# Multilocus selection in subdivided populations II. Maintenance of polymorphism under weak or strong migration

Reinhard Bürger

Department of Mathematics, University of Vienna, Austria

**Address for correspondence:**

Reinhard Bürger
Institut für Mathematik
Universität Wien
Nordbergstrasse 15
A-1090 Wien
Austria
E-mail: reinhard.buerger@univie.ac.at

# Abstract

The potential of maintaining multilocus polymorphism by migration-selection balance is studied. A large population of diploid individuals is distributed over finitely many demes connected by migration. Generations are discrete and nonoverlapping, selection may vary across demes, and loci are multiallelic. It is shown that if migration and recombination are strong relative to selection, then with weak or no epistasis and intermediate dominance at every locus and in every deme, arbitrarily many alleles can be maintained at arbitrarily many loci at a stable equilibrium. If migration is weak relative to selection and recombination, then with weak or no epistasis and intermediate dominance at every locus and in every deme, as many alleles as there are demes can be maintained at arbitrarily many loci at equilibrium. In both cases open sets of such parameter combinations are constructed, thus the results are robust with respect to small, but arbitrary, perturbations in the parameters. For weak migration, the number of demes is, in fact, a generic upper bound to the number of alleles that can be maintained at any locus. Thus, several scenarios are identified under which multilocus polymorphism can be maintained by migration-selection balance when this is impossible in a panmictic population.



**Key words:** selection, migration, recombination, multilocus polymorphism, multiple alleles, weak epistasis, intermediate dominance, dispersal, genetic variation

**2000 Mathematics Subject Classification:** 92D10, 92D15; 37N25

**Running head:** Multilocus migration-selection models II

# 1 Introduction

How important is population subdivision in maintaining genetic diversity? The study of this and related questions has engaged evolutionary biologists for many decades. Providing satisfactory answers requires comparative studies of empirical data of genetic variation in subdivided and in panmictic populations along with solid information about the contributing factors and, especially, a proper understanding of the structure and amount of variation expected under such conditions.

It has been known for long that a single diallelic locus can be maintained polymorphic without displaying overdominance or underdominance in any subpopulation (Levene 1953, Moran 1959, Maynard Smith 1970, Eyland 1971, Bulmer 1972). Stable polymorphism without overdominance is possible even in a homogeneous environment, for instance, if there is underdominance in each deme (Karlin and McGregor 1972). Thus, there clearly are conditions under which polymorphism can be maintained in a subdivided population but not in a panmictic population.

A very useful concept for studying maintenance of variation at a single diallelic locus is that of a protected polymorphism, meaning that none of the alleles will be lost if rare (Prout 1968). Explicit and instructive conditions have been worked out in important special cases, such as two demes or the Levene model, in which individuals disperse independently of their deme of origin (see references above and Karlin 1982, Nagylaki 1992, Chap. 6, for reviews). Although existence of a protected polymorphism ensures the maintenance of genetic variation, it does not yield any information on the polymorphic equilibria themselves. Even worse, it does not even guarantee that an equilibrium exists because limit cycles or complex dynamics can, in general, not be ruled out. For more than two demes, seemingly simple questions, such as "Does more migration favor or disfavor the maintenance of a stable polymorphism?", have no simple answer (e.g., Karlin 1982).

The effects of migration and selection on a single multiallelic locus were studied only recently. In Nagylaki and Lou (2001) it was shown that without dominance, generically,

the number of alleles present at equilibrium cannot exceed the number of demes. For the Levene model, they showed that if there is an internal equilibrium, then it is globally asymptotically stable. Moreover, sufficient conditions for an internal equilibrium were given. Again for the Levene model, Nagylaki and Lou (2006b) derived sufficient conditions for the nonexistence of an internal equilibrium and, in the absence of dominance, a necessary and sufficient condition for an isolated internal equilibrium. Many illustrative special cases and examples were worked out. In Nagylaki and Lou (2007), among other topics, maintenance of polymorphism is studied under weak migration, strong migration, and uniform selection. We will discuss these results further below.

In the present paper, the potential of maintaining polymorphism at multiple, multiallelic loci in a geographically structured population is explored. Individuals are diploid and subpopulations are sufficiently large such that random genetic drift can be neglected. Because for a panmictic population, polymorphic equilibria cannot exist in the absence of epistasis and of overdominance or underdominance, the questions we will address are the following:

(1) If there is intermediate dominance at every locus and in every deme, how many loci can be maintained polymorphic at a (stable) equilibrium?

(2) If there is intermediate dominance at every locus and in every deme, how many alleles per locus can be maintained at a (stable) equilibrium for a given number of loci? Is there a generic upper bound?

As already indicated, answers to these questions seem to be of most interest if there is no or weak epistasis. In addition, and for reasons that will become clear below, we also pose the following problem:

(3) In the absence of epistasis and of dominance, i.e., for purely additive selection, is the number of demes a generic upper bound for the number of alleles at a (stable) equilibrium?

We shall provide answers to these three questions for strong and for weak migration. In Section 2, we establish that if there are at least two demes, then for an open set of

parameters, an arbitrary number of alleles at an arbitrary number of loci can be maintained at a stable equilibrium. Such an open set is constructed under the assumption that selection is weak relative to migration and recombination. Section 3 deals with weak migration. First we demonstrate that in a panmictic population, with no or weak epistasis and with partial dominance, generically, a single gamete is fixed. Then, for sufficiently weak migration and sufficiently weak epistasis, and with partial dominance at every locus and in each deme, we construct an open set of parameters such that at arbitrarily many loci as many alleles can be maintained at a globally asymptotically stable equilibrium as there are number of demes. In addition, we prove that the number of demes is a generic upper bound to the number of alleles maintained at any locus at any equilibrium.

This paper should be read together with part I (Bürger 2009). The present results are applications of the main theorems there, and we refer to results, remarks, and equations there by the preceding roman numeral I. We use the same notation as in part I, mostly without reintroducing it because it can be found in Table 1 of that paper. We advise the reader to consult I-Section 2 for the basic notation and the general model. Here, we only give a very brief outline.

We use the following recursion relations to describe selection, recombination, and migration:

$$p'_{i,\alpha} = \sum_{\beta} m_{\alpha\beta} p^{\#}_{i,\beta} \,, \tag{1.1a}$$

where

$$p^{\#}_{i,\alpha} = p_{i,\alpha} \frac{w_{i,\alpha}}{\bar{w}_{\alpha}} - D_{i,\alpha} \,; \tag{1.1b}$$

cf. I-(2.9), I-(2.14). Here, $p_{i,\alpha}$, $p^{\#}_{i,\alpha}$, and $p'_{i,\alpha}$ denote the frequency of gamete $i$ in deme $\alpha$ before selection, its frequency after selection and recombination, and after migration in the next generation, respectively. Further, $w_{i,\alpha}$ is the marginal fitness of gamete $i$ in deme $\alpha$, $\bar{w}_{\alpha}$ the mean fitness in deme $\alpha$, $D_{i,\alpha}$ the linkage disequilibrium in gamete $i$ in deme $\alpha$, and $m_{\alpha\beta}$ the probability that an adult individual in deme $\alpha$ migrated from deme $\beta$. Because we assume constant backward migration rates, our results apply in

particular to soft selection. We view (1.1) as a dynamical system on $\Delta_I^\Gamma$, the $\Gamma$-fold cartesian product of the simplex $\Delta_I$ of gamete frequencies in a single deme (see I-Section 2 for details)

We call an equilibrium $\hat{p}$ of (1.1) *fully polymorphic* if at every locus all alleles occur at positive frequency. This does not imply that the equilibrium is internal, i.e., $\hat{p} \in \text{int}\Delta_I^\Gamma$. For a fully polymorphic equilibrium of (1.1), neither all allele nor all gamete frequencies need to be positive in every deme. We call an equilibrium *polymorphic* if at every locus at least two alleles are segregating.

**Remark 1.1.** Because all pairwise recombination rates are assumed positive, despite possible interference, all gametes are generated in a randomly mating population after at most $L-1$ rounds of mating if initially all alleles are present. If migration is ergodic, i.e., irreducible and aperiodic, it takes at most $\Gamma-1$ generations to introduce every allele in each deme. Therefore, it takes at most $\Gamma+L-2$ generations to generate all gametes in every deme. In particular, at every fully polymorphic equilibrium, all gametes will be present in every deme, hence a fully polymorphic equilibrium is internal if migration is ergodic.

As in part I, we abbreviate references to Nagylaki et al. (1999) by NHB, and references to Nagylaki and Lou (2007) by NL07. In contrast to part I, here we assume that there is no position effect. With position effects, analogous results to those below can be derived with only trivial changes.

# 2 Polymorphism under strong migration

Throughout this section we assume that migration and recombination are strong relative to selection. We achieve this by writing fitnesses of genotypes as

$$w_{ij,\alpha} = 1 + \epsilon r_{ij,\alpha} \tag{2.1}$$

and assuming that $\epsilon \geq 0$ is sufficiently small and $|r_{ij,\alpha}| \leq 1$; cf. I-(4.1).

A key role will be played by what is called the weak-selection limit in I-Sect. 4.1, i.e., eq. I-(4.9). Because of its central role, as well as for convenience, we restate it:

$$\frac{d}{dt}P_{i_n}^{(n)} = P_{i_n}^{(n)}\left[\omega_{i_n}^{(n)}(\pi) - \bar{\omega}(\pi)\right] . \tag{2.2}$$

Here, $P_{i_n}^{(n)}$ is the average frequency (over suitably weighted demes) of allele $\mathcal{A}_{i_n}^{(n)}$; $\pi$ is the vector holding all $P_{i_n}^{(n)}$; $\omega_{i_n}^{(n)}$ and $\bar{\omega}$ may be interpreted as the marginal (Malthusian) fitness of allele $\mathcal{A}_{i_n}^{(n)}$ and the mean fitness of the population, each suitably averaged over all demes. This weak-selection limit is best viewed as a dynamical system on the submanifold $\Psi_0$ of $\Delta_I^\Gamma$ on which global linkage equilibrium holds and allele frequencies in the demes are identical and, for allele $\mathcal{A}_{i_n}^{(n)}$, equal to $P_{i_n}^{(n)}$; cf. I-(3.14). It describes evolution of gene frequencies on $\Psi_0$ and is formally closely related to a panmictic multi-allelic, single-locus selection model. In particular, I-(4.10) shows that $\bar{\omega}$ is a Lyapunov function.

The analysis below is based on I-Theorem 4.3. Essentially, it states that if selection is sufficiently weak relative to migration and recombination, then every equilibrium of the full dynamics (1.1) is a perturbation of an equilibrium of (2.2), these pairs of equilibria have the same stability properties, and every trajectory of the full dynamics converges to an equilibrium point. The only assumptions needed are that the backward migration matrix is ergodic, I-(E) in I-Sect. 2, and that all equilibria of (2.2) are hyperbolic, I-(H.1) in I-Sect. 4.1. Both assumptions are fulfilled generically (I-Remark 4.1). With this theorem, we will be able to derive our main result on the maintenance of multilocus polymorphism by proving it first for the much simpler weak-selection limit and then applying the perturbation result of I-Theorem 4.3.

We need the following definitions. There is *no epistasis* if we can write

$$r_{ij,\alpha} = \sum_n u_{i_n j_n,\alpha}^{(n)} \tag{2.3}$$

for every $i, j \in \mathsf{I}$, and every $\alpha \in \mathsf{G}$.

There is *partial dominance* (at every locus and in every deme) if

$$u_{i_n i_n,\alpha}^{(n)} > u_{i_n j_n,\alpha}^{(n)} > u_{j_n j_n,\alpha}^{(n)} \quad \text{or} \quad u_{i_n i_n,\alpha}^{(n)} < u_{i_n j_n,\alpha}^{(n)} < u_{j_n j_n,\alpha}^{(n)} \tag{2.4}$$

holds for every $\alpha, n, i_n$ and for every $j_n \neq i_n$. Unless otherwise mentioned, partial dominance always applies to every locus and to every deme.

Subsequently, we collect the fitnesses $w_{ij,\alpha}$ of all genotypes in every deme in a vector $W$. Because we ignore position effects, fitnesses need to be assigned only to $H = \prod_{n\in\mathsf{L}} \binom{I_n+1}{2}$ unordered genotypes. Thus, $W = (w_{ij,\alpha}) \in (0,\infty)^{H\Gamma}$. If there is no epistasis, fitness in each deme is determined by the $J = \sum_{n\in\mathsf{L}} \binom{I_n+1}{2}$ parameters $u^{(n)}_{i_n j_n,\alpha}$, which we collect in the $J\Gamma$-dimensional vector $U$. We introduce the following sets of parameters, cf. I-(5.9):

$$\widetilde{\mathsf{W}} = \{W = (w_{ij,\alpha}) \in (0,\infty)^{H\Gamma}\}\,, \tag{2.5a}$$

$$\mathsf{R}_1 = \{R = (r_{ij,\alpha}) \in [-1,1]^{H\Gamma}\}\,, \tag{2.5b}$$

$$\mathsf{U} = \{U = (u^{(n)}_{i_n j_n,\alpha}) \in (0,\infty)^{J\Gamma}\}\,, \tag{2.5c}$$

$$\mathsf{U}_1 = \Big\{U = (u^{(n)}_{i_n j_n,\alpha}) \in [-1,1]^{J\Gamma} \text{ such that } \Big|\sum_n u^{(n)}_{i_n j_n,\alpha}\Big| \leq 1\Big\}\,, \tag{2.5d}$$

$$\mathsf{V} = \{U \in \mathsf{U} : U \text{ satisfies (2.4)}\}\,, \tag{2.5e}$$

$$\mathsf{V}_1 = \{U \in \mathsf{U}_1 : U \text{ satisfies (2.4)}\}\,, \tag{2.5f}$$

$$\mathsf{M} = \{M = (m_{\alpha\beta}) \in \mathsf{M}_\Gamma : M \text{ is a nonnegative stochastic matrix}\}\,, \tag{2.5g}$$

$$\mathsf{M}_E = \{M \in \mathsf{M} : M \text{ is ergodic}\}\,. \tag{2.5h}$$

Here, $\mathsf{M}_\Gamma$ denotes the set of all real matrices of dimension $\Gamma \times \Gamma$ and $\mathsf{M}$ is the set of all backward migration matrices. Thus, the dynamics (1.1) is fully determined by specifying $W \in \widetilde{\mathsf{W}}$, $M \in \mathsf{M}$, and the recombination distribution $\{c_\mathsf{K} : \mathsf{K} \subseteq \mathsf{L}\}$. In the case of no epistasis, $U \in \mathsf{U}$ instead of $W \in \widetilde{\mathsf{W}}$ is required. The weak-selection limit (2.2) is fully determined by specifying $R \in \mathsf{R}_1$ and $M \in \mathsf{M}$. In the case of no epistasis, $U \in \mathsf{U}_1$ instead of $R \in \mathsf{R}_1$ is required. If, in addition, there is partial dominance, we choose $U \in \mathsf{V}_1$. We note that $\mathsf{V}_1$ is an open subset of $\mathsf{U}_1$ and call the elements in these sets parameter combinations.

First, we study the maintenance of polymorphism in the weak-selection limit.

**Proposition 2.1.** *Let $L \geq 1$, $\Gamma = 2$, $I_n \geq 2$ for every $n \in \mathsf{L}$, and suppose the weak-selection limit* (2.2)*.*

*(a) There exists an open set of migration and selection parameters, $\mathsf{O} \subset \mathsf{R}_1 \times \mathsf{M}_E$, such that for every parameter combination in $\mathsf{O}$, there is a unique internal, asymptotically stable equilibrium point. This equilibrium attracts all trajectories with internal initial condition. The set of chain-recurrent points (see I-Remark 4.2) of* (2.2) *consists of finitely many hyperbolic equilibria.*

*(b) Such an open set, $\mathsf{O}'$, also exists if the set of all fitnesses is restricted to be nonepistatic and to display partial dominance, i.e., $\mathsf{O}' \subset \mathsf{V}_1 \times \mathsf{M}_E$ is open.*

*Proof.* We start by assuming absence of epistasis, (2.3). To each single locus, we apply a slightly generalized version of the procedure in Remark 4.15 in NL07. We choose the backward migration matrix

$$M = \begin{pmatrix} 1-m_1 & m_1 \\ m_2 & 1-m_2 \end{pmatrix} , \tag{2.6}$$

in which $0 < m_1, m_2 < 1$. Clearly, $M$ is ergodic and it has the maximal left eigenvector $\mu = (\frac{m_2}{m_1+m_2}, \frac{m_1}{m_1+m_2})^T$. We assume that for every locus $n$, the additive fitness components $u^{(n)}_{i_n j_n, \alpha}$ satisfy

$$u^{(n)}_{i_n i_n, 1} < u^{(n)}_{i_n j_n, 1} < u^{(n)}_{j_n j_n, 1} , \tag{2.7a}$$

$$u^{(n)}_{i_n j_n, 2} = -\frac{m_2}{m_1} u^{(n)}_{i_n j_n, 1} , \tag{2.7b}$$

$$u^{(n)}_{i_n i_n, 2} < -\frac{m_2}{m_1}\, u^{(n)}_{i_n i_n, 1} , \tag{2.7c}$$

$$-\frac{m_2}{m_1}\, u^{(n)}_{i_n j_n, 1} < u^{(n)}_{i_n i_n, 2} \tag{2.7d}$$

for every $n$, every $i_n$ and every $j_n > i_n$. Choosing selection coefficients in the order $\{u^{(n)}_{i_n i_n, 1}\}$, $\{u^{(n)}_{i_n j_n, 1}\}$, $\{u^{(n)}_{i_n j_n, 2}\}$, $\{u^{(n)}_{i_n i_n, 2}\}$ shows that (2.7) can be satisfied. By (2.7a), there is partial dominance in deme 1. That partial dominance also holds in deme 2 follows by applying successively (2.7c) with $j_n$ instead of $i_n$, (2.7a), (2.7b), again (2.7b), and

(2.7d). Equations (2.7b) and (2.7c) show that

$$\omega^{(n)}_{i_n j_n} = \sum_{\alpha=1}^{2} \mu_\alpha u^{(n)}_{i_n j_n,\alpha} = 0 \text{ and } \omega^{(n)}_{i_n i_n} < 0\,, \tag{2.8}$$

respectively, for every $n$ and every $i_n$, $j_n$ such that $i_n \neq j_n$. Therefore, in the weak-selection limit at each (single) locus, called strong-migration limit in NL07, a unique internal equilibrium exists, with allele frequencies at locus $n$ given by eq. (4.41) in NL07:

$$\hat{P}^{(n)}_{i_n} = \frac{1}{\omega^{(n)}_{i_n i_n}} \Bigg/ \sum_{j_n} \frac{1}{\omega^{(n)}_{j_n j_n}}\,.$$

Convergence to this equilibrium occurs from the interior of the simplex $\Delta_{I_n}$ because

$$\bar{\omega}_n(\pi^{(n)}) = \sum_{i_n, j_n} \omega^{(n)}_{i_n j_n} P^{(n)}_{i_n} P^{(n)}_{j_n} = \sum_{\alpha} \sum_{i_n, j_n} \mu_\alpha u^{(n)}_{i_n j_n,\alpha} P^{(n)}_{i_n} P^{(n)}_{j_n}\,,$$

where $\pi^{(n)} = (P^{(n)}_1, \ldots, P^{(n)}_{I_n})^T \in \Delta_{I_n}$, is strictly increasing along trajectories except at equilibria. By (2.8), $\bar{\omega}_n(\pi^{(n)})$ is strictly concave on $\Delta_{I_n}$ for every $n$. Hence, the mean fitness, $\bar{\omega}(\pi) = \sum_n \bar{\omega}_n(\pi^{(n)})$, is a strictly concave quadratic form on $\Delta_{I_1} \times \cdots \times \Delta_{I_L}$. Its local (and global) maximum is given by

$$\hat{\pi} = \left(\hat{P}^{(1)}_1, \ldots, \hat{P}^{(1)}_{I_1}, \ldots, \hat{P}^{(L)}_1, \ldots, \hat{P}^{(L)}_{I_L}\right)^T \in \Delta_{I_1} \times \cdots \times \Delta_{I_L}\,.$$

Therefore, $\hat{\pi}$ is the only fully polymorphic equilibrium of the $L$-locus weak-selection limit (2.2). Moreover, because $\bar{\omega}$ is a strict Lyapunov function and migration is ergodic, $\hat{\pi}$ is globally asymptotically stable with respect to all internal states.

Because strict concavity of a quadratic form is structurally stable with respect to small perturbations of the parameters and by the implicit function theorem, a fully polymorphic, asymptotically stable equilibrium of (2.2) persists under sufficiently small perturbations of the $\omega^{(n)}_{i_n j_n}$. Because $\bar{\omega}$ is a strict Lyapunov function, the perturbed equilibrium is also globally asymptotically stable for the interior of the state space. Perturbations of the $\omega^{(n)}_{i_n j_n}$ are caused by perturbations of $R = (r_{ij,\alpha})$ (or $M$, which we don't consider further). We consider two cases: (1) small but arbitrary perturbations of

$R \in \mathsf{R}_1$ and (2) perturbations of $U \in \mathsf{V}_1$ that are small enough such that partial dominance is preserved (i.e., the perturbed parameter combination is still in $\mathsf{V}_1$). Clearly, admissible perturbations of $R$ or $U$ depend continuously on the migration matrix $M$. It follows that, both in case (1) and (2), open subsets of migration and fitness parameters exist such that conclusions (a) and (b) of the proposition hold for the dynamics (2.2). By further restricting these subsets to include only parameter combinations such that the hyperbolicity assumption I-(H.1) is satisfied (which is a generic property both in $\mathsf{R}_1$ and in $\mathsf{U}$; see Appendices B and A, respectively, in NHB, and I-Remark 4.1), we obtain our desired open subsets $\mathsf{O} \subset \mathsf{R}_1 \times \mathsf{M}_E$ and $\mathsf{O}' \subset \mathsf{V}_1 \times \mathsf{M}_E$, respectively. Because (2.2) has gradient-like dynamics, the set of chain-recurrent points consists of the (finitely many) hyperbolic equilibria (NHB, p. 116) if parameter combinations are choosen either in $\mathsf{O}$ or in $\mathsf{O}'$. Note also that $\mathsf{O}$ and $\mathsf{O}'$ are independent of the recombination distribution. □

Now, we are able to formulate and prove the main result of this section.

**Theorem 2.2.** *Let $L \geq 1$, $\Gamma = 2$, $I_n \geq 2$ for every $n \in \mathsf{L}$, let all recombination rates $c_{\mathsf{K}}$ be positive and fixed, and suppose the full migration-selection dynamics* (1.1).

*(a) There exists an open set $\mathsf{Q}$ of migration and selection parameters, $\mathsf{Q} \subset \widetilde{\mathsf{W}} \times \mathsf{M}_E$, such that for every parameter combination in $\mathsf{Q}$, there is a unique, internal, asymptotically stable equilibrium point. This equilibrium is spatially quasi-homogeneous, is in quasi-linkage equilibrium, and attracts all trajectories with internal initial condition. Furthermore, every trajectory converges to an equilibrium point as $t \to \infty$.*

*(b) Such an open set, $\mathsf{Q}'$, also exists if the set of all fitnesses is restricted to be nonepistatic and to display partial dominance, i.e., $\mathsf{Q}' \subset \mathsf{V} \times \mathsf{M}_E$.*

*Proof.* According to Proposition 2.1, we choose a parameter combination $(R, M) \in \mathsf{O}$ or $(U, M) \in \mathsf{O}'$. Every $M$ in (2.6) is ergodic anyway. I-Theorem 4.3 shows that for sufficiently small $\epsilon$ in (2.1) and $(W, M) = (\mathbb{I} + \epsilon R, M)$ or $(W, M) = (\mathbb{I} + \epsilon U, M)$, the dynamics (1.1) has an internal, asymptotically stable equilibrium. (Here, $\mathbb{I}$ denotes the

matrix of dimension $H\Gamma$ with all entries 1.) I-Theorem 4.3 also establishes quasi-linkage equilibrium; cf. I-Remark 4.9. By Proposition 2.1, for both $(R, M)$ and $(U, M)$, the set of chain-recurrent points of (2.2) consists of finitely many hyperbolic equilibria. Hence, I-Theorem 4.3 shows that for each of the two choices of $(W, M)$, (1.1) has no other stable equilibria, and the set of chain-recurrent points consists again of finitely many hyperbolic equilibria (I-Remark 4.5 (i)). In particular, for both choices of $(W, M)$, we obtain convergence to the internal equilibrium for all internal initial conditions as well as convergence of every trajectory to some equilibrium point. By Corollary 32 on p. 244 of Akin (1993), i.e., robustness of the set of chain-recurrent points under sufficiently small $C_1$ perturbations of the dynamics, (1.1) has the same properties for every sufficiently small perturbation of $(W, M)$, either in $\widetilde{\mathsf{W}} \times \mathsf{M}_E$ or in $\mathsf{V} \times \mathsf{M}_E$. Therefore, we obtain open sets $\mathsf{Q} \subset \widetilde{\mathsf{W}} \times \mathsf{M}_E$ or $\mathsf{Q}' \subset \mathsf{V} \times \mathsf{M}_E$ with the desired properties. □

**Remark 2.3.** (i) By Remark 1.1, the asymptotically stable internal equilibria in Proposition 2.1 and Theorem 2.2 attract all trajectories such that initially every allele is present in at least one deme.

(ii) The proof shows that such an open set exists in the parameter range where migration is sufficiently strong relative to selection.

(iii) Theorem 2.2 remains valid if $\Gamma \geq 2$, because additional demes can be added provided the migration matrix is ergodic and selective differences in these demes are sufficiently small, i.e., of order $O(\epsilon^2)$.

(iv) In a single panmictic population, polymorphic equilibria do not exist in the absence of epistasis unless there is overdominance or underdominance (or no selection at all); see Corollary 3.4.

**Remark 2.4.** Proposition 2.1 and Theorem 2.2 hold if migration is restricted to the Levene model, i.e., to matrices satisfying $m_1 + m_2 = 1$.

**Remark 2.5.** In the absence of epistasis, the conclusions of the above theorem do not hold if dominance is absent at every locus and in every deme, because the open

set $\mathsf{O}'$ constructed in the proof of Proposition 2.1 cannot contain fitness matrices with complete absence of dominance. For given $n$, the conditions (2.7) are incompatible with the assumption of no dominance in every deme (NL07, Remark 4.15).

In fact, the following result can be shown.

**Proposition 2.6.** *If selection is nonepistatic and sufficiently weak relative to migration and recombination, then, generically, there exists a globally asymptotically stable monomorphic state if single-locus fitnesses are*

*(a) additive, i.e., there is no dominance, or*

*(b) multiplicative.*

*In particular, no polymorphism is maintained.*

*Proof.* (a) In view of I-Theorem 4.3 it is sufficient to prove that the weak-selection limit (2.2) has a globally asymptotically stable monomorphic state. Let us assume absence of epistasis and of dominance, i.e., additive genes. Then, ignoring degenerate cases, there exist $v^{(n)}_{i_n,\alpha} > 0$ such that, instead of (2.3),

$$r_{ij,\alpha} = \sum_n \left( v^{(n)}_{i_n,\alpha} + v^{(n)}_{j_n,\alpha} \right) \tag{2.9}$$

holds for every $i, j \in \mathsf{I}$, and every $\alpha \in \mathsf{G}$. Then, a straightforward calculation shows that the weak-selection limit (2.2) becomes

$$\frac{dP^{(n)}_{i_n}}{dt} = P^{(n)}_{i_n} \left[ \nu^{(n)}_{i_n} - \bar{\nu}^{(n)}(\pi) \right] , \tag{2.10}$$

where

$$\nu^{(n)}_{i_n} = \sum_\alpha \mu_\alpha v^{(n)}_{i_n,\alpha} , \quad \bar{\nu}^{(n)}(\pi) = \sum_{i_n} \nu^{(n)}_{i_n} P^{(n)}_{i_n} , \tag{2.11}$$

and the $\nu^{(n)}_{i_n}$ are constants, i.e., independent of $\pi$.

Apparently, for every $n$, the dynamics (2.10) is decoupled from all other loci. For each single $n$, the dynamics on $\Delta_{I_n}$ is just the well known selection dynamics for $I_n$ alleles in a haploid population (e.g., Bürger 2000, p. 26). Therefore, no internal equilibrium

exists if, as is generically the case, for given $n$ all $\nu_{i_n}^{(n)}$ are different. More precisely, suppose without loss of generality that for every $n$ we have $\nu_1^{(n)} > \nu_{i_n}^{(n)}$ for every $i_n \neq 1$. Then, $\nu_1^{(n)} > \bar{\nu}^{(n)}$ unless $P_1 = 0$, hence

$$\frac{dP_1^{(n)}}{dt} > 0\,, \tag{2.12}$$

except when $P_1^{(n)} = 0$. Therefore, at locus $n$, allele $\mathcal{A}_1^{(n)}$ becomes fixed if it is initially present. This proves our assertion.

(b) Instead of (2.9), we assume

$$r_{ij,\alpha} = \sum_n v_{i_n,\alpha}^{(n)}\, v_{j_n,\alpha}^{(n)}\,. \tag{2.13}$$

Then, another straightforward calculation shows that instead of (2.10) we obtain

$$\frac{dP_{i_n}^{(n)}}{dt} = P_{i_n}^{(n)} \sum_\alpha \mu_\alpha \bar{v}_\alpha^{(n)} \left[ v_{i_n,\alpha}^{(n)} - \bar{v}_\alpha^{(n)}(\pi) \right]\,, \tag{2.14}$$

where $\bar{v}_\alpha^{(n)}(\pi) = \sum_{i_n} v_{i_n,\alpha}^{(n)} P_{i_n}^{(n)}$. Generically, for every $n$, the values $\sum_\alpha \mu_\alpha \bar{v}_\alpha^{(n)} v_{i_n,\alpha}^{(n)}$ are different, and we assume without loss of generality that $\mathcal{A}_1^{(n)}$ is the allele for which this average fitness is maximized. Hence, (2.12) follows. □

The above proposition applies in particular if selection acts on haploids. Essentially, part (b) follows from the Corollary to Theorem 4 in Kirzhner and Lyubich (1997). Proposition 3.2 below suggests further generalizations of Proposition 2.6 by requiring (3.3) – (3.5) for the $\omega_{i_n j_n}^{(n)}$ instead of (2.9) or (2.13).

**Remark 2.7.** Statement (a) of Proposition 2.6 can be extended to deme-independent degree of dominance; cf. Remark 2.2 in Nagylaki (2008), which also refers to the precise definition of deme-independent degree of dominance.

# 3 Polymorphism under weak migration

Before treating weak migration, we need to establish the equilibrium structure for a single randomly mating population if there is partial dominance and no epistasis.

## 3.1 Directional selection under panmixis

For a panmictic population, we demonstrate that in the absence of epistasis and with partial dominance, the gamete with the highest fitness becomes fixed and no equilibria other than fixation of a single gamete exist. Although this is highly intuitive, it seems that it has never been formulated or proved. In fact, in Proposition 3.2, we derive a more general result. It is also valid for sufficiently weak epistasis.

We use notation analogous to the multi-deme case but omit subscripts $\alpha$. From (1.1), the dynamics of gamete frequencies in a panmictic population is

$$p_i' = p_i \frac{w_i}{\bar{w}} - D_i \,. \tag{3.1}$$

We assume absence of epistasis,

$$w_{ij} = \sum_n u^{(n)}_{i_n j_n} \,, \tag{3.2}$$

i.e., $R = 0$ in I-(5.11), and order alleles at each locus such that

$$u^{(n)}_{i_n i_n} \geq u^{(n)}_{i_n+1, i_n+1} \tag{3.3}$$

for every $n \in \mathsf{L}$ and $i_n = 1, \ldots, I_n - 1$.

We shall need the following well known result (Kun and Lyubich 1980; Lyubich 1992, Theorem 9.6.13; NHB, Lemma 2.1; Bürger 2000, p. 77).

**Remark 3.1.** In a panmictic population and in the absence of epistasis, every trajectory converges to an equilibrium point. A point $p$ is an equilibrium of (3.1) if and only if it is both a selection equilibrium for each locus and it is in linkage equilibrium.

**Proposition 3.2.** *Let $L \geq 1$, $\Gamma = 1$, suppose* (3.1), (3.2), *and* (3.3). *In addition, we assume*

$$u^{(n)}_{11} > u^{(n)}_{22} \tag{3.4}$$

*for every $n$ and*

$$u^{(n)}_{1 i_n} \geq u^{(n)}_{i_n j_n} \tag{3.5}$$

*for every $n$, every $i_n$ and every $j_n > i_n$.*

*(a) Then, the gamete $(1,\dots,1)$ is fixed as $t \to \infty$. The corresponding equilibrium is asymptotically stable and attracts all trajectories such that initially every allele is present.*

*(b) For generic $U$ satisfying the above conditions, all equilibria of (3.1) are hyperbolic and the only chain recurrent points are these (finitely many) equilibria.*

*Proof.* (a) By Remark 3.1, it is sufficient to show that on the linkage equilibrium manifold, no other point than $(1,\dots,1)$ is a stable equilibrium and no allele $\mathcal{A}_1^{(n)}$ can be lost. We define

$$u_{i_n}^{(n)} = \sum_{j_n} u_{i_n j_n}^{(n)} p_{j_n}^{(n)} , \tag{3.6a}$$

$$\bar{u}^{(n)} = \sum_{i_n, j_n} u_{i_n j_n}^{(n)} p_{i_n}^{(n)} p_{j_n}^{(n)} = \sum_{i_n} u_{i_n}^{(n)} p_{i_n}^{(n)} , \tag{3.6b}$$

and the marginal fitness of allele $i_n$,

$$w_{i_n}^{(n)} = \frac{1}{p_{i_n}^{(n)}} \sum_{i | i_n} \sum_{j} w_{ij} p_i p_j .$$

Allele frequencies accross generations change according to (e.g., Bürger 2000, p. 72)

$$p_{i_n}^{(n)\prime} = p_{i_n}^{(n)} w_{i_n}^{(n)} / \bar{w} , \tag{3.7}$$

where

$$\bar{w} = \sum_{n} \bar{u}^{(n)} \tag{3.8}$$

is population mean fitness.

Assuming linkage equilibrium, a simple calculation shows that

$$w_{i_n}^{(n)} = u_{i_n}^{(n)} + \sum_{k \neq n} \bar{u}^{(k)} . \tag{3.9}$$

From (3.3), (3.4), and (3.5), we find

$$u_1^{(n)} > u_{i_n}^{(n)} \tag{3.10}$$

for every $n$, $i_n \geq 2$, and $0 < p_1^{(n)} < 1$. Then, from (3.7), (3.9) and (3.8), (3.6b), and (3.10), we obtain for every $n$:

$$\begin{aligned}\Delta p_1^{(n)} &= p_1^{(n)}(w_1^{(n)} - \bar{w})/\bar{w} \\ &= p_1^{(n)}(u_1^{(n)} - \bar{u}^{(n)})/\bar{w} \\ &= p_1^{(n)}\Big[\sum_{i_n} p_{i_n}(u_1^{(n)} - u_{i_n}^{(n)})\Big]\Big/\bar{w} \geq 0\,, \end{aligned} \tag{3.11a}$$

and

$$\Delta p_1^{(n)} > 0 \tag{3.11b}$$

if $0 < p_1^{(n)} < 1$. This proves our assertion because (3.11b) shows that no allele $\mathcal{A}_1^{(n)}$ can be lost and no other equilibrium than fixation of every $\mathcal{A}_1^{(n)}$ can be stable. Obviously, the latter is asymptotically stable.

(b) Generic hyperbolicity of all equilibria follows from Appendix A in NHB together with the fact that strict inequalities preserve openness. The second statement is I-Lemma 5.3. □

**Remark 3.3.** Part (a) of the above proposition was inspired by Corollary 2.14 of Nagylaki and Lou (2006a) which it generalizes to multiple loci under nonepistatic selection.

**Corollary 3.4.** *Suppose that instead of* (3.3), (3.4), *and* (3.5), *$U$ satisfies*

$$u_{i_n i_n}^{(n)} > u_{i_n j_n}^{(n)} > u_{j_n j_n}^{(n)} \quad \textit{or} \quad u_{i_n i_n}^{(n)} < u_{i_n j_n}^{(n)} < u_{j_n j_n}^{(n)} \tag{3.12}$$

*for every $n$, $i_n$, and every $j_n \neq i_n$, i.e., there is partial dominance at every locus.*

*(a) Conclusions (a) and (b) of Proposition 3.2 hold.*

*(b) At every equilibrium (stable or not), there is exactly one gamete fixed.*

*Proof.* (a) If alleles are ordered according to (3.3), the assertion follows immediately from $u_{1i_n}^{(n)} > u_{i_n i_n}^{(n)} > u_{i_n j_n}^{(n)}$ if $j_n > i_n$.

(b) The strict inequalities in (3.12) together with the ordering (3.3) yield $u_{i_n i_n}^{(n)} > u_{i_n+1, i_n+1}$ for every $n$ and $i_n$. Hence, (3.10) generalizes to $u_{i_n}^{(n)} > u_{i_n+1}^{(n)}$ for every $n$ and

$i_n$. Therefore, instead of (3.11b), we obtain

$$\Delta p_{i_n}^{(n)} > 0 \tag{3.13}$$

if $0 < p_{i_n}^{(n)} < 1$ and $p_{j_n}^{(n)} = 0$ for every $j_n > i_n$. Thus, equilibrium allele frequencies can only be 0 or 1. □

**Corollary 3.5.** *Proposition 3.2 and Corollary 3.4 hold if epistasis is sufficiently weak.*

*Proof.* This is an immediate consequence of Theorem 2.3 in NHB, or I-Theorem 5.4 with $\Gamma = 1$. □

## 3.2 Weak migration

Our first goal is to construct an open set of parameters exhibiting weak migration, partial dominance, and no or weak epistasis, for which the number of alleles maintained at any of arbitrarily many loci equals the number of demes. Then we show that under these assumptions, generically, no equilibrium exists at which more alleles are maintained at a locus.

The starting point of our analysis is a population distributed over $\Gamma$ demes with no migration between them. According to I-(5.1), the dynamics is given by

$$p'_{i,\alpha} = p_{i,\alpha}\frac{w_{i,\alpha}}{\bar{w}_\alpha} - D_{i,\alpha} \tag{3.14}$$

for every $i \in I$ and every $\alpha \in \mathsf{G}$.

We assume partial dominance, (2.4), and that at each locus every allele is the fittest in at least one deme, i.e., for every $n$ we have

$$\mathsf{G}_{i_n}^{(n)} = \left\{\alpha \in \mathsf{G} : u_{i_n i_n,\alpha}^{(n)} > \max_{j_n : j_n \neq i_n} u_{j_n j_n,\alpha}^{(n)}\right\}, \quad \mathsf{G} = \bigcup_{i_n \in \mathsf{I}_n} \mathsf{G}_{i_n}^{(n)}, \tag{3.15}$$

$\mathsf{G}_{i_n}^{(n)} \neq \emptyset$, and $\mathsf{G}_{i_n}^{(n)} \cap \mathsf{G}_{j_n}^{(n)} = \emptyset$ for every $i_n, j_n \in \mathsf{I}_n$ with $i_n \neq j_n$. This assumption requires $\Gamma \geq \max_n I_n$. Partial dominance implies that in every deme and at each locus one allele

is the fittest, i.e.,

$$\text{for every } \alpha \text{ and for every } n \text{ there exists } \iota_n = \iota_n(\alpha) \in \mathsf{I}_n$$
$$\text{such that } u^{(n)}_{\iota_n\iota_n,\alpha} > u^{(n)}_{j_nj_n,\alpha} \text{ for every } j_n \neq \iota_n\,. \tag{3.16}$$

Let $\mathsf{Y} \subset \mathsf{V}$ denote the open subset of parameter combinations satisfying (2.4), (3.15), and (3.16), i.e., for $U \in \mathsf{Y}$ there is no epistasis, partial dominance, every allele is the fittest in at least one deme, and $\mathcal{A}^{(n)}_{\iota_n\iota_n,\alpha}$ is the allele with highest fitness at locus $n$ in deme $\alpha$. The following result concerns multiple demes in the absence of migration and epistasis.

**Proposition 3.6.** *Let $L \geq 1$, $\Gamma \geq \max_n I_n$, $U \in \mathsf{Y}$, and suppose* (3.2) *and* (3.14)*.*

*(a) Then, there exists a unique asymptotically stable equilibrium point. It is fully polymorphic, but not internal, and attracts all trajectories such that initially every allele is present in every deme.*

*(b) At every equilibrium, exactly one gamete is fixed in each deme.*

*(c) For generic $U \in \mathsf{Y}$, all equilibria of* (3.14) *are hyperbolic and the only chain recurrent points are these (finitely many) equilibria.*

*Proof.* (a) We apply Corollary 3.4 to each deme $\alpha$ by renumbering alleles, i.e., allele 1 at locus $n$ in the corollary corresponds to allele $\iota_n(\alpha)$ defined in (3.16). Then, within each deme $\alpha$, the equilibrium $p_{\iota,\alpha} = 1$, $\iota = (\iota_1, \ldots, \iota_L)$, is globally asymptotically stable (for all states such that initially every allele is present). Hence, (3.15) implies that in the full system (3.14), there is an asymptotically stable, fully polymorphic equilibrium. It corresponds to fixation of $\iota(\alpha)$ in deme $\alpha$ for every $\alpha$ and attracts all trajectories such that initially every allele is present in every deme.

(b) This follows immediately from Corollary 3.4.

(c) Generic hyperbolicity of all equilibria follows from Appendix A in NHB. The second statement is I-Lemma 5.3. □

To state our first main result, we define

$$\mathsf{M}_\epsilon = \{M \in \mathsf{M} : M = I + A, \text{ where } A \in \mathsf{M}_{w,\epsilon}\}\,; \tag{3.17}$$

cf. I-(5.10), I-(5.9b).

**Theorem 3.7.** *Let $L \geq 1$, $\Gamma \geq \max_n I_n$, let all recombination rates $c_{\mathsf{K}}$ be positive and fixed, and suppose the (full) migration-selection dynamics* (1.1).

*(a) There exists an open set $\mathsf{P}_\epsilon$ of migration and selection parameters, $\mathsf{P}_\epsilon \subset \widetilde{\mathsf{W}} \times \mathsf{M}_\epsilon$, such that a unique, fully polymorphic, asymptotically stable equilibrium exists for every $(W, M) \in \mathsf{P}_\epsilon$. This equilibrium exhibits weak linkage disequilibrium within every deme and attracts all trajectories with internal initial condition. Furthermore, every trajectory converges to an equilibrium point as $t \to \infty$.*

*(b) The set $\mathsf{P}_\epsilon$ can be chosen such that either $w_{ii,\alpha} < w_{ij,\alpha} < w_{jj,\alpha}$ or $w_{ii,\alpha} > w_{ij,\alpha} > w_{jj,\alpha}$ holds for every $i$, every $j \neq i$, and every $\alpha$.*

*(c) Such an open set, $\mathsf{P}'_\epsilon$, also exists if the set of all fitnesses is restricted to be nonepistatic and to display partial dominance, i.e., $\mathsf{P}'_\epsilon \subset \mathsf{V} \times \mathsf{M}_\epsilon$.*

*(d) If $\mathsf{M}_\epsilon$ is restricted to ergodic migration matrices, then the equilibrium in (a), (b), (c) is internal.*

*Proof.* (a) follows directly from I-Theorem 5.4, I-Remark 5.5 (iv), and Proposition 3.6 by choosing $U \in \mathsf{Y}$, such that (c) of Proposition 3.6 is satisfied, and setting

$$\mathsf{P}_\epsilon = \{(W, M) : W \text{ is given by I-(5.11) with } R \in \mathsf{R}_\eta \text{ and } M \in \mathsf{M}_\epsilon\},$$

where $\epsilon$ is sufficiently small, $\mathsf{R}_\eta$ is as in I-(5.9a) and $\eta$ as in I-(5.12). We note that since here, in contrast to part I, position effects are ignored, we have $W \in \mathbb{R}^{H\Gamma}$ instead of $W \in \mathbb{R}^{I^2\Gamma}$.

(b) is obvious because restriction to (genotypic) partial dominance maintains openness and the construction leading to Proposition 3.6 ensures that this set is not empty.

(c) The map $\psi : \mathsf{U} \times \mathsf{M}_\epsilon \to \widetilde{\mathsf{W}} \times \mathsf{M}_\epsilon$ defined by $\psi(U, M) = (W, M)$, where $w_{ij,\alpha} = \sum_n u^{(n)}_{i_n j_n, \alpha}$, is continuous. Therefore, the preimage $\psi^{-1}(\mathsf{P}_\epsilon)$ is open in $\mathsf{U} \times \mathsf{M}_\epsilon$. Now we can chose $\mathsf{P}'_\epsilon = \psi^{-1}(\mathsf{P}_\epsilon) \cap (\mathsf{V} \times \mathsf{M}_\epsilon)$ which is open and nonempty by the construction leading to Proposition 3.6.

(d) follows from Remark 1.1. □

**Remark 3.8.** In contrast to the case of strong migration and weak selection (Theorem 2.2), here the number of alleles that is maintained at a locus cannot exceed the number of demes. Below, we show that this is true generically. We also note that, again in contrast to Theorem 2.2, in the above theorem, absence of dominance at every locus and in every deme is admitted. Indeed, any intermediate level of dominance is admitted.

**Theorem 3.9.** *For arbitrary number of loci, sufficiently weak migration and epistasis, and partial dominance, the number of demes is the generic maximum for the number of alleles that can be maintained at any locus at any equilibrium (stable or not) of* (1.1)*.*

*Proof.* By Theorem 3.7, this upper bound is assumed on an open set. Corollary 3.4, applied to every deme, shows that in the absence of migration and with partial dominance, only equilibria exist with at most $\Gamma$ different gametes present. Thus, the number of alleles per locus is also at most $\Gamma$. Consequently, under the generic hyperbolicity assumption I-(H.2), I-Theorem 5.4 can be applied, and this yields the assertion. □

**Remark 3.10.** For a single locus, Theorem 3.9 was proved in NL07 (Remark 4.4). According to Theorem 2.4 of Nagylaki and Lou (2001), for a single locus and in the absence of dominance, the number of demes is a generic upper bound to the number of alleles maintained at any, i.e., stable or unstable, equilibrium. It is an open problem if this holds for multiple loci with no epistasis.

# 4 Discussion

We have established that in a subdivided population, migration-selection balance can maintain multiallelic polymorphism at arbitrarily many loci under conditions for which in a panmictic population no polymorphism at all can be maintained (Theorem 2.2, Theorem 3.7, Corollaries 3.4 and 3.5). These conditions are no or weak epistasis and intermediate levels of dominance at every locus (and in every deme). Because it is well known that high levels of genetic variation can be maintained in a panmictic population if there is either (strong) epistasis or overdominance, we restricted our attention to no

or weak epistasis and intermediate dominance. We ignored underdominance because, at a single locus, weak migration can maintain (multiple) stable polymorphic equilibria in the presence of heterozygote inferiority in every deme (Karlin and McGregor 1972).

The polymorphisms determined in the present work are robust with respect to small perturbations of the parameters, i.e., of the backward migration rates and the genotypic fitnesses, and do not depend on the recombination distribution provided all pairwise recombination rates are positive. Since we assume constant backward migration rates, our results apply in particular to soft selection. The open sets of parameters, for which stable, multiallelic multilocus polymorphism exists, have been constructed by extending single-locus results of Nagylaki and Lou (2007). These constructions, carried out in the proofs of Proposition 2.1 and Proposition 3.6, provide insight into the conditions required for maintaining genetic variation in subdivided populations.

Interestingly, for strong migration and only two demes, an arbitrary number of alleles can be maintained at each of arbitrarily many loci, whereas for weak migration, the number of demes is a generic upper bound for the number of alleles that can be maintained (Theorem 3.9). At first, this contrast appears counterintuitive given the frequently expressed view that, in general, it should be easier to maintain polymorphism under weak migration than under strong migration. This view derives from the fact that for a single diallelic locus and two demes, the parameter region for which a protected polymorphism exists increases with decreasing migration rate, as is immediate from the conditions for protection (e.g., Bulmer 1972, Nagylaki 1992, pp. 148-151). Additional support for this expectation comes from the study of weak migration in homogeneous and heterogeneous environments (Karlin and McGregor 1972, Christiansen and Feldman 1975, Christiansen 1999), from an analysis of the Deakin (1966) model, in which a fixed fraction of individuals does not migrate whereas the rest is distributed randomly over the demes in proportion to their size (also called random outbreeding and site homing, and generalizing the Levene model) for a single locus (Christiansen 1974, Karlin 1982) and two loci (Christiansen and Feldman 1975), as well as from recent

numerical studies of symmetric migration between two demes. Spichtig and Kawecki (2004) explored a soft-selection model with multiple diallelic loci, Star et al. (2007a,b) one with a single multiallelic locus. In contrast, Karlin (1982, p. 128) noted that for the non-homogeneous Deakin model, in which the 'homing probabilities' vary among demes, "it is possible to increase a single homing rate and reduce or even abrogate the event of $A$-protection".

Obviously, the present results, as well as those of Nagylaki and Lou (2001, 2006, 2007) for a single locus, are not at variance with those mentioned above, but complimentary. In particular, they yield a deeper understanding of the conditions under which variation can be maintained by migration-selection balance. With strong migration, a stable multiallelic polymorphism requires some form of overdominance at this locus for suitably averaged fitnesses. The reason is that strong migration leads to strong mixing, so that gamete and allele frequencies become similar among demes. Therefore, we expect that the constraints on selection for maintaining multilocus polymorphism under strong migration are quite stringent, i.e., some form of average overdominance is required. If there is no dominance (additive genes), or fitnesses within loci are multiplicative (haploid selection), then only monomorphic equilibria exist if migration is strong (Corollary 3.5), and no polymorphism can be maintained.

The conditions for maintaining loci polymorphic under weak migration are much weaker. Basically, with arbitrary intermediate dominance, to maintain all alleles it is sufficient that at each locus every allele is the fittest in at least one niche. This, obviously, limits the number of alleles that can be maintained.

It seems remarkable that these results hold for any number of loci and, thus, provide a simple answer to question (1) in the Introduction. As will be shown elsewhere, also in the Levene model arbitrarily many loci can be maintained polymorphic. But there, the maximum number of polymorphic loci depends on the pattern of dominance and the number of demes (Nagylaki unpubl., Bürger unpubl.). It is an open problem if for every (ergodic) migration scheme, arbitrarily many loci can be maintained polymorphic

in the absence of epistasis and of overdominance and underdominance.

Concerning question (2), our results demonstrate that the maximum number of alleles that can be maintained at a locus depends on the strength of migration. Since the proofs of Proposition 2.1 and Theorem 2.2 remain valid if migration is restricted to yield the Levene model (Remark 2.4), also in this special case arbitrarily many alleles at arbitrarily many loci can be maintained at a stable equilibrium if selection is weak and dominance is intermediate. Of course, Theorem 3.7 does not apply to the Levene model because the Levene model precludes weak migration.

For a single locus, arbitrary migration, and in the absence of dominance, the number of demes is the generic maximum for the number of alleles that can be maintained. This is also true for Levene model (Nagylaki and Lou 2001). Although we have extended this result to multiple (nonepistatic) loci if selection is weak or strong (in the latter case, the upper bound is 1), question (3) remains open for arbitrary migration. It will be shown elsewhere that it holds for the multilocus Levene model. It would be of interest to study these problems for other special migration models, such as the Deakin or the stepping-stone model.

Although the range of parameters, in which the conditions for maintenance of multiallelic multilocus polymorphisms are satisfied, decreases rapidly in proportion to the full parameter space as the number of alleles or loci increases, for a small or moderate number of alleles or loci, stable multiallelic polymorphism does not seem unlikely. What is required if dispersal is weak, basically, is that there is a mosaic of directional selection pressures and different genes or genotypes that are locally well adapted. Essentially, weak epistasis and intermediate dominance are expected under many forms of directional selection. Under strong epistatic directional selection (whether concave or convex), the number of polymorphic loci at a stable equilibrium can be severely reduced relative to nonepistatic selection (cf. Spichtig and Kawecki 2004).

All polymorphic equilibria constructed by the methods in this paper are in (quasi) linkage equilibrium. As demonstrated by Li and Nei (1974), this is not the case for mi-

gration of arbitrary strength, even if fitnesses are completely additive (i.e., no epistasis and no dominance). Their figure shows that linkage disequilibrium can be substantial and is maximized at intermediate migration rates. In a completely additive model, linkage disequilibrium is expected whenever there is differentiation between the niches (unequal allele frequencies) and migration is not very weak (Christiansen and Feldman 1975). Clearly, differentiation requires that migration is not too strong.

Finally, we briefly point out a difference to another mechanism that has the potential of maintaining high levels of polymorphism. This is negative frequency-dependent selection, in particular, intraspecific competition for resources that is sufficiently strong to induce disruptive selection (see Bürger 2005 and references therein). At least in the case of competition for a unimodal, continuously distributed resource, the pattern of multilocus polymorphism differs from those described above for migration-selection balance. In the first case, at most two alleles can be maintained at a locus, namely the alleles with most extreme effects (Schneider 2006), whereas in the second case we have seen that multiallelic polymorphism is possible, and even likely if genotypes carrying different alleles are the fittest in different niches. Competition for multivariate resources or several discrete resources could lead to different results.

*Acknowledgments.* The author is grateful to Thomas Nagylaki for inspiring discussions and helpful comments on the manuscript. The author also acknowledges helpful discussions with Josef Hofbauer and useful comments on the manuscript by Yuan Lou and Kristan Schneider. This work was supported by a grant of the Vienna Science and Technology Fund (WWTF). Part of this work was done while the author was a visiting researcher at the Program for Evolutionary Dynamics at Harvard University. Its hospitality is gratefully acknowledged.

# References

Akin, E. 1993. The General Topology of Dynamical Systems. Providence, R.I.: Amer. Math. Soc.

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. 2005. A multilocus analysis of intraspecific competition and stabilizing selection on a quantitative trait. J. Math. Biol. 50, 355-396.

Bürger, R. 2009. Multilocus selection in subdivided populations I. Convergence properties for weak or strong migration. J. Math. Biol., this issue.

Christiansen, F.B. 1974. Sufficient conditions for protected polymorphism in a subdivided population. Amer. Natur. 108, 157-166.

Christiansen, F.B. 1999. Population Genetics of Multiple Loci. Wiley, Chichester.

Christiansen, F.B., Feldman, M. 1975. Subdivided populations: A review of the one- and two-locus deterministic theory. Theor. Popul. Biol. 7, 13-38.

Deakin, M.A.B. 1966. Sufficient conditions for genetic polymorphism. Amer. Natur. 100, 690-692.

Eyland, E.A. 1971. Moran's island model. Genetics 69, 399-403.

Ewens, W.J. 1969. Mean fitness increases when fitnesses are additive. Nature 221, 1076.

Karlin. S., 1982. Classification of selection-migration structures and conditions for a protected polymorphism. Evol. Biol. 14, 61-204.

Karlin, S., McGregor, J. 1972. Application of method of small parameters to multi-niche population genetics models. Theor. Popul. Biol. 3, 186-208.

Kirzhner, V., Lyubich, Yu. 1997. Multilocus dynamics under haploid selection. J. Math. Biol. 35, 391-408.

Kun, L.A., Lyubich, Yu.I. 1980. Convergence to equilibrium in a polylocus polyallele population with additive selection. Probl. Inform. Transmiss. 16, 152-161.

Levene, H. 1953. Genetic equilibrium when more than one ecological niche is available. Amer. Natur. 87, 331-333.

Li, W.-H., Nei, M. 1974. Stable linkage disequilibrium without epistasis in subdivided

populations. Theor. Popul. Biol. 6, 173-183.

Lyubich, Yu.I. 1992. Mathematical Structures in Population Genetics. Berlin Heidelberg New York: Springer.

Maynard Smith, J. 1970. Genetic polymorphism in a varied environment. Amer. Natur. 104, 487-490.

Moran, P.A.P. 1959. The theory of some genetical effects of population subdivison. Austr. J. Biol. Sci. 12, 109-116.

Nagylaki, T. 1992. Introduction to Theoretical Population Genetics. Berlin: Springer

Nagylaki, T. 2008. Polymorphism in multiallelic migration-selection models with dominance. Submitted.

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

Nagylaki, T., Lou, Y. 2001. Patterns of multiallelic poylmorphism maintained by migration and selection. Theor. Popul. Biol. 59, 297-33.

Nagylaki, T., Lou, Y. 2006a. Multiallelic selection polymorphism. Theor. Popul. Biol. 69, 217-229.

Nagylaki, T., Lou, Y. 2006b. Evolution under the multiallelic Levene model. Theor. Popul. Biol. 70, 401-411.

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. Lect. Notes Math. 1922, pp. 119 - 172. Berlin Heidelberg New York: Springer.

Prout, T. 1968. Sufficient conditions for multiple niche polymorphism. Amer. Natur. 102, 493-496.

Schneider, K.A. 2006. A multilocus-multiallele analysis of frequency-dependent selection induced by intraspecific competition. J. Math. Biol. 52, 483-523.

Spichtig, M., Kawecki, T.J. 2004. The maintenance (or not) of polygenic variation by

soft selection in a heterogeneous environment. Amer. Natur. 164, 70-84.

Star, B., Stoffels, R.J., Spencer, H.G. 2007a. Single-locus polymorphism in a heterogeneous two-deme model. Genetics 176, 1625-1633.

Star, B., Stoffels, R.J., Spencer, H.G. 2007b. Evolution of fitnesses and allele frequencies in a population with spatially heterogeneous selection pressures. Genetics 177, 1743-1751.