mu_1^B. \quad (\text{A.39g})$$

The following relations describe when  $F_1$  leaves the state space via collision with a boundary equilibrium:

$F_1 \rightarrow E_{A,1}$  if and only if one of the following two cases holds:

$$-\frac{1}{3} \leq \vartheta < 1 \quad \text{and} \quad -\frac{1}{3} \leq \sigma < 1 \quad \text{and} \quad \mu_1^B < \mu_1^A \quad \text{and} \quad \mu \uparrow \mu_1^B, \quad \text{or} \quad (\text{A.40a})$$

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -\frac{1}{3} \leq \sigma < 1 \quad \text{and} \quad \mu_1^B < \mu_2^A \quad \text{and} \quad \mu \uparrow \mu_1^B. \quad (\text{A.40b})$$

$F_1 \rightarrow E_{B,1}$  holds if and only if one of the following two cases holds:

$$-\frac{1}{3} \leq \vartheta < 1 \quad \text{and} \quad -\frac{1}{3} \leq \sigma < 1 \quad \text{and} \quad \mu_1^A < \mu_1^B \quad \text{and} \quad \mu \uparrow \mu_1^A, \quad \text{or} \quad (\text{A.40c})$$

$$-\frac{1}{3} \leq \vartheta < 1 \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu_1^A < \mu_2^B \quad \text{and} \quad \mu \uparrow \mu_1^A. \quad (\text{A.40d})$$

$F_1 \rightarrow M$  if and only if one of the following conditions holds:

$$-\frac{1}{3} \leq \vartheta < 1 \quad \text{and} \quad -\frac{1}{3} \leq \sigma < 1 \quad \text{and} \quad \mu \uparrow \mu_1^A = \mu_1^B, \quad \text{or} \quad (\text{A.40e})$$

$$-\frac{1}{3} \leq \vartheta < 1 \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu \uparrow \mu_1^A = \mu_2^B, \quad \text{or} \quad (\text{A.40f})$$

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -\frac{1}{3} \leq \sigma < 1 \quad \text{and} \quad \mu \uparrow \mu_2^A = \mu_1^B. \quad (\text{A.40g})$$

$$\text{and} \quad \mu \uparrow \mu_2^A = \mu_1^B. \quad (\text{A.40g})$$

$F_1$  never leaves the state space via  $E_{A,2}$  or  $E_{B,2}$ .

Also,  $F_2$  and  $F_3$  may leave the state space via collision with a boundary equilibrium. This occurs in the following cases:

$F_2 \rightarrow E_{B,2}$  if and only if

$$-\frac{1}{3} \leq \vartheta < 1 \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu_1^B < \mu_1^A < \mu_2^B \quad \text{and} \quad \mu \uparrow \mu_1^A. \quad (\text{A.41a})$$

$$\text{and} \quad \mu_1^B < \mu_1^A < \mu_2^B \quad \text{and} \quad \mu \uparrow \mu_1^A. \quad (\text{A.41a})$$

$F_3 \rightarrow E_{A,2}$  if and only if

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -\frac{1}{3} \leq \sigma < 1 \quad \text{and} \quad \mu_1^A < \mu_1^B < \mu_2^A \quad \text{and} \quad \mu \uparrow \mu_1^B. \quad (\text{A.41b})$$

$$\text{and} \quad \mu_1^A < \mu_1^B < \mu_2^A \quad \text{and} \quad \mu \uparrow \mu_1^B. \quad (\text{A.41b})$$

The following relations describe when internal equilibria collide and annihilate each other:

$F_1 \rightarrow F_2$  if and only if one of the following two conditions holds:

$$-\frac{1}{3} \leq \vartheta < 1 \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu_2^B < \mu_1^A \quad \text{and} \quad \mu \uparrow \mu_2^B, \quad \text{or} \quad (\text{A.42a})$$

$$\text{and} \quad \mu_2^B < \mu_1^A \quad \text{and} \quad \mu \uparrow \mu_2^B, \quad \text{or} \quad (\text{A.42a})$$

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu_2^B < \mu_2^A \quad \text{and} \quad \mu \uparrow \mu_2^B. \quad (\text{A.42b})$$

$$\text{and} \quad \mu_2^B < \mu_2^A \quad \text{and} \quad \mu \uparrow \mu_2^B. \quad (\text{A.42b})$$

$F_1 \rightarrow F_3$  if and only if one of the following two conditions holds:

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -\frac{1}{3} \leq \sigma < 1 \quad \text{and} \quad \mu_2^A < \mu_1^B \quad \text{and} \quad \mu \uparrow \mu_2^A, \quad \text{or} \quad (\text{A.42c})$$

$$\text{and} \quad \mu \uparrow \mu_2^A, \quad \text{or} \quad (\text{A.42c})$$

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu_2^A < \mu_2^B \quad (\text{A.42d})$$

and  $\mu \uparrow \mu_2^A$ .

$F_4 \rightarrow F_2$  if and only if

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \max\{\mu_1^A, \mu_1^B\} < \mu_2^A < \mu_2^B \quad \text{and} \quad \mu \uparrow \mu_2^A. \quad (\text{A.42e})$$

$F_4 \rightarrow F_3$  if and only if

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad -1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \max\{\mu_1^A, \mu_1^B\} < \mu_2^B < \mu_2^A \quad \text{and} \quad \mu \uparrow \mu_2^B. \quad (\text{A.42f})$$

If  $-1 < \vartheta < -\frac{1}{3}$  and  $-1 < \sigma < -\frac{1}{3}$  and  $\mu_2^A = \mu_2^B$  hold, all internal equilibria  $F_i$  ( $i = 1, 2, 3, 4$ ) collide and annihilate each other at  $m \uparrow \mu_2^A = \mu_2^B$ .

The stability of the equilibria of the full system (4.11) is obtained as a Cartesian product of the stability of the equilibria in the two single-locus system derived in Nagylaki (1975). Whenever equilibria are admissible, they have real eigenvalues. As already stated below (4.13),  $F_1$  is asymptotically stable whenever admissible, and the remaining full polymorphisms  $F_i$  ( $i = 2, 3, 4$ ) are always unstable.  $E_{A,1}$  is asymptotically stable if and only if one of the following relations holds:

$$-\frac{1}{3} \leq \vartheta < 1 \quad \text{and} \quad \mu_1^B < \mu < \mu_1^A, \quad \text{or} \quad (\text{A.43a})$$

$$-1 < \vartheta < -\frac{1}{3} \quad \text{and} \quad \mu_1^B < \mu < \mu_2^A, \quad (\text{A.43b})$$

where  $-1 < \sigma < 1$  is arbitrary. Analogously,  $E_{B,1}$  is asymptotically stable if and only if one of the following relations holds:

$$-\frac{1}{3} \leq \sigma < 1 \quad \text{and} \quad \mu_1^A < \mu < \mu_1^B, \quad \text{or} \quad (\text{A.44a})$$

$$-1 < \sigma < -\frac{1}{3} \quad \text{and} \quad \mu_1^A < \mu < \mu_2^B, \quad (\text{A.44b})$$

where  $-1 < \vartheta < 1$  is arbitrary. As  $E_{A,2}$  and  $E_{B,2}$  arise from the Cartesian product of unstable equilibria in the one-locus model, they are always unstable within the full dynamics (4.11). From (4.6) and (A.18d) it follows that M is asymptotically stable if and only if

$$\mu > \max\{\mu_1^A, \mu_1^B\}. \quad (\text{A.45})$$

Now we have all ingredients to finish the proof. In the following, we assume w.l.o.g. that (4.15) holds. Then, by (A.20a) we need to consider the case  $\mu_1^A < \mu_1^B$  only. Under the assumption of (4.15), by (A.45) and (A.20a) it follows that

$$M \text{ is asymptotically stable} \iff \mu > \mu_1^B. \quad (\text{A.46})$$

Assume  $-\frac{1}{3} < \vartheta < 1$  and  $-\frac{1}{3} < \sigma < 1$  hold.  $E_{A,1}$  and  $E_{B,1}$  are admissible if and only if (A.22a) and (A.26a) holds, respectively. By (A.23) and (A.27),  $E_{A,2}$  and  $E_{B,2}$  are not admissible. By (A.38),  $F_1$  is the only full polymorphism that may be admissible in this case. Applying (4.15) and (A.20a) to (A.25a) and (A.40c) it follows that  $F_1 \rightarrow E_{B,1}$  and  $E_{A,1} \rightarrow M$  if  $\mu \uparrow \mu_1^A$ . By (A.44a), at collision,  $F_1$  exchanges stability with  $E_{B,1}$ .  $E_{B,1}$  leaves the state space via M if  $\mu \uparrow \mu_1^B$ ; cf. (A.29a). By (A.46), it exchanges stability with M. This gives (4.16) in case (i) of Theorem 4.1 and Fig. 1(a).

The remaining cases of Theorem 4.1 can be derived similarly. They follow by combining the relations given in (A.20) with results given in the proof above.

### A.9. The dynamics at a linked neutral locus

In Section 5.3, we derive the effective migration rate at a neutral locus N linked to the loci under selection. Recombination between locus A (B) and N occurs with rate  $\rho_1$  ( $\rho_2$ ), such that  $\rho = \rho_1 + \rho_2$  is the recombination rate between locus A and B. Evolution at the three loci is described by a system of seven ordinary differential equations for the allele frequencies and linkage disequilibria ( $p, q, D_{AB}, n, D_{AN}, D_{NB}, D_{ANB}$ ). The equations for the change of  $p, q$ , and  $D_{AB}$  are given by (4.2). The allele frequency of  $N_1$  at the neutral locus evolves according to

$$\dot{n} = \alpha(1 + \vartheta(1 - 2p))D_{AN} + \beta(1 + \sigma(1 - 2q))D_{NB} + \mu(n_c - n), \quad (\text{A.47a})$$

where  $n_c$  denotes the frequency of allele  $N_1$  within immigrants. The rates of change of LD involving the neutral locus are given by

$$\dot{D}_{AN} = -\rho_1 D_{AN} + \alpha(1 - 2p)(1 + \vartheta(1 - 2p))D_{AN} + \beta(1 + \sigma(1 - 2q))D_{ANB} - \mu p(n_c - n), \quad (\text{A.47b})$$

$$\dot{D}_{NB} = -\rho_2 D_{NB} + \beta(1 - 2q)(1 + \sigma(1 - 2q))D_{NB} + \alpha(1 + \vartheta(1 - 2q))D_{ANB} - \mu q(n_c - n), \quad (\text{A.47c})$$

$$\begin{aligned} \dot{D}_{ANB} = & -\rho D_{ANB} + [\alpha(1 - 2p)(1 - \vartheta(1 - 2p)) \\ & + \beta(1 - 2q)(1 + \sigma(1 - 2q))]D_{ANB} \\ & - 2[\alpha(1 + \vartheta(1 - 2p))D_{AN} \\ & + \beta(1 + \sigma(1 - 2q))D_{NB}]D_{AB} \\ & - \mu(D_{ANB} - pD_{NB} - qD_{AB}) \\ & + \mu(pq - D_{AB})(n_c - n). \end{aligned} \quad (\text{A.47d})$$

Genic selection is obtained if  $\vartheta = \sigma = 0$ , a case studied in BA11. The equations describing evolution at the neutral locus therein, i.e., Eqs. (4.25), (4.26) and (4.28), contain errors. Instead of their Eqs. (4.25) and (4.26), the correct dynamics are obtained from (A.47) with  $\vartheta = \sigma = 0$ . We point out that formula (4.30) in BA11 for the effective migration rate is not affected by these errors.

### References

Aeschbacher, S., Bürger, R., 2014. The effect of linkage on establishment and survival of locally beneficial mutations. *Genetics* 197, <http://dx.doi.org/10.1534/genetics.114.163477>.

Akerman, A., Bürger, R., 2014. The consequences of gene flow for local adaptation and differentiation: a two-locus two-deme model. *J. Math. Biol.* 68, 1135–1198.

Avery, P.J., Hill, W.G., 1978. The effect of linkage disequilibrium on the genetic variance of a quantitative trait. *Adv. Appl. Probab.* 10, 4–6.

Bank, C., Bürger, R., Hermisson, J., 2012. The limits to parapatric speciation: Dobzhansky–Muller incompatibilities in a continent–island model. *Genetics* 191, 845–865.

Barton, N.H., 1983. Multilocus clines. *Evolution* 37, 454–471.

Barton, N.H., 2010. What role does natural selection play in speciation? *Phil. Trans. R. Soc. B* 365, 1825–1840.

Barton, N.H., Bengtsson, B.O., 1986. The barrier to genetic exchange between hybridising populations. *Heredity* 56, 357–376.

Bengtsson, B.O., 1985. The flow of genes through a genetic barrier. In: Greenwood, J.J., Harvey, P.H., Slatkin, M. (Eds.), *Evolution Essays in Honour of John Maynard Smith*. University Press, Cambridge, pp. 31–42.

Bulmer, M.G., 1972. Multiple Niche polymorphism. *Amer. Natur.* 106, 254–257.

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

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

Bürger, R., 2009b. Multilocus selection in subdivided populations II. Maintenance of polymorphism under weak or strong migration. *J. Math. Biol.* 58, 979–997.

Bürger, R., 2009c. Polymorphism in the two-locus Levene model with nonepistatic directional selection. *Theor. Popul. Biol.* 76, 214–228.

Bürger, R., 2010. Evolution and polymorphism in the multilocus Levene model with no or weak epistasis. *Theor. Popul. Biol.* 78, 123–138.

Bürger, R., 2014. A survey of migration–selection models in population genetics. *Discrete Contin. Dyn. Syst. B* 19, 883–959.

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

Charlesworth, B., Nordborg, M., Charlesworth, D., 1997. The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet. Res.* 70, 155–174.

- Chasnov, J.R., 2012. Equilibrium properties of a multi-locus, haploid-selection, symmetric-viability model. *Theor. Popul. Biol.* 81, 119–130.
- Crow, J.F., Kimura, M., 1970. *An Introduction to Population Genetics Theory*. Harper and Row, New York (Reprinted by The Blackburn Press, Caldwell, NJ, 2009).
- Ewens, W.J., 1969. Mean fitness increases when fitnesses are additive. *Nature* 221, 1076.
- Goudet, J., Martin, G., 2007. Under neutrality,  $Q_{ST} \leq F_{ST}$  when there is dominance in an island model. *Genetics* 176, 1371–1374.
- Karlin, S., 1977. Gene frequency patterns in the Levene subdivided population model. *Theor. Popul. Biol.* 11, 356–385.
- Karlin, S., 1982. Classification of selection-migration structures and conditions for a protected polymorphism. *Evol. Biol.* 14, 61–204.
- Karlin, S., Campbell, R.B., 1980. Selection-migration regimes characterized by a globally stable equilibrium. *Genetics* 94, 1065–1084.
- Karlin, S., Liberman, U., 1978. The two-locus multi-allele additive viability model. *J. Math. Biol.* 5, 201–211.
- Karlin, S., McGregor, J., 1972. Polymorphism for genetic and ecological systems with weak coupling. *Theor. Popul. Biol.* 3, 210–238.
- Kawecki, T.J., Ebert, D., 2004. Conceptual issues in local adaptation. *Ecol. Lett.* 7, 1225–1241.
- Kobayashi, Y., Hammerstein, P., Telschow, A., 2008. The neutral effective migration rate in a mainland-island context. *Theor. Popul. Biol.* 74, 84–92.
- Lande, R., 1992. Neutral theory of quantitative genetic variance in an island model with local extinction and colonization. *Evolution* 46, 381–389.
- LaSalle, J.P., 1976. *The Stability of Dynamics Systems*. SIAM, Philadelphia.
- Lenormand, T., 2002. Gene flow and the limits to natural selection. *Trends Ecol. Evol.* 17, 183–189.
- Leviyang, S., Hamilton, M.B., 2011. Properties of Weir and Cockerham's  $F_{ST}$  estimators and associated bootstrap confidence intervals. *Theor. Popul. Biol.* 79, 39–52.
- Li, W.-H., Nei, M., 1974. Stable linkage disequilibrium without epistasis in subdivided populations. *Theor. Popul. Biol.* 6, 173–183.
- Maynard Smith, J., 1970. Genetic polymorphism in a varied environment. *Amer. Natur.* 104, 487–490.
- Nagylaki, T., 1975. Conditions for the existence of clines. *Genetics* 80, 595–615.
- Nagylaki, T., 1992. *Introduction to Theoretical Population Genetics*. Springer, Berlin Heidelberg.
- Nagylaki, T., 1998. Fixation indices in subdivided populations. *Genetics* 148, 1325–1332.
- Nagylaki, T., 2009a. Polymorphism in multiallelic migration-selection models with dominance. *Theor. Popul. Biol.* 75, 239–259.
- Nagylaki, T., 2009b. Evolution under the multilocus Levene model. *Theor. Popul. Biol.* 76, 197–213.
- Nagylaki, T., Lou, Y., 2001. Patterns of multiallelic polymorphism maintained by migration and selection. *Theor. Popul. Biol.* 59, 297–313.
- Nagylaki, T., Lou, Y., 2007. Evolution under multiallelic migration-selection models. *Theor. Popul. Biol.* 72, 21–40.
- Nagylaki, T., Lou, Y., 2008. The dynamics of migration-selection models. In: Friedman, A. (Ed.), *Tutorials in Mathematical Biosciences IV*. In: *Lect. Notes Math.*, 1922. Springer, Berlin Heidelberg New York, pp. 119–172.
- Nordborg, M., Tavaré, S., 2002. Linkage disequilibrium: what history has to tell us. *Trends Genet.* 18, 83–90.
- Novak, S., 2011. The number of equilibria in the diallelic Levene model with multiple demes. *Theor. Popul. Biol.* 79, 97–101.
- Peischl, S., 2010. Dominance and the maintenance of polymorphism in multiallelic migration-selection models with two demes. *Theor. Popul. Biol.* 78, 12–25.
- Petry, D., 1983. The effect on neutral gene flow of selection at a linked locus. *Theor. Popul. Biol.* 23, 300–313.
- Prout, T., 1968. Sufficient conditions for multiple niche polymorphism. *Amer. Natur.* 102, 493–496.
- Slatkin, M., 2008. Linkage disequilibrium – understanding the evolutionary past and mapping the medical future. *Nat. Rev. Genet.* 9, 477–485.
- Spitze, K., 1993. Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics* 135, 367–374.
- Weir, B., Cockerham, C., 1984. Estimating  $F$  statistics for the analysis of population structure. *Evolution* 38, 1358–1370.
- Whitlock, M.C., 1999. Neutral additive genetic variance in a metapopulation. *Genet. Res.* 74, 215–221.
- Whitlock, M.C., 2008. Evolutionary inference from  $Q_{ST}$ . *Molecular Ecology* 17, 1885–1896.
- Wolfram, S., 2010. *Mathematica*, 8. Wolfram Research Inc., Champaign, IL.