J. Math. Biol.
DOI 10.1007/s00285-008-0248-1

Mathematical Biology

Diffusion approximations for one-locus multi-allele kin selection, mutation and random drift in group-structured populations: a unifying approach to selection models in population genetics

Sabin Lessard

Received: 14 March 2008 / Revised: 11 December 2008

Abstract Diffusion approximations are ascertained from a two-time-scale argument in the case of a group-structured diploid population with scaled viability parameters depending on the individual genotype and the group type at a single multi-allelic locus under recurrent mutation, and applied to the case of random pairwise interactions within groups. The main step consists in proving global and uniform convergence of the distribution of the group types in an infinite population in the absence of selection and mutation, using a coalescent approach. An inclusive fitness formulation with coefficient of relatedness between a focal individual J affecting the reproductive success of an individual I, defined as the expected fraction of genes in I that are identical by descent to one or more genes in J in a neutral infinite population, given that J is allozygous or autozygous, yields the correct selection drift functions. These are analogous to the selection drift functions obtained with pure viability selection in a population with inbreeding. They give the changes of the allele frequencies in an infinite population without mutation that correspond to the replicator equation with fitness matrix expressed as a linear combination of a symmetric matrix for allozygous individuals and a rank-one matrix for autozygous individuals. In the case of no inbreeding, the mean inclusive fitness is a strict Lyapunov function with respect to this deterministic dynamics. Connections are made between dispersal with exact replacement (proportional dispersal), uniform dispersal, and local extinction and recolonization. The timing of dispersal (before or after selection, before or after mating) is shown to have an effect on group competition and the effective population size.

In memory of Sam Karlin.

S. Lessard (✉)
Département de mathématiques et de statistique,
Université de Montréal, Montréal, QC H3C 3J7, Canada
e-mail: lessards@dms.umontreal.ca

Published online: 21 January 2009

Keywords Diffusion approximation · Kin selection · Island model · Covariance formulas · Relatedness coefficients · Coalescent process · Group selection · Effective population size · Replicator equation · Fundamental theorem of natural selection

Mathematics Subject Classification (2000) Primary 92D10; Secondary 60J70

1 Introduction

Ethier and Nagylaki (1980) provided key conditions for diffusion approximations of discrete-time Markov chains with two time scales based on weak-convergence results for perturbed operator semigroups (Kurtz 1975). They applied their results to population genetics models for random drift at a multi-allelic locus under selection and mutation in a finite diploid random-mating population. Considering three different schemes for reproduction, namely, multinomial sampling, overlapping generations, and general progeny distributions, they showed convergence in distribution to the standard, Hardy–Weinberg, one-locus diffusion model for gene frequencies (Kimura 1964), provided that the mutation rates and the selection intensities are appropriately scaled as the population size goes to infinity. The random-mating limit was also established in the case of strong migration for a Wright–Fisher population geographically structured into a fixed finite number of colonies, which exchange migrants according to a fixed ergodic pattern as the sizes of the colonies go to infinity (Nagylaki 1980). A diffusion approximation was derived later on for migration and selection in a population subdivided into an integer lattice of panmictic colonies, assuming a symmetric, finite-range migration pattern (Nagylaki 1989, 1996). In this case, mutation and random drift were neglected, and space and time scaled with respect to the intensity of selection so that the partial diffusion equation of the diffusion limit could be obtained as the intensity of selection vanishes. The results were extended to a partially selfing plant population (Nagylaki 1997).

More recently, Wakeley (2003) and Wakeley and Takahashi (2004) showed that the conditions stated in Ethier and Nagylaki (1980) hold for models of the island type (Wright 1931; Moran 1959) with demes keeping the same fixed finite size, N, and individuals the same fixed uniform migration rate, $m > 0$, as the number of demes goes to infinity. This was done in the context of two alleles under frequency-independent viability selection in a haploid population reproducing according to a birth-and-death process of the Moran type (Moran 1958) or a binomial sampling scheme of the Wright–Fisher type (Fisher 1930; Wright 1931). In both cases, mutation was neglected and migration was supposed to occur just after reproduction. In the latter, it was noticed that the model can be applied to a diploid population with genic selection and gametic migration. The main result was that, apart from a change in time scale, the many-deme diffusion limit for the allele frequencies under selection and drift was identical to the usual Wright–Fisher diffusion limit with genic selection in a large random-mating population.

In the case of reproduction of the Wright–Fisher type, Wakeley (2003) result relied on a conjecture which has been shown to be true (Lessard 2007). Actually, the conjecture reduces to saying that the probability of finding any given number of genes

of a given type in a deme chosen at random in an equilibrium population made of an infinite number of demes of N genes and in the absence of selection and mutation can be expressed as a polynomial of degree N with respect to the frequency of the given type in the whole population. This probability is ancillary to the exact probability distribution for the genetic configuration of a focal deme with respect to its ancestral genes once they are in different demes. This distribution extends the Ewens sampling formula for the infinitely-many-alleles model (Ewens 1972), owing to the analogy between a migration event to a new deme and a mutation event to a new allele.

The island model is an ideal framework to examine kin selection theory. Let us recall that kin selection theory (Hamilton 1964) suggested that all effects on individual reproductive success resulting from interactions between relatives should be transferred to the individuals that caused them, weighted by coefficients of relatedness to the individuals affected by them, originally defined as Wright (1922) coefficient of relationship. This inclusive fitness was conjectured to play the same role as Darwinian fitness in populations without interactions. Accordingly, with reference to Fisher (1930) fundamental theorem of natural selection interpreted as the increase of the mean fitness (see, e.g., Kingman 1961, for a proof in the case of one-locus multiallele viability selection in an infinite diploid random-mating population undergoing discrete, non-overlapping generations; see also Ewens 1989; Lessard 1997, for other interpretations under more general conditions), the mean inclusive fitness should tend to maximize at least in random mating populations. In particular, altruism should evolve if the coefficient of relatedness of donors to recipients exceeds the ratio of the cost to a donor over the benefit to a recipient per donor, which is known as Hamilton's rule.

Adaptive topographies and covariance approaches to predict equilibrium and stability in additive models of sibling altruism in an infinite, random-mating population seemed to support Hamilton's rule in the case of no dominance for altruism propensity (Uyenoyama and Feldman 1981; Uyenoyama et al. 1981), provided relatedness is measured by a regression coefficient, as suggested by Hamilton (1970), instead of a correlation coefficient. Evolution in family-structured populations can also be studied in the framework of fertility-viability selection, for which approximate adaptive topographies have been established in the case of weak selection (Nagylaki 1987; Lessard 1993).

Considering a wide range of group structures and local fitness functions, and assuming an infinite, random-mating population, Karlin and Matessi (1983) showed that Hamilton's rule for initial increase in frequency and stability at fixation of an altruistic allele holds if the penetrance of this allele is small, which is tantamount to conditions of very weak selection.

Developing Price (1970, 1972) covariance formula for the change in gene frequency from one generation to the next, Taylor (1989) predicted the initial increase or decrease of a mutant allele coding for a small behavioral deviation based on an inclusive fitness formulation using a covariance ratio under neutrality to measure relatedness between a donor and a recipient (Michod and Hamilton 1980), and applied his findings to a population organized into an infinite number of discrete patches within which interactions occur at random. In the absence of inbreeding, the covariance ratio reduces to

a pedigree index, namely, the expected fraction of genes in the recipient identical by descent (IBD) to one or more genes in the donor, or twice Malécot (1946) coefficient of kinship. This is not the case with inbreeding unless there is no dominance in gene action or special relationships between allele frequencies, since then the covariance ratio depends not only on identity-by-descent measures between pairs of genes carried by two individuals at the same locus (Gillois 1965), but also on average allelic effects in autozygous individuals (carrying IBD genes) and in allozygous individuals (carrying non IBD genes). In such a case, it has been suggested to measure relatedness by resorting to two pedigree indices corresponding to the situations where the donor is autozygous and allozygous, respectively (Grafen 1985; Lessard 1992).

Group selection involving extinction and recolonization of demes (Lewontin 1965; Levins 1970), or understood more generally as differential contribution of groups to the whole population through differential growth or expansion (Wright 1970, and references therein), plays a central role in population genetics to tackle evolutionary problems. Group selection is closely related to kin selection theory (see, e.g., Maynard Smith 1964; Eshel 1972; Boorman and Levitt 1973, 1980; Uyenoyama and Feldman 1980; Wade 1980). Kimura (1984), for instance, used the diffusion equation method for a large deme size and a small migration rate in the case of an infinite number of demes to describe the evolution of an altruistic trait through group selection. Assuming inter-deme competition proportional to the frequency of an altruistic allele contained in the deme, which induces a selective genic detrimental effect on its carrier within the deme, he deduced the steady-state in the case of recurrent mutation and confirmed Aoki (1982) condition on Wright (1922) fixation index for "group selection to prevail over individual selection" when mutation is neglected. Note that this conclusion relies on a decomposition of selection into an individual component within demes and a group component between demes.

Assuming diffusion approximations supported by simulations under an island model or a stepping-stone model, several authors (Cherry 2003a,b; Cherry and Wakeley 2003; Whitlock 2003) studied the effects of dominance, inbreeding, hard versus soft selection, local frequency dependence and local extinction on the probability of fixation of a mutant allele. Using a direct fitness approach for a continuous phenotype (Taylor and Frank 1996), Roze and Rousset (2003, 2004), extending Rousset and Billiard (2000), expressed this probability under those different effects in terms of probabilities of genetic identity in the neutral model and partial derivatives of the fitness function defined as the expected number of successful gametes, with respect to the phenotype of the individual, the mean phenotype in the deme and the mean phenotype in the whole population. Furthermore, extending a heuristic argument of Maruyama (1983) to compute fixation probabilities and exploiting a separation of time scales, Rousset (2006) examined the effect of isolation by distance in a haploid population. Earlier, Taylor et al. (2000) had addressed the proper definition of inclusive fitness in predicting the expected change in the frequency of a mutant allele in such a population with either a deme structure or a one-dimensional stepping-stone structure.

In this paper, we will ascertain diffusion approximations for group-structured diploid populations undergoing discrete, non-overlapping generations, with recurrent mutation and viability parameters depending on the individual genotype and the group type at a single multi-allelic locus. For the use of diffusion approximations in

population genetics, we refer to Crow and Kimura (1970), Karlin and Taylor (1981), Ewens (2004), and Lessard (2005a). A group will be made of a finite number of mating pairs and the group size will be kept fixed as the number of groups tends to infinity. In particular, the results will apply to family-structured populations. Dispersal at any fixed rate will be assumed to take place either before viability differences within groups, in which case selection will be soft, or after viability differences, and then selection will be hard, since the groups will contribute equally to the next generation in the former case and proportionally to the mean viabilities in the groups in the latter (see, e.g., Christiansen 1975). In the latter case, uniform dispersal, with groups receiving the same number of migrants, proportional dispersal, with exact replacement of all migrants, and local extinction and recolonization will be compared. The particular case of random pairwise interactions within groups will be considered.

The diffusion approximations will result from Ethier and Nagylaki (1980) conditions. The main preliminary step will be to prove that the frequencies of the group types in the case of an infinite number of groups in the absence of selection and mutation, which obey a non-linear system of recurrence equations, converge globally and uniformly to stationary limits that depend only on allele frequencies and identity measures. Since genes descending from the same ancestral gene will be identical by state (IBS) and their allelic distribution given by the allele frequencies in the whole population in the absence of mutation and selection, the key argument in the proof of this lemma will be to partition the genes in a focal group according to the ancestral genes t generations back and let t go to infinity. This shares some similarity with the coalescent for partition structures (Kingman 1982), with the possibility of simultaneous coalescence events (Pitman 1999; Sagitov 1999) and mutation events (Griffiths and Lessard 2005; Lessard 2007), but with migration playing the role of mutation and without the usual condition of exchangeability.

The diffusion approximations obtained for the different models considered in this paper in the case of a large number of small groups will make possible to address the following questions: How does the diffusion approximations compare with the classical one-locus multi-allele diffusion for a panmictic population (Kimura 1964) and the replicator equation for a deterministic game-theoretic model (Taylor and Jonker 1978)? What are the exact roles of individual competition, group competition and inbreeding in the selection drift functions and diffusion functions? Are the selection drift functions in agreement with an inclusive fitness formulation in a diploid population and what are the appropriate coefficients of relatedness coming into play?

With respect to previous works in kin selection theory in a finite population as the population size goes to infinity, the present treatment is more rigorous. Moreover, it integrates kin selection in relation with individual selection and group selection into the classical mathematical population genetics theory in a general multi-allele diploid setting and the more recent evolutionary game theory.

In Sect. 2, the group-structured population with proportional dispersal after selection and recurrent mutation is described, and the key lemma on global and uniform convergence in the case of an infinite number of groups in the absence of selection and mutation is stated. The diffusion approximation method as the number of groups becomes large is introduced in Sect. 3 and applied to the case of random pairwise interactions within groups in Sect. 4. The effects of uniform dispersal after selection,

dispersal before selection, dispersal after mating, and local extinction and recolonization after selection are considered and compared in Sect. 5. The proof of the lemma stated in Sect. 2 and some technical details in the proof of the diffusion approximation are presented in Sects. 6 and 7. Concluding remarks and perspectives follow in Sect. 8.

2 Group population structure

We consider a group structure for a finite diploid population. The population is assumed to be subdivided into a finite number D of groups made of a finite number N of mating pairs. Therefore, with respect to a single autosomal locus at which L alleles $A_1, \ldots, A_L$ are segregating, there is a finite number of group types. A group will be said to be of type i if ordering arbitrarily the mating pairs in the group, the individuals in the mating pairs and the genes at the given locus in the individuals yields the $(4N)$-dimensional vector

$$G_i = (G_{i,1}, \ldots, G_{i,4N}), \tag{2.1}$$

where $G_{i,v}$ is a specific allelic type among $A_1, \ldots, A_L$, for $i = 1, \ldots, n$, with $n = L^{4N}$.

We assume discrete, non-overlapping generations, and let $z_i = D_i/D$ be the frequency of the groups of type i, for $i = 1, \ldots, n$, in the initial generation, $t = 0$. Following reproduction, the frequency of the ordered genotype $A_k A_l$ among the offspring within a group of type i, assumed to be in infinite number, is given by $x_{kl,i}$ for $k, l = 1, \ldots, L$. The frequency of allele A_k among these offspring, which is also the frequency of A_k in their N parental pairs assuming Mendelian segregation at meiosis, is then

$$x_{k,i} = \sum_{l=1}^{L} x_{kl,i}. \tag{2.2}$$

Note that the frequency of the ordered genotype $A_k A_l$ among the offspring in the whole population is

$$x_{kl} = \sum_{i=1}^{n} z_i x_{kl,i}, \tag{2.3}$$

while the corresponding frequency of allele A_k, which is the same in all the offspring and all their parents, is

$$x_k = \sum_{i=1}^{n} z_i x_{k,i}. \tag{2.4}$$

This allele frequency can also be represented as

$$x_k = E_{\mathbf{z}}(q_{k,I}), \tag{2.5}$$

where $E_{\mathbf{z}}$ designates an expectation with respect to the current distribution of the group types, $\mathbf{z} = (z_1, \ldots, z_n)$, and $q_{k,I}$ stands for the frequency of allele A_k in an offspring I chosen at random, which is 1 if the ordered genotype of I is A_kA_k, 1/2 if it is A_kA_l or A_lA_k for $l \neq k$, and 0 otherwise.

In order to model kin selection, the viability of an offspring from conception to maturity will be assumed to depend not only on its genotype but also on its group type. It will take the general form

$$w_{kl,i} = 1 + s\sigma_{kl,i} \geq 0, \tag{2.6}$$

for an A_kA_l offspring in a group of type i. The parameter $\sigma_{kl,i} = \sigma_{lk,i}$ represents a scaled selection coefficient with respect to an intensity of selection $s \geq 0$. In the case $s = 0$, selection is neutral.

Selection changes the frequency of the ordered genotype A_kA_l and the frequency of allele A_k in a group of type i into

$$x^*_{kl,i} = \frac{w_{kl,i}x_{kl,i}}{\bar{w}_i} \tag{2.7}$$

and

$$x^*_{k,i} = \frac{w_{k,i}x_{k,i}}{\bar{w}_i}, \tag{2.8}$$

respectively, where

$$\bar{w}_i = \sum_{k,l=1}^{L} w_{kl,i}x_{kl,i} \tag{2.9}$$

and

$$w_{k,i} = \sum_{l=1}^{L} \frac{w_{kl,i}x_{kl,i}}{x_{k,i}} \tag{2.10}$$

are the mean viability and marginal viability of allele A_k, respectively, in a group of type i. The mean viability in a group of type i corresponds also to the relative size of a group of type i after selection.

To incorporate migration, it is assumed that a fixed proportion of offspring in each group, $m > 0$, disperse and are replaced with as many offspring chosen at random among all migrant offspring. Following this dispersal, called proportional dispersal, the relative sizes of the groups remain unchanged, but the frequencies of the ordered genotype A_kA_l and allele A_k in a group of type i are transformed into

$$\tilde{x}^*_{kl,i} = (1 - m)x^*_{kl,i} + mx^*_{kl,\bullet} \tag{2.11}$$

and

$$\tilde{x}^*_{k,i} = (1-m)x^*_{k,i} + mx^*_{k,\bullet}, \tag{2.12}$$

respectively, where

$$x^*_{kl,\bullet} = \frac{\sum_{j=1}^n z_j \bar{w}_j x^*_{kl,j}}{\bar{w}} \tag{2.13}$$

and

$$x^*_{k,\bullet} = \frac{\sum_{j=1}^n z_j \bar{w}_j x^*_{k,j}}{\bar{w}} \tag{2.14}$$

are the corresponding frequencies among all migrant offspring, with

$$\bar{w} = \sum_{i=1}^n z_i \bar{w}_i \tag{2.15}$$

being the mean viability in the whole population. This gives

$$\tilde{x}^*_k = (1-m)x^*_k + mx^*_{k,\bullet}, \tag{2.16}$$

where

$$x^*_k = \sum_{i=1}^n z_i x^*_{k,i}, \tag{2.17}$$

for the frequency of A_k among all offspring in the population after selection and migration.

Note that the frequency of allele A_k among all offspring in the population after selection and migration can be written as

$$\tilde{x}^*_k = x_k + (1-m)E_{\mathbf{z}}\left\{\frac{\mathrm{Cov}_i(w_I, q_{k,I})}{\bar{w}_i}\right\} + m\frac{\mathrm{Cov}_{\mathbf{z}}(w_I, q_{k,I})}{\bar{w}}, \tag{2.18}$$

where

$$\mathrm{Cov}_{\mathbf{z}}(w_I, q_{k,I}) = \sum_{i=1}^n \sum_{l=1}^L z_i w_{kl,i} x_{kl,i} - \bar{w}x_k = s\mathrm{Cov}_{\mathbf{z}}(\sigma_I, q_{k,I}) \tag{2.19}$$

and

$$E_{\mathbf{z}}\left\{\frac{\mathrm{Cov}_i(w_I, q_{k,I})}{\bar{w}_i}\right\} = \sum_{i=1}^n z_i \left\{\frac{\mathrm{Cov}_i(w_I, q_{k,I})}{\bar{w}_i}\right\}, \tag{2.20}$$

with

$$\mathrm{Cov}_i(w_I, q_{k,I}) = \sum_{l=1}^{L} w_{kl,i} x_{kl,i} - \bar{w}_i x_{k,i} = s\mathrm{Cov}_i(\sigma_I, q_{k,I}) \tag{2.21}$$

and

$$\bar{w}_i = 1 + s\bar{\sigma}_i, \tag{2.22}$$

where $\bar{\sigma}_i = \sum_{k,l=1}^{L} \sigma_{kl,i} x_{kl,i}$ is the mean scaled selection coefficient in a group of type i, for $i = 1, \ldots, n$. In the absence of selection ($s = 0$), migration has no effect on the allele frequencies in the population, and then $\tilde{x}_k^* = x_k$ for $k = 1, \ldots, L$.

Mutation is assumed to follow selection and migration with

$$u_{kl} = s\mu_{kl} \geq 0 \tag{2.23}$$

being the probability that each gene of type A_k, independently of all other genes, mutates to a gene of type A_l, for $k, l = 1, \ldots, L$, so that μ_{kl} is a scaled mutation coefficient with respect to an intensity of mutation $s \geq 0$. Using the convention $u_{kk} = 0$ for $k = 1, \ldots, L$, and defining

$$u_{kl}^* = \left(1 - \sum_{k'=1}^{L} u_{kk'}\right)\delta_{kl} + u_{kl}, \tag{2.24}$$

with δ_{kl} being the Kronecker delta, that is, 1 if $k = l$, and 0 otherwise, for $k, l = 1, \ldots, L$, the frequencies of the ordered genotype $A_k A_l$ and allele A_k in a group of type i become

$$x_{kl,i}^{**} = \sum_{k',l'=1}^{L} u_{k'k}^* u_{l'l}^* \tilde{x}_{k'l',i}^* \tag{2.25}$$

and

$$x_{k,i}^{**} = \sum_{k'=1}^{L} u_{k'k}^* \tilde{x}_{k',i}^*, \tag{2.26}$$

respectively, for $k, l = 1, \ldots, L$. This gives

$$x_k^{**} = \tilde{x}_k^* - s\sum_{l=1}^{L} \mu_{kl}\tilde{x}_k^* + s\sum_{l=1}^{L} \mu_{lk}\tilde{x}_l^* \tag{2.27}$$

for the frequency of A_k in the whole population after selection, migration and mutation, for $k = 1, \ldots, L$. In the case $s = 0$, there is no selection nor mutation, and $x_k^{**} = \tilde{x}_k^* = x_k$ for $k = 1, \ldots, L$.

The next generation, $t = 1$, is obtained by random mating and random sampling of N mating pairs within each group. Therefore, the frequency of A_k in the next generation will be

$$X_k^D(1) = \frac{1}{D} \sum_{i=1}^{n} \sum_{d_i=1}^{D_i} X_{k,i,d_i}^D(1), \tag{2.28}$$

where $X_{k,i,d_i}^D(1)$, for $d_i = 1, \ldots, D_i$ and $i = 1, \ldots, n$, are independent random variables in the form

$$X_{k,i,d_i}^D(1) = \frac{1}{2N} \sum_{\nu=1}^{2N} q_{k,I_\nu^{**}(i)}, \tag{2.29}$$

with $I_1^{**}(i), \ldots, I_{2N}^{**}(i)$ representing $2N$ offspring chosen at random and independently in a group of type i after selection, migration and mutation. For the expected value, we have

$$E_{\mathbf{z}}(X_k^D(1)) = \sum_{i=1}^{n} z_i x_{k,i}^{**} = x_k^{**}, \tag{2.30}$$

which reduces to x_k in the absence of selection and mutation. Therefore, in the limit case of an infinite number of groups, the allele frequencies in the whole population will not change if $s = 0$.

More generally, a group of type i will be transformed into a group of type j in the next generation with some probability that depends on the intensity of selection and the distribution of the group types in the current generation, $P_{ij}(s, \mathbf{z})$. Then, the frequency of the groups of type j in the next generation can be expressed as

$$Z_j^D(1) = \frac{1}{D} \sum_{i=1}^{n} \sum_{d_i=1}^{D_i} Z_{j,i,d_i}^D(1), \tag{2.31}$$

for $j = 1, \ldots, n$, where the random vectors $(D_i Z_{1,i,d_i}^D(1), \ldots, D_i Z_{n,i,d_i}^D(1))$, for $d_i = 1, \ldots, D_i$ and $i = 1, \ldots, n$, are independent and follow multinomial distributions of parameters D_i and $P_{i1}(s, \mathbf{z}), \ldots, P_{in}(s, \mathbf{z})$, respectively. In particular, the expected value is given by

$$E_{\mathbf{z}}(Z_j^D(1)) = \sum_{i=1}^{n} z_i P_{ij}(s, \mathbf{z}). \tag{2.32}$$

Note that the recurrence equations

$$z_j(t+1) = \sum_{i=1}^{n} z_i(t) P_{ij}(s, \mathbf{z}(t)), \tag{2.33}$$

for $j = 1, \ldots, n$, give the deterministic transformation for the frequencies of the group types under the assumption of an infinite number of groups.

In matrix notation, we have the transformation

$$\mathbf{z}(t+1) = \mathbf{z}(t)^T P(s, \mathbf{z}(t)), \tag{2.34}$$

where $\mathbf{z}(t) = (z_1(t), \ldots, z_n(t))$ is a frequency vector, and $P(s, \mathbf{z}(t)) = (P_{ij}(s, \mathbf{z}(t)))_{i,j=1}^n$ is a stochastic matrix. If $s = 0$, then the iterates of this transformation will converge to a limit that is a function of the allele frequencies in the population, which remain constant over time. This result will provide a key argument for a diffusion approximation in the case of a finite number of groups as this number tends to infinity.

Lemma 1 *Consider the recurrence system of equations*

$$z_j(t+1) = \sum_{i=1}^{n} z_i(t) P_{ij}(0, \mathbf{z}(t)), \tag{2.35}$$

for $j = 1, \ldots, n$, or in matrix notation

$$\mathbf{z}(t+1) = \mathbf{z}(t)^T P(0, \mathbf{z}(t)), \tag{2.36}$$

for $t \geq 0$, where $\mathbf{z}(t) = (z_1(t), \ldots, z_n(t))$ represents the frequency vector for the group types in the case of an infinite number of groups and $P_{ij}(0, \mathbf{z}(t))$ is the probability of transition from a group of type i to a group of type j in the absence of selection and mutation, for $i, j = 1, \ldots, n$. Let s_{jk} be the number of genes of type A_k in a group of type j and define $\mathbf{s}_j = (s_{j1}, \ldots, s_{jL})$ with $s_{jk} \geq 0$ for $k = 1, \ldots, L$ and $\sum_{k=1}^{L} s_{jk} = 4N$, for $j = 1, \ldots, n$. Then,

$$\lim_{t \to \infty} z_j(t) = \hat{z}_j(\mathbf{x}), \tag{2.37}$$

where

$$\hat{z}_j(\mathbf{x}) = \sum_{\mathbf{0} < \mathbf{r} \leq \mathbf{s}_j} c_j(\mathbf{r}) x_1^{r_1} \cdot \cdots \cdot x_L^{r_L}, \tag{2.38}$$

for some non negative coefficients $c_j(\mathbf{r})$ for $\mathbf{r} = (r_1, \ldots, r_L) \neq (0, \ldots, 0)$ with $0 \leq r_k \leq s_{jk}$ for $k = 1, \ldots, L$ and $j = 1, \ldots, n$, and

$$x_k = \sum_{i=1}^{n} z_i(0) x_{k,i}, \tag{2.39}$$

for $k = 1, \ldots, L$. Moreover, the convergence of $\mathbf{z}(t)$ as t goes to infinity is uniform with respect to $\mathbf{z}(0)$.

The coefficient $c_j(\mathbf{r})$ represents the number of ways that the genes in a group of type j can have r_k unrelated ancestral genes of type A_k, for $k = 1, \ldots, L$. This r_k is between 1 and s_{jk} for each type A_k represented in a group of type j. The types of unrelated ancestral genes are independent and identically distributed with probability distribution given by x_k, for $k = 1, \ldots, L$. The genes having the same ancestral gene are IBD and therefore, in the absence of mutation, IBS. The proof of Lemma 1 is relegated to Sect. 6.

3 Diffusion approximation

In this section, we shall apply a general diffusion approximation theorem for Markov chains in the case of two time scales that is due to Ethier and Nagylaki (1980).

Let $\mathbf{Z}^D(t) = (Z_1^D(t), \ldots, Z_n^D(t))$ be the frequency vector for the group types $i = 1, \ldots, n$ among D groups of N mating pairs at the beginning of some generation $t \geq 0$. The vector $\mathbf{Z}^D(t)$ has all elements that are multiples of D^{-1} and belongs to the simplex

$$\Delta_n = \left\{ \mathbf{z} = (z_1, \ldots, z_n) : z_i \geq 0, i = 1, \ldots, n, \sum_{i=1}^{n} z_i = 1 \right\}. \tag{3.1}$$

Assume a frequency $x_{kl,i}$ and a viability in the form

$$w_{kl,i} = 1 + \frac{\sigma_{kl,i}}{4ND}, \tag{3.2}$$

for offspring of genotype $A_k A_l$ in a group of type i, so that the scaled selection coefficient of $A_k A_l$ is $\sigma_{kl,i}$ with respect to an intensity of selection given by the inverse of the total number of genes, that is, $s = (4ND)^{-1}$, for $k, l = 1, \ldots, L$ and $i = 1, \ldots, n$. Similarly, let

$$u_{kl} = \frac{\mu_{kl}}{4ND} \tag{3.3}$$

be the probability of mutation from A_k to A_l in one generation, so that the scaled mutation coefficient of A_k to A_l is μ_{kl} with respect to an intensity of mutation $s = (4ND)^{-1}$, for $k, l = 1, \ldots, L$.

Define

$$X_k^D(t) = \sum_{i=1}^{n} Z_i^D(t) x_{k,i}, \tag{3.4}$$

for $k = 1, \ldots, L$, and

$$Y_j^D(t) = Z_j^D(t) - \hat{Z}_j^D(t), \tag{3.5}$$

for $j = 1, \ldots, n$, where

$$\hat{Z}_j^D(t) = \sum_{\mathbf{0}<\mathbf{r}\leq\mathbf{s}_j} c_j(\mathbf{r})(X_1^D(t))^{r_1} \cdot \cdots \cdot (X_L^D(t))^{r_L}, \tag{3.6}$$

with $c_j(\mathbf{r})$ defined in Lemma 1. Let us introduce the vectors $\mathbf{X}^D(t) = (X_1^D(t), \ldots, X_L^D(t))$ and $\mathbf{Y}^D(t) = (Y_1^D(t), \ldots, Y_n^D(t))$. Note that $\mathbf{X}^D(t)$ takes its values in the simplex

$$\Delta_L = \left\{ \mathbf{x} = (x_1, \ldots, x_L) : x_k \geq 0, k = 1, \ldots, L, \sum_{k=1}^{L} x_k = 1 \right\}. \tag{3.7}$$

Consider

$$\Delta X_k^D = X_k^D(1) - X_k^D(0) \tag{3.8}$$

and

$$\Delta Y_j^D = Y_j^D(1) - Y_j^D(0), \tag{3.9}$$

for $k = 1, \ldots, L$ and $j = 1, \ldots, n$, with $\mathbf{Z}^D(0) = \mathbf{z}$, $\mathbf{X}^D(0) = \mathbf{x}$, and $\mathbf{Y}^D(0) = \mathbf{y} = \mathbf{z} - \hat{\mathbf{z}}$, where $\hat{\mathbf{z}} = \hat{\mathbf{z}}(\mathbf{x})$ is given in Lemma 1.

The following conditions will hold:

$$\text{I. } E_{\mathbf{z}}(\Delta X_k^D) = \frac{b_k(\mathbf{x}, \mathbf{y})}{4ND} + o(D^{-1}), \tag{3.10}$$

$$\text{II. } E_{\mathbf{z}}((\Delta X_k^D)(\Delta X_l^D)) = \frac{a_{kl}(\mathbf{x}, \mathbf{y})}{4ND} + o(D^{-1}), \tag{3.11}$$

$$\text{III. } E_{\mathbf{z}}((\Delta X_k^D)^4) = o(D^{-1}), \tag{3.12}$$

$$\text{IV. } E_{\mathbf{z}}(\Delta Y_j^D) = c_j(\mathbf{x}, \mathbf{y}) + o(1), \tag{3.13}$$

$$\text{V. } \mathrm{Var}_{\mathbf{z}}(\Delta Y_j^D) = o(1), \tag{3.14}$$

as $D \to \infty$, uniformly in $\mathbf{z}$, where

$$b_k(\mathbf{x}, \mathbf{y}) = \mathrm{Cov}_{\mathbf{z}}(\sigma_I, q_{k,I}) - (1-m)\mathrm{Cov}_{\mathbf{z}}(\sigma_J, q_{k,I}) + \sum_{l=1}^{L} \mu_{lk} x_l - \sum_{l=1}^{L} \mu_{kl} x_k, \tag{3.15}$$

$$a_{kl}(\mathbf{x}, \mathbf{y}) = 2\left\{ \mathrm{Cov}_{\mathbf{z}}(q_{k,I}, q_{l,I}) - (1-m)^2 \mathrm{Cov}_{\mathbf{z}}(q_{k,I}, q_{l,J}) \right\}, \tag{3.16}$$

and

$$c_j(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{n} z_i P_{ij}(0,\mathbf{z}) - z_j, \tag{3.17}$$

with σ_I and $q_{k,I}$ being respectively the scaled selection coefficient and the frequency of allele A_k in an offspring I chosen at random in the population before selection, migration and mutation, and similarly for an offspring J chosen at random in the same group as I. Explicitly, we have

$$\mathrm{Cov}_{\mathbf{z}}(\sigma_I, q_{k,I}) = \sum_{i=1}^{n} z_i \sigma_{k,i} x_{k,i} - \bar{\sigma} x_k, \tag{3.18}$$

$$\mathrm{Cov}_{\mathbf{z}}(\sigma_J, q_{k,I}) = \sum_{i=1}^{n} z_i \bar{\sigma}_i x_{k,i} - \bar{\sigma} x_k, \tag{3.19}$$

$$\mathrm{Cov}_{\mathbf{z}}(q_{k,I}, q_{l,I}) = \sum_{i=1}^{n} z_i \left(\frac{x_{kl,i} + \delta_{kl} x_{k,i}}{2} \right) - x_k x_l, \tag{3.20}$$

$$\mathrm{Cov}_{\mathbf{z}}(q_{k,I}, q_{l,J}) = \sum_{i=1}^{n} z_i x_{k,i} x_{l,i} - x_k x_l, \tag{3.21}$$

where $\bar{\sigma}_i = \sum_{k=1}^{L} \sigma_{k,i} x_{k,i}$ with $\sigma_{k,i} = x_{k,i}^{-1} \sum_{l=1}^{L} \sigma_{kl,i} x_{kl,i}$, and $\bar{\sigma} = \sum_{i=1}^{n} \bar{\sigma}_i z_i$.

Condition I is obtained by expanding the expected change in the frequency of allele A_k, given by $(x_k^{**} - x_k)$, with respect to D^{-1} using (2.18) and (2.27) with viabilities and mutation rates in the form (3.2) and (3.3), respectively, and by noting that

$$E_{\mathbf{z}}\left\{\mathrm{Cov}_i(\sigma_I, q_{k,I})\right\} = \mathrm{Cov}_{\mathbf{z}}(\sigma_I, q_{k,I}) - \mathrm{Cov}_{\mathbf{z}}(\sigma_J, q_{k,I}), \tag{3.22}$$

where I and J are two offspring chosen at random in the same group.

Moreover, using the representation (2.28) for the allele frequencies, we find that

$$E_{\mathbf{z}}((\Delta X_k^D)(\Delta X_l^D)) = \mathrm{Cov}_{\mathbf{z}}(X_k^D(1), X_l^D(1)) + E_{\mathbf{z}}(\Delta X_k^D) E_{\mathbf{z}}(\Delta X_l^D), \tag{3.23}$$

where

$$\mathrm{Cov}_{\mathbf{z}}(X_k^D(1), X_l^D(1)) = \frac{1}{4N^2 D} \sum_{i=1}^{n} \sum_{\nu,\mu=1}^{2N} z_i \mathrm{Cov}_{\mathbf{z}}(q_{k,I_\nu^{**}(i)}, q_{l,I_\mu^{**}(i)}), \tag{3.24}$$

with $I_1^{**}(i), \ldots, I_{2N}^{**}(i)$ being $2N$ offspring chosen at random in a group of type i after selection, migration and mutation. The product of the expected changes in the frequencies of A_k and A_l is a function $o(D^{-1})$ owing to condition I. On the other hand, we have

$$\mathrm{Cov}_{\mathbf{z}}(q_{k,I_\nu^{**}(i)}, q_{l,I_\mu^{**}(i)}) = 0, \tag{3.25}$$

for all $\mu \neq \nu$, while

$$E_{\mathbf{z}}(q_{k,I_\nu^{**}(i)}) = (1-m)x_{k,i} + mx_k + o(1) \tag{3.26}$$

and

$$E_{\mathbf{z}}(q_{k,I_\nu^{**}(i)} q_{l,I_\nu^{**}(i)}) = (1-m)\left(\frac{x_{kl,i}+\delta_{kl}x_{k,i}}{2}\right) + m\left(\frac{x_{kl}+\delta_{kl}x_k}{2}\right) + o(1). \tag{3.27}$$

Then, we get

$$\mathrm{Cov}_{\mathbf{z}}(X_k^D(1), X_l^D(1)) = \frac{1}{2ND}\left\{\sum_{i=1}^{n} z_i\left(\frac{x_{kl,i}+\delta_{kl}x_{k,i}}{2}\right) - x_k x_l\right\} - \frac{(1-m)^2}{2ND}\left\{\sum_{i=1}^{n} z_i x_{k,i} x_{l,i} - x_k x_l\right\} + o(D^{-1}), \tag{3.28}$$

which is equivalent to condition II.

When $\mathbf{y} = \mathbf{0}$, that is, $\mathbf{z} = \hat{\mathbf{z}}(\mathbf{x})$, the function $a_{kl}(\mathbf{x}, \mathbf{y})$ in condition II takes a simple form.

Lemma 2 *For every $\boldsymbol{x} = (x_1, \ldots, x_L)$ in Δ_L, we have*

$$a_{kl}(\mathbf{x}, \mathbf{0}) = (1 - f_I)x_k(\delta_{kl} - x_l), \tag{3.29}$$

for $k, l = 1, \ldots, L$, where

$$f_I = \frac{(1-m)^2}{4Nm(2-m) + (1-m)^2} \tag{3.30}$$

is the inbreeding coefficient of an offspring I chosen at random before migration, namely, the probability for the two genes of I to be IBD, in the case of an infinite number of groups without selection nor mutation.

Proof We have

$$a_{kl}(\mathbf{x}, \mathbf{0}) = 2\left\{\mathrm{Cov}_{\hat{\mathbf{z}}}(q_{k,I}, q_{l,I}) - (1-m)^2 \mathrm{Cov}_{\hat{\mathbf{z}}}(q_{k,I}, q_{l,J})\right\}, \tag{3.31}$$

where I and J are two randomly chosen offspring in the same group before migration. The distribution $\hat{\mathbf{z}} = \hat{\mathbf{z}}(\mathbf{x})$ for the group types is the equilibrium distribution in the case of an infinite number of groups without selection nor mutation and with constant allele frequencies given by $\mathbf{x}$.

Then, we find

$$\mathrm{Cov}_{\hat{\mathbf{z}}}(q_{k,I}, q_{l,J}) = f_{IJ}x_k(\delta_{kl} - x_l), \tag{3.32}$$

where f_{IJ} is the probability that two genes chosen randomly and independently, one in I and one in J, are IBD. An analogous formula holds if J is replaced with I. Therefore, we get

$$a_{kl}(\mathbf{x}, \mathbf{0}) = 2\{f_{II} - (1-m)^2 f_{IJ}\}x_k(\delta_{kl} - x_l). \tag{3.33}$$

Moreover, the identity measures f_{II} and f_{IJ} are related by the equations

$$f_{II} = \frac{1}{2} + \frac{1}{2}(1-m)^2 f_{IJ}, \tag{3.34}$$

$$f_{IJ} = \frac{1}{2N} f_{II} + \left(1 - \frac{1}{2N}\right)(1-m)^2 f_{IJ}, \tag{3.35}$$

from which

$$2\{f_{II} - (1-m)^2 f_{IJ}\} = 2\{1 - f_{II}\} = 1 - f_I, \tag{3.36}$$

where f_I is the probability for the two genes of I to be IBD, as given in the statement of Lemma 2.

The above lemmas and conditions will ascertain the convergence result below (see Sect. 7 for the remaining details of the proof).

Proposition 1 *Consider the vector of allele frequencies* $\mathbf{X}^D([4ND\tau])$ *at time* $\tau \geq 0$ *in number of* $4ND$ *generations in a diploid population subdivided into* D *groups of* N *mating pairs with scaled selection coefficient* $\sigma_{kl,i}$ *with respect to an intensity of selection* $(4ND)^{-1}$ *for an offspring of genotype* A_kA_l *in a group of type* i*, a fixed proportion* $m > 0$ *of offspring in each group replaced by migrants each generation after selection as a result of proportional dispersal, and scaled mutation coefficient* μ_{kl} *for* A_k *to* A_l *with respect to an intensity of mutation* $(4ND)^{-1}$*, for* $i = 1, \ldots, n$ *and* $k, l = 1, \ldots, L$*, where* $[4ND\tau]$ *designates the integer part of* $4ND\tau$*. Then, as* $D \to \infty$*, the process* $\mathbf{X}^D([4ND\tau])$ *converges in distribution to a diffusion* $\mathbf{X}(\tau)$ *in* Δ_L *whose generator is*

$$\mathcal{L} = \frac{1}{2}\sum_{k,l=1}^{L} a_{kl}(\mathbf{x}, \mathbf{0})\frac{\partial^2}{\partial x_k \partial x_l} + \sum_{k=1}^{L} b_k(\mathbf{x}, \mathbf{0})\frac{\partial}{\partial x_k}, \tag{3.37}$$

with diffusion functions $a_{kl}(\mathbf{x}, 0)$ *given in Lemma 2, for* $k, l = 1, \ldots, L$*, and drift functions*

$$b_k(\mathbf{x}, \mathbf{0}) = \sigma_k(\mathbf{x}) + \mu_k(\mathbf{x}), \tag{3.38}$$

where

$$\mu_k(\mathbf{x}) = \sum_{l=1}^{L} \mu_{lk} x_l - \sum_{l=1}^{L} \mu_{kl} x_k, \tag{3.39}$$

and

$$\sigma_k(\mathbf{x}) = \mathrm{Cov}_{\hat{\mathbf{z}}}(\sigma_I, q_{k,I}) - (1-m)\mathrm{Cov}_{\hat{\mathbf{z}}}(\sigma_J, q_{k,I}), \tag{3.40}$$

with I and J being two offspring chosen at random in the same group before migration and $\hat{\mathbf{z}} = \hat{\mathbf{z}}(\mathbf{x})$ being given in Lemma 1, are respectively the mutation and selection drift functions for allele A_k, both with respect to the allele frequency vector $\mathbf{x}$, for $k = 1, \ldots, L$. Moreover, $\mathbf{Z}^D([\tau_D]) - \hat{\mathbf{Z}}^D([\tau_D])$ converges in probability to $\mathbf{0}$ whenever $\tau_D \to \infty$, where $\mathbf{Z}^D([\tau_D])$ is the vector of the group type frequencies at time τ_D in the model with D groups and $\hat{\mathbf{Z}}^D([\tau_D])$ the corresponding equilibrium vector in the model with an infinite number of groups, without selection nor mutation, and with allele frequency vector given by $\mathbf{X}^D([\tau_D])$.

4 Intra-group pairwise interactions

In this section, we suppose a scaled selection coefficient in the previous model determined by the individual genotype and the group type in the form

$$\sigma_{kl,i} = h_{kl} + \bar{v}_i, \tag{4.1}$$

where

$$\bar{v}_i = \sum_{k,l=1}^{L} v_{kl} x_{kl,i}, \tag{4.2}$$

with symmetric parameters $h_{kl} = h_{lk}$ and $v_{kl} = v_{lk}$, for $k, l = 1, \ldots, L$ and $i = 1, \ldots, n$. This models random pairwise interactions occurring between the offspring within the same group and having additive effects on fitness. More precisely, the scaled selection coefficient of an offspring is a sum of two different effects: one that depends on its own genotype and one that depends on the genotype of an offspring chosen at random in the same group. Then, we have

$$\mathrm{Cov}_{\hat{\mathbf{z}}}(\sigma_I, q_{k,I}) = \mathrm{Cov}_{\hat{\mathbf{z}}}(h_I, q_{k,I}) + \mathrm{Cov}_{\hat{\mathbf{z}}}(v_J, q_{k,I}) \tag{4.3}$$

and

$$\mathrm{Cov}_{\hat{\mathbf{z}}}(\sigma_J, q_{k,I}) = \mathrm{Cov}_{\hat{\mathbf{z}}}(h_J, q_{k,I}) + \mathrm{Cov}_{\hat{\mathbf{z}}}(v_J, q_{k,I}), \tag{4.4}$$

for two offspring I and J chosen at random in the same group after migration.

Conditioning on the number of genes in J having ultimately the same ancestral gene as the gene chosen at random in I in the case of an infinite number of groups, without selection nor mutation, and using the fact that unrelated ancestral genes are independent under the distribution $\hat{\mathbf{z}}$ for the group types yield (Lessard 1992, see also Grafen 1985)

$$\mathrm{Cov}_{\hat{\mathbf{z}}}(h_J, q_{k,I}) = \mathrm{Cov}_{\hat{\mathbf{z}}}(h_J, q_{k,J}|J\text{allo})\gamma_{JI} + \mathrm{Cov}_{\hat{\mathbf{z}}}(h_J, q_{k,J}|J\text{auto})\delta_{JI}, \quad (4.5)$$

where γ_{JI} is the probability for a gene chosen at random in I to be IBD to one and only one gene in J, and δ_{JI} the probability for a gene chosen at random in I to be IBD to both genes in J. In the former event, J is allozygous (abbreviated by allo), while in the latter, J is autozygous (abbreviated by auto). In the case $J = I$, we have $\delta_{II} = f_I$ and $\gamma_{II} = (1 - f_I)$. Direct calculation leads to

$$\mathrm{Cov}_{\hat{\mathbf{z}}}(h_J, q_{k,J}|J\text{allo}) = x_k h_k^{\bullet}(\mathbf{x}), \quad (4.6)$$

where

$$h_k^{\bullet}(\mathbf{x}) = \sum_{l=1}^{L} h_{kl} x_l - \sum_{l,m=1}^{L} h_{lm} x_l x_m \quad (4.7)$$

is the average effect of A_k on the genotypic value h in an allozygous offspring, and

$$\mathrm{Cov}_{\hat{\mathbf{z}}}(h_J, q_{k,J}|J\text{auto}) = 2x_k h_k^{\bullet\bullet}(\mathbf{x}), \quad (4.8)$$

where

$$h_k^{\bullet\bullet}(\mathbf{x}) = \frac{h_{kk} - \sum_{l=1}^{L} h_{ll} x_l}{2} \quad (4.9)$$

is the average effect of A_k on h in an autozygous offspring (see, e.g., Lessard 1992, 1997, and references therein, for more details).

Finally, defining the relatedness coefficients

$$\rho_{J\to I}^{\bullet} = \frac{\gamma_{JI}}{\gamma_{JJ}}, \quad \rho_{J\to I}^{\bullet\bullet} = \frac{\delta_{JI}}{\delta_{JJ}}, \quad (4.10)$$

which are the expected fractions of genes in I that are IBD to one or more genes in J at the same locus, given that J is allozygous or autozygous, respectively, Proposition 1 yields the following result.

Corollary 1 *With scaled selection coefficient for genotype $A_k A_l$ in a group of type i in the form $\sigma_{kl,i} = h_{kl} + \bar{v}_i$, for $k, l = 1, \ldots, L$ and $i = 1, \ldots, n$, the selection drift function for A_k in Proposition 1 can be expressed as*

$$\sigma_k(\mathbf{x}) = \mathrm{Cov}_{\hat{\mathbf{z}}}(h_I, q_{k,I}) - (1 - m)\mathrm{Cov}_{\hat{\mathbf{z}}}(h_J, q_{k,I}) + m\mathrm{Cov}_{\hat{\mathbf{z}}}(v_J, q_{k,I}), \quad (4.11)$$

where

$$\mathrm{Cov}_{\hat{\mathbf{z}}}(h_J, q_{k,I}) = x_k \left\{(1 - f_J)\rho^{\bullet}_{J\to I} h^{\bullet}_k(\mathbf{x}) + 2 f_J \rho^{\bullet\bullet}_{J\to I} h^{\bullet\bullet}_k(\mathbf{x})\right\}, \tag{4.12}$$

and similarly for the other covariances, with $h^{\bullet}_k(\mathbf{x})$ *and* $h^{\bullet\bullet}_k(\mathbf{x})$ *being the average effects of* A_k *on the genotypic value* h *in allozygous and autozygous individuals, respectively, with respect to the allele frequency vector* $\mathbf{x}$*, while* f_J *is the inbreeding coefficient of* J*, and* $\rho^{\bullet}_{J\to I}$ *and* $\rho^{\bullet\bullet}_{J\to I}$ *are relatedness coefficients of* J *to* I *given that* J *is allozygous and autozygous, respectively,* J *and* I *being two offspring chosen at random in the same group before migration in an infinite population without selection nor mutation.*

Note that

$$\rho^{\bullet}_{J\to I} = \rho^{\bullet\bullet}_{J\to I} = \frac{f_{JI}}{f_{JJ}}, \tag{4.13}$$

when $N = 1$, which corresponds to a partial sib-mating model (Lessard 1992). This is also the case in the limit of N large and m small such that $M = 8Nm$ is kept constant, since then the Ewens sampling formula (Ewens 1972) with migration playing the role of mutation and identity by descent that of identity by state (see, e.g., Lessard 2007, and references therein) tells us that the probability for a third gene chosen at random to be IBD to at least one of the previous two genes chosen at random is given by

$$\frac{2}{2+M} = \frac{1}{1+4Nm} = \frac{f_{JI}}{f_{JJ}}, \tag{4.14}$$

independently of the IBD status of the two previous genes.

5 Other population assumptions

In this section, we consider other assumptions about the different factors of evolution in the previous model and their ordering.

5.1 Uniform dispersal after selection

One of the main assumptions of the previous model is that the number of migrant offspring entering a group is equal to the number of migrant offspring leaving the group. Under this proportional dispersal, the group size is unaffected by migration. An alternative assumption is that each group receives the same number of migrant offspring as a result of uniform dispersal of migrants among all groups. Then, the relative size of a group of type i after migration of a proportion $m > 0$ of offspring will be $(1 - m)\bar{w}_i + m\bar{w}$, where $\bar{w}_i$ is its relative size before migration, which corresponds to the mean viability in a group of type i, and $\bar{w}$ is the mean viability in all groups. Moreover, the frequency of the ordered genotype $A_k A_l$ in a group of type i after selection and migration becomes

$$\tilde{x}^*_{kl,i} = \frac{(1-m)\bar{w}_i x^*_{kl,i} + m\sum_{j=1}^n z_j \bar{w}_j x^*_{kl,j}}{(1-m)\bar{w}_i + m\bar{w}}. \tag{5.1}$$

Summing over $l = 1, \ldots, L$ and developing with respect to the intensity of selection yield

$$\begin{aligned}\tilde{x}^*_{k,i} = (1-m)x_{k,i} + mx_k + s\left\{(1-m)\sigma_{k,i}x_{k,i} + m\sum_{j=1}^n z_j\sigma_{k,j}x_{k,j}\right\} \\ - s\left\{(1-m)\bar{\sigma}_i + m\bar{\sigma}\right\}\left\{(1-m)x_{k,i} + mx_k\right\} + o(s).\end{aligned} \tag{5.2}$$

Then, multiplying by z_i and summing over $i = 1, \ldots, n$ give

$$\tilde{x}^*_k = x_k + s\left\{\left(\sum_{i=1}^n z_i\sigma_{k,i}x_{k,i} - \bar{\sigma}x_k\right) - (1-m)^2\left(\sum_{i=1}^n z_i\bar{\sigma}_i x_{k,i} - \bar{\sigma}x_k\right)\right\} + o(s), \tag{5.3}$$

for the frequency of allele A_k in the whole population after selection and migration. The other arguments in the proof of Proposition 1 can be applied *mutatis mutandis*. Therefore, we have the following result.

Proposition 2 *Under the assumptions of Proposition 1 but with uniform dispersal, so that a fixed proportion $m > 0$ of offspring disperse uniformly among all groups each generation after selection, the same conclusion holds but with the selection drift function for A_k given by*

$$\sigma_k(\mathbf{x}) = \mathrm{Cov}_{\hat{\mathbf{z}}}(\sigma_I, q_{k,I}) - (1-m)^2\mathrm{Cov}_{\hat{\mathbf{z}}}(\sigma_J, q_{k,I}), \tag{5.4}$$

in the general case, and by

$$\sigma_k(\mathbf{x}) = \mathrm{Cov}_{\hat{\mathbf{z}}}(h_I, q_{k,I}) - (1-m)^2\mathrm{Cov}_{\hat{\mathbf{z}}}(h_J, q_{k,I}) + m(2-m)\mathrm{Cov}_{\hat{\mathbf{z}}}(v_J, q_{k,I}) \tag{5.5}$$

in the particular case of Corollary 1, for $k = 1, \ldots, L$.

5.2 Dispersal before selection

With dispersal of offspring before selection, there is no difference between uniform dispersal and proportional dispersal. Following such a dispersal, the genotype and allele frequencies in the population remain the same and the groups keep the same relative sizes. Only the frequencies within the groups are changed. If $x_{kl,i}$ and $x_{k,i}$ represent the genotype and allele frequencies, respectively, in a group of type i after migration, for $k, l = 1, \ldots, L$ and $i = 1, \ldots, n$, then the formulas for the changes due to selection apply with $m = 0$. The covariances, however, are computed with respect to the state of the population after migration. In particular, we have

$$a_{kl}(\mathbf{x}, \mathbf{0}) = 2\left\{\mathrm{Cov}_{\hat{\mathbf{z}}}(q_{k,I}, q_{l,I})\right\} = (1 - f_I)x_k(\delta_{kl} - x_l), \tag{5.6}$$

with I being an offspring chosen at random after migration in an infinite population without selection nor mutation. Note that the inbreeding coefficient f_I is then the same for I chosen after migration or before migration. Finally, we conclude the following.

Proposition 3 *Under the assumptions of Proposition 1 but with dispersal preceding selection, so that a fixed proportion $m > 0$ of offspring disperse uniformly among all groups each generation just after reproduction, the same conclusion holds, but with the selection drift function for A_k given by*

$$\sigma_k(\mathbf{x}) = \mathrm{Cov}_{\hat{\mathbf{z}}}(\sigma_I, q_{k,I}) - \mathrm{Cov}_{\hat{\mathbf{z}}}(\sigma_J, q_{k,I}) \tag{5.7}$$

in the general case, and by

$$\sigma_k(\mathbf{x}) = \mathrm{Cov}_{\hat{\mathbf{z}}}(h_I, q_{k,I}) - \mathrm{Cov}_{\hat{\mathbf{z}}}(h_J, q_{k,I}) \tag{5.8}$$

in the particular case of Corollary 1, for $k = 1, \ldots, L$, with I and J being two offspring chosen at random in the same group after migration in an infinite population without selection nor mutation.

Therefore, the only effect of dispersal before selection is to change the identity measures between the individuals in the population.

5.3 Dispersal after mating

With dispersal of a proportion $m > 0$ of mating pairs before population regulation, the changes in allele frequencies from one generation to the next have the same means as the changes with dispersal of a proportion $m > 0$ of offspring after selection. However, the covariances are not the same, since $q_{k,I_\nu^{**}(i)}$ and $q_{l,I_\mu^{**}(i)}$ are not independent random variables for $\nu \neq \mu$ with $I_\nu^{**}(i)$ and $I_\mu^{**}(i)$ in the same mating pair. In this case, we have

$$E_{\mathbf{z}}(q_{k,I_\nu^{**}(i)}q_{l,I_\mu^{**}(i)}) = (1 - m)x_{k,i}x_{l,i} + m\sum_{j=1}^{n} z_j x_{k,j}x_{l,j} + o(1), \tag{5.9}$$

from which

$$\sum_{i=1}^{n} z_i \mathrm{Cov}_{\mathbf{z}}(q_{k,I_\nu^{**}(i)}, q_{l,I_\mu^{**}(i)}) = m(2 - m)\left\{\sum_{i=1}^{n} z_i x_{k,i}x_{l,i} - x_k x_l\right\} + o(1). \tag{5.10}$$

Incorporating these terms in the covariance of the allele frequencies in the whole population from one generation to the next and ignoring terms of order $o(D^{-1})$, we find

the approximation

$$\mathrm{Cov}_{\mathbf{z}}(X_k^D(1), X_l^D(1)) \approx \frac{\mathrm{Cov}_{\mathbf{z}}(q_{k,I}, q_{l,I}) - (1 - 4m + 2m^2)\mathrm{Cov}_{\mathbf{z}}(q_{k,I}, q_{l,J})}{2ND}, \tag{5.11}$$

with I and J being two offspring chosen at random in the same group before selection, mating and migration.

On the other hand, in an infinite population without selection nor mutation, the identity measures f_{IJ} and f_{II} satisfy the relationships

$$f_{II} = \frac{1}{2} + \frac{1}{2} f_{IJ}, \tag{5.12}$$

$$f_{IJ} = \frac{1}{2N} f_{II} + \frac{1}{2N} f_{IJ} + \left(1 - \frac{1}{N}\right)(1 - m)^2 f_{IJ}, \tag{5.13}$$

from which

$$2\{f_{II} - (1 - 4m + 2m^2) f_{IJ}\} = 1 - (1 - 8m + 4m^2) f_I, \tag{5.14}$$

where $f_I = f_{IJ}$.

We conclude the following.

Proposition 4 *Under the assumptions of Proposition 1 and Corollary 1, or Proposition 2, but with $m > 0$ being the proportion of mating pairs dispersing after selection and local mating, the same conclusions hold, except that the diffusion functions are given by*

$$a_{kl}(\mathbf{x}, \mathbf{0}) = \left\{1 - (1 - 8m + 4m^2) f_I\right\} x_k(\delta_{kl} - x_l), \tag{5.15}$$

for $k, l = 1, \ldots, L$, where

$$f_I = \frac{1}{4Nm(2 - m) + (1 - m)^2} \tag{5.16}$$

is the inbreeding coefficient of an offspring I chosen at random before migration, in an infinite population without selection nor mutation.

5.4 Local extinction and recolonization after selection

If $m > 0$ represents the probability of extinction of a group after selection, independently of what happens to all other groups, and if all groups contribute in proportion to their size to colonize any extinct group, even the extinct group itself, which will not make any difference in the limit of a large number of groups, then the frequency of the ordered genotype $A_k A_l$ in a group of type i after selection, recolonization, and

mutation will be

$$x^{**}_{kl,i} = \sum_{k',l'=1}^{L} u_{k'k} u_{l'l} x^{*}_{k'l',i} \tag{5.17}$$

with probability $(1 - m)$, and

$$x^{**}_{kl,\bullet} = \sum_{k',l'=1}^{L} u_{k'k} u_{l'l} x^{*}_{k'l',\bullet} \tag{5.18}$$

with probability m. Therefore, the frequency $q_{k,I^{**}_{\nu}(i)}$ of allele A_k in an offspring $I^{**}_{\nu}(i)$ chosen at random in such a group will take the value

$$x^{**}_{k,i} = \sum_{l=1}^{L} x^{**}_{kl,i} \tag{5.19}$$

with probability $(1 - m)$, and the value

$$x^{**}_{k,\bullet} = \sum_{l=1}^{L} x^{**}_{kl,\bullet} \tag{5.20}$$

with probability m. This gives the same approximations for the expected values of $q_{k,I^{**}_{\nu}(i)}$ and $q_{k,I^{**}_{\nu}(i)} q_{l,I^{**}_{\nu}(i)}$ as before, given by Eqs. (3.26) and (3.27), respectively. But this time, $q_{k,I^{**}_{\nu}(i)}$ and $q_{l,I^{**}_{\mu}(i)}$ are not independent random variables for every $\nu \neq \mu$ $(\nu, \mu = 1, \ldots, 2N)$, since either all offspring in the same group are migrants or none of the offspring in the same group is a migrant. In this case, we have

$$E_{\mathbf{z}}(q_{k,I^{**}_{\nu}(i)} q_{l,I^{**}_{\mu}(i)}) = (1 - m) x_{k,i} x_{l,i} + m x_k x_l + o(1), \tag{5.21}$$

from which

$$\sum_{i=1}^{n} z_i \mathrm{Cov}_{\mathbf{z}}(q_{k,I^{**}_{\nu}(i)}, q_{l,I^{**}_{\mu}(i)}) = m(1 - m) \left\{ \sum_{i=1}^{n} z_i x_{k,i} x_{l,i} - x_k x_l \right\} + o(1). \tag{5.22}$$

Incorporating these terms in the covariance of the allele frequencies in the whole population from one generation to the next and ignoring terms of order $o(D^{-1})$ yield the approximation

$$\mathrm{Cov}_{\mathbf{z}}(X^{D}_{k}(1), X^{D}_{l}(1)) \approx \frac{\mathrm{Cov}_{\mathbf{z}}(q_{k,I}, q_{l,I}) + (1 - m)(2Nm - 1)\mathrm{Cov}_{\mathbf{z}}(q_{k,I}, q_{l,J})}{2ND}, \tag{5.23}$$

with I and J being two offspring chosen at random in the same group before selection, extinction and recolonization.

On the other hand, in an infinite population without selection nor mutation, the identity measures f_{IJ} and f_{II} satisfy the relationships

$$f_{II} = \frac{1}{2} + \frac{1}{2}(1-m) f_{IJ}, \tag{5.24}$$

$$f_{IJ} = \frac{1}{2N} f_{II} + \left(1 - \frac{1}{2N}\right)(1-m) f_{IJ}, \tag{5.25}$$

from which

$$2\{f_{II} + (1-m)(2Nm-1) f_{IJ}\} = (2-m)(1-f_I), \tag{5.26}$$

where $f_I = 2f_{II} - 1$.

The changes in the allelic frequencies take the same form as previously and all the other arguments for the diffusion approximation can be applied in the same way.

Proposition 5 *Under the assumptions of Proposition 1 and Corollary 1 but with $m > 0$ being the probability for a group to go extinct after selection, independently of all others, and to be recolonized proportionally by all groups, the same conclusions hold but with diffusion functions given by*

$$a_{kl}(\mathbf{x}, \mathbf{0}) = (2-m)(1-f_I) x_k (\delta_{kl} - x_l), \tag{5.27}$$

for $k, l = 1, \ldots, L$, where

$$f_I = \frac{(1-m)}{4Nm + (1-m)} \tag{5.28}$$

is the inbreeding coefficient of an offspring I chosen at random before extinction and recolonization, in an infinite population without selection nor mutation.

According to Proposition 5, the only difference between the diffusion approximations with local extinction after selection and with proportional migration after selection is a multiplicative factor in the diffusion term.

6 Proof of Lemma 1

Assume an infinite number of groups, neutral selection and no mutation, and take a focal group at random in generation $t \geq 0$. Consider the ancestral genes of the focal group in the whole population in generation 0 obtained by tracing the history of the genes of the focal group t generations back. The number of the ancestral genes is necessarily between 1 and $4N$. Some are carried by mating pairs in the focal group and others by mating pairs in other groups. Let us label the focal group with 0 and number arbitrarily the other groups that contain ancestral genes with the next positive integers. Moreover, in each group, order arbitrarily the mating pairs, the individuals in the mating pairs and the genes in the individuals. This labels the genes from 1 to

$4N$ in each group. Then, an ancestral gene can be represented by a couple of integers $\alpha = (\alpha_1, \alpha_2)$, where $1 \leq \alpha_1 \leq 4N$ is a gene label and $0 \leq \alpha_2 \leq 4N$ is a group label.

Let $\mathcal{A}(t)$ be the set of ancestral genes of the focal group t generations back. The ancestral process $\{\mathcal{A}(t)\}_{t\geq 0}$ is a Markov chain on a finite state space with initial state

$$\mathcal{A}(0) = \{(\alpha_1, 0) : \alpha_1 = 1, \ldots, 4N\}. \tag{6.1}$$

With probability one, the ancestral set $\mathcal{A}(t)$ can be partitioned into three subsets. The subset of ancestral genes in the focal group,

$$\mathcal{A}_0(t) = \{\alpha \epsilon \mathcal{A}(t) : \alpha_2 = 0\}, \tag{6.2}$$

the subset of ancestral genes that are single in groups different from the focal group,

$$\mathcal{A}_1(t) = \{\alpha \epsilon \mathcal{A}(t) : \beta_2 \neq \alpha_2 \geq 1, \forall \beta \neq \alpha, \beta \epsilon \mathcal{A}(t)\}, \tag{6.3}$$

and the subset of ancestral genes that are in pairs in groups different from the focal group,

$$\mathcal{A}_2(t) = \{\alpha \epsilon \mathcal{A}(t) : \exists! \beta \neq \alpha, \beta \epsilon \mathcal{A}(t), \beta_2 = \alpha_2 \geq 1\}. \tag{6.4}$$

With probability one, there are at most two ancestral genes in groups different from the focal group, since every migrant carries two genes, and comes from a group chosen at random among an infinite number of groups.

Let $n_0(t)$, $n_1(t)$, and $n_2(t)$ be the numbers of ancestral genes in $\mathcal{A}_0(t)$, $\mathcal{A}_1(t)$, and $\mathcal{A}_2(t)$, respectively, so that

$$n(t) = n_0(t) + n_1(t) + n_2(t) \tag{6.5}$$

is the total number of ancestral genes in $\mathcal{A}(t)$. With probability one, the number $n_0(t)$ does not increase as t increases, since any migrant from another group chosen at random almost surely does not carry ancestral genes of the focal group. Actually, this number decreases to 0 with probability one as t increases as a result of coalescence events (when two or more ancestral genes are copied from the same parental gene one generation back) or migration events (when one or more ancestral genes are carried by migrants one generation back), which occur with some positive probability as long as there remain ancestral genes in the focal group. In the model considered, we have the uniform bound

$$P(n_0(t+1) < n_0(t)) \geq m, \tag{6.6}$$

as long as $n_0(t) \geq 1$, where m is the probability for a particular individual carrying ancestral genes to be a migrant one generation back. Therefore, with probability one,

$$n_0(t) \downarrow 0 \quad \text{as } t \uparrow \infty. \tag{6.7}$$

Then, the stopping time $t_0 = \inf\{t \geq 0 : n_0(t) = 0\}$ is almost surely finite.

Similarly, with probability one,

$$n_2(t_0 + t) \downarrow 0 \quad \text{as } t \uparrow \infty. \tag{6.8}$$

In this case, we have

$$P(n_2(t_0 + t + 1) \leq n_2(t_0 + t)) = 1 \tag{6.9}$$

and

$$P(n_2(t_0 + t + 2) < n_2(t_0 + t)) \geq m, \tag{6.10}$$

for all $t \geq 0$, as long as $n_2(t_0 + t) \geq 2$. The decrease in the last inequality occurs with probability at least m in one generation in the model considered if there is at least one group at time $t_0 + t$, other than the focal group, that contains two ancestral genes in different individuals, and then one of them in particular is a migrant one generation back. Otherwise, it occurs in two generations, since it takes only one generation back for two ancestral genes in a single individual to be in different parental individuals. Note that $n_2(t)$ is always even and bounded by $4N$. Then, the stopping time $t_2 = \inf\{t \geq 0 : n_2(t) = 0\}$ is almost surely finite.

Actually, the stopping time $t_1 = \inf\{t \geq 0 : n_0(t) = n_2(t) = 0\}$ is almost surely finite and, with probability one,

$$n_1(t) \uparrow n_1(t_1) \quad \text{as } t \uparrow \infty, \tag{6.11}$$

where the variable $n_1(t_1)$ represents the ultimate number of ancestral genes in different groups other than the focal group. Therefore, with probability one,

$$n(t) \to n(t_1) \quad \text{as } t \to \infty. \tag{6.12}$$

In other words, with probability one, the ancestral genes end up in different groups other than the focal group after considering enough generations back, and then their number remains fixed. The above arguments can be adapted to the case of dispersal of mating pairs instead of individuals.

For each α in $\mathcal{A}(t)$, let $\xi_\alpha(t) \subseteq \{1, \ldots, 4N\}$ be the subset of genes in the focal group in generation t that have α as ancestral gene in generation 0. Then,

$$\xi(t) = \{\xi_\alpha(t) : \alpha \epsilon \mathcal{A}(t)\} \tag{6.13}$$

belongs to the set of partitions of $S = \{1, \ldots, 4N\}$, represented by $\mathcal{P}(S)$, since

$$\bigcup_{\alpha \epsilon \mathcal{A}(t)} \xi_\alpha(t) = S \tag{6.14}$$

and

$$\xi_\alpha(t) \bigcap \xi_\beta(t) = \emptyset, \tag{6.15}$$

for every $\beta \neq \alpha$ in $\mathcal{A}(t)$. Note that the initial partition is

$$\xi(0) = \{\{1\}, \ldots, \{4N\}\}. \tag{6.16}$$

On the other hand, with probability one,

$$\xi(t) \to \xi(t_1) \quad \text{as } t \to \infty. \tag{6.17}$$

This means that the partition $\xi(t)$ does not change almost surely backward in time, once all ancestral genes are in different groups other than the focal group. Let

$$p(\xi) = P(\xi(t_1) = \xi), \tag{6.18}$$

for every ξ in $\mathcal{P}(S)$, be the probability mass function of $\xi(t_1)$.

Note that $z_j(t)$ in Lemma 1 corresponds to the probability for a focal group chosen at random in generation t to be of type j defined by

$$G_j = (G_{j,1}, \ldots, G_{j,4N}), \tag{6.19}$$

where $G_{j,v}$ is the allelic type of gene v in the focal group, for $v = 1, \ldots, 4N$ and $j = 1, \ldots, n$. Let

$$\eta_j = \{\eta_{jk} : k = 1, \ldots, L\}, \tag{6.20}$$

where

$$\eta_{jk} = \{v \epsilon S : G_{j,v} \text{ is } A_k\}, \tag{6.21}$$

be the partition of S determined by the allelic types in a group of type j, for $j = 1, \ldots, n$. Denoting the type of the focal group in generation t by $G(t)$, we have

$$z_j(t) = P(G(t) = G_j), \tag{6.22}$$

which can be decomposed into

$$z_j(t) = P(G(t) = G_j, t \geq t_1) + P(G(t) = G_j, t < t_1). \tag{6.23}$$

First, we have

$$P(G(t) = G_j, t < t_1) \leq P(t < t_1), \tag{6.24}$$

where

$$\lim_{t \to \infty} P(t < t_1) = 0. \tag{6.25}$$

On the other hand,

$$P(G(t) = G_j, t \geq t_1) = P(G(t) = G_j | t \geq t_1) P(t \geq t_1), \tag{6.26}$$

where

$$\lim_{t \to \infty} P(t \geq t_1) = 1 \tag{6.27}$$

and

$$P(G(t) = G_j | t \geq t_1) = \sum_{\xi \epsilon \mathcal{P}(S)} P(G(t) = G_j | \xi(t) = \xi, t \geq t_1) P(\xi(t) = \xi | t \geq t_1). \tag{6.28}$$

Since all genes that are copies of the same ancestral gene are of the same allelic type, and ancestral genes that are in different groups other than the focal group in generation 0 are independently of allelic type A_k with probability given by its frequency in the whole population in generation 0, given by x_k, we have

$$P(G(t) = G_j | \xi(t) = \xi, t \geq t_1) = \prod_{k=1}^{L} x_k^{r_k}, \tag{6.29}$$

if $\xi = \{\xi_\alpha : \alpha \epsilon \mathcal{A}(\xi)\}$ is a finer partition of S than η_j, denoted by $\xi \prec \eta_j$, with

$$\mathcal{A}(\xi) = \bigcup_{k=1}^{L} \mathcal{A}_k(\xi), \tag{6.30}$$

such that

$$\eta_{jk} = \bigcup_{\alpha \epsilon \mathcal{A}_k(\xi)} \xi_\alpha, \tag{6.31}$$

and $\mathcal{A}_k(\xi)$ contains r_k elements, for $k = 1, \ldots, L$. Of course, we must have

$$r_k \leq s_{jk}, \tag{6.32}$$

where s_{jk} is the number of elements in η_{jk}, for $k = 1, \ldots, L$. Otherwise, the conditional probability is 0. Moreover, we have

$$\lim_{t \to \infty} P(\xi(t) = \xi | t \geq t_1) = p(\xi). \tag{6.33}$$

Then, we conclude that

$$\lim_{t \to \infty} P(G(t) = G_j) = \sum_{\mathbf{0} < \mathbf{r} \leq \mathbf{s}_j} c_j(\mathbf{r}) x_1^{r_1} \cdot \cdots \cdot x_L^{r_L}, \tag{6.34}$$

for $\mathbf{r} = (r_1, \ldots, r_L) \neq (0, \ldots, 0)$, with $0 \leq r_k \leq s_{jk}$ for $k = 1, \ldots, L$, where

$$c_j(\mathbf{r}) = \sum_{\xi \prec \eta_j, \mathbf{r}_j(\xi) = \mathbf{r}} p(\xi), \tag{6.35}$$

and $\mathbf{r}_j(\xi) = (r_{j1}(\xi), \ldots, r_{jL}(\xi))$, with $r_{jk}(\xi)$ being the number of non void subsets of ξ in η_{jk}, for $k = 1, \ldots, L$. Moreover, this convergence does not depend on the distribution of the group types in generation 0.

7 Proof of Proposition 1

Note that $\mathbf{X}^D(t) = \Phi(\mathbf{Z}^D(t))$ and $\mathbf{Y}^D(t) = \Psi(\mathbf{Z}^D(t))$, where $\Phi(\mathbf{z}) = \mathbf{x} = (x_1, \ldots, x_L)$ and $\Psi(\mathbf{z}) = \mathbf{z} - \hat{\mathbf{z}}(\mathbf{x}) = (z_1 - \hat{z}_1(\mathbf{x}), \ldots, z_n - \hat{z}_n(\mathbf{x}))$ are defined by Eqs. (2.4) and (2.38), respectively, for every $\mathbf{z} = (z_1, \ldots, z_n)$ in Δ_n. The applications $\Phi : \Delta_n \to \mathbb{R}^L$ and $\Psi : \Delta_n \to \mathbb{R}^n$ are continuous, and the set

$$E = \{(\mathbf{x}, \mathbf{y}) \epsilon \mathbb{R}^L \times \mathbb{R}^n : \mathbf{x} = \Phi(\mathbf{z}), \mathbf{y} = \Psi(\mathbf{z}), \mathbf{z} \epsilon \Delta_n\} \tag{7.1}$$

is compact. Moreover, the application $(\Phi, \Psi) : \Delta_n \to E$ is one-to-one and $\Phi(\Delta_n) = \Delta_L$. For every $(\mathbf{x}, \mathbf{y})$ in E, the difference equation

$$\mathbf{y}(t+1) - \mathbf{y}(t) = \mathbf{c}(\mathbf{x}, \mathbf{y}(t)), \tag{7.2}$$

for $t \geq 0$, with $\mathbf{y}(0) = \mathbf{y}$, and

$$\mathbf{c}(\mathbf{x}, \mathbf{y}(t)) = \mathbf{z}(t)^T P(0, \mathbf{z}(t)) - \mathbf{z}(t), \tag{7.3}$$

for $\mathbf{z}(t) = \mathbf{y}(t) + \hat{\mathbf{z}}(\mathbf{x})$ in Δ_n, is equivalent to the recurrence system in Lemma 1. Then, $(\mathbf{x}, \mathbf{y}(t))$ belongs to E for all $t \geq 0$, and $\mathbf{y}(t)$ converges to $\mathbf{0}$ as $t \to \infty$, uniformly with respect to $(\mathbf{x}, \mathbf{y}(0)) = (\mathbf{x}, \mathbf{y})$ in E. This ascertains the conclusion of Lemma 3.2 in Ethier and Nagylaki (1980).

On the other hand, the expression for $a_{kl}(\mathbf{x}, \mathbf{0})$ given in Lemma 2, and the fact that $b_k(\mathbf{x}, \mathbf{0}) = \sum_{l=1}^{L} \mu_{lk} x_l \geq 0$ when $x_k = 0$, guarantee that the closure of $\mathcal{L}$ defined on $\mathcal{C}^2(\Delta_L)$ generates a strongly continuous semigroup on $\mathcal{C}(\Delta_L)$ corresponding to a diffusion process in Δ_L (Ethier 1976).

Conditions I and II have already been checked in Sect. 3, and therefore it remains to ensure that conditions III, IV and V are satisfied to apply Theorem 3.3 in Ethier and Nagylaki (1980).

For the fourth moment of the change in the frequency of A_k, we have

$$\begin{aligned} E_{\mathbf{z}}((\Delta X_k^D)^4) = {} & E_{\mathbf{z}}((X_k^D(1) - x_k^{**})^4) + 3(x_k^{**} - x_k) E_{\mathbf{z}}((X_k^D(1) - x_k^{**})^3) \\ & + o(|x_k^{**} - x_k|), \end{aligned} \tag{7.4}$$

where $o(|x_k^{**} - x_k|) = o(D^{-1})$. Using (2.28) and denoting the frequency of allele A_k in any given group of type i at time $t = 1$ by $X_{k,i}^D(1)$, we deduce that

$$E_{\mathbf{z}}((X_k^D(1) - x_k^{**})^3) = \frac{1}{D^2}\sum_{i=1}^{n} z_i\, E_{\mathbf{z}}((X_{k,i}^D(1) - x_{k,i}^{**})^3), \tag{7.5}$$

which is a function $o(D^{-1})$, since the random variables $(X_{k,i,d_i}^D(1) - x_{k,i}^{**})$, for $d_i = 1, \ldots, D_i$ and $i = 1, \ldots, n$, are independent, centered at 0, and uniformly bounded in absolute value by 2. For the same reason, we have

$$\begin{aligned} E_{\mathbf{z}}((X_k^D(1) - x_k^{**})^4) = {} & \frac{1}{D^3}\sum_{i=1}^{n} z_i\, E_{\mathbf{z}}((X_{k,i}^D(1) - x_{k,i}^{**})^4) \\ & - \frac{1}{D^3}\sum_{i=1}^{n} z_i \left[E_{\mathbf{z}}((X_{k,i}^D(1) - x_{k,i}^{**})^2)\right]^2 \\ & + \frac{1}{D^2}\left[\sum_{i=1}^{n} z_i\, E_{\mathbf{z}}((X_{k,i}^D(1) - x_{k,i}^{**})^2)\right]^2, \end{aligned} \tag{7.6}$$

which is a function $o(D^{-1})$. This establishes condition III.

Using Lemma 1, Eq. (3.6) for $t = 1$ can be rewritten in the form

$$\begin{aligned} \hat{Z}_j^D(1) &= \sum_{\mathbf{0}<\mathbf{r}\le\mathbf{s}_j} c_j(\mathbf{r})\left(\prod_{k=1}^{L}(\Delta X_k^D + x_k)^{r_k}\right) \\ &= \hat{z}_j(\mathbf{x}) + \sum_{\mathbf{0}<\mathbf{v}\le\mathbf{r}\le\mathbf{s}_j} c_j(\mathbf{r},\mathbf{v},\mathbf{x})\left(\prod_{k=1}^{L}(\Delta X_k^D)^{v_k}\right), \end{aligned} \tag{7.7}$$

for some uniformly bounded non negative coefficients $c_j(\mathbf{r},\mathbf{v},\mathbf{x})$, for $\mathbf{r} = (r_1, \ldots, r_L)$ and $\mathbf{v} = (v_1, \ldots, v_L) \neq (0, \ldots, 0)$ satisfying $0 \le v_k \le r_k \le s_{jk}$, for $k = 1, \ldots, L$ and $j = 1, \ldots, n$. Therefore, recalling (2.32), we find that

$$E_{\mathbf{z}}(\Delta Y_j^D) = \sum_{i=1}^{n} z_i\, P_{ij}(s,\mathbf{z}) - z_j - \sum_{\mathbf{0}<\mathbf{v}\le\mathbf{r}\le\mathbf{s}_j} c_j(\mathbf{r},\mathbf{v},\mathbf{x})\, E_{\mathbf{z}}\left(\prod_{k=1}^{L}(\Delta X_k^D)^{v_k}\right), \tag{7.8}$$

where $s = (4ND)^{-1}$, and

$$P_{ij}(s,\mathbf{z}) = P_{ij}(0,\mathbf{z}) + o(1), \tag{7.9}$$

as D goes to infinity. In order to get condition IV, it remains to show that

$$E_{\mathbf{z}}\left(\prod_{k=1}^{L}(\Delta X_k^D)^{v_k}\right) = o(1), \tag{7.10}$$

as D goes to infinity, for every $\mathbf{0} < \mathbf{v} \leq \mathbf{r} \leq \mathbf{s}_j$. Note that

$$E_{\mathbf{z}}\left(\prod_{k=1}^{L}(\Delta X_k^D)^{v_k}\right) = E_{\mathbf{z}}\left(\prod_{k=1}^{L}(X_k^D(1) - x_k^{**})^{v_k}\right) + o(1), \tag{7.11}$$

since $(x_k^{**} - x_k) = o(1)$, for $k = 1, \ldots, L$. Using the representation (2.28) for the allele frequencies, the above product can be expressed as a mean of $D^{\sum_{k=1}^{L} v_k}$ products of $\sum_{k=1}^{L} v_k$ terms in the form $(X_{k,i,d_i}^D(1) - x_{k,i}^{**})$, each one bounded in absolute value by 2. Hence, all these products are bounded by $2^{\sum_{k=1}^{L} v_k}$. We claim that at least D of these products have an expected value 0. As a matter of fact, if $v_k > 0$, then the product of each term in the form $(X_{k,i,d_i}^D(1) - x_{k,i}^{**})$, whose expected value is 0, with all other independent terms in the form $(X_{l,j,d_j}^D(1) - x_{l,j}^{**})$ with the same (j, d_j) different from (i, d_i) will have an expected value 0. We conclude that the mean of the expected values of the products is of order D^{-1}, and therefore a function $o(1)$ as D goes to infinity.

Finally, we have

$$\mathrm{Var}_{\mathbf{z}}(\Delta Y_j^D) = \mathrm{Var}_{\mathbf{z}}(Z_j^D(1) - \hat{Z}_j^D(1)) \leq 2\mathrm{Var}_{\mathbf{z}}(Z_j^D(1)) + 2\mathrm{Var}_{\mathbf{z}}(\hat{Z}_j^D(1)), \tag{7.12}$$

where, owing to (2.31), (7.7) and (7.10),

$$\mathrm{Var}_{\mathbf{z}}(Z_j^D(1)) = \frac{1}{D}\sum_{i=1}^{n} z_i P_{ij}(s, \mathbf{z})(1 - P_{ij}(s, \mathbf{z})), \tag{7.13}$$

with $s = (4ND)^{-1}$, and

$$\begin{aligned}\mathrm{Var}_{\mathbf{z}}(\hat{Z}_j^D(1)) &= \mathrm{Var}_{\mathbf{z}}\left(\sum_{\mathbf{0}<\mathbf{v}\leq\mathbf{r}\leq\mathbf{s}_j} c_j(\mathbf{r}, \mathbf{v}, \mathbf{x})\left(\prod_{k=1}^{L}(\Delta X_k^D)^{v_k}\right)\right), \\ &= \sum_{\mathbf{0}<\mathbf{v}\leq\mathbf{r}\leq\mathbf{s}_j}\ \sum_{\mathbf{0}<\mathbf{v}'\leq\mathbf{r}'\leq\mathbf{s}_j} c_j(\mathbf{r}, \mathbf{v}, \mathbf{x}) c_j(\mathbf{r}', \mathbf{v}', \mathbf{x}) \\ &\quad \cdot \mathrm{Cov}_{\mathbf{z}}\left(\prod_{k=1}^{L}(\Delta X_k^D)^{v_k}, \prod_{k=1}^{L}(\Delta X_k^D)^{v'_k}\right),\end{aligned} \tag{7.14}$$

which are both $o(1)$ as D goes to infinity. This shows condition V.

Note that all functions $o(1)$ and $o(D^{-1})$ above are uniform in $\mathbf{z}$, since all parameters and all variables, which are finite in number, are uniformly bounded. This completes the proof of Proposition 1.

8 Concluding remarks and perspectives

The group structure model for a diploid population with migration of offspring after selection considered in Proposition 1 encompasses, and extends, the classical one-locus multi-allele viability model for a panmictic population, which corresponds to the case of complete dispersal ($m = 1$, also known as the Levene model; Levene 1953) and scaled selection coefficient depending only on the individual genotype ($\sigma_{kl,i} = h_{kl}$ for genotype A_kA_l in a group of type i for $k, l = 1, \ldots, L$ and $i = 1, \ldots, n$). This leads to the selection drift function for allele A_k given by

$$\sigma_k(\mathbf{x}) = x_k h_k^{\bullet}(\mathbf{x}), \tag{8.1}$$

where $h_k^{\bullet}(\mathbf{x}) = \sum_{l=1}^{L} h_{kl}x_l - \sum_{l,m=1}^{L} h_{lm}x_l x_m$ stands for the average effect of A_k on the genotypic value h in allozygous individuals, with respect to the allele frequency vector $\mathbf{x} = (x_1, \ldots, x_L)$. Neglecting mutation and drift, which is equivalent to considering the limit of a pure selection model in an infinite population with the inverse of the intensity of selection $(1/s)$ as unit of time as the intensity of selection goes to 0, leads to the replicator equation (Taylor and Jonker 1978; see Hofbauer and Sigmund 1998, 2003, and references therein)

$$\dot{x}_k = x_k \Big((H\mathbf{x})_k - \mathbf{x}^T H \mathbf{x} \Big), \tag{8.2}$$

for $k = 1, \ldots, L$, with symmetric fitness matrix $H = \|h_{kl}\|$ and increasing strict Lyapunov function $\mathbf{x}^T H \mathbf{x} = \sum_{k,l=1}^{L} h_{kl}x_k x_l$.

With random pairwise interactions within groups having additive effects on viability ($\sigma_{kl,i} = h_{kl} + \bar{v}_i$), and complete dispersal of offspring after selection ($m = 1$), so that there is no inbreeding, the selection drift function for allele A_k according to Corollary 1 is given by

$$\sigma_k(\mathbf{x}) = x_k(h_k^{\bullet}(\mathbf{x}) + \rho_{J\to I}^{\bullet} v_k^{\bullet}(\mathbf{x})), \tag{8.3}$$

where $\rho_{J\to I}^{\bullet} = 2f_{JI} = 1/2$ is the expected fraction of genes in I that are IBD to genes in J with I and J being two offspring chosen at random in the same group before dispersal in an infinite population without selection nor mutation. This leads to a symmetric fitness matrix

$$A = \|h_{kl} + \rho_{J\to I}^{\bullet} v_{kl}\|. \tag{8.4}$$

This is in agreement with an inclusive fitness formulation for kin selection (Hamilton 1964), a genotype A_kA_l in an offspring J having an effect h_{kl} on the fitness of J itself and an effect v_{kl} weighted by a coefficient of relatedness $\rho_{J\to I}^{\bullet}$ on the fitness of an offspring I chosen at random in the same group. More importantly, this supports the conjecture of an increase of the mean inclusive fitness in the context of a deterministic model with an infinite number of groups in the absence of inbreeding, if mutation is neglected and selection is weak enough. The result can be extended to a wider variety of group structures without inbreeding as those considered in Karlin and Matessi (1983).

Local mating and complete dispersal of mating pairs ($m = 1$) after selection introduce inbreeding and local mate competition (Hamilton 1967). For a pure viability model, that is, a viability model without interactions ($\sigma_{kl,i} = h_{kl}$), Proposition 4 yields

$$\sigma_k(\mathbf{x}) = x_k((1 - f_I)h_k^{\bullet}(\mathbf{x}) + 2f_I h_k^{\bullet\bullet}(\mathbf{x})), \tag{8.5}$$

where $f_I = 1/(4N)$ is the inbreeding coefficient. This corresponds to a fitness matrix

$$A = \|(1 - f_I)h_{kl} + f_I h_{kk}\|, \tag{8.6}$$

which is not symmetric in general. However, evolutionary game theory (see, e.g., Hofbauer and Sigmund 1998, 2003, and references therein) tells us that an evolutionarily stable strategy $\mathbf{x} = (x_1, \ldots, x_L)$ for the game matrix A in Maynard Smith and Price (1973) sense, that is, a Nash equilibrium such that $\xi^T A\xi = \xi^T H\xi < 0$, for all $\xi = (\xi_1, \ldots, \xi_L)$ with $\sum_{i=1}^{L} \xi_i = 0$ and $\xi_i = 0$ if $x_i = 0$, is an asymptotically stable rest point of the replicator dynamics. In the case of two alleles segregating in the population, the replicator dynamics can be described by Wright (1942) adaptive topography (see Nagylaki 1997, for a detailed stability analysis).

With dispersal before selection rather than after selection, complete or partial, there is pure competition within groups and only the viability differences between the offspring in the same group come into play, even in the case of interactions within the group. This corresponds to a situation of soft selection (Christiansen 1975) or local resource competition (Clark 1978), which is caused by population regulation within groups. In this case, Proposition 3 reveals a fitness matrix

$$A = \|(1 - f_J)(1 - \rho_{J\to I}^{\bullet})h_{kl} + f_J(1 - \rho_{J\to I}^{\bullet\bullet})h_{kk}\|, \tag{8.7}$$

with interactions stemming only from viability differences. This is in agreement with Whitlock (2003) when the dispersal rate m is small and the group size N large, in which case $\rho_{J\to I}^{\bullet} = \rho_{J\to I}^{\bullet\bullet} = f_{JI}/f_{JJ}$. A similar effect of competition has already been pointed out for family-structured populations (Lessard 2005b), which can also be studied from approximate adaptive topographies in fertility-viability selection models (Nagylaki 1987; Lessard 1993). Let us stress that the inbreeding coefficient f_J in Proposition 3, as well as the coefficients of relatedness $\rho_{J\to I}^{\bullet}$ and $\rho_{J\to I}^{\bullet\bullet}$ defined as the expected fractions of genes in I that are IBD to genes in J given that J is allozygous or autozygous, respectively, are computed after dispersal rather than before dispersal, in an infinite population without selection nor mutation. Moreover, they are all different from 0 if dispersal is incomplete.

Note that Wakeley (2003) model corresponds to dispersal before selection with $h_I = 2\gamma q_{l,I}$ for some mutant allele A_l and some constant $\gamma \geq 0$, in which case

$$\sigma_k(\mathbf{x}) = (1 - f_I)\gamma x_k(\delta_{kl} - x_l), \tag{8.8}$$

with inbreeding coefficient f_I given in Lemma 2.

Group selection comes into play with partial dispersal after selection ($0 < m < 1$), which corresponds to the Deakin model; Deakin 1966). In the case of intra-group pairwise interactions and proportional dispersal, for instance, Corollary 1 of Proposition 1 yields a fitness matrix in the form

$$A = \|(1 - f_J)a_{kl} + f_J b_{kk}\|, \tag{8.9}$$

with

$$a_{kl} = h_{kl} - (1 - m)\rho^{\bullet}_{J\to I} h_{kl} + m\rho^{\bullet}_{J\to I} v_{kl} \tag{8.10}$$

and

$$b_{kk} = h_{kk} - (1 - m)\rho^{\bullet\bullet}_{J\to I} h_{kk} + m\rho^{\bullet\bullet}_{J\to I} v_{kk}. \tag{8.11}$$

These parameters represent the inclusive fitness of an individual J, allozygous with genotype $A_k A_l$ or autozygous with genotype $A_k A_k$, respectively, in a group with individual competition in a fraction $1 - m$ of the group, and group competition in the complementary fraction m. Proposition 2 asserts that uniform dispersal of a fraction m of offspring has the same effect as proportional dispersal of a fraction $m(2 - m) = 1 - (1 - m)^2$ of offspring. Proposition 2 in the case of two alleles with fitness parameters in the form $h_{kl} = -cg_{kl}$ and $v_{kl} = bg_{kl}$, for $k, l = 1, 2$, and some constants b and c, corresponds to results given in Roze and Rousset (2003, 2004).

Dispersal after mating, proportional or uniform, have the same effect on the allelic selection drift functions as dispersal of offspring after selection, proportional or uniform, respectively. As ascertained in Proposition 4, the only difference is found in the effect of the population structure on the diffusion functions for a random mating population. While the population structure decreases these functions by a multiplicative factor $(1 - f_I)$, where f_I is the inbreeding coefficient, in the case of dispersal of offspring in agreement with Wakeley (2003) and Whitlock (2003), it decreases less (or even increases if $m > 1 - \sqrt{3}/2$) the same functions by a multiplicative factor $(1 - (1 - 8m + 4m^2) f_I)$ in the case of dispersal of mating pairs. This suggests an effective population size $(2ND)/(1 - f_I)$ in the former case (Wright 1943), and $(2ND)/(1 - (1 - 8m + 4m^2) f_I)$ in the latter, compared to $2ND$ for a random mating population. Therefore, drift in a group-structured population is stronger with dispersal of mating pairs than with dispersal of offspring, and even stronger than drift in a random mating population if the dispersal rate is high enough. These will be the only differences in the diffusion approximations in the absence of selection.

Group selection is often modeled with local extinction and recolonization. The effect of these factors on the allelic selection drift functions is the same as proportional dispersal after selection, if local extinction occurs after selection with probability $m > 0$ equal to the dispersal rate, and all groups, including groups going extinct, contribute to recolonization in the same proportions as their relative sizes after selection. In this case, the population structure decreases the diffusion functions for a random mating population by a multiplicative factor $(2 - m)(1 - f_I)$, but with inbreeding coefficient f_I given in Proposition 5. This factor is smaller than the factor $(1 - f_I)$

for proportional dispersal with f_I given in Lemma 2, leading to a larger effective population size, if and only if $4Nm < 1$. The effective population size given by $(2ND)/((2-m)(1-f_I))$ is in agreement with previous works under corresponding assumptions at least in the case where the extinction probability is small (see, e.g., Slatkin 1977; Wade and McCauley 1988; Wakeley and Aliacar 2001; Rousset 2003; Whitlock 2003).

Note that the diffusion approximation studied by Kimura (1984) concerns the density of the frequency of an altruistic allele among an infinite number of demes. It assumes an island model with a large deme size even though a small deme size seems to be more appropriate for extinction and recolonization, and genic selection at both the individual and group levels, whose intensity is proportional to time. Group competition is modeled by an extinction-recolonization rate that is a linear function of the frequency of the altruistic allele. The model also incorporates mutation and dispersion occurring before group competition and acting with the same intensity as selection. Similar assumptions have been made for other diffusion equations modeling group selection (see, e.g., Levins 1970; Boorman and Levitt 1973). Under such assumptions, a two-time-scale argument cannot be applied, since there is no evolutionary force that dominates the others. Moreover, the variability between groups, which is a necessary ingredient for group selection to occur, is then a result of local isolation that is maintained by weak migration.

On the other hand, a two-time-scale argument has been used for a population geographically structured into a fixed finite number of demes that exchange migrants according to a fixed ergodic scheme as the deme sizes go to infinity (Nagylaki 1980). Then, the limiting diffusion result relies on the fact that the allele frequencies within the demes converge globally and uniformly to the same limits under the mixing effects of migration, before these frequencies change under the evolutionary effects of selection and mutation. The result is a consequence of the Perron–Frobenius theory (see, e.g., Gantmacher 1959, or Karlin and Taylor 1975) applied to a constant backward migration matrix, at least in the case of soft viability selection determined by the individual genotype. Actually, the limiting diffusion, known as the strong-migration limit, is analogous to the one obtained for a panmictic population provided the effective population size and the selection coefficients are appropriately defined with respect to the deme sizes and the stationary distribution of the backward migration matrix, which is not given by the relative deme sizes unless migration is conservative. Note that, in this model, the variability in the genetic structure of the groups vanishes in the limit of a large population.

The diffusion approximations ascertained in Propositions 1–5 hold for groups of a fixed finite number of mating pairs, either going extinct and being recolonized by offspring produced in the whole population with a fixed probability each generation, or producing a fixed proportion of migrant offspring each generation that disperse in the whole population. The variability between groups is maintained in this case by the finite size of the groups as the number of groups tends to infinity. This can be seen as an extension of individual selection, which arises from groups of single individuals producing gametes that disperse in the whole population. The proofs rely on applying Ethier and Nagylaki (1980) diffusion result on two-time-scale Markov chains and the main step has been to show global and uniform convergence of the distribution

of the group types in an infinite population in the absence of selection and mutation. The transformation of this distribution involves a stochastic matrix that depends on the population state, and convergence to an equilibrium distribution resorts to a coalescence approach (Kingman 1982; Pitman 1999; Sagitov 1999), extended to a single deme in an infinite collection of demes of finite size with the possibility of multiple simultaneous migration events and coalescence events. The equilibrium distribution given in Lemma 1 is obtained by conditioning on the relatedness structure of the deme at stationarity, that is, the number of ancestral genes and the multiple ways that their descendants can be distributed within the deme. In the particular case of genic selection and gametic migration, the genes in the deme are exchangeable and the relatedness structure of the deme has a probability distribution given by an extension of the Ewens sampling formula to an exact Wright–Fisher population with the mutation rate to a new allele corresponding to the migration rate to a new deme (Lessard 2007).

The diffusion approximations in Propositions 1–5 are valid for general selection coefficients in a group-structure model for a diploid population that depend on the individual genotype and the group type. Interactions within groups, not necessarily pairwise, may occur between parents and offspring, and have different effects, not necessarily additive, according to sex, ploidy, or the degree of relatedness between the interacting individuals. In such cases, however, extended coefficients of relatedness as in Lessard and Rocheleau (2004) for family-structured populations may have to be considered.

References

Aoki K (1982) A condition for group selection to prevail over counteracting individual selection. Evolution 36:832–842

Boorman S, Levitt PR (1973) Group selection on the boundary of a stable population. Theor Popul Biol 4:85–128

Boorman S, Levitt PR (1980) The genetics of altruism. Academic Press, New York

Cherry JL (2003a) Selection in a subdivided population with dominance or local frequency dependence. Genetics 163:1511–1518

Cherry JL (2003b) Selection in a subdivided population with local extinction and recolonization. Genetics 164:789–795

Cherry JL, Wakeley J (2003) A diffusion approximation for selection and drift in a subdivided population. Genetics 163:421–428

Christiansen FB (1975) Hard and soft selection in a subdivided population. Am Nat 109:11–16

Clark AB (1978) Sex ratio and local resource competition in a prosimian primate. Science 201:163–165

Crow JF, Kimura M (1970) An introduction to population genetics theory. Harper and Row, New York

Deakin MAB (1966) Sufficient conditions for genetic polymorphism. Am Nat 100:690–692

Eshel I (1972) On the neighbor effect and the evolution of altruistic traits. Theor Popul Biol 3:258–277

Ethier SN (1976) A class of degenerate diffusion processes occurring in population genetics. Comm Pure Appl Math 29:483–493

Ethier SN, Nagylaki T (1980) Diffusion approximations of Markov chains with two time scales and applications to population genetics. Adv Appl Prob 12:14–49

Ewens WJ (1972) The sampling theory of selectively neutral alleles. Theor Popul Biol 3:87–112

Ewens WJ (1989) An interpretation and proof of the fundamental theorem of natural selection. Theor Popul Biol 36:167–180

Ewens WJ (2004) Mathematical population genetics, 2nd edn. Springer, New York

Fisher RA (1930) The genetical theory of natural selection. Clarendon Press, Oxford

Gantmacher FR (1959) The theory of matrices, vol 2. Chelsea, New York

Gillois M (1965) Relation d'identité en génétique I.—Postulats et axiomes mendéliens. Ann Inst Henri Poincaré B2:1–94
Grafen A (1985) A geometric view of relatedness. Oxford Surv Evol Biol 2:28–89
Griffiths WJ, Lessard S (2005) Ewens' sampling formula and related formulae: combinatorial proofs, extensions to variable population size and applications to ages of alleles. Theor Popul Biol 68:167–177
Hamilton WD (1964) The genetical evolution of social behaviour I and II. J Theor Biol 7:1–52
Hamilton WD (1967) Extraordinary sex ratios. Science 156:477–488
Hamilton WD (1970) Selfish and spiteful behaviour in an evolutionary model. Nature 228:1218–1220
Hofbauer J, Sigmund K (1998) Evolutionary games and population dynamics. Cambridge University Press, Cambridge
Hofbauer J, Sigmund K (2003) Evolutionary game dynamics. Bull Am Math Soc 40:479–519
Karlin S, Taylor HM (1975) A first course in stochastic processes, 2nd edn. Academic Press, New York
Karlin S, Taylor HM (1981) A second course in stochastic processes. Academic Press, New York
Karlin S, Matessi C (1983) Kin selection and altruism. Proc R Soc Lond Ser B 219:327–353
Kimura M (1964) Diffusion models in population genetics. J Appl Prob 1:177–232
Kimura M (1984) Evolution of an altruistic trait through group selection as studied by the diffusion equation method. IMA J Math Appl Med Biol 1:1–15
Kingman JFC (1961) A matrix inequality. Quart J Math 12:78–80
Kingman JFC (1982) The coalescent. Stoch Proc Appl 13:235–248
Kurtz TG (1975) Semigroups of conditioned shifts and approximation of Markov processes. Ann Prob 3:618–642
Lessard S (1992) Relatedness and inclusive fitness with inbreeding. Theor Popul Biol 42:284–307
Lessard S (1993) Adaptive topography in fertility-viability selection models: an alternative to inclusive fitness in kin selection models. Theor Popul Biol 43:281–309
Lessard S (1997) Fisher's fundamental theorem of natural selection revisited. Theor Popul Biol 52:119–136
Lessard S (2005a) Long-term stability from fixation probabilities in finite populations: New perspectives for ESS theory. Theor Popul Biol 68:19–27
Lessard S (2005b) Kin selection is implicated in partial sib-mating populations with constant viability differences before mating. Genetics 171:407–413
Lessard S (2007) An exact sampling formula for the Wright–Fisher model and a conjecture about the finite-island model. Genetics 177:1249–1254
Lessard S, Rocheleau G (2004) Kin selection and coefficients of relatedness in family-structured populations with inbreeding. Theor Popul Biol 66:287–306
Levene H (1953) Genetic equilibrium when more than one ecological niche is available. Am Nat 87:311–313
Levins R (1970) Extinction. In: Gerstenhaber M (ed) Some mathematical questions in biology. Lectures on Mathematics in the Life Sciences, vol 2. American Mathematical Society, Providence, pp 75–107
Lewontin RC (1965) Selection in and of populations. In: Moore JA (ed) Ideas in modern biology. Natural History Press, New York, pp 292–311
Malécot G (1946) La consanguinité dans une population limitée. C R Acad Sci Paris 222:841–843
Maruyama T (1983) Stochastic theory of population genetics. Bull Math Biol 45:521–554
Maynard Smith J (1964) Group selection and kin selection. Nature 201:1145–1147
Maynard Smith J, Price G (1973) The logic of animal conflicts. Nature 246:15–18
Michod RE, Hamilton WD (1980) Coefficients of relatedness in sociobiology. Nature 288:694–697
Moran PAP (1958) Random processes in genetics. Proc Camb Phil Soc 54:60–71
Moran PAP (1959) The theory of some genetical effects of population subdivision. Austr J Biol Sci 12:109–116
Nagylaki T (1980) The strong-migration limit in geographically structured populations. J Math Biol 9:101–114
Nagylaki T (1987) Evolution under fertility and viability selection. Genetics 115:367–375
Nagylaki T (1989) The diffusion model for migration and selection. In: Hastings A (ed) Some mathematical questions in biology: models in population biology. Lectures on Mathematics in the Life Sciences, vol 20. American Mathematical Society, Providence, pp 55–75
Nagylaki T (1996) The diffusion model for migration and selection in a dioecious population. J Math Biol 34:334–360
Nagylaki T (1997) The diffusion model for migration and selection in a plant population. J Math Biol 35:409–431
Pitman J (1999) Coalescents with multiple collisions. Ann Probab 27:1870–1902

Price GR (1970) Selection and covariance. Nature (London) 227:520–521
Price GR (1972) Extension of covariance selection mathematics. Am Hum Genet Lond 35:485–490
Rousset F (2003) Effective size in simple metapopulation models. Heredity 91:107–111
Rousset F (2006) Separation of time scales, fixation probabilities and convergence to evolutionarily stable states under isolation by distance. Theor Popul Biol 69:165–179
Rousset F, Billiard S (2000) A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. J Evol Biol 13:814–825
Roze D, Rousset F (2003) Selection and drift in subdivided populations: a straightforward method for deriving diffusion approximations and applications involving dominance, selfing and local extinctions. Genetics 165:2153–2166
Roze D, Rousset F (2004) The robustness of Hamilton's rule with inbreeding and dominance: Kin selection and fixation probabilities under partial sib-mating. Am Nat 164:214–231
Sagitov S (1999) The general coalescent with asynchronous mergers of ancestral lines. J Appl Probab 36:1116–1125
Slatkin M (1977) Gene flow and genetic drift in a species subject to frequent local extinctions. Theor Popul Biol 12:253–262
Taylor PD (1989) Evolutionary stability in one-parameter models under weak selection. Theor Popul Biol 36:125–143
Taylor PD, Frank SA (1996) How to make a kin selection model. J Theor Biol 180:27–37
Taylor PD, Irwin AJ, Day T (2000) Inclusive fitness in finite deme-structured and stepping-stone populations. Selection 1:153–163
Taylor PD, Jonker L (1978) Evolutionarily stable strategies and game dynamics. Math Biosc 40:145–156
Uyenoyama MK, Feldman MW (1980) Theories of kin and group selection: a population genetics perspective. Theor Popul Biol 17:380–414
Uyenoyama MK, Feldman MW (1981) On relatedness and adaptive topography in kin selection. Theor Popul Biol 19:87–123
Uyenoyama MK, Feldman MW, Mueller LD (1981) Population genetic theory of kin selection: multiple alleles at one locus. Proc Natl Acad Sci USA 78:5036–5040
Wade MJ (1980) Kin selection: its components. Science 210:665–667
Wade M, McCauley D (1988) Extinction and recolonization: their effects on the genetic differentiation of local populations. Evolution 42:995–1005
Wakeley J (2003) Polymorphism and divergence for island-model species. Genetics 163:411–420
Wakeley J, Aliacar N (2001) Gene genealogies in a metapopulation. Genetics 159:893–905
Wakeley J, Takahashi T (2004) The many-demes limit for selection and drift in a subdivided population. Theor Popul Biol 66:83–91
Whitlock MC (2003) Fixation probability and time in subdivided populations. Genetics 164:767–779
Wright S (1922) Coefficients of inbreeding and relationship. Am Nat 56:330–338
Wright S (1931) Evolution in Mendelian populations. Genetics 16:97–159
Wright S (1942) Statistical genetics and evolution. Bull Am Math Soc 48:223–246
Wright S (1943) Isolation by distance. Genetics 28:114–138
Wright S (1970) Random drift and the shifting balance theory of evolution. In: Kojima K (ed) Mathematical topics in population genetics. Springer, Berlin, pp 1–31