Theoretical Population Biology 82 (2012) 131–145

Contents lists available at SciVerse ScienceDirect

Theoretical Population Biology

journal homepage: www.elsevier.com/locate/tpb

Effect of epistasis and linkage on fixation probability in three-locus models: An ancestral recombination–selection graph approach

Amir R. Kermany [a,b], Sabin Lessard [a,*]

[a] Département de mathématiques et de statistique, Université de Montréal, Montréal (Québec) H3C 3J7, Canada

[b] Department of Human Genetics, University of Chicago, Chicago, IL, USA

ARTICLE INFO

Article history:
Received 29 February 2012
Available online 23 May 2012

Keywords:
Evolution of recombination
Ancestral recombination graph
Ancestral selection graph
Ancestral influence graph
Modifier theory
Hill–Robertson effect

ABSTRACT

We study the probability of ultimate fixation of a single new mutant arising in an individual chosen at random at a locus linked to two other loci carrying previously arisen mutations. This is done using the Ancestral Recombination–Selection Graph (ARSG) in a finite population in the limit of a large population size, which is also known as the Ancestral Influence Graph (AIG). An analytical expansion of the fixation probability with respect to population-scaled recombination rates and selection intensities is obtained. The coefficients of the expansion are expressed in terms of the initial state of the population and the epistatic interactions among the selected loci. Under the assumption of weak selection at tightly linked loci, the sign of the leading term, which depends on the signs of epistasis and initial linkage disequilibrium, determines whether an increase in recombination rates increases the chance of ultimate fixation of the new mutant. If mutants are advantageous, this is the case when epistasis is positive or null and the initial linkage disequilibrium is negative, which is an expected state in a finite population under directional selection. Moreover, this is also the case for a neutral mutant modifier coding for higher recombination rates if the same conditions hold at the selected loci. Under the same conditions, deleterious mutants are disfavored for ultimate fixation and neutral modifiers for higher recombination rates still favored. The recombination rates between the modifier locus and the selected loci do not come into play in the leading terms of the approximation for the fixation probability, but they do in higher-order terms.

One of the first formal arguments for the evolutionary advantage of recombination is given in the early works of Fisher (1930) and Muller (1932). They argued that the evolutionary advantage of recombination is to bring beneficial mutants arising at different loci together on the same chromosome. As a result, recombination would increase the rate of evolution (Crow and Kimura, 1965). However, Maynard Smith (1968) showed that in an infinite population initially at linkage equilibrium and in the absence of epistasis, which then maintains this equilibrium later on, recombination has no effect on the rate of evolution.

Since recombination can affect the rate of evolution only by diminishing linkage disequilibrium (LD) in the population, this effect should be positive if and only if linkage disequilibrium is negative (see, e.g., Eshel and Feldman, 1970). In the case of a two-locus model, LD is negative when the frequency of the double mutant is strictly smaller than the product of the frequencies of the mutant alleles. This situation is arguably likely to happen when mutant alleles are rare (Crow and Kimura, 1969). Moreover, negative LD could be produced by synergistic epistasis among deleterious mutants, which corresponds to a situation of negative epistasis. In a population exhibiting negative LD, recombination renders directional selection more effective by increasing the variance in fitness.

In a finite population subject to selection and recombination and in initial linkage equilibrium, genetic drift creates random instances of linkage disequilibrium. Although drift can generate both positive and negative disequilibrium without any *a priori* bias, selection promotes the latter. As shown by Hill and Robertson (1966), the joint effect of drift and selection leads to an average accumulation of negative LD, even in the absence of epistasis. As a consequence, provided that epistasis is non-negative, responses to selection at different loci are expected to interfere with each other, even in the absence of direct interaction between the loci. This phenomenon, known as the Hill–Robertson (HR) effect, predicts that an increased recombination rate, by reducing the interference caused by the randomly-generated linkage disequilibrium, enhances the rate of fixation of favorable mutants. The relationship with the Fisher–Muller theory for the advantage of higher recombination rates and Muller's ratchet mechanism

* Correspondence to: Département de mathématiques et de statistique, C.P. 6128, Succursale Centre-ville, Université de Montréal, Montréal (Québec) H3C 3J7, Canada.

E-mail addresses: kermany@uchicago.edu (A.R. Kermany), lessards@dms.umontreal.ca (S. Lessard).

0040-5809/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.tpb.2012.05.002

(Muller, 1964) for a greater accumulation of deleterious mutants in the absence of recombination was pointed out by Felsenstein (1974).

The amount and effect of linkage disequilibrium and epistasis in a finite population with tight linkage and weak selection were first studied by diffusion approximations (Hill and Robertson, 1966; Ohta, 1968; Ohta and Kimura, 1969). The first-order effects of recombination and selection are obtained by differentiating with respect to population-scaled parameters, as in Robertson (1962) for weak selection relative to drift. In the case of initial linkage equilibrium and no epistasis, the first-order effects of selection and recombination on the probability of fixation of an allele vanish. Higher-order effects were first exhibited by simulations (Hill and Robertson, 1966).

Assuming an infinite random mating population with no recurrent mutation, Feldman et al. (1980) considered the initial change in the frequency of a new modifier allele for the rate of recombination among the loci under selection. Considering various regimes for epistatic interactions they concluded that the frequency of a new modifier for recombination which is introduced into a population initially at a polymorphic genetic (viability-analogous) equilibrium, increases in frequency only if it reduces the amount of recombination. This conclusion was later extended to include modifiers of migration or mutation parameters. If the modifier allele is tightly linked to the other loci, an allele increasing the parameter value decreases in frequency when it is rare. This is known as the reduction principle (Feldman and Liberman, 1986). Actually, it was later shown (Zhivotovsky et al., 1994) that this is also true in the case of a multi-locus model where the modifiers are loosely linked to the other loci.

The reduction principle is valid as long as genetic drift and recurrent mutations as well as environmental fluctuations are ignored, and the population is initially at a polymorphic genetic equilibrium. If the population is at a balance between selection against deleterious alleles and mutation towards them, then a modifier increasing the recombination rate succeeds in invading the population if linkage disequilibrium is negative (Feldman et al., 1980; Otto and Feldman, 1997).

It has been argued that a modifier allele that increases the recombination rate would be promoted in a population due to its role in reducing the negative effect of background selection on the probability of fixation of a new favorable mutant, at least in the absence of epistasis (Otto and Barton, 1997). In this approach a branching process forward in time is applied to selected mutant lines in an infinite population as in Barton (1995) and modifier alleles hitchhike along as the mutants rise in frequency.

Simulations have indicated that the conclusion is true across a broad range of epistatic interactions, from weak negative epistasis to positive epistasis, at least, provided that the population size is small enough and selection is strong enough (Otto and Barton, 2001).

A perturbation method to track fluctuations in linkage disequilibrium during the spread of beneficial alleles and to measure the impact on a modifier allele of recombination has been proposed (Barton and Otto, 2005). The method consists in considering only the first and second moments of random sampling effects on the deterministic dynamics for the allele frequencies and linkage disequilibria determining the genetic background in an infinite population.

Another perturbation technique to approximate the probability of ultimate fixation of an allele in a multilocus setting assumes small selection effects at different loci in a population of fixed finite size (Lehman and Rousset, 2009). This is an extension of a direct Markov chain approach for one-locus models based on expected changes in allele frequencies (Rousset, 2003; Lessard and Ladret, 2007; Lessard and Lahaie, 2009). Then, the first-order effect of selection can be expressed in terms of mean times that lineages for sampled genes take to merge backward in time, under neutrality. The calculation of these mean times for one-locus models in the limit of a large population size makes use of the coalescent (Kingman, 1982), and its extension to incorporate multiple mergers in the case of highly skewed reproduction schemes (Pitman, 1999; Sagitov, 1999; Möhle and Sagitov, 2001).

In the case of multilocus Wright–Fisher models with recombination and no epistasis, Lehman and Rousset (2009) deduce exact linear systems of equations to trace gamete frequencies backward in time. They use matrix theory to sum iterates of these equations in order to approximate the fixation probability. They conclude that a third-order approximation with respect to the intensity of selection is necessary in order to detect the Hill–Robertson effect.

In this paper, we focus on a discrete-time three-locus selection Moran model (Moran, 1958) and we use the ancestral recombination–selection graph for sampled sequences in the limit of a large population size, known as the ancestral influence graph, AIG (Donnelly and Kurtz, 1999), in order to study the fixation probability. This provides a supra-genealogy that combines the ancestral selection graph, ASG (Krone and Neuhauser, 1997), and the ancestral recombination graph, ARG (Griffiths, 1981; Hudson, 1983; Griffiths and Marjoram, 1996). The objective is to obtain an analytic approximation for the probability of ultimate fixation of a mutant allele in a finite population under weak selection and tight linkage, and investigate the effect of recombination on the fixation probability. This is an extension of a previous analysis for two-locus selection models (Lessard and Kermany, 2012).

First we consider the case of directional selection at three loci with fitness parameters depending only on the number of mutants. A single mutant allele is introduced at random at the first locus, into a population where previous mutants are already segregating at the other two loci. We then derive a polynomial approximation for the probability of ultimate fixation of the new mutant allele with respect to small population-scaled recombination and selection parameters, up to the leading term involving recombination. Exact conditions for a small increase in the recombination rate to increase the fixation probability are deduced.

Then we investigate the probability of ultimate fixation of a neutral modifier allele that increases the rate of recombination among the loci under consideration assuming that recombination is parent-dependent. We derive the fixation probability in terms of epistatic effects and initial linkage disequilibrium among the loci under selection, and the dominance effect at the modifier locus. In doing so, we extend the ancestral recombination–selection graph to accommodate the possibility of perturbed recombination rates among the loci.

1. The model

Assume a discrete-time Moran model (Moran, 1958) for a haploid population of size N. Without loss of generality, we assume that each individual of the population occupies a site unique to that individual. This corresponds to an arbitrary labeling of the N individuals in the population and does not imply any population structure. At each time step two individuals are chosen at random with replacement and they produce one offspring. At the same time one individual is chosen at random to be replaced by the offspring. Replacement takes place with some probability given by the mortality of the individual chosen to be replaced. With the complementary probability, replacement does not occur and the offspring dies.

Consider three ordered loci, 1, 2 and 3, with two alleles at each locus. The mutant (more recent) allele at each locus is denoted by 1 and the resident (older) allele by 0. An individual type is represented by a binary sequence in the form

$$\mathbf{i} = i_1 i_2 i_3,$$

where $i_l = 1$ or 0 depending on which allele is at locus l for $l = 1, 2, 3$. The number of mutant alleles in type $\mathbf{i}$ is represented by $|\mathbf{i}| = i_1 + i_2 + i_3$.

Let $R(\mathbf{i}|\mathbf{j}, \mathbf{k})$ be the transmission probability that an offspring produced by parents of types $\mathbf{j} = j_1j_2j_3$ and $\mathbf{k} = k_1k_2k_3$ in this order is of type $\mathbf{i} = i_1i_2i_3$ with $i_1 = j_1$. By convention the first parent is the one that transmits its allele at locus 1 to the offspring.

Assuming no interference, the transmission probability can be expressed in terms of the recombination fraction between loci 1 and 2, denoted by $r_1(\mathbf{j}, \mathbf{k})$, and the recombination fraction between loci 2 and 3, represented by $r_2(\mathbf{j}, \mathbf{k})$. We assume that

$$r_1(\mathbf{j}, \mathbf{k}) = N^{-1}\nu_{j_1+k_1}, \quad (1)$$
$$r_2(\mathbf{j}, \mathbf{k}) = N^{-1}\rho_{j_1+k_1},$$

for $0 \le \nu_0 \le \nu_1 \le \nu_2$ and $0 \le \rho_0 \le \rho_1 \le \rho_2$.

Therefore, the recombination fractions are assumed to be increasing functions with respect to the number of mutant alleles at locus 1. The parameters ν_l and ρ_l, for $l = 0, 1, 2$, represent population-scaled recombination rates.

Assuming N large so that the loci are tightly linked and terms of order $O(N^{-2})$ can be ignored, the offspring is an exact copy of the first parent with probability $1 - r_1(\mathbf{j}, \mathbf{k}) - r_2(\mathbf{j}, \mathbf{k})$, and a recombinant with the complementary probability. Therefore, we have

$$R(\mathbf{i}|\mathbf{j}, \mathbf{k}) \approx (1 - r_1(\mathbf{j}, \mathbf{k}) - r_2(\mathbf{j}, \mathbf{k}))\,\delta_{\mathbf{i},\mathbf{j}} + r_1(\mathbf{j}, \mathbf{k})\delta_{\mathbf{i},j_1k_2k_3} + r_2(\mathbf{j}, \mathbf{k})\delta_{\mathbf{i},j_1j_2k_3}. \quad (2)$$

Here $\delta_{\mathbf{i},\mathbf{j}}$ denotes the Kronecker delta (1 if $\mathbf{i} = \mathbf{j}$, and 0 otherwise).

Now, if the individual chosen to be replaced by the offspring produced is of type $\mathbf{i}$, then replacement actually occurs with some probability denoted by $m_{\mathbf{i}}$. This mortality parameter is expressed in the form

$$m_{\mathbf{i}} = 1 - c_{\mathbf{i}}s, \quad (3)$$

where $0 \le c_{\mathbf{i}} \le 1$ represents the coefficient of selection with respect to an intensity of selection $0 < s = \sigma N^{-1} < 1$. The parameter σ represents the population-scaled intensity of selection.

In the case of advantageous mutations, an individual of type $\mathbf{1} = 111$ has the lowest mortality, while an individual of type $\mathbf{0} = 000$ has the highest, since then

$$c_{111} \ge \max(c_{110}, c_{101}, c_{011}) \ge \min(c_{100}, c_{010}, c_{001}) \ge c_{000}.$$

The opposite inequalities

$$c_{111} \le \min(c_{110}, c_{101}, c_{011}) \le \max(c_{100}, c_{010}, c_{001}) \le c_{000}$$

hold in the case of deleterious mutations.

There is epistatic interaction among the mutant alleles if their combined effects on fitness deviates from independent effects. In order to measure this interaction, the relative effects of the mutants taken separately are compared to their relative effects taken together.

The relative mortality of type $\mathbf{i}$ with respect to the mortality of the wild type $\mathbf{0}$ is given by

$$\tilde{m}_{\mathbf{i}} = \frac{m_{\mathbf{i}}}{m_{\mathbf{0}}} = \frac{1 - sc_{\mathbf{i}}}{1 - sc_{\mathbf{0}}} = 1 - s(c_{\mathbf{i}} - c_{\mathbf{0}}) + O(s^2),$$

where $c_{\mathbf{0}} = c_{000}$ denotes the coefficient of selection of the wild type. In particular this implies that

$$\frac{\tilde{m}_{001}\tilde{m}_{010}}{\tilde{m}_{011}} = 1 + s\epsilon_{2,3} + O(s^2),$$

where

$$\epsilon_{2,3} = (c_{011} - c_{000}) - (c_{001} - c_{000}) - (c_{010} - c_{000}). \quad (4)$$

This quantity measures the epistatic interaction between the mutant alleles at loci 2 and 3 in the case of weak selection, given no mutant at locus 1. If $\epsilon_{2,3} > 0$, then the combined effect of the mutant alleles at loci 2 and 3 on reducing mortality in individuals carrying no mutant at locus 1 is higher at least under weak selection than what it would be if the mutant alleles had independent effects. The quantities $\epsilon_{1,2}$ and $\epsilon_{1,3}$ which measure the epistatic interactions between the mutant alleles at the other pairs of loci under weak selection, given no mutant at the other locus, are defined analogously.

In order to measure the epistatic interaction among the mutant alleles at all three loci, 1, 2 and 3, under weak selection we use the quantity

$$\epsilon_{1,2,3} = (c_{111} - c_{000}) - (c_{100} - c_{000}) - (c_{010} - c_{000}) - (c_{001} - c_{000}). \quad (5)$$

Note that

$$\epsilon_{1,2,3} - \epsilon_{1,2} - \epsilon_{1,3} = c_{111} - c_{110} - c_{101} + c_{100} \quad (6)$$

measures the epistatic interaction between the mutant alleles at loci 2 and 3 under weak selection, given a mutant at locus 1. Actually, the four parameters $\epsilon_{1,2}$, $\epsilon_{1,3}$, $\epsilon_{2,3}$ and $\epsilon_{1,2,3}$ allow a full description of the epistatic interactions in a bi-allelic three-locus setting under weak selection.

The current frequency of type $\mathbf{i}$ in the population is denoted by $x_{\mathbf{i}}$ and the corresponding population state by

$$\mathbf{x} = (x_{\mathbf{i}}) = (x_{000}, x_{001}, x_{010}, x_{011}, x_{100}, x_{101}, x_{110}, x_{111})\,. \quad (7)$$

On the other hand the current frequency of the mutant allele 1 at locus l is denoted by x_l for $l = 1, 2, 3$. This means that

$$x_1 = x_{111} + x_{110} + x_{101} + x_{100},$$
$$x_2 = x_{111} + x_{110} + x_{011} + x_{010},$$
$$x_3 = x_{111} + x_{101} + x_{011} + x_{001}.$$

Similarly the current frequency of the double mutant at loci j and l is denoted by x_{jl} for $j, l = 1, 2, 3$ and $j \ne l$. Therefore, we have

$$x_{13} = x_{101} + x_{111},$$
$$x_{12} = x_{110} + x_{111},$$
$$x_{23} = x_{011} + x_{111}.$$

The amount of linkage disequilibrium (LD) between the mutant alleles at two different loci is defined as the difference between the frequency of the double mutant, and the product of the frequencies of the mutant alleles. Then linkage disequilibrium between the mutant alleles at loci j and l is given by

$$D_{jl} = x_{jl} - x_jx_l,$$

for $j, k = 1, 2, 3$, and $j \ne l$. In a three-locus model, the quantity

$$D_{123} = x_{111} - x_1x_2x_3 - x_1D_{23} - x_2D_{13} - x_3D_{12} \quad (8)$$

is used to measure linkage disequilibrium among the three mutant alleles.

In this paper we assume that the mutant alleles at loci 2 and 3 were introduced some time ago into a population that was originally made of all 000 individuals. Moreover, as a result of recombination events, the double mutant at loci 2 and 3 may now be present in the population. Therefore, there are four possible types, namely 000, 001, 010, and 011, segregating in the population when the mutant allele at locus 1 is introduced. These types are represented in the frequencies given by

$$\mathbf{x} = (x_{000}, x_{001}, x_{010}, x_{011}, 0, 0, 0, 0)\,.$$

The mutation at locus 1 creates one of the new types 100, 101, 110, or 111. This means that there are four possible initial states of the

population following the introduction of the mutant allele at locus 1. These initial states are given by

$$\begin{aligned}
\mathbf{x}_1 &= \left(x_{000} - N^{-1}, x_{001}, x_{010}, x_{011}, N^{-1}, 0, 0, 0\right), \\
\mathbf{x}_2 &= \left(x_{000}, x_{001} - N^{-1}, x_{010}, x_{011}, 0, N^{-1}, 0, 0\right), \\
\mathbf{x}_3 &= \left(x_{000}, x_{001}, x_{010} - N^{-1}, x_{011}, 0, 0, N^{-1}, 0\right), \\
\mathbf{x}_4 &= \left(x_{000}, x_{001}, x_{010}, x_{011} - N^{-1}, 0, 0, 0, N^{-1}\right).
\end{aligned} \tag{9}$$

They occur with probabilities $x_{000}, x_{001}, x_{010}, x_{011}$, respectively, since the mutant allele at locus 1 is introduced at random. Note also that the frequency of this allele is given in all cases by the inverse of the population size, namely N^{-1}.

With probability $x_2 = x_{011} + x_{010}$, the mutant allele at locus 1 arises in linkage with the mutant allele at locus 2. In this case the linkage disequilibrium between these mutant alleles is initially given by

$$D_{12}^{(1)} = \frac{1}{N} - \frac{1}{N}x_2.$$

With the complementary probability $1 - x_2 = x_{000} + x_{001}$, the mutant allele at locus 1 arises in linkage with the resident allele at locus 2, and then the linkage disequilibrium between these mutant alleles is initially negative and is equal to

$$D_{12}^{(2)} = -\frac{1}{N}x_2.$$

This yields an initial average linkage disequilibrium between the mutant alleles at loci 1 and 2 given by

$$\bar{D}_{12} = x_2 D_{12}^{(1)} + (1 - x_2)D_{12}^{(2)} = 0.$$

Analogously the initial average linkage disequilibrium between the mutant alleles at loci 1 and 3, denoted by $\bar{D}_{13}$, is equal to 0. Note that this is the value of this disequilibrium before the introduction of allele 1 at locus 1.

In a similar way, the linkage disequilibrium between the mutant alleles at loci 2 and 3 is not affected on average by the introduction of allele 1 at locus 1, and

$$\bar{D}_{23} = x_{011} - x_2 x_3 = D_{23} = D.$$

Finally, by considering the four possible states of the population following the introduction of the mutant allele at locus 1, given in (9), and by calculating the value of the linkage disequilibrium among the mutant alleles at the three loci in each case, given by (8), we get

$$\bar{D}_{123} = N^{-1}\left(x_{011} - x_2 x_3 - D_{23}\right) = 0,$$

as the initial average linkage disequilibrium among the three mutant alleles.

2. Conditional expected change in the frequency of an allele

Let the vector of the individual type frequencies at time step τ be $\mathbf{x}(\tau) = (x_{\mathbf{i}}(\tau))$. Then the frequency of type $\mathbf{i}$ at the next time step is

$$x_{\mathbf{i}}(\tau + 1) = \begin{cases} x_{\mathbf{i}}(\tau) + N^{-1} & \text{with probability } \eta_{\mathbf{i}}(\tau), \\ x_{\mathbf{i}}(\tau) - N^{-1} & \text{with probability } \gamma_{\mathbf{i}}(\tau), \\ x_{\mathbf{i}}(\tau) & \text{with probability } 1 - \eta_{\mathbf{i}}(\tau) - \gamma_{\mathbf{i}}(\tau), \end{cases}$$

where

$$\eta_{\mathbf{i}}(\tau) = \left(\sum_{\mathbf{j},\mathbf{k}} x_{\mathbf{j}}(\tau)x_{\mathbf{k}}(\tau)R(\mathbf{i}|\mathbf{j},\mathbf{k})\right)\left(\sum_{\mathbf{j}\neq\mathbf{i}} m_{\mathbf{j}}x_{\mathbf{j}}(\tau)\right),$$

$$\gamma_{\mathbf{i}}(\tau) = \left(1 - \sum_{\mathbf{j},\mathbf{k}} x_{\mathbf{j}}(\tau)x_{\mathbf{k}}(\tau)R(\mathbf{i}|\mathbf{j},\mathbf{k})\right) m_{\mathbf{i}}x_{\mathbf{i}}(\tau).$$

Now let

$$\bar{m}(\tau) = \sum_{\mathbf{i}} m_{\mathbf{i}}x_{\mathbf{i}}(\tau)$$

be the average mortality in the population at time step τ. Then the change in the frequency of type $\mathbf{i}$ from time step τ to time step $\tau+1$, namely $\Delta x_{\mathbf{i}}(\tau) = x_{\mathbf{i}}(\tau + 1) - x_{\mathbf{i}}(\tau)$, has conditional expectation given by

$$\begin{aligned}
&E\left(\Delta x_{\mathbf{i}}(\tau)|\mathbf{x}(\tau)\right) \\
&\quad = N^{-1}\left(\bar{m}(\tau)\sum_{\mathbf{j},\mathbf{k}} x_{\mathbf{j}}(\tau)x_{\mathbf{k}}(\tau)R(\mathbf{i}|\mathbf{j},\mathbf{k}) - m_{\mathbf{i}}x_{\mathbf{i}}(\tau)\right).
\end{aligned}$$

Summing over all $\mathbf{i} = i_1 i_2 i_3$ such that $i_1 = 1$ yields

$$\begin{aligned}
&E\left(\Delta x_1(\tau)|\mathbf{x}(\tau)\right) \\
&\quad = N^{-1}\left(\bar{m}(\tau)\sum_{\mathbf{j},\mathbf{k}} x_{\mathbf{j}}(\tau)x_{\mathbf{k}}(\tau)\sum_{\mathbf{i}:i_1=1} R(\mathbf{i}|\mathbf{j},\mathbf{k}) - \sum_{\mathbf{i}:i_1=1} m_{\mathbf{i}}x_{\mathbf{i}}(\tau)\right),
\end{aligned}$$

for the conditional expected change in the frequency of the mutant allele at locus 1. From the definition of the transmission probability we have

$$\sum_{\mathbf{i}:i_1=1} R(\mathbf{i}|\mathbf{j},\mathbf{k}) = \begin{cases} 1 & \text{if } j_1 = 1, \\ 0 & \text{if } j_1 = 0. \end{cases}$$

This is the frequency of allele 1 at locus 1 in the first parent, which is represented by j_1. It follows that

$$E\left(\Delta x_1(\tau)|\mathbf{x}(\tau)\right) = N^{-1}\left(\bar{m}(\tau)\sum_{\mathbf{i}:i_1=1} x_{\mathbf{i}}(\tau) - \sum_{\mathbf{i}:i_1=1} m_{\mathbf{i}}x_{\mathbf{i}}(\tau)\right).$$

Using the expression of the mortality parameter, this leads to

$$E\left(\Delta x_1(\tau)|\mathbf{x}(\tau)\right) = N^{-1}s\sum_{\mathbf{i}:i_1=1}(c_{\mathbf{i}} - \bar{c}(\tau))x_{\mathbf{i}}(\tau),$$

where

$$\bar{c}(\tau) = \sum_{\mathbf{j}} c_{\mathbf{j}}x_{\mathbf{j}}(\tau)$$

is the mean coefficient of selection, and

$$c_{\mathbf{i}} - \bar{c}(\tau) = \sum_{\mathbf{j}}(c_{\mathbf{i}} - c_{\mathbf{j}})x_{\mathbf{j}}(\tau).$$

This yields the following standard result.

Proposition 1. *For the three-locus, two-allele discrete-time Moran model with recombination and selection, the conditional expected change in the frequency of the mutant allele* 1 *at locus* 1 *from time step* τ *to the time step* $\tau + 1$ *is given by*

$$E\left(\Delta x_1(\tau)|\mathbf{x}(\tau)\right) = \sigma N^{-2}\sum_{\mathbf{i}:i_1=1}\sum_{\mathbf{j}} x_{\mathbf{i}}(\tau)x_{\mathbf{j}}(\tau)(c_{\mathbf{i}} - c_{\mathbf{j}}), \tag{10}$$

where N *is the population size and* $\sigma = sN$ *is the population-scaled intensity of selection with coefficient* $c_{\mathbf{i}}$ *for an individual of type* $\mathbf{i} = i_1 i_2 i_3$ *with* $i_1, i_2, i_3 = 0$ *or* 1.

3. Probability of ultimate fixation of an allele

The random vector $\mathbf{x}(\tau) = (x_{\mathbf{i}}(\tau))$ for $\tau \geq 0$ is a Markov chain on the finite state space of frequency vectors of the form $N^{-1}(K_{\mathbf{i}})$ for non-negative integers $K_{\mathbf{i}}$ satisfying $\sum_{\mathbf{i}} K_{\mathbf{i}} = N$. There are as many absorbing states as individual types. The state corresponding to fixation of type $\mathbf{i}$ is represented by

$$\mathbf{e}_{\mathbf{i}} = \left(\delta_{\mathbf{i},\mathbf{j}}\right) = \left(\delta_{\mathbf{i},000}, \ldots, \delta_{\mathbf{i},111}\right). \tag{11}$$

All other states are transient.

Let $P_{\mathbf{xy}}(\tau)$ be the transition probability that the chain starting from state $\mathbf{x}$ is in state $\mathbf{y}$ after τ time steps, i.e.

$$P_{\mathbf{xy}}(\tau) = P\big(\mathbf{x}(\tau) = \mathbf{y}|\mathbf{x}(0) = \mathbf{x}\big).$$

By the ergodic theorem for Markov chains (see, e.g., Karlin and Taylor, 1975), this transition probability converges to some limit as the number of time steps tends to ∞, represented by $P_{\mathbf{xy}}(\infty)$. This limit is 0 unless $\mathbf{y}$ is an absorbing state, i.e. $\mathbf{y} = \mathbf{e_i}$ for some type $\mathbf{i}$. Using

$$x_1(\tau) = \sum_{\mathbf{i}:i_1=1} x_{\mathbf{i}}(\tau)$$

and defining

$$E_{\mathbf{x}}\big(x_1(\tau)\big) = E\big(x_1(\tau)|\mathbf{x}(0) = \mathbf{x}\big),$$

one obtains that

$$E_{\mathbf{x}}\big(x_1(\tau)\big) = \sum_{\mathbf{i}:i_1=1} \sum_{\mathbf{y}} y_{\mathbf{i}} P_{\mathbf{xy}}(\tau).$$

From any initial state $\mathbf{x}$, as the number of time steps τ tends to ∞, $x_1(\tau)$ converges almost surely and in mean to a random variable $x_1(\infty)$, which takes the value 1 with some probability $u(\mathbf{x})$, and 0 otherwise. Here

$$u(\mathbf{x}) = E_{\mathbf{x}}\big(x_1(\infty)\big) = \sum_{\mathbf{i}:i_1=1} P_{\mathbf{xe_i}}(\infty). \tag{12}$$

The quantity $u(\mathbf{x})$ represents the probability of ultimate fixation of the mutant allele at locus 1.

On the other hand, we have

$$x_1(\infty) = x_1(0) + \sum_{\tau=0}^{\infty} \Delta x_1(\tau),$$

from which

$$\begin{aligned} E_{\mathbf{x}}\big(x_1(\infty)\big) &= x_1(0) + \sum_{\tau=0}^{\infty} E_{\mathbf{x}}\big(\Delta x_1(\tau)\big) \\ &= x_1(0) + \sum_{\tau=0}^{\infty} E_{\mathbf{x}}\Big(E\big(\Delta x_1(\tau)|\mathbf{x}(\tau)\big)\Big). \end{aligned}$$

Owing to Proposition 1, the following less standard result ensues.

Proposition 2. *Let $u(\mathbf{x})$ be the probability of ultimate fixation of the mutant allele 1 at locus 1 in the discrete-time Moran model of Proposition 1, given that the initial type frequency vector is $\mathbf{x}(0) = \mathbf{x}$. Then*

$$u(\mathbf{x}) = x_1 + \frac{\sigma}{2} \sum_{\mathbf{i}:i_1=1} \sum_{\mathbf{j}} (c_{\mathbf{i}} - c_{\mathbf{j}}) \psi_{\mathbf{ij}}(\mathbf{x}), \tag{13}$$

where x_1 is the initial frequency of allele 1 at locus 1, and

$$\psi_{\mathbf{ij}}(\mathbf{x}) = 2N^{-2} \sum_{\tau=0}^{\infty} E_{\mathbf{x}}\big(x_{\mathbf{i}}(\tau) x_{\mathbf{j}}(\tau)\big). \tag{14}$$

This is the expected time in number of $N^2/2$ time steps that two individuals chosen at random with replacement in the population at the same time step will be of types $\mathbf{i}$ and $\mathbf{j}$ in this order, given a population state $\mathbf{x}$ at the initial time step.

4. Ancestral recombination–selection graph

In this section we consider that the recombination rates among the loci depend on the allele at locus 1. The population-scaled recombination rates between loci 1 and 2 and between loci 2 and 3 are represented by ν_l and ρ_l, respectively, for $l = 0, 1, 2$ corresponding to the number of mutants carried by the parents at locus 1. Also we assume that $0 \le \rho_0 \le \rho_1 \le \rho_2$ and $0 \le \nu_0 \le \nu_1 \le \nu_2$, as defined in (1).

The reproduction process in the discrete-time Moran model can be described as follows. At each time step an ordered pair of individuals is chosen at random to reproduce. Their offspring is an exact copy of the first parent irrespective of the parental types with probability

$$\Big(1 - \frac{\nu_2}{N}\Big)\Big(1 - \frac{\rho_2}{N}\Big).$$

This is a minimum for the probability of a non-recombinant offspring. With the complementary probability

$$1 - \Big(1 - \frac{\nu_2}{N}\Big)\Big(1 - \frac{\rho_2}{N}\Big) = \frac{\nu_2 + \rho_2}{N} + \frac{\nu_2 \rho_2}{N^2},$$

parent-dependent recombination occurs. Then, given types $\mathbf{j} = j_1 j_2 j_3$ and $\mathbf{k} = k_1 k_2 k_3$ in this order for the parents, locus 1 is inherited from the first parent and loci 2 and 3 from the second parent with the conditional probability

$$\frac{\nu_{j_1+k_1}}{\nu_2 + \rho_2} + O(N^{-1}).$$

Similarly, loci 1 and 2 are inherited from the first parent and locus 3 from the second parent with the conditional probability

$$\frac{\rho_{j_1+k_1}}{\nu_2 + \rho_2} + O(N^{-1}).$$

Otherwise, ignoring simultaneous recombination events between loci 1, 2 and between loci 2, 3 which are of probabilities of order $O(N^{-2})$, all three loci are inherited from the first parent.

The replacement process following the production of an offspring is defined in a similar way. A randomly selected individual is chosen to be replaced by the offspring. Replacement occurs in all cases with probability $(1 - N^{-1}\sigma)$, where $\sigma = Ns$ is the population-scaled intensity of selection. Type-dependent replacement occurs with the complementary probability $N^{-1}\sigma$. Given this event, replacement occurs with the conditional probability $1 - c_{\mathbf{i}}$ if the individual to be replaced is of type $\mathbf{i}$. Therefore, given type $\mathbf{i}$ for the individual to be replaced, the total probability of replacement is given by

$$1 - N^{-1}\sigma + N^{-1}\sigma(1 - c_{\mathbf{i}}) = 1 - sc_{\mathbf{i}} = m_{\mathbf{i}},$$

which is the mortality associated to type $\mathbf{i}$.

Starting from a random sample at a given time step and going backward in time, the above description of the model leads to an ancestral recombination–selection graph (ARSG). It corresponds to an ancestral influence graph, AIG (Donnelly and Kurtz, 1999), in the limit of a large population size with time measured in units of $N^2/2$ time steps. It provides a supra-genealogy that combines the ancestral selection graph, ASG (Krone and Neuhauser, 1997), and the ancestral recombination graph, ARG (Griffiths, 1981; Hudson, 1983; Griffiths and Marjoram, 1996).

Consider a random sample of size n taken without replacement at a given time step. The sample size is assumed to be small compared to the population size N. Let $i = 1, \ldots, n$ be arbitrary labels for the sampled individuals. Label arbitrarily the other $N - n$ individuals in the population at the same time step with the integers $i = n + 1, \ldots, N$. In one time step back, there may be a pure coalescence event of i and $j \ne i$ for $i, j = 1, \ldots, n$. This occurs if j produced an exact copy of itself irrespective of its own type and the type of the other parent (with probability $N^{-1}(1-\nu_2 N^{-1})(1-\rho_2 N^{-1})$), and this copy replaced the individual that occupied the site of i irrespective of its type (with probability $N^{-1}(1 - \sigma N^{-1})$), or vice versa. Therefore, the probability of each

pure coalescence event in the sample of size n in one time step back is given by

$$p(C_n) = \frac{2}{N^2}\left(1-\frac{\sigma}{N}\right)\left(1-\frac{\nu_2}{N}\right)\left(1-\frac{\rho_2}{N}\right) = \frac{2}{N^2}(1+O(N^{-1})) \leq \frac{2}{N^2}. \quad (15)$$

In this case the sample size is reduced by one by merging the lineages of two sampled individuals.

On the other hand a pure recombination event in the ancestry of the sample in one time step back occurs when an individual $i = 1,\ldots,n$ is a parent-dependent-recombinant offspring of parents $k, l = n+1,\ldots,N$ for $k \neq l$, which occurs with probability

$$\frac{(N-n)(N-n-1)}{N^2}\left(\frac{\nu_2+\rho_2}{N}+\frac{\nu_2\rho_2}{N^2}\right),$$

and that this offspring replaced the individual that occupied the site of i irrespective of its type, which occurs with probability $N^{-1}(1-\sigma N^{-1})$. Therefore, we have

$$p(R_n) = \frac{(N-n)(N-n-1)}{N^3}\left(\frac{\nu_2+\rho_2}{N}+\frac{\nu_2\rho_2}{N^2}\right)\left(1-\frac{\sigma}{N}\right) = \frac{\nu_2+\rho_2}{N^2}(1+O(N^{-1})) \leq \frac{\nu_2+\rho_2}{N^2}, \quad (16)$$

for the probability of each pure recombination event in the sample of size n in one time step back. In this case the lineage of one sampled individual splits into two ordered lineages. The first lineage is ancestral to a subset of loci and the second lineage ancestral to the complementary set. The conditional probabilities associated to the different sets depend on the types of the individuals carrying these lineages once they are known.

Finally there is a pure selection event in the ancestry of the sample in one time step back when the offspring produced was an exact copy of a first parent not in the sample, irrespective of the types of the parents (with probability $(N-n)N^{-1}(1-\nu_2N^{-1})(1-\rho_2N^{-1})$), and the individual chosen for replacement according to its type occupied the site of $i = 1,\ldots,n$ (with probability σN^{-2}). Therefore, the probability of each pure selection event in a sample of size n in one time step back is given by

$$p(S_n) = \frac{\sigma(N-n)}{N^3}\left(1-\frac{\nu_2}{N}\right)\left(1-\frac{\rho_2}{N}\right) \quad (17)$$

$$= \frac{\sigma}{N^2}(1+O(N^{-1})) \leq \frac{\sigma}{N^2}. \quad (18)$$

In this case the sample size is increased by one: the lineage of one sampled individual branches into two, each one being potentially ancestral to the sampled individual at all loci. The incoming lineage is the lineage of the offspring produced one time step back, while the continuing lineage is the lineage of the individual chosen to be replaced by the offspring. One of these lineages is real with some conditional probability, the other virtual with the complementary probability, but both lineages have to be traced back until ancestors of known types are reached. Only then can the conditional probabilities be determined.

Note that pure coalescence, recombination or selection events cannot occur simultaneously and their probabilities are all of order $O(N^{-2})$. In comparison, probabilities of multiple events are of order $O(N^{-3})$. They can be neglected in the limit of a large population size for a sample of any given size.

Note also that pure recombination or selection events are potential events which may or may not be realized depending on the types of the individuals involved.

In a sample of size n, there are $n(n-1)/2$ possible pure coalescence events and n possible pure selection or recombination events. Therefore, the total probability of change due to a pure coalescence, selection or recombination event in one time step back is given by

$$p_n = \binom{n}{2}p(C_n) + np(R_n) + np(S_n) = 2\lambda_n N^{-2} + O(N^{-3}),$$

where

$$\lambda_n = \frac{n(n-1+\sigma+\nu_2+\rho_2)}{2}.$$

This parameter is the total rate of change in the limit of large population size with $N^2/2$ time steps as the unit of time. Moreover, given that a change in one time step back has occurred, the conditional probability of each coalescence, recombination or selection event is given by

$$P(C_n) = \frac{1}{\lambda_n} + O(N^{-1}), \quad (19)$$

$$P(R_n) = \frac{\nu_2+\rho_2}{2\lambda_n} + O(N^{-1}), \quad (20)$$

or

$$P(S_n) = \frac{\sigma}{2\lambda_n} + O(N^{-1}), \quad (21)$$

respectively. Note that

$$\binom{n}{2}P(C_n) + nP(R_n) + nP(S_n) = 1 + O(N^{-1}).$$

This means that in the limit of large population size, the total probability of pure events adds up to one.

Let τ_n to be the waiting time back, in units of time steps, until the next event of coalescence, recombination or selection changes the lineages of a sample of size n. This sojourn time is a geometric random variable which is independent of all previous events of change and all previous waiting times between these events. We have

$$E(\tau_n) = \sum_{k=0}^{\infty} P(\tau_n > k),$$

where

$$P(\tau_n > k) = (2\lambda_n N^{-2} + O(N^{-3}))^k.$$

Now if we measure time in units of $N^2/2$ time steps, the distribution of the scaled waiting time $T_n = 2N^{-2}\tau_n$ converges to the distribution of an exponential random variable with parameter λ_n in the limit of large population size. As a matter of fact

$$P(T_n > t) = P(\tau_n > \lfloor tN^2/2 \rfloor),$$

where $\lfloor x \rfloor$ denotes the integer value of x, and

$$\lim_{N\to\infty} P(\tau_n > \lfloor tN^2/2 \rfloor) = \exp(-\lambda_n t),$$

for every $t > 0$. Moreover

$$E(T_n) = \int_0^{\infty} P(T_n > t)dt,$$

where

$$P(T_n > t) \leq \left(1-\lambda_n N^{-2}\right)^{\frac{tN^2}{2}-1} \leq 2\exp(-\lambda_n t/2),$$

for N large enough. Then

$$\lim_{N\to\infty} E(T_n) = \int_0^{\infty} \exp(-\lambda_n t)dt = \lambda_n^{-1}., \quad (22)$$

as ascertained by the dominated convergence theorem.

Let us summarize the above findings.

Proposition 3. *Consider the discrete-time Moran model of Proposition 1 for a population of size N, in which the rates of recombination are determined by the alleles j_1 and k_1 at locus 1 in ordered parents of types $\mathbf{j} = j_1j_2j_3$ and $\mathbf{k} = k_1k_2k_3$. Assume that the population-scaled recombination rate is ν_0, ν_1 or ν_2 ($0 \leq \nu_0 \leq \nu_1 \leq \nu_2$) between loci 1 and 2 and ρ_0, ρ_1 or ρ_2 ($0 \leq \rho_0 \leq \rho_1 \leq \rho_2$) between loci 2 and 3, corresponding to $i_1 + j_1 = 0$, 1 or 2, respectively. Let σ be the population-scaled intensity of selection with coefficient $0 \leq c_{\mathbf{i}} \leq 1$ for type $\mathbf{i}$. Consider an ordered sample of size n, taken without replacement in the population at a given time step. Going backward in time and ignoring events of probability of order $O(N^{-3})$, there are three kinds of events that may affect the lineages of the sampled individuals:*

1. *Each pair of lineages merges as a result of a pure coalescence event with approximate probability $2N^{-2}$.*
2. *Each lineage splits into two ordered lineages due to a pure recombination event with approximate probability $N^{-2}(\nu_2 + \rho_2)$. Given this event and alleles j_1 and k_1 at locus 1 in the potential ancestors carrying the first and second lineages respectively, the first lineage is ancestral to locus 1 and the second lineage ancestral to loci 2 and 3 with approximate conditional probability*

   $$\frac{\nu_{j_1+k_1}}{\nu_2+\rho_2},$$

   while the first lineage is ancestral to loci 1 and 2 and the second lineage ancestral to locus 3 with approximate probability

   $$\frac{\rho_{j_1+k_1}}{\nu_2+\rho_2},$$

   and the first lineage is ancestral to all three loci and the second ancestral to none with the complementary probability.
3. *With approximate probability σN^{-2}, each lineage branches into two ordered lineages as a result of a pure selection event, each one being potentially ancestral to the sampled individual at all loci. One is the lineage of the offspring produced one time step back (incoming lineage), while the other is the lineage of the individual to be potentially replaced by the offspring (continuing lineage). Given this event and a potential ancestor of type $\mathbf{i}$ carrying the continuing lineage, the incoming branch is real (i.e. ancestral to the sample) with conditional probability $1 - c_{\mathbf{i}}$, and virtual with the complementary probability.*

The total probability of change is approximately $2N^{-2}\lambda_n = n(n - 1 + \nu_2 + \rho_2 + \sigma)N^{-2}$. Moreover, in the limit of a large population size with time measured in units of $N^2/2$ time steps, the expected time for a change is given by λ_n^{-1}.

5. Expected time in a sample state

In this section, we derive an expression for the expected time in an ordered sample state that will be used later on to compute $\psi_{\mathbf{ij}}(\mathbf{x})$ in (14).

The n-dimensional vector of an ordered sample of n individuals is denoted by $\mathbf{z} = (z_1, \ldots, z_n)$, where z_i is the type of individual $i = 1, \ldots, n$. The sample configuration is represented by the vector $\mathbf{n} = (n_{\mathbf{i}})$, where $n_{\mathbf{i}}$ is the number of individuals of type $\mathbf{i}$ in the sample and $\sum_{\mathbf{i}} n_{\mathbf{i}} = n$.

Let $\mathbf{z}(\tau)$ be an ordered sample of n individuals chosen at random without replacement at time step $\tau \geq 0$. The probability distribution of this sample depends on the ancestral recombination–selection graph from time step τ to time step 0, denoted by $G(\tau)$, and the type frequencies at time step 0, given by $\mathbf{x}(0) = \mathbf{x}$. In fact, what is important is the genealogy of the graph (which lineages merge, split or branch) from time step τ to time step 0. It is represented by a sequence of events backward in time written in the form

$$G = (G_1, \ldots, G_m),$$

where $m = m(G) > 0$ is the total number of events. These correspond to pure or multiple events of coalescence, recombination or selection in one time step back.

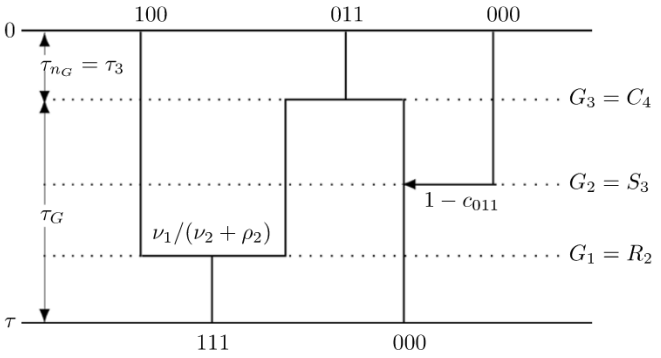

Fig. 1. Possible ancestral recombination–selection graph with a genealogy represented by $G = (R_2, S_3, C_4)$ for two ordered individuals at time step τ to be of types 111 and 000 in this order.

Let n_G be the number of potential ancestors after the occurrence of the last event of G backward in time. This last event takes place at some time step back τ_G (see Fig. 1). On the other hand, the time with these n_G ancestors is represented by τ_{n_G}. For G to be admissible as a genealogy of the graph from time step τ to time step 0, it is necessary that $\tau_G \leq \tau < \tau_G + \tau_{n_G}$. We define

$$G(\tau) = \{G, \tau_G \leq \tau < \tau_G + \tau_{n_G}\}.$$

The random variables τ_G and τ_{n_G} are independent. Moreover, τ_{n_G} has a geometric distribution with parameter p_{n_G}, while τ_G is a sum of independent random variables of geometric distribution.

The conditional probability of the event $\mathbf{z}(\tau) = \mathbf{z}$, given that $\mathbf{x}(0) = \mathbf{x}$, can be written in the form

$$P\big(\mathbf{z}(\tau) = \mathbf{z}|\mathbf{x}(0) = \mathbf{x}\big) = \sum_{G(\tau)} P\big(\mathbf{z}(\tau) = \mathbf{z}|G(\tau), \mathbf{x}(0) = \mathbf{x}\big)P\big(G(\tau)\big). \tag{23}$$

We define

$$P_G\big(\mathbf{z}|\mathbf{x}\big) = P\big(\mathbf{z}(\tau) = \mathbf{z}|G(\tau), \mathbf{x}(0) = \mathbf{x}\big),$$

which is the conditional sample probability given G. It does not actually depend on time step τ at which the sample is taken. On the other hand, we have

$$P(G(\tau)) = P(G)P(\tau_G \leq \tau < \tau_G + \tau_{n_G}),$$

where

$$P(G) = \prod_{k=1}^{m} P(G_k). \tag{24}$$

Here $P(G_k) = P(C_{n_k})$, $P(R_{n_k})$ or $P(S_{n_k})$ as defined in (19)–(21), if G_k is a pure event of coalescence, recombination or selection, respectively, and n_k is the number of potential ancestors when this event occurs. Moreover, we have

$$P(\tau_G \leq \tau < \tau_G + \tau_{n_G}) = P(\tau_G + \tau_{n_G} > \tau) - P(\tau_G > \tau),$$

from which

$$\sum_{\tau=0}^{\infty} P(\tau_G \leq \tau < \tau_G + \tau_{n_G}) = E(\tau_G + \tau_{n_G}) - E(\tau_G) = E(\tau_{n_G}).$$

Therefore, summing over $\tau \geq 0$ in (23) yields the following result.

Proposition 4. *Let $\mathbf{z}(\tau)$ be an ordered sample of n individuals chosen at random without replacement at time step $\tau \geq 0$ and $\mathbf{x}(0) = \mathbf{x}$ be the vector of the individual type frequencies at time step 0 in the discrete-time Moran model of Proposition 1. Then the expected time in a given ordered sample state $\mathbf{z}$ is*

$$\sum_{\tau=0}^{\infty} P(\mathbf{z}(\tau) = \mathbf{z}|\mathbf{x}(0) = \mathbf{x}) = \sum_{G} P_G(\mathbf{z}|\mathbf{x})P(G)E(\tau_{n_G}),$$

where G is a sequence of (pure or multiple) coalescence, recombination or selection events from time step τ to time step 0, while n_G is the number of potential ancestors at time step 0 and τ_{n_G} is the number of time steps back with this number of ancestors.

6. Approximation of the fixation probability

We are now ready to approximate the probability of ultimate fixation of the mutant allele 1 at locus 1, under the assumptions that the population size is large and the population-scaled recombination and selection parameters are small. First note that

$$E_{\mathbf{x}}(x_{\mathbf{i}}(\tau)x_{\mathbf{j}}(\tau)) = (1 - N^{-1})P(\mathbf{z}(\tau) = \mathbf{z}_{\mathbf{ij}}|\mathbf{x}(0) = \mathbf{x}),$$

where $\mathbf{z}_{\mathbf{ij}} = (\mathbf{i}, \mathbf{j})$ is an ordered sample of size two, the first sampled individual being of type $\mathbf{i}$ and the second individual of type $\mathbf{j} \neq \mathbf{i}$. From (14) we deduce that

$$\psi_{\mathbf{ij}}(\mathbf{x}) = 2N^{-2}(1 - N^{-1}) \sum_{\tau=0}^{\infty} P(\mathbf{z}(\tau) = \mathbf{z}_{\mathbf{ij}}|\mathbf{x}(0) = \mathbf{x}).$$

Then Proposition 4 leads to the expression

$$\psi_{\mathbf{ij}}(\mathbf{x}) = (1 - N^{-1}) \sum_{G} P_G(\mathbf{z}_{\mathbf{ij}}|\mathbf{x})P(G)E(T_{n_G}).$$

Here $T_{n_G} = 2\tau_{n_G}N^{-2}$ is the time back spent with the n_G potential ancestors of G measured in units of $N^2/2$ time steps. Note that the above summation is over an infinite number of G. However, approximation results can be obtained by considering only a finite number of G. In the following we provide a bound for the order of the error.

Without loss of generality we focus attention on $(\mathbf{i}, \mathbf{j}) = (\mathbf{0}, \mathbf{1})$. Let $|G|$ denote the minimum number of potential ancestors along a sequence of events G for the ordered sample $\mathbf{z}_{\mathbf{0},\mathbf{1}} = (000, 111)$. If $|G| = 1$, then

$$P_G(\mathbf{z}_{\mathbf{0},\mathbf{1}}|\mathbf{x}) = 0,$$

since alleles 1 and 0 at any locus cannot have a common ancestor. If $|G| \geq 2$, actually equal to the sample size $n = 2$ in the case at hand, then

$$P_G(\mathbf{z}_{\mathbf{0},\mathbf{1}}|\mathbf{x}) \leq x_1,$$

since allele 1 at locus 1 must be present in at least one of the ordered potential ancestors at time step 0. This leads to the inequality

$$\psi_{\mathbf{0},\mathbf{1}}(\mathbf{x}) \leq x_1 E(W_2),$$

where

$$E(W_2) = \sum_{G:|G|\geq 2} P(G)E(T_{n_G}).$$

Actually, W_2 is the time back in number of $N^2/2$ time steps before the number of potential ancestors of a sample of size $n = 2$ reaches one for the first time. This occurs when the most recent ultimate ancestor of the sample is found. It can be shown that $E(W_2)$ is finite and bounded by a constant that does not depend on N. This is also true for $E(W_n) \geq E(W_{n-1})$, which is defined analogously for a sample of any size $n \geq 3$. A proof is presented in Lessard and Kermany (2012).

Now, assume that the population-scaled intensities of recombination and selection, ν_l, ρ_l ($l = 0, 1, 2$) and σ, are of the same small order, so that $\nu_l = \alpha_l\sigma$ and $\rho_l = \beta_l\sigma$ for $\sigma \ll 1$ and some constants $\alpha_l, \beta_l > 0$ for $l = 0, 1, 2$.

Let n_G^+ denote the sum of all increases in the number of potential ancestors along a sequence of events G starting from the ordered sample $\mathbf{z}_{\mathbf{0},\mathbf{1}}$ and going backward in time. This number increases when recombination or selection events occur. Note that, in the exact Moran model, the number of potential ancestors can increase by at most 2 at a time, when recombination and selection events occur simultaneously at the same time step. In the limit of a large population size, this number can increase only by one at a time with probability one. Note also that $n_G \leq n_G^+ + n$, where n is the sample size, here $n = 2$.

If $n_G^+ > k$ for some integer $k \geq 0$, then the sequence of events can be divided into two parts, namely

$$G = (F, H).$$

Here F is the first part of G with the smallest number of events such that $n_F^+ > k$ and $n_F \leq k + 4$ (this upper bound that takes into account the possibility of increases in the number of potential ancestors by 2 at a time corrects an error in Lessard and Kermany, 2012). On the other hand, H is the remaining part of G and is such that $n_H = n_G$. Note that n_F is the sample size at the beginning of H, and corresponds to the number of potential ancestors just after the time step backward in time that brings the sum of all increases in the number of potential ancestors above k for the first time. In Fig. 1, for instance, we have $F = (R_2, S_3)$ and $H = (C_4)$ for $k = 1$, with $n_F^+ = 2$ and $n_F = 4$.

In general, the probability of F can be neglected compared to σ^k, since the probability of each of its n_F^+ increases is of order $O(\sigma)$. Actually

$$P(F) = O(\sigma^{k+1}).$$

On the other hand,

$$\sum_{|H|\geq 2} P(H)E(T_{n_H}) \leq E(W_{k+4}),$$

and this is a finite bound. Since $P(G) = P(F)P(H)$ and $n_G = n_H$, one obtains that

$$\sum_{G:n_G^+>k} P_G(\mathbf{z}_{\mathbf{0},\mathbf{1}}|\mathbf{x})P(G)E(T_{n_G}) \leq x_1 O(\sigma^{k+1}),$$

in which the summation is actually over all G such that the number of potential ancestors never goes down to 1. This gives a bound for the error in the approximation

$$\psi_{\mathbf{0},\mathbf{1}}(\mathbf{x}) \approx (1 - N^{-1}) \sum_{G:n_G^+\leq k} P_G(\mathbf{z}_{\mathbf{0},\mathbf{1}}|\mathbf{x})P(G)E(T_{n_G}).$$

Moreover, all terms in the sum above are approximated by their leading terms in the case of a large population size. Then the coefficient of $\sigma^m\nu_2^l\rho_2^{r-l}$ for $m + r \leq k$, for instance, can be obtained by considering only the sequences of events G involving exactly m pure selection events and r pure recombination events. Similar arguments can be applied to approximate $\psi_{\mathbf{ij}}(\mathbf{x})$ in (14) for every $\mathbf{i} \neq \mathbf{j}$. Therefore using (13) we have the following approximation for the fixation probability in Proposition 2.

Proposition 5. *Let the population-scaled recombination rates in Proposition 3 be of the same order as the intensity of selection, that is, $\nu_l = \alpha_l\sigma$ and $\rho_l = \beta_l\sigma$ for some constants α_l and β_l, $l = 0, 1, 2$. Then the probability of ultimate fixation of allele 1 at locus 1 in Proposition 2 in the case of a large population size given*

an initial state $\mathbf{x}(0) = \mathbf{x}$ *for type frequencies corresponding to a frequency* x_1 *for allele* 1 *at locus* 1 *can be approximated as*

$$u(\mathbf{x}) \approx x_1 + \frac{\sigma}{2}\sum_{\mathbf{i}:i_1=1}\sum_{\mathbf{j}}(c_{\mathbf{i}} - c_{\mathbf{j}})\psi^{(k)}_{\mathbf{ij}}(\mathbf{x}) + O(\sigma^{k+1}),$$

where

$$\psi^{(k)}_{\mathbf{ij}}(\mathbf{x}) = \sum_{G:n_G^+\le k-1} P_G(\mathbf{z}_{\mathbf{ij}}|\mathbf{x})P(G)E(T_{n_G}). \tag{25}$$

In this summation only terms for sequences of pure coalescence, recombination or selection events represented by G *and such that the number of potential ancestors never goes down to* 1 *are considered, and they are approximated by their leading terms for a large population size.*

For $k = 2$, for instance, four representations of G have to be considered in (25), namely

$$G = (R_2),\ (R_2, C_3),\ (S_2),\ \text{and}\ (S_2, C_3).$$

Here R_n, S_n and C_n represent pure recombination, selection and coalescence events, respectively, when there are n potential ancestors. Note that two possible events are represented by R_2 and by S_2, while three possible events are represented by C_3. Note also that $G = (C_2)$, which diminishes the number of ancestors to 1, and $G = (R_2, S_3)$ or $G = (S_2, R_3)$, for instance, which increases the number of potential ancestors twice, are ignored.

The probability $P(G)$ is given by (24), while the limit of $E(T_{n_G})$ is given by (22). It remains to use the formula that is given in Proposition 8 of Appendix A, to calculate the conditional probability $P_G(\mathbf{z}_{\mathbf{ij}}|\mathbf{x})$ in each case.

In the next two sections we use the approximation given in Proposition 5 to calculate the fixation probability for a new mutant in a three-locus model in two different situations. First we consider the case where all three loci are under selection and we assume that the effect of a mutation on mortality is independent of the locus at which it occurs. Also the recombination rates are assumed to be parent-independent. Second we look at the case where general mortality parameters are determined at two loci and a new neutral mutant at another locus modifies the recombination rates among the three loci.

7. Ultimate fixation of a single mutant in a three-locus selection model

In this section, we assume that the mortality of an individual depends only on the number of mutants that it carries at three loci, 1, 2 and 3 in this order, so that each mutant has the same effect on mortality. In addition we consider parent-independent recombination, i.e. $\nu_0 = \nu_1 = \nu_2 = \nu$ and $\rho_0 = \rho_1 = \rho_2 = \rho$.

Then the coefficient of selection for type $\mathbf{i} = i_1i_2i_3$ defined in (3) is given by

$$c_{\mathbf{i}} = b_0, b_1, b_2 \text{ or } b_3,$$

according to

$$|\mathbf{i}| = i_1 + i_2 + i_3 = 0, 1, 2 \text{ or } 3,$$

respectively. Moreover, the epistatic interaction between the mutant alleles at two loci, given no mutant at the third locus, as defined in (4) takes the form

$$\epsilon_0 = b_2 - 2b_1 + b_0.$$

Similarly, according to (6), the epistatic interaction between the mutant alleles at two loci, given a mutant at the third locus, is

$$\epsilon_1 = b_3 - 2b_2 + b_1.$$

If b_i is convex as a function of i, then $\epsilon_0, \epsilon_1 > 0$. On the other hand, if b_i is concave, then $\epsilon_0, \epsilon_1 < 0$. Finally, if b_i is linear, then $\epsilon_0 = \epsilon_1 = 0$.

In the limit of a large population size N with $N^2/2$ time steps as unit of time, each lineage branches into two lineages as a result of a selection event at rate $\sigma/2$. The conditional probability that the continuing lineage rather than the incoming lineage is real, given that the continuing lineage carries type $\mathbf{i}$, is

$$1 - b_{|\mathbf{i}|}.$$

On the other hand, a lineage splits into two lineages following a recombination event at a rate $(\nu + \rho)/2$. Given a recombination event affecting a lineage of type $\mathbf{i} = i_1i_2i_3$, and assuming parent-independent recombination rates, the parental types $\mathbf{j} = j_1j_2j_3$ and $\mathbf{k} = k_1k_2k_3$ in this order are such that

$$\mathbf{i} = \begin{cases} j_1k_2k_3 & \text{with conditional probability } \dfrac{\nu}{\nu+\rho}, \\ j_1j_2k_3 & \text{with conditional probability } \dfrac{\rho}{\nu+\rho}. \end{cases} \tag{26}$$

We define

$$Q_{\mathbf{i}}(\mathbf{j}, \mathbf{k}) = \frac{\nu\delta_{\mathbf{i},j_1k_2k_3} + \rho\delta_{\mathbf{i},j_1j_2k_3}}{\nu+\rho}. \tag{27}$$

This is the conditional probability that ordered parental types $\mathbf{j}$ and $\mathbf{k}$ are compatible with an offspring of type $\mathbf{i}$ given a recombination event.

When a single mutant 1 is introduced at random at locus 1, four possible types are already present in the population, namely 000, 001, 010 and 011. Their respective frequencies are x_{000}, x_{001}, x_{010} and x_{011}. These are the probabilities of the values

$$\begin{aligned} \mathbf{x}_1 &= (x_{000} - N^{-1}, x_{001}, x_{010}, x_{011}, N^{-1}, 0, 0, 0), \\ \mathbf{x}_2 &= (x_{000}, x_{001} - N^{-1}, x_{010}, x_{011}, 0, N^{-1}, 0, 0), \\ \mathbf{x}_3 &= (x_{000}, x_{001}, x_{010} - N^{-1}, x_{011}, 0, 0, N^{-1}, 0), \\ \mathbf{x}_4 &= (x_{000}, x_{001}, x_{010}, x_{011} - N^{-1}, 0, 0, 0, N^{-1}), \end{aligned} \tag{28}$$

respectively, for the initial type frequency vector $\mathbf{x}(0)$. The corresponding probabilities of ultimate fixation of allele 1 at locus 1 are $u(\mathbf{x}_1)$, $u(\mathbf{x}_2)$, $u(\mathbf{x}_3)$, and $u(\mathbf{x}_4)$, respectively. The average fixation probability is then given by

$$u = x_{000}u(\mathbf{x}_1) + x_{001}u(\mathbf{x}_2) + x_{010}u(\mathbf{x}_3) + x_{011}u(\mathbf{x}_4).$$

Note that

$$\begin{aligned} x_{010} + x_{011} &= x_2, \\ x_{001} + x_{011} &= x_3, \end{aligned}$$

are the initial frequencies of allele 1 at loci 2 and 3, respectively, while

$$x_{011} - x_2x_3 = D_{23} = D$$

is the initial linkage disequilibrium between the mutants at loci 2 and 3.

Approximations of the average fixation probability u up to terms or order $O(\sigma^k)$ for $k = 3, 4$ or 5 are given in Appendix B. The calculations are based on Proposition 5, under the assumption that the population-scaled recombination rates ν and ρ are both of order $O(\sigma)$, and its implementation as described in Appendix A. The corresponding leading terms with respect to ν and ρ are denoted by $L_k(\nu)$ and $L_k(\rho)$, respectively. For an approximation of the form $u = a_0\sigma + a_1\sigma\rho + O(\sigma^3)$, for instance, we have $L_2(\nu) = 0$ and $L_2(\rho) = a_1\sigma\rho$. The following result gives expressions for these leading terms under various assumptions on the epistatic parameters ϵ_0, ϵ_1 and the initial linkage disequilibrium at loci 2 and 3 represented by D.

Proposition 6. *Let a single mutant* 1 *at locus* 1 *be introduced at random into a population of size N, so that its initial frequency is* $x_1 = N^{-1}$*, when previously introduced mutant alleles at loci* 2 *and* 3 *exhibit frequencies* x_2 *and* x_3*, respectively, and linkage disequilibrium* $D = x_{23} - x_2x_3$*. Assume the discrete-time Moran model of Proposition 1 with the mortality of type* $\mathbf{i}$ *depending only on the number of mutants* $|\mathbf{i}|$*, so that* $m_{\mathbf{i}} = 1 - b_{|\mathbf{i}|}\sigma N^{-1}$*, and population-scaled recombination rates* ν *and* ρ *of order* $O(\sigma)$ *between loci* 1, 2 *and loci* 2, 3*, respectively, that do not depend on the types of the parents. Then the leading terms with respect to* ν *and* ρ *in the average probability of ultimate fixation of allele* 1 *at locus* 1*, ignoring terms of order* $O(N^{-2})$ *and using the notation* $d = b_1 - b_0$*, are given by the following expressions:*

1. *for* $\epsilon_0 \neq \epsilon_1$ *and* $D \neq 0$

$$L_2(\nu) \approx 0, \qquad (29)$$
$$L_2(\rho) \approx -\frac{\sigma\rho}{6N}(\epsilon_1 - \epsilon_0)D;$$

2. *for* $\epsilon_0 = \epsilon_1 = \epsilon \neq 0$

$$L_3(\nu) \approx \frac{\sigma^2\nu\epsilon}{432N}\Big(\epsilon D(16 - 5x_2 - 5x_3) + 5d\big(x_2(1-x_2) + x_3(1-x_3) + 2D\big) + \epsilon\big(x_2(1-x_2) \times (3 + 5x_3) + x_3(1-x_3)(3 + 5x_2)\big)\Big), \qquad (30)$$

$$L_3(\rho) \approx \frac{\sigma^2\rho\epsilon}{432N}\Big(\epsilon D\,(11x_2 + 6x_3 - 52) - 44dD + x_3(1-x_3)\,(\epsilon\,(5x_2 + 3) + 5d)\Big); \qquad (31)$$

3. *for* $\epsilon_0 \neq \epsilon_1$ *and* $D = 0$

$$L_3(\nu) \approx \frac{\sigma^2\nu}{432N}\Big(\epsilon_0(1-x_2)(1-x_3)((x_2+x_3)(5d + 3\epsilon_0) + 2x_2x_3\epsilon_0) + \epsilon_1x_2x_3\big(5d(2 - x_2 - x_3) + (11 - 6x_2 - 6x_3 + x_2x_3)\epsilon_0\big) + 3\epsilon_1^2x_2x_3(1 - x_2x_3)\Big), \qquad (32)$$

$$L_3(\rho) \approx \frac{\sigma^2\rho}{432N}x_3(1-x_3)\Big((1-x_2)\epsilon_0(5d + (3-x_2)\epsilon_0) + x_2\epsilon_1(5d + (6-x_2)\epsilon_0) + 3x_2\epsilon_1^2\Big); \qquad (33)$$

4. *for* $\epsilon_0 = \epsilon_1 = 0$

$$L_4(\nu) \approx \frac{19\nu\sigma^3d^3}{432N}(2D + x_2(1-x_2) + x_3(1-x_3)), \qquad (34)$$
$$L_4(\rho) \approx \frac{\rho\sigma^3d^3}{432N}\,(31D + 19x_3(1-x_3))\,.$$

8. Ultimate fixation of a single neutral modifier for recombination

In this section we consider the probability of ultimate fixation of a new neutral modifier allele for recombination introduced at random at one of three loci, say locus 1 among loci 1, 2 and 3 in this order. The rates of recombination between loci 1, 2 and loci 2, 3 are defined in (1) for parents of types $\mathbf{j}$ and $\mathbf{k}$ in this order. With $N^2/2$ time steps as unit of time in a Moran model in the limit of a large population of size N, each lineage splits into two lineages as a result of a recombination event at a rate $(\nu_2 + \rho_2)/2$. Given a recombination event, parental types $\mathbf{j} = j_1j_2j_3$ and $\mathbf{k} = k_1k_2k_3$ in this order are compatible with an offspring of type $\mathbf{i} = i_1i_2i_3$ with conditional probability given by

$$Q_{\mathbf{i}}(\mathbf{j},\mathbf{k}) = \frac{q_1(\mathbf{j},\mathbf{k})\delta_{\mathbf{i},j_1k_2k_3} + q_2(\mathbf{j},\mathbf{k})\delta_{\mathbf{i},j_1j_2k_3}}{\nu_2 + \rho_2}, \qquad (35)$$

where

$$q_1(\mathbf{j},\mathbf{k}) = \nu_0\delta_{0,j_1+k_1} + \nu_1\delta_{1,j_1+k_1} + \nu_2\delta_{2,j_1+k_1}$$

and

$$q_2(\mathbf{j},\mathbf{k}) = \rho_0\delta_{0,j_1+k_1} + \rho_1\delta_{1,j_1+k_1} + \rho_2\delta_{2,j_1+k_1}.$$

Moreover, since the modifier allele is assumed to have no effect on mortality, the coefficient of selection $c_{\mathbf{i}}$ for an individual of type $\mathbf{i} = i_1i_2i_3$ satisfies

$$\begin{aligned} c_{100} &= c_{000} = c_{00},\\ c_{001} &= c_{101} = c_{01},\\ c_{010} &= c_{110} = c_{10},\\ c_{011} &= c_{111} = c_{11}. \end{aligned} \qquad (36)$$

Then the quantity

$$\epsilon = \epsilon_{2,3} = c_{11} - c_{10} - c_{01} + c_{00} \qquad (37)$$

measures the epistatic interaction between the mutant alleles at loci 2 and 3, given no mutant at locus 1.

Let u be the probability of ultimate fixation of the new modifier allele, averaged over all possible initial type frequency vectors as explained in the previous section. Using Proposition 5 for population-scaled recombination rates of order $O(\sigma)$ and its implementation described in Appendix A, approximations of order $O(\sigma^k)$ for u can be obtained for any $k \geq 1$.

Proposition 7. *Assume that a single neutral recombination modifier* 1 *at locus* 1 *is introduced at random into a population of size N, so that its initial frequency is* $x_1 = N^{-1}$*, when previously introduced mutant alleles under selection at loci* 2 *and* 3 *exhibit frequencies* x_2 *and* x_3*, respectively, and linkage disequilibrium* $D = x_{23} - x_2x_3$*. The mortality of type* $\mathbf{i} = i_1i_2i_3$ *in the discrete-time Moran model of Proposition 1 is given by* $m_{\mathbf{i}} = 1 - c_{i_2i_3}\sigma N^{-1}$*. The population-scaled recombination rate between loci* 2 *and* 3 *is* $\rho_{j_1+k_1}$ *with* $0 \leq \rho_0 \leq \rho_1 \leq \rho_2$*, if the parental types are* $\mathbf{j} = j_1j_2j_3$ *and* $\mathbf{k} = k_1k_2k_3$*. Similarly the corresponding rate between loci* 1 *and* 2 *is* $\nu_{i_1+j_1}$ *with* $0 \leq \nu_0 \leq \nu_1 \leq \nu_2$*. Assuming that* ν_l *and* ρ_l $(l = 0, 1, 2)$ *are of order* $O(\sigma)$ *for* $\sigma \ll 1$*, and ignoring terms of order* $O(N^{-2})$*, we have the following approximations for the average probability of ultimate fixation of allele* 1 *at locus* 1*:*

1. *for* $D \neq 0$ *and* $\epsilon \neq 0$

$$u = \frac{1}{N}\left(1 - \frac{\sigma(\rho_1 - \rho_0)}{12}\epsilon D\right) + O(\sigma^3); \qquad (38)$$

2. *for* $D = 0$ *and* $\epsilon \neq 0$

$$u = \frac{1}{N}\left(1 + \frac{\sigma^2(\rho_1 - \rho_0)\epsilon^2}{432}x_2(1-x_2)x_3(1-x_3)\right) + O(\sigma^4); \qquad (39)$$

3. *for* $D \neq 0$ *and* $\epsilon = 0$

$$u = \frac{1}{N}\left(1 - \frac{\sigma^2(\rho_1 - \rho_0)}{36}D(c_{10} - c_{00})(c_{01} - c_{00})\right) + O(\sigma^4); \qquad (40)$$

4. *for* $D = 0$ *and* $\epsilon = 0$ *with* $|c_{10} - c_{00}| = |c_{01} - c_{00}| = 1/2$

$$u = \frac{1}{N}\left(1 + \frac{61(\rho_1 - \rho_0)\sigma^4}{1036800}x_3(1-x_3)x_2(1-x_2)\right) + O(\sigma^6). \qquad (41)$$

9. Discussion

9.1. Effect of recombination on ultimate fixation of mutants under selection in a three-locus model

The effect of recombination on the probability of ultimate fixation of a new mutant allele indicates how recombination influences the rate of adaptation. This effect can be studied through the sign of the coefficient of the leading recombination terms in approximations for the fixation probability as given in Proposition 6. When mortality selection in a three-locus, two-allele discrete-time Moran model depends on the number of mutants and recombination is parent-independent, then (29) predicts that an increase in the recombination rate between loci 2 and 3 increases the average probability of ultimate fixation of a new single mutant introduced at random at locus 1 if

$$D(\epsilon_1 - \epsilon_0) < 0,$$

where D is the linkage disequilibrium between the mutants at loci 2 and 3 prior to the introduction of the single mutant at locus 1, and $\epsilon_1 - \epsilon_0$ is the difference that this latter mutant makes in the epistatic interaction between the former mutants.

Suppose that all mutations are beneficial so that selection is directional with the coefficient of selection increasing as the number of mutants increases, that is, $b_0 < b_1 < b_2 < b_3$. Suppose also that the mutants at loci 2 and 3 were introduced at random one at a time, so that they were in average linkage equilibrium the first time that they were both present in the population. Then the expected value of D at any later time as long as the mutants are segregating with the wild types should be negative as the result of the combined effects of random drift and selection. This holds in the absence of epistasis, that is

$$\epsilon_0 = (b_2 - b_1) - (b_1 - b_0) = 0,$$

so that the beneficial effect of a second mutant on a first mutant is equal to the beneficial effect of the first mutant (Hill and Robertson, 1966). This holds also in the presence of positive epistasis ($\epsilon_0 > 0$) or weak negative epistasis ($\epsilon_0 < 0$ and small in absolute value) as shown in Otto and Barton (2001) and (Lessard and Kermany, 2012), and not necessarily for any level of negative epistasis as could be expected from what is known for an infinite population (see, e.g., Eshel and Feldman, 1970). Then a higher recombination rate between loci 2 and 3 renders selection more effective at these loci by increasing the variance in fitness. This in turn favors the ultimate fixation of a single mutant introduced at random at locus 1 if

$$\epsilon_1 - \epsilon_0 = (b_3 - b_2) - 2(b_2 - b_1) + (b_1 - b_0) > 0.$$

This condition means that the epistatic effect of a third beneficial mutant given at least one beneficial mutant, is larger than the epistatic effect of a second beneficial mutant.

A symmetric conclusion can be reached for deleterious mutants so that $b_0 > b_1 > b_2 > b_3$. As long as $D < 0$, the ultimate fixation of a single mutant introduced at random at locus 1 is disfavored by an increase in the recombination rate between loci 2 and 3 if $\epsilon_1 - \epsilon_0 < 0$. This means that the epistatic effect of a third deleterious mutant given at least one deleterious mutant, is smaller than the epistatic effect of a second deleterious mutant.

Note that the first-order effect of recombination between loci 1 and 2 is negligible compared to the first-order effect of recombination between loci 2 and 3. The reason is likely that the average LD between the alleles at these loci is zero when the mutant allele is introduced at random at locus 1.

If $\epsilon_0 = \epsilon_1 = \epsilon$ (this is the case when $b_i = a + bi + ci^2$ with $\epsilon = 2c$), then higher-order terms in the approximation of the fixation probability have to be considered to detect the effect of recombination. Assuming that epistasis is weak so that terms of order $O(\epsilon^2)$ can be ignored, Eqs. (30) and (31) for the leading terms with respect to the population-scaled recombination rates ν and ρ between loci 1, 2 and loci 2, 3, respectively, simplify to

$$L_3(\nu) \approx \frac{5\sigma^2\nu\epsilon}{432N} d\Big(2D + x_2(1-x_2) + x_3(1-x_3)\Big),$$

$$L_3(\rho) \approx \frac{\sigma^2\rho\epsilon}{432N} d\Big(5x_3(1-x_3) - 44D\Big).$$

Since $d = b_1 - b_0$ is positive in the case of advantageous mutants and negative in the case of deleterious mutants, increasing the recombination rates will have a favorable effect on the rate of adaptation under the condition

$$-\frac{x_2(1-x_2) + x_3(1-x_3)}{2} < D < \frac{5x_3(1-x_3)}{44},$$

in the case $\epsilon > 0$, and under the condition

$$-\frac{5x_3(1-x_3)}{44} < D < \frac{x_2(1-x_2) + x_3(1-x_3)}{2}$$

in the case $\epsilon < 0$. Note that the former condition is more easily satisfied under the constraint $D < 0$.

When $D = 0$ and epistasis is weak with $\epsilon_0 \neq \epsilon_1$, then the leading recombination terms given in (32) and (33) simplify to

$$L_3(\nu) \approx \frac{5d\sigma^2\nu}{432N}\Big(\epsilon_0(1-x_2)(1-x_3)(x_2+x_3) + \epsilon_1 x_2 x_3\big(2 - x_2 - x_3\big)\Big),$$

$$L_3(\rho) \approx \frac{5d\sigma^2\rho}{432N} x_3(1-x_3)\Big((1-x_2)\epsilon_0 + x_2\epsilon_1\Big).$$

Hence, positive epistasis between the mutants at loci 2 and 3 regardless of the allele at locus 1 ($\epsilon_0, \epsilon_1 > 0$) provides a sufficient condition for more recombination to be evolutionary advantageous.

Note that if $D = 0$ with $x_3 = 0$, then the average probability of ultimate fixation of the mutant allele at locus 1 as given in Eq. (54) in Appendix B takes the form

$$\begin{aligned} u = {} & \frac{1}{N} + \frac{\sigma}{2N}\Big(d + x_2\epsilon_0\Big) \\ & + \frac{\sigma^2}{12N}\Big(d^2 + 2d\epsilon_0 x_2\,(2 - x_2) + \epsilon_0^2 x_2\Big) \\ & - \frac{\sigma^3}{24N} d^2 x_2(1-x_2)\big(b_1 - b_0 + x_2\epsilon_0\big) \\ & + \frac{\sigma^2\nu}{432N}\epsilon_0 x_2(1-x_2)(5d + 3\epsilon_0) + O(\sigma^4). \end{aligned} \tag{42}$$

This is consistent with previous results for the two-locus model (see Eq. (70) in Lessard and Kermany, 2012) in the special case where the mortality of an individual depends only on the number of mutants that it carries.

Finally, in the absence of epistasis with $\epsilon_0 = \epsilon_1 = 0$, higher-order recombination terms exhibited in Eq. (34) reveal that more recombination is favored (in either case of all advantageous mutations or all deleterious mutations) if linkage disequilibrium is not too negative, actually if

$$D \geq \max\left(-\frac{19x_3(1-x_3)}{31}, -\frac{x_2(1-x_2) + x_3(1-x_3)}{2}\right).$$

It might be that increasing recombination between loci 2 and 3 when D is too negative causes the mean fitness of the background genotypes to increase, reducing the relative fitness (and hence the fixation probability) of the focal allele at locus 1. Note that the condition for more recombination to be favored is weaker when $x_2(1-x_2)$ and $x_3(1-x_3)$ take larger values, that is, when the frequencies of the resident mutations are far from 0 or 1. Moreover, recombination is always advantageous when $D = 0$, in agreement with the Hill–Robertson effect.

9.2. Ultimate fixation of neutral modifiers for recombination among loci under selection

The previous conclusions are based on the effects that recombination rates have on the fixation probability of new mutants under directional selection. The sign of the effects tells us if recombination is advantageous in enhancing the incorporation rate of beneficial mutants (or impeding the fixation rate of deleterious mutants). In order to study the evolution of recombination based on genetic mechanisms, the fate of a neutral mutant modifier whose sole role is to increase the recombination rates among three loci, including the modifier locus, is considered.

In the absence of selection, a single modifier has probability of ultimate fixation given by the inverse of the population size, namely N^{-1}. According to Proposition 7, the leading selection term in the average probability of ultimate fixation of a single modifier introduced at random at locus 1 is positive if

$$\epsilon D < 0,$$

where ϵ is the epistatic interaction between the mutant alleles under selection at loci 2 and 3, which exhibit linkage disequilibrium D just before the introduction of the modifier. This provides a sufficient condition for weak selection to favor ultimate fixation of the modifier. Note that this condition requires a mismatch between the sign of disequilibrium present and the sign of epistasis, contrary to the current view from studies in infinite populations that both must be negative for recombination to be favored (Eshel and Feldman, 1970).

If $D = 0$ and $\epsilon \neq 0$, then according to Eq. (39) the modifier is always favored to ultimately fix in the population under weak selection regardless of the sign of ϵ. Note that the effect of selection on the fixation probability of the modifier increases with the magnitude of ϵ and the values of $x_2(1-x_2)$ and $x_3(1-x_3)$, where x_2 and x_3 are the initial frequencies of the mutants at the selected loci 2 and 3, respectively.

If $D \neq 0$ and $\epsilon = 0$, with the mutants at the selected loci 2 and 3 being either both deleterious or both advantageous, then Eq. (40) predicts that weak selection favors ultimate fixation of the modifier if $D < 0$. This is the most likely initial condition on linkage disequilibrium in the population just prior to the introduction of the new modifier, since as a result of genetic drift and natural selection linkage disequilibrium between the mutants at loci 2 and 3 as measured by D is expected to be negative on average. However, if one of the mutants at the selected loci 2 and 3 is advantageous and the other deleterious, then positive initial LD provides a favorable condition for the modifier of recombination to fix under weak selection.

In the case $D = 0$ and $\epsilon = 0$, with coefficients of selection defined in (36) and satisfying $|c_{10} - c_{00}| = |c_{01} - c_{00}| = 1/2$, the ultimate fixation of the modifier is always favored by weak selection according to Eq. (41). Relabeling the alleles if necessary, the above conditions on the coefficients of selection in the absence of epistasis mean that allele 1 at each of the loci 2 and 3 has the same additive effect.

In Proposition 7, we assumed that the population-scaled recombination rates between loci 1, 2 and loci 2, 3 increase with the number of mutant modifiers in the parents, that is $\nu_0 \leq \nu_1 \leq \nu_2$ and $\rho_0 \leq \rho_1 \leq \rho_2$, respectively. However, the results stated remain valid if the inequalities are reversed, so that the new modifier reduces the recombination rates. In this case, it suffices to take $(\nu_0 + \rho_0)/2$ instead of $(\nu_2 + \rho_2)/2$ as rate of recombination of each lineage in the limiting ancestral process. The rest of the analysis follows *mutatis mutandis*.

In all results on the fixation probability of a mutant modifier increasing the rates of recombination, the leading effect of weak selection increases with $\rho_1 - \rho_0$. Given the inequalities $\rho_0 \leq \rho_1 \leq \rho_2$, the effect is maximum when $\rho_2 = \rho_1$ and $\rho_0 = 0$. This is the case when the mutant modifier is dominant and the resident modifier suppresses recombination between loci 2 and 3.

Moreover, the leading effect of weak selection on the fixation probability of a mutant modifier does not depend on the population-scaled recombination rates ν_0, ν_1, ν_2 between locus 1, which is the modifier locus, and locus 2, which is the closer locus under selection. This is likely due to the fact that the modifier locus is neutral and in initial linkage equilibrium on average with the selected loci.

Note, however, that the population-scaled recombination rates between the modifier locus and the selected loci come into play in higher-order terms in the approximation of the fixation probability. When the mutant modifier is recessive so that $\rho_1 = \rho_0$, for instance, then the leading effect of weak selection in the case $\epsilon D \neq 0$ is captured by the approximation

$$u = \frac{1}{N}\left(1 - \frac{\sigma \nu_1(\rho_2 - \rho_0)}{72}\epsilon D\right) + O(\sigma^4), \tag{43}$$

for the probability of ultimate fixation of a single mutant modifier. The effect may not be of the same order, but its sign is consistent.

Moreover, in the case of recombination modifiers at locus 2 and selected alleles at loci 1 and 3, so that the modifier locus is between the loci under selection, it can be checked that the results of Proposition 7 are valid with ρ_0 and ρ_1 replaced by $\rho_0 + \nu_0$ and $\rho_1 + \nu_1$, respectively. Again, only the recombination rates between the loci under selection seem to matter, but this is ascertained only for the leading terms.

9.3. Final thoughts

It has often been argued that negative linkage disequilibrium is a necessary condition for recombination to be favored in the presence of directional selection. On the other hand negative epistasis, by giving a selective advantage to pairs of single mutants over pairs of a double mutant and a wild type, contributes to promote the frequencies of the former over the latter (Felsenstein, 1965; Eshel and Feldman, 1970; Barton, 1995; Otto and Barton, 1997). This seems to be a key factor in an infinite population, where it is difficult to find another force capable of creating negative linkage disequilibrium. However, in a finite population random negative associations may build up even in the absence of epistasis due to the Hill–Robertson interference between selected loci (Hill and Robertson, 1966; Felsenstein, 1974). As pointed out by Barton (2010), simulations suggest that drift in the presence of any kind of directional selection is more efficient than negative epistasis in creating negative associations (Otto and Barton, 2001; Keightley and Otto, 2006). Previous results (Lessard and Kermany, 2012) even suggest that positive epistasis rather than negative epistasis enhances the effect of drift on the average linkage disequilibrium.

Therefore, the condition $\epsilon D < 0$ for weak selection to favor a new modifier which increases the recombination rate to ultimately fix in the population is more likely to occur when $D < 0$ and $\epsilon > 0$. The alternate scenario with $D > 0$ and $\epsilon < 0$ is less likely to occur, since the initial LD at the selected loci when the new modifier is introduced into the population is expected to be negative.

Note, however, that our conclusions are based on approximations for the fixation probability in the case of population-scaled intensity of selection and recombination rates that are of the same small order. This means not only that selection and recombination are weak, but also that they are weaker than drift whose strength is measured by the inverse of the population size.

Moreover, the results have been deduced from a discrete-time Moran model with mortality selection in the limit of a large population size. A viability model with a probability of replacement depending on the type of the offspring produced instead of the

type of the individual chosen to be replaced could be treated analogously. More importantly, the same results are expected to hold for a wide range of models including the Wright–Fisher model in the limit of a large population size as long as the corresponding ancestral recombination–selection graph, ARSG, converges to the same limiting ancestral process known as the ancestral influence graph, AIG (Donnelly and Kurtz, 1999). This generalization would raise further technical difficulties. An alternative approach would be a derivation carried through a perturbation of the diffusion equation for fixation probability. This does not look simpler though.

The approach is computationally intensive and this limits the order of the approximations that can be reached. For an approximation of order k with respect to the population-scaled recombination and selection parameters, ancestral recombination–selection graphs with up to $k-1$ splittings or branchings of lineages corresponding to recombination or selection events, respectively, have to be considered. The number of possible graphs increases exponentially as the number of these events increases.

On the other hand, the approach can be extended to approximate the fixation probability of an allele or a gamete in a multilocus setting, given any initial state of the population.

A Mathematica package for calculating the fixation probabilities presented in this article is archived as supplementary material.

Acknowledgments

We thank two anonymous reviewers for helpful comments to improve this paper. This research was supported in part by the Natural Sciences and Engineering Research Council of Canada.

Appendix A. Conditional sample probabilities given genealogies

In this appendix we present an algorithm for calculating the conditional probability $P_G(\mathbf{z}|\mathbf{x})$ for an ordered sample of size n represented by $\mathbf{z} = (z_1, \ldots, z_n)$, given a sequence of events backward in time $G = (G_1, \ldots, G_m)$, where G_k represents a pure coalescence, selection or recombination event for $k = 1, \ldots, m$, and given individual types represented initially in the past in the population with frequencies $\mathbf{x} = (x_\mathbf{i})$.

The sample configuration corresponding to $\mathbf{z} = (z_1, \ldots, z_n)$ is given by $\mathbf{n} = (n_\mathbf{i})$ with $|\mathbf{n}| = \sum_\mathbf{i} n_\mathbf{i} = n$, where $n_\mathbf{i}$ represents the number of individuals of type $\mathbf{i}$ in the sample, that is, the number of j such that $z_j = n_\mathbf{i}$ for $j = 1, \ldots, n$. For example, if $\mathbf{z} = (101, 000, 000, 111)$, then $\mathbf{n} = (2, 0, 0, 0, 0, 1, 0, 1)$ and $|\mathbf{n}| = 4$.

If $G = \Phi$, which means no change in the ordered sample $\mathbf{z}$, then

$$P_\Phi(\mathbf{z}|\mathbf{x}) = N^{-n} \prod_\mathbf{i} (N_\mathbf{i})_{n_\mathbf{i}}, \tag{44}$$

where $(N)_n = N \times (N-1) \times \cdots \times (N-n+1)$ denotes a falling factorial, while

$$N\mathbf{x} = (N_\mathbf{i})$$

represents the population configuration at time step 0. In the limit of large population size

$$P_\Phi(\mathbf{z}|\mathbf{x}) = P_\mathbf{x}(\mathbf{n}) = \prod_\mathbf{i} x_\mathbf{i}^{n_\mathbf{i}}, \tag{45}$$

which depends only on the sample configuration and the population state.

Now consider $G = G_1$. Denote by $\mathbf{z}' = (z'_1, \ldots, z'_{n'})$ the ancestral ordered sample obtained as a result of the pure coalescence, recombination or selection event represented by G_1 affecting the ordered sample $\mathbf{z}$. The corresponding sample configuration is $\mathbf{n}' = (n'_\mathbf{i})$ with $|\mathbf{n}'| = \sum_\mathbf{i} n'_\mathbf{i} = n'$.

The result of a pure coalescence event C_n involving the lineages of the sampled individuals i and j for $i < j$ is

$$z'_k = \begin{cases} z_k & \text{if } k < j, \\ z_{k+1} & \text{if } k \geq j, \end{cases}$$

for $k = 1, \ldots, n-1$. In this case we write $\mathbf{z}' = \mathcal{C}_{ij}(\mathbf{z})$. For example, if $\mathbf{z} = (000, 111, 000, 101, 111)$, then $\mathcal{C}_{13}(\mathbf{z}) = (000, 111, 101, 111)$ and $\mathcal{C}_{25}(\mathbf{z}) = (000, 111, 000, 101)$.

The result of a pure selection event S_n involving the lineage of the sampled individual i as continuing lineage and a lineage carrying type $\mathbf{j}$ as incoming lineage is given by

$$z'_k = \begin{cases} z_k & \text{if } k \leq i, \\ \mathbf{j} & \text{if } k = i+1, \\ z_{k-1} & \text{if } k > i+1, \end{cases}$$

for $k = 1, \ldots, n+1$. Then we write $\mathbf{z}' = \mathcal{S}_{i,\mathbf{j}}(\mathbf{z})$. If $\mathbf{z} = (000, 111, 000, 101, 111)$, for instance, then $\mathcal{S}_{2,010}(\mathbf{z}) = (000, 111, 010, 000, 101, 111)$.

Similarly the result is

$$z'_k = \begin{cases} z_k & \text{if } k < i, \\ \mathbf{j} & \text{if } k = i, \\ z_{k-1} & \text{if } k \geq i+1, \end{cases}$$

for $k = 1, \ldots, n+1$, in the case of an incoming lineage carried by i and a continuing lineage carrying $\mathbf{j}$. In this case we write $\mathbf{z}' = \mathcal{S}_{\mathbf{j},i}$. With $\mathbf{z} = (000, 111, 000, 101, 111)$ as previously, we have $\mathcal{S}_{010,2}(\mathbf{z}) = (000, 010, 111, 000, 101, 111)$.

Finally, the result of a pure recombination event R_n involving the sampled individual i with parents of types $\mathbf{j}$ and $\mathbf{k}$ in this order is

$$z'_k = \begin{cases} z_k & \text{if } k < i, \\ \mathbf{j} & \text{if } k = i, \\ \mathbf{k} & \text{if } k = i+1, \\ z_{k-1} & \text{if } k > i+1, \end{cases}$$

for $k = 1, \ldots, n+1$. Then we write $\mathbf{z}' = \mathcal{R}_{\mathbf{j},i,\mathbf{k}}(\mathbf{z})$. For example, if $\mathbf{z} = (000, 111, 000, 101, 111)$, then $\mathcal{R}_{100,2,101}(\mathbf{z}) = (000, 100, 011, 000, 101, 111)$.

Using the above definitions, the conditional probability of an ordered sample $\mathbf{z} = (z_1, \ldots, z_n)$ given a pure event of coalescence, recombination or selection G_1 and an ancestral ordered sample $\mathbf{z}'$ is

$$P_{G_1}(\mathbf{z}|\mathbf{z}') = \begin{cases} \delta_{z_i,z_j} & \text{if } G_1 = C_n \text{ and } \mathbf{z}' = \mathcal{C}_{ij}(\mathbf{z}), \\ c_{z_i} & \text{if } G_1 = S_n \text{ and } \mathbf{z}' = \mathcal{S}_{i,\mathbf{j}}(\mathbf{z}), \\ 1 - c_\mathbf{j} & \text{if } G_1 = S_n \text{ and } \mathbf{z}' = \mathcal{S}_{\mathbf{j},i}(\mathbf{z}), \\ Q_{z_i}(\mathbf{j}, \mathbf{k}) & \text{if } G_1 = R_n \text{ and } \mathbf{z}' = \mathcal{R}_{\mathbf{j},i,\mathbf{k}}(\mathbf{z}), \\ 0 & \text{otherwise,} \end{cases}$$

for $i, j = 1, \ldots, n$ with $i < j$, where $Q_{z_i}(\mathbf{j}, \mathbf{k})$ is defined by (27) and (35).

Now define

$$\Psi_{G_1}(\mathbf{n}, \mathbf{n}') = \sum_{\mathbf{z}':\mathbf{n}'} P_{G_1}(\mathbf{z}|\mathbf{z}'), \tag{46}$$

where the summation is over all ancestral ordered samples $\mathbf{z}'$ with a given sample configuration $\mathbf{n}'$. One obtains that

$$\Psi_{G_1}(\mathbf{n}, \mathbf{n}') = \begin{cases} n_\mathbf{i}(n_\mathbf{i} - 1)/2 & \text{if } G_1 = C_n \\ & \text{and } \mathbf{n}' = \mathbf{n} - \mathbf{e}_\mathbf{i}, \\ \sum_\mathbf{i} n_\mathbf{i} c_\mathbf{i} + n(1 - c_\mathbf{j}) & \text{if } G_1 = S_n \\ & \text{and } \mathbf{n}' = \mathbf{n} + \mathbf{e}_\mathbf{j}, \\ n_\mathbf{i} Q_\mathbf{i}(\mathbf{j}, \mathbf{k}) & \text{if } G_1 = R_n \\ & \text{and } \mathbf{n}' = \mathbf{n} - \mathbf{e}_\mathbf{i} + \mathbf{e}_\mathbf{j} + \mathbf{e}_\mathbf{k}, \\ \sum_\mathbf{i} n_\mathbf{i} Q_\mathbf{i}(\mathbf{i}, \mathbf{k}) & \text{if } G_1 = R_n \\ & \text{and } \mathbf{n}' = \mathbf{n} + \mathbf{e}_\mathbf{k}, \\ 0 & \text{otherwise.} \end{cases} \tag{47}$$

Here, $\mathbf{e}_\mathbf{j}, \mathbf{e}_\mathbf{k} \neq \mathbf{e}_\mathbf{i}$ with $\mathbf{e}_\mathbf{i}$ being the unit vector defined in (11).

We are now ready to calculate $P_{G_1}(\mathbf{z}|\mathbf{x})$. Conditioning on $\mathbf{z}'$, we have

$$P_{G_1}(\mathbf{z}|\mathbf{x}) = \sum_{\mathbf{z}'} P_{G_1}(\mathbf{z}|\mathbf{z}')P_{\Phi}(\mathbf{z}'|\mathbf{x}),$$

from which

$$P_{G_1}(\mathbf{z}|\mathbf{x}) = \sum_{\mathbf{n}'} \Psi_{G_1}(\mathbf{n}, \mathbf{n}')P_{\mathbf{x}}(\mathbf{n}'),$$

using the notation introduced in (45) and (46).

Similarly for $G = (G_1, G_2)$, we find that

$$P_G(\mathbf{z}|\mathbf{x}) = \sum_{\mathbf{n}''} \Psi_G(\mathbf{n}, \mathbf{n}'')P_{\mathbf{x}}(\mathbf{n}''),$$

where

$$\Psi_G(\mathbf{n}, \mathbf{n}'') = \sum_{\mathbf{z}'':\mathbf{n}''} P_G(\mathbf{z}|\mathbf{z}''), \tag{48}$$

with $\mathbf{z}''$ representing the ancestral ordered sample obtained as a result of the pure events G_1 and G_2 affecting the ordered sample $\mathbf{z}$, and $\mathbf{n}''$ being the corresponding sample configuration. Conditioning on the ancestral ordered sample $\mathbf{z}'$ of configuration $\mathbf{n}'$ following G_1, one obtains that

$$\begin{aligned}\Psi_G(\mathbf{n}, \mathbf{n}'') &= \sum_{\mathbf{z}'':\mathbf{n}''} \sum_{\mathbf{z}'} P_{G_1}(\mathbf{z}|\mathbf{z}')P_{G_2}(\mathbf{z}'|\mathbf{z}'') \\ &= \sum_{\mathbf{z}'} P_{G_1}(\mathbf{z}|\mathbf{z}')\Psi_{G_2}(\mathbf{n}', \mathbf{n}'') \\ &= \sum_{\mathbf{n}'} \Psi_{G_1}(\mathbf{n}, \mathbf{n}')\Psi_{G_2}(\mathbf{n}', \mathbf{n}''). \end{aligned} \tag{49}$$

Induction on $m \geq 1$ for any $G = (G_1, \ldots, G_m)$ leads to the following result.

Proposition 8. *Let $\mathbf{z} = (z_1, \ldots, z_n)$ be an ordered random sample of size n and of configuration $\mathbf{n} = (n_{\mathbf{i}})$ taken from a population of large size N at the current time step. Consider a given sequence of pure events of coalescence, recombination or selection backward in time represented by $G = (G_1, \ldots, G_m)$ affecting the lineages of the sampled individuals up to the initial time step. Let $\mathbf{z}^{(0)} = \mathbf{z}, \mathbf{z}^{(1)}, \ldots, \mathbf{z}^{(m)}$ and $\mathbf{n}^{(0)} = \mathbf{n}, \mathbf{n}^{(1)}, \ldots, \mathbf{n}^{(m)}$ be the corresponding sequence of ancestral ordered samples and configurations. Given G and the type frequency vector at the initial time step represented by $\mathbf{x} = (x_{\mathbf{i}})$, the conditional probability of $\mathbf{z}$ is*

$$P_G(\mathbf{z}|\mathbf{x}) = \sum_{\mathbf{n}^{(m)}} \Psi_G(\mathbf{n}^{(0)}, \mathbf{n}^{(m)})P_{\mathbf{x}}(\mathbf{n}^{(m)}),$$

where

$$P_{\mathbf{x}}(\mathbf{n}^{(m)}) = \prod_{\mathbf{i}} x_{\mathbf{i}}^{n_{\mathbf{i}}^{(m)}} \tag{50}$$

and

$$\Psi_G(\mathbf{n}^{(0)}, \mathbf{n}^{(m)}) = \sum_{\mathbf{n}^{(1)}} \cdots \sum_{\mathbf{n}^{(m-1)}} \prod_{k=1}^{m} \Psi_{G_k}(\mathbf{n}^{(k-1)}, \mathbf{n}^{(k)}), \tag{51}$$

with

$$\Psi_{G_k}(\mathbf{n}, \mathbf{n}') = \begin{cases} n_{\mathbf{i}}(n_{\mathbf{i}} - 1)/2 & \text{if } G_k = C_n \\ & \text{and } \mathbf{n}' = \mathbf{n} - \mathbf{e}_{\mathbf{i}}, \\ \sum_{\mathbf{i}} n_{\mathbf{i}} c_{\mathbf{i}} + n(1 - c_{\mathbf{j}}) & \text{if } G_k = S_n \\ & \text{and } \mathbf{n}' = \mathbf{n} + \mathbf{e}_{\mathbf{j}}, \\ n_{\mathbf{i}} Q_{\mathbf{i}}(\mathbf{j}, \mathbf{k}) & \text{if } G_k = R_n \\ & \text{and } \mathbf{n}' = \mathbf{n} - \mathbf{e}_{\mathbf{i}} + \mathbf{e}_{\mathbf{j}} + \mathbf{e}_{\mathbf{k}}, \\ \sum_{\mathbf{i}} n_{\mathbf{i}} Q_{\mathbf{i}}(\mathbf{i}, \mathbf{k}) & \text{if } G_k = R_n \\ & \text{and } \mathbf{n}' = \mathbf{n} + \mathbf{e}_{\mathbf{k}}, \\ 0 & \text{otherwise,} \end{cases}$$

in which $\mathbf{j}, \mathbf{k} \neq \mathbf{i}$ represent individual types and $Q_{\mathbf{i}}(\mathbf{j}, \mathbf{k})$ is defined in (27) and (35).

Appendix B. Fixation probabilities for the three-locus model

Consider the conditions of Proposition 6 and define $d = b_1 - b_0$. In the case $\epsilon_1 \neq \epsilon_0$ and $D \neq 0$, the average probability of ultimate fixation of allele 1 at locus 1 is approximated by

$$\begin{aligned} u = {} & \frac{1}{N} + \frac{\sigma}{2N}\Big(d + x_2 + x_3\epsilon_0 + (\epsilon_1 - \epsilon_0)x_{011}\Big) \\ & + \frac{\sigma^2}{12N}\Big(d^2 + 2d\epsilon_0(x_2 + x_3)(2 - x_2 - x_3) \\ & + \epsilon_0^2(x_2 + x_3) + x_{011}\big(\epsilon_0^2(1 - 2(x_2 + x_3)) + 4\epsilon_0\epsilon_1 + \epsilon_1^2 \\ & - 2d\epsilon_0(1 - x_2 - x_3) - 2d\epsilon_1(3 - x_2 - x_3)\big) \\ & + 2\epsilon_0(\epsilon_0 - \epsilon_1)x_{011}^2\Big) - \frac{\sigma\rho}{6N}(\epsilon_1 - \epsilon_0)D + O(\sigma^3). \end{aligned} \tag{52}$$

In the case $\epsilon_0 = \epsilon_1 = \epsilon \neq 0$, we have the approximation

$$\begin{aligned} u = {} & \frac{1}{N} + \frac{\sigma}{2N}(d + \epsilon(x_2 + x_3)) \\ & + \frac{\sigma^2}{12N}\Big(2\epsilon x_{011}\,(2d + (3 - x_2 - x_3)\epsilon) \\ & + 2\epsilon d(x_2 + x_3)(2 - x_2 - x_3) + d^2 + (x_2 + x_3)\epsilon^2\Big) \\ & + \frac{\sigma^3}{24N}\Big((d + \epsilon(x_2 + x_3))\,\big(x_{011}\big(\epsilon^2 x_{011} \\ & + 2d\epsilon(x_2 + x_3 - 2) - 2d^2 - \epsilon^2\big) \\ & + d^2(x_2 + x_3 - 1)(x_2 + x_3)\big)\Big) \\ & + \frac{\sigma^2\rho\epsilon}{432N}\Big(D\Big(\epsilon\,(11x_2 + 6x_3 - 52) - 44d\Big) \\ & + x_3(1 - x_3)\,(\epsilon\,(5x_2 + 3) + 5d)\Big) \\ & + \frac{\sigma^2\nu\epsilon}{432N}\Big(D\big(10d + \epsilon(16 - 5x_2 - 5x_3)\big) \\ & + 5d(x_2(1 - x_2) + x_3(1 - x_3)) \\ & + \epsilon\big(x_2(1 - x_2)(3 + 5x_3) + x_3(1 - x_3)(3 + 5x_2)\big)\Big) \\ & + O(\sigma^4). \end{aligned} \tag{53}$$

If $\epsilon_1 \neq \epsilon_0$ and $D = 0$, then we have

$$\begin{aligned} u = {} & \frac{1}{N} + \frac{\sigma}{2N}\Big(d + (x_2 + x_3 - x_2x_3)\epsilon_0 + x_2x_3\epsilon_1\Big) \\ & + \frac{\sigma^2}{12N}\Big(d^2 - 2d\big(x_2^2(1 - x_3) \\ & - x_2(2 - x_3)(1 - x_3) - (2 - x_3)x_3\big)\epsilon_0 \\ & + \big(x_3 + x_2(1 - x_3)(1 + 2(1 - x_2)x_3)\big)\epsilon_0^2 \\ & + 2dx_2x_3(3 - x_2 - x_3)\epsilon_1 \\ & + x_2x_3\epsilon_1^2 + 2x_2x_3(2 - x_2x_3)\epsilon_0\epsilon_1\Big) \\ & - \frac{\sigma^3}{24N}\Big(d + (x_2 + x_3)\epsilon_0 + x_2x_3(\epsilon_1 - \epsilon_0)\Big) \\ & \times \Big(d^2(x_2(1 - x_2) + x_3(1 - x_3)) + 2dx_2x_3(2 - x_2 - x_3)\epsilon_0 \\ & + x_2x_3(1 - x_2x_3)\epsilon_0^2\Big) \\ & + \frac{\sigma^2\rho}{432N}x_3(1 - x_3)\Big((1 - x_2)\epsilon_0(5d + (3 - x_2)\epsilon_0) \end{aligned}$$

$$+x_2(5d+(6-x_2)\epsilon_0)\epsilon_1+3x_2\epsilon_1^2\Big)$$
$$+\frac{\sigma^2\nu}{432N}\Big(\epsilon_0(1-x_2)(1-x_3)(5d(x_2+x_3)+3(x_2+x_3)\epsilon_0$$
$$+2x_2x_3\epsilon_0)$$
$$+\epsilon_1x_2x_3\big(5d(2-x_2-x_3)+(11-6x_2-6x_3+x_2x_3)\epsilon_0\big)$$
$$+3\epsilon_1^2x_2x_3(1-x_2x_3)\Big)+O(\sigma^4). \quad (54)$$

Finally, if $\epsilon_0=\epsilon_1=0$, the approximation is given by

$$u=\frac{1}{N}+\frac{\sigma d}{2}+\frac{(d\sigma)^2}{12N}$$
$$-\frac{(\sigma d)^3}{24N}\big(2D+x_2(1-x_2)+x_3(1-x_3)\big)$$
$$-\frac{(\sigma d)^4}{720N}\big(1+x_2(1-x_2)(15-14x_2)$$
$$+x_3(1-x_3)(15-14x_3)+2D(29-21(x_2+x_3))\big)$$
$$+\frac{\rho(\sigma d)^3}{432N}(31D+19x_3(1-x_3))$$
$$+\frac{19\nu(\sigma d)^3}{432N}(2D+x_2(1-x_2)+x_3(1-x_3))+O(\sigma^5). \quad (55)$$

In the case $D=0$, the approximation simplifies to

$$u=\frac{1}{N}+\frac{\sigma d}{2}+\frac{(d\sigma)^2}{12N}-\frac{(\sigma d)^3}{24N}\big(x_2(1-x_2)+x_3(1-x_3)\big)$$
$$-\frac{(\sigma d)^4}{720N}\big(1+15(x_2(1-x_2)+x_3(1-x_3))$$
$$-14(x_2x_2(1-x_2)+x_3x_3(1-x_3))\big)$$
$$+\frac{\rho(\sigma d)^3}{432N}(19x_3(1-x_3))$$
$$+\frac{19\nu(\sigma d)^3}{432N}(x_2(1-x_2)+x_3(1-x_3))+O(\sigma^5). \quad (56)$$

Appendix C. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.tpb.2012.05.002.

References

Barton, N.H., 1995. Linkage and the limits to natural selection. Genetics 140, 821–841.

Barton, N.H., 2010. Genetic linkage and natural selection. Philosophical Transactions of the Royal Society B: Biological Sciences 365 (1552), 2559–2569.

Barton, N.H., Otto, S.P., 2005. Evolution of recombination due to random drift. Genetics 169, 2353–2370.

Crow, J.F., Kimura, M., 1965. Evolution in sexual and asexual populations. American Naturalist 99, 439–450.

Crow, J.F., Kimura, M., 1969. Evolution in sexual and asexual populations, a reply. American Naturalist 103, 89–91.

Donnelly, P., Kurtz, T.G., 1999. Genealogical processes for Fleming–Viot models with selection and recombination. The Annals of Applied Probability 9, 1091–1148.

Eshel, I., Feldman, M.W., 1970. On the evolutionary effect of recombination. Theoretical Population Biology 1, 88–100.

Feldman, M.W., Christiansen, F.B., Brooks, L.D., 1980. Evolution of recombination in a constant environment. Proceedings of the National Academy of Sciences USA 77, 4838–4841.

Feldman, M.W., Liberman, U., 1986. An evolutionary reduction principle for genetic modifiers. Proceedings of the National Academy of Sciences USA 83, 4824–4827.

Felsenstein, J., 1965. The effect of linkage on directional selection. Genetics 52, 349–363.

Felsenstein, J., 1974. The evolutionary advantage of recombination. Genetics 78, 737–756.

Fisher, R.A., 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

Griffiths, R.C., 1981. Neutral two-locus multiple allele models with recombination. Theoretical Population Biology 19, 169–186.

Griffiths, R.C., Marjoram, P., 1996. Ancestral inference from samples of dna sequences with recombination. Journal of Computational Biology 3, 479–502.

Hill, W.G., Robertson, A., 1966. The effect of linkage on limits to artificial selection. Genetical Research 8, 269–294.

Hudson, R.R., 1983. Properties of a neutral allele with intragenic recombination. Theoretical Population Biology 23, 183–201.

Karlin, S., Taylor, H.M., 1975. A First Course in Stochastic Processes, Second ed. Academic Press, New York.

Keightley, P.D., Otto, S.P., 2006. Interference among deleterious mutations favours sex and recombination in finite populations. Nature 443, 89–92.

Kingman, J.F.C., 1982. The coalescent. Stochastic Processes and their Applications 13, 235–248.

Krone, S.M., Neuhauser, C., 1997. Ancestral processes with selection. Theoretical Population Biology 51, 210–237.

Lehman, L., Rousset, F., 2009. Perturbation expansions of multilocus fixation probabilities for frequency-dependent selection with applications to the Hill–Robertson effect and to the joint evolution of helping and punishment. Theoretical Population Biology 76, 35–51.

Lessard, S., Kermany, A.R., 2012. Fixation probability in a two-locus model by the ancestral recombination-selection graph. Genetics 190 (2), 691–707.

Lessard, S., Ladret, V., 2007. The probability of fixation of a single mutant in an exchangeable selection model. Journal of Mathematical Biology 54, 721–744.

Lessard, S., Lahaie, P., 2009. Fixation probability with multiple alleles and projected average allelic effect on selection. Theoretical Population Biology 75, 266–277.

Maynard Smith, J., 1968. Evolution in sexual and asexual populations. American Naturalist 102, 469–673.

Möhle, M., Sagitov, S., 2001. A classification of coalescent processes for haploid exchangeable population models. Annals of Probability 29 (4), 1547–1562.

Moran, P.A.P., 1958. Random processes in genetics. Mathematical Proceedings of the Cambridge Philosophical Society 54, 60–71.

Muller, H.J., 1932. Some genetic aspects of sex. American Naturalist 66, 118–138.

Muller, H.J., 1964. The relation of recombination to mutational advance. Mutation Research 1, 2–9.

Ohta, T., 1968. Effect of initial linkage disequilibrium and epistasis on fixation probability in a small population, with two segregating loci. Theoretical and Applied Genetics 38, 243–248.

Ohta, T., Kimura, M., 1969. Linkage disequilibrium due to random genetic drift. Genetical Research 13, 47–55.

Otto, S.P., Barton, N.H., 1997. The evolution of recombination: removing the limits to natural selection. Genetics 147, 879–906.

Otto, S.P., Barton, N.H., 2001. Selection for recombination in small populations. Evolution 55, 1921–1931.

Otto, S.P., Feldman, M.W., 1997. Deleterious mutations, variable epistatic interactions and the evolution of recombination. Theoretical Population Biology 51, 134–147.

Pitman, J., 1999. Coalescents with multiple collisions. Annals of Probability 27, 1870–1902.

Robertson, A., 1962. Selection for heterozygotes in small populations. Genetics 47, 1291–1300.

Rousset, F., 2003. A minimal derivation of convergence stability measures. Journal of Theoretical Biology 221, 665–668.

Sagitov, S., 1999. The general coalescent with asynchronous mergers of ancestral lines. Journal of Applied Probability 36, 1116–1125.

Zhivotovsky, L., Feldman, M., Christiansen, F., 1994. Evolution of recombination among multiple selected loci, a generalized reduction principle. Proceedings of the National Academy of Sciences USA 91 (3), 1079–1083.