



HAL
open science

Fixation probability in a two-locus intersexual selection model

Guillermo Durand, Sabin Lessard

► **To cite this version:**

Guillermo Durand, Sabin Lessard. Fixation probability in a two-locus intersexual selection model. *Theoretical Population Biology*, 2016, 109, pp.75-87. <10.1016/j.tpb.2016.03.004>. <hal-01361671>

HAL Id: hal-01361671

<https://hal.science/hal-01361671v1>

Submitted on 7 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Fixation probability in a two-locus intersexual selection model

Guillermo Durand^a and Sabin Lessard^b

^aLaboratoire de Mathématiques d'Orsay
Université Paris-Saclay
Bâtiment 307, rue Michel Magat
F-91405 Orsay Cedex France

^bDépartement de mathématiques et de statistique
Université de Montréal
C.P. 6128, succursale Centre-ville
Montréal (Québec) H3C 3J7 Canada

2018-07-03

Running head: Intersexual selection

Keywords and phrases: Ancestral recombination-selection graph; Fixation probability; Sexual selection.

Corresponding author:

Sabin Lessard

Département de mathématiques et de statistique

Université de Montréal

C.P. 6128, succursale Centre-ville

Montréal (Québec) H3C 3J7 Canada

email: lessards@dms.umontreal.ca

Abstract

We study a two-locus model of intersexual selection in a finite haploid population reproducing according to a discrete-time Moran model with a trait locus expressed in males and a preference locus expressed in females. We show that the probability of ultimate fixation of a single mutant allele for a male ornament introduced at random at the trait locus given any initial frequency state at the preference locus is increased by weak intersexual selection and recombination, weak or strong. Moreover, this probability exceeds the initial frequency of the mutant allele even in the case of a costly male ornament if intersexual selection is not too weak. On the other hand, the probability of ultimate fixation of a single mutant allele for a female preference towards a male ornament introduced at random at the preference locus is increased by weak intersexual selection and weak recombination if the female preference is not costly, and is strong enough in the case of a costly male ornament. The analysis relies on an extension of the ancestral recombination-selection graph for samples of haplotypes to take into account events of intersexual selection, while the symbolic calculation of the fixation probabilities is made possible in a reasonable time by an optimizing algorithm.

1 Introduction

Darwin (1871) considered sexual selection to be an evolutionary process as important as natural selection. The concept arises from the observation that some weapons or ornaments known as secondary sexual traits that have evolved in the animal kingdom, as antlers in deer or tails and colors in birds, appear not to be advantageous for survival or even to be deleterious. A possible explanation is that these traits enhance the reproductive success of their carriers either by deterring rivals of the same sex, usually males (intrasexual selection), or by making them more attractive to individuals of the opposite sex, usually females (intersexual selection). These two forms of sexual selection may be responsible for a large amount of gender differences in structure or appearance in nature.

This paper concerns intersexual selection. As for intrasexual selection, let us just mention that it has been studied, e.g., by applying game theory to evolution in the context of conflicts between males for access to females and reproduction with the key concept of evolutionarily stable strategy (Maynard Smith, 1982).

Fisher (1930, 1958) described the runaway process by which a preferred conspicuous trait in one sex and a preference for this trait in the other sex could evolve in tandem with the preference being genetically transmitted along with the preferred trait. Taking a plumage character as an example, Fisher (1930, p.137) writes: *Moreover, as long as there is a net advantage in favour of further plumage development, there will also be a net advantage in favour of giving to it a more decided preference.* This runaway mechanism could reinforce or accelerate phenotypic difference and speciation.

With the above notable exception, however, sexual selection involving mate choice in one sex, most often the female but not always (see, e.g., Edward and Chapman, 2011), was largely disregarded by most of the leading evolutionary biologists up to the mid 20th century (see, e.g., Maynard Smith, 2000).

O'Donald (1962, 1980) was one of the first to propose, and study numerically, two-locus models of sexual selection with one locus coding for trait variation in males and one locus influencing mating preferences in females. Lande (1981) analysed models in the case of polygenic inheritance and Kirkpatrick (1982) exact haploid population genetic models. They both showed that, in the absence of direct selection on female preferences, these can evolve as a correlated response to changes in the male trait associated with positive linkage disequilibrium. Moreover, in the framework of an infinite population, they exhibited curves of stable equilibria at which neutral preferences for less viable traits can counterbalance the deleterious effects of the traits. The evolutionary outcome along the line of stable equilibria is left to weaker forces such as genetic drift. The preference function (see, e.g., Charlesworth and Charlesworth, 1981, Seger, 1985, Carrier, 1995) as well as the recombination rate (see, e.g., Kirkpatrick, 1982, Barton and Turelli, 1991) seem to have little qualitative effects on the conclusions.

Similar conclusions for diploid populations with some discrepancies due to dominance at the trait locus have been reached from simulations (Heisler and Curtsinger, 1990), local stability analyses near fixation either at the trait locus

(Gomulkiewicz and Hastings, 1990) or the preference locus (Otto, 1991), as well as quasi-linkage equilibrium analyses away from fixation boundaries (Green-spoon and Otto, 2009).

With selection on female preferences, the line of stable equilibria may collapse to a single, stable point and lead to a balance between selection and drift in finite populations under recurrent mutation. This may be the case for instance with a preference-dependent risk that females remain unmated as supported by simulations of polygenic models (De Jong and Sabelis, 1991). More simulations have shown that adding a cost to female preferences may restrict, but not eliminate, the possibility of ornament diversification and speciation (Mead and Arnold, 2004, Uyeda *et al.*, 2009, Bergen *et al.*, 2012). One of the surprising consequences of sexual selection may be to create a positive correlation between the female preference and the progeny sex ratio (Fawcett *et al.*, 2007).

Note also that there may be a benefit associated with female preferences which is to enhance species recognition and, therefore, fertility (see, e.g., Carrier, 1995). Indeed, sexual preference of females for given traits in males is akin to assortative mating and, combined with drift, can help answer questions related to species divergence (Otto *et al.* 2008, Servedio, 2011).

Recently, Miller (2000) drew attention to some of Darwin's neglected ideas about human behaviours not clearly connected to survival, such as humour, creativity, and some forms of altruism, that may have been favoured through sexual selection. See Puts (2010) for more references on this subject.

As for a recent review of mathematical models of sexual selection, we refer to Kuijper *et al.* (2012).

In this paper, we consider a two-locus model of intersexual selection in a finite haploid population with a trait locus expressed in males and a preference locus expressed in females. We use a discrete-time Moran model so that, at each time step, there is one individual produced according to the types of the parents chosen to reproduce and one individual replaced according to the type of the individual chosen to be replaced. We study the probability of fixation of a single mutant introduced at the trait locus given any initial frequency state at the preference locus, and the other way around. We consider the case of a preference for a deleterious trait and the case of a preference for a beneficial trait, but describe the method at length only in the former. We deduce the leading effects of weak selection and weak recombination using an ancestral recombination-selection graph in the limit of a large population size. This extends a similar approach for a two-locus viability selection model to study the Hill-Robertson effect in favor of recombination (Lessard and Kermany, 2012). The leading effects of weak selection under the assumption of free recombination, actually of any fixed recombination rate as the population size goes to infinity so that recombination and selection events occur at different timescales backwards in time, are also studied. We provide an algorithm which strikingly reduces the time of symbolic calculation.

2 Model

Suppose a large but finite population of N haploid individuals. The population is assumed to be monoecious (hermaphroditic) so that each individual can act as either the male or the female in sexual reproduction. In particular, an individual can reproduce with itself. Consider two genes at two different loci, each one with two possible alleles. The first gene at a trait locus, denoted by T , has an effect on viability. It is assumed that a mutant allele T_2 decreases the viability of a male carrying it compared to a resident allele T_1 .

On the other hand, the second gene at the other locus, denoted by P , codes for sexual preference. It is assumed that a female carrying a mutant allele P_2 shows a preference for males carrying the mutant allele T_2 , while a female carrying a resident allele P_1 does not show any preference. Therefore, the individuals can be of four possible haplotypes, called types for simplicity: (T_1P_1) , (T_1P_2) , (T_2P_1) and (T_2P_2) , or in vector notation $(1, 1)$, $(1, 2)$, $(2, 1)$ and $(2, 2)$, respectively. These are represented by **1**, **2**, **3** and **4**, respectively.

We use a discrete-time Moran model (Moran, 1958). At each time step, two individuals are sampled at random with replacement to mate and to produce an offspring (hence the possibility of reproduction of an individual with itself). The first one is assumed to act as the female and the second one as the male. The sampled individuals reproduce with some probability depending on their types. Actually, a P_2 -female reproduces with a T_2 -male with probability 1 and with a T_1 -male with probability $1 - As$. If the female carries allele P_1 , however, reproduction takes place with probability $1 - \frac{As}{2}$ whatever the type of the male is. Therefore, there is a reproduction cost associated not only with a female being choosy that depends on the type of the male but also with a female not being choosy irrespective of this type.

Here, $A > 0$ represents a coefficient of intersexual selection with respect to an intensity of selection $s > 0$. Weak selection in a large population is modeled by assuming

$$s = \frac{\sigma}{N},$$

where $\sigma > 0$ represents a population-scaled intensity.

If reproduction takes place, then the type of the offspring produced depends on the parental types and the recombination rate r . With probability $1 - r$ the type of the offspring is one of the two parental haplotypes chosen at random (probability $\frac{1}{2}$ for each), while it is one of the two recombinant haplotypes chosen at random (probability $\frac{1}{2}$ for each) with the complementary probability r . Weak recombination in a large population is modeled by assuming

$$r = \frac{R\sigma}{N}, \tag{1}$$

for some coefficient of recombination $R > 0$. Note that weak recombination is scaled at the same order of magnitude as weak selection.

If the two individuals chosen to reproduce does not actually reproduce, the population state does not change. If they produce an offspring, then an individual is sampled at random in the population to be replaced by the offspring. The individual is actually replaced with some probability depending on its haplotype. Replacement occurs with probability

$$1 - cs = 1 - \frac{c\sigma}{N}, \quad (2)$$

where $c \geq 0$ is a coefficient of viability selection that depends on the type of the individual to be replaced. In this study we set $c_1 = c_2 = 1$ and $c_3 = c_4 = 0$. This models a selective advantage in favor of allele T_1 . If the individual chosen to be replaced is not actually replaced, then the offspring does not survive so that the population state does not change.

Let $x_i(\tau)$ be the frequency of type \mathbf{i} at time step $\tau \geq 0$ and define the type frequency vector $\mathbf{x}(\tau) = (x_1(\tau), x_2(\tau), x_3(\tau), x_4(\tau))$. Initially (time step 0), let allele T_2 be introduced as a mutant at the first locus in a single individual of the population (hence with an initial frequency N^{-1}), where alleles P_2 and P_1 are currently segregating at the second locus with frequencies x and $1 - x$, respectively. Therefore, the initial frequencies of the haplotypes (T_1P_1) , (T_1P_2) , (T_2P_1) and (T_2P_2) are given respectively either by

$$x_1(0) = 1 - x, \quad x_2(0) = x - \frac{1}{N}, \quad x_3(0) = 0, \quad x_4(0) = \frac{1}{N}, \quad (3)$$

or by

$$x_1(0) = 1 - x - \frac{1}{N}, \quad x_2(0) = x, \quad x_3(0) = \frac{1}{N}, \quad x_4(0) = 0. \quad (4)$$

If the mutant allele T_2 arises in an individual chosen at random, then the former event has probability x and the latter event probability $1 - x$. We focus in the following sections on the probability of ultimate fixation of T_2 under these assumptions.

Note that the probability of ultimate fixation of P_2 introduced as a single mutant at the second locus while T_2 and T_1 are already segregating at the first locus with given frequencies can be studied in a similar way. The assumption that the preferred trait T_2 is beneficial instead of being deleterious can also be analogously dealt with. Only the results are stated in these cases in Section 6.

Finally, the analysis under the assumption of free recombination which corresponds to the condition $r = 1/2$, and actually of any fixed recombination rate r in the limit of a large population size which corresponds to strong recombination, is slightly different since it involves two timescales, a fast one for recombination and a slow one for selection. Some details and results for this case are given in Appendix E and Section 6.

3 Fixation probability

From time step $\tau \geq 0$ to time step $\tau + 1$, the frequency $x_i(\tau)$ increases by N^{-1} if an offspring of type \mathbf{i} is produced and an individual of type different from \mathbf{i} is replaced by the offspring. For $\mathbf{i} = \mathbf{4}$, for instance, which corresponds to haplotype (T₂P₂), we list below the probabilities for a female of type \mathbf{j} and a male of type \mathbf{k} to mate and to produce an offspring, and then for this offspring to be of type $\mathbf{4}$, respectively.

$x_4(\tau)x_4(\tau)$	1	1
$x_3(\tau)x_4(\tau)$	$1 - \frac{A\sigma}{2N}$	$\frac{1}{2}$
$x_4(\tau)x_3(\tau)$	1	$\frac{1}{2}$
$x_2(\tau)x_4(\tau)$	1	$\frac{1}{2}$
$x_4(\tau)x_2(\tau)$	$1 - \frac{A\sigma}{N}$	$\frac{1}{2}$
$x_2(\tau)x_3(\tau)$	1	$\frac{R\sigma}{2N}$
$x_3(\tau)x_2(\tau)$	$1 - \frac{A\sigma}{2N}$	$\frac{R\sigma}{2N}$
$x_1(\tau)x_4(\tau)$	$1 - \frac{A\sigma}{2N}$	$\frac{1}{2} \left(1 - \frac{R\sigma}{N}\right)$
$x_4(\tau)x_1(\tau)$	$1 - \frac{A\sigma}{N}$	$\frac{1}{2} \left(1 - \frac{R\sigma}{N}\right)$

Moreover, an individual of type \mathbf{j} is chosen to be replaced with probability $x_j(\tau)$ and then the replacement occurs with probability $(1 - c_j\sigma N^{-1})$ for $\mathbf{j} = \mathbf{1}, \mathbf{2}, \mathbf{3}$. Therefore, the conditional probability that $x_4(\tau)$ increases by N^{-1} is

$$\begin{aligned}
 & \mathbb{P} \left(x_4(\tau + 1) = x_4(\tau) + \frac{1}{N} \mid \mathbf{x}(\tau) \right) \\
 &= \left(\left(1 - \frac{R\sigma}{N}\right) \left(1 - \frac{3A\sigma}{4N}\right) x_1(\tau)x_4(\tau) + \frac{R\sigma}{N} \left(1 - \frac{A\sigma}{4N}\right) x_2(\tau)x_3(\tau) \right. \\
 &+ \left. \left(1 - \frac{A\sigma}{2N}\right) x_2(\tau)x_4(\tau) + \left(1 - \frac{A\sigma}{4N}\right) x_3(\tau)x_4(\tau) + x_4(\tau)^2 \right) \\
 &\times \left(\left(1 - \frac{c_1\sigma}{N}\right) x_1(\tau) + \left(1 - \frac{c_2\sigma}{N}\right) x_2(\tau) + \left(1 - \frac{c_3\sigma}{N}\right) x_3(\tau) \right). \tag{5}
 \end{aligned}$$

Likewise, the conditional probability that $x_4(\tau)$ decreases by N^{-1} is

$$\begin{aligned}
& \mathbb{P}\left(x_4(\tau+1) = x_4(\tau) - \frac{1}{N} \mid \mathbf{x}(\tau)\right) \\
&= \left(\left(1 - \frac{A\sigma}{2N}\right) x_1(\tau)^2 + 2 \left(1 - \frac{3A\sigma}{4N}\right) x_1(\tau)x_2(\tau) + 2 \left(1 - \frac{A\sigma}{2N}\right) x_1(\tau)x_3(\tau) \right. \\
&+ \left(1 + \frac{R\sigma}{N}\right) \left(1 - \frac{3A\sigma}{4N}\right) x_1(\tau)x_4(\tau) + \left(1 - \frac{A\sigma}{N}\right) x_2(\tau)^2 \\
&+ \left(2 - \frac{R\sigma}{N}\right) \left(1 - \frac{A\sigma}{4N}\right) x_2(\tau)x_3(\tau) + \left(1 - \frac{A\sigma}{2N}\right) x_2(\tau)x_4(\tau) \\
&+ \left(1 - \frac{A\sigma}{2N}\right) x_3(\tau)^2 + \left(1 - \frac{A\sigma}{4N}\right) x_3(\tau)x_4(\tau) \Big) \\
&\times \left(1 - \frac{c_4\sigma}{N}\right) x_4(\tau). \tag{6}
\end{aligned}$$

For $\mathbf{i} = \mathbf{3}$ the corresponding conditional probabilities are

$$\begin{aligned}
& \mathbb{P}\left(x_3(\tau+1) = x_3(\tau) + \frac{1}{N} \mid \mathbf{x}(\tau)\right) \\
&= \left(\left(1 - \frac{A\sigma}{2N}\right) x_1(\tau)x_3(\tau) + \frac{R\sigma}{N} \left(1 - \frac{3A\sigma}{4N}\right) x_1(\tau)x_4(\tau) \right. \\
&+ \left(1 - \frac{R\sigma}{N}\right) \left(1 - \frac{A\sigma}{4N}\right) x_2(\tau)x_3(\tau) \\
&+ \left(1 - \frac{A\sigma}{2N}\right) x_3(\tau)^2 + \left(1 - \frac{A\sigma}{4N}\right) x_3(\tau)x_4(\tau) \Big) \\
&\times \left(\left(1 - \frac{c_1\sigma}{N}\right) x_1(\tau) + \left(1 - \frac{c_2\sigma}{N}\right) x_2(\tau) + \left(1 - \frac{c_4\sigma}{N}\right) x_4(\tau) \right) \tag{7}
\end{aligned}$$

and

$$\begin{aligned}
& \mathbb{P}\left(x_3(\tau+1) = x_3(\tau) - \frac{1}{N} \mid \mathbf{x}(\tau)\right) \\
&= \left(\left(1 - \frac{A\sigma}{2N}\right) x_1(\tau)^2 + 2 \left(1 - \frac{3A\sigma}{4N}\right) x_1(\tau)x_2(\tau) + \left(1 - \frac{A\sigma}{2N}\right) x_1(\tau)x_3(\tau) \right. \\
&+ \left(2 - \frac{R\sigma}{N}\right) \left(1 - \frac{3A\sigma}{4N}\right) x_1(\tau)x_4(\tau) + \left(1 - \frac{A\sigma}{N}\right) x_2(\tau)^2 \\
&+ \left(1 + \frac{R\sigma}{N}\right) \left(1 - \frac{A\sigma}{4N}\right) x_2(\tau)x_3(\tau) + 2 \left(1 - \frac{A\sigma}{2N}\right) x_2(\tau)x_4(\tau) \\
&+ \left(1 - \frac{A\sigma}{4N}\right) x_3(\tau)x_4(\tau) + x_4(\tau)^2 \Big) \\
&\times \left(1 - \frac{c_3\sigma}{N}\right) x_3(\tau). \tag{8}
\end{aligned}$$

The frequency of allele T_2 represented by $x_{T_2}(\tau) = x_{\mathbf{3}}(\tau) + x_{\mathbf{4}}(\tau)$ has a conditional expected change given by

$$\mathbb{E}(\Delta x_{T_2}(\tau) \mid \mathbf{x}(\tau)) = \mathbb{E}(\Delta x_{\mathbf{3}}(\tau) \mid \mathbf{x}(\tau)) + \mathbb{E}(\Delta x_{\mathbf{4}}(\tau) \mid \mathbf{x}(\tau)), \quad (9)$$

where

$$\begin{aligned} \mathbb{E}(\Delta x_{\mathbf{i}}(\tau) \mid \mathbf{x}(\tau)) &= \frac{1}{N} \mathbb{P} \left(x_{\mathbf{i}}(\tau+1) = x_{\mathbf{i}}(\tau) + \frac{1}{N} \mid \mathbf{x}(\tau) \right) \\ &\quad - \frac{1}{N} \mathbb{P} \left(x_{\mathbf{i}}(\tau+1) = x_{\mathbf{i}}(\tau) - \frac{1}{N} \mid \mathbf{x}(\tau) \right) \end{aligned} \quad (10)$$

is the conditional expected value of the change $\Delta x_{\mathbf{i}}(\tau) = x_{\mathbf{i}}(\tau+1) - x_{\mathbf{i}}(\tau)$ in the frequency of type $\mathbf{i} = \mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}$. This leads to

$$\mathbb{E}(\Delta x_{T_2}(\tau) \mid \mathbf{x}(\tau)) = \frac{2}{N^2} \sum_{\mathbf{i}, \mathbf{j}, \mathbf{k}} Q_{\mathbf{i}, \mathbf{j}, \mathbf{k}} x_{\mathbf{i}}(\tau) x_{\mathbf{j}}(\tau) x_{\mathbf{k}}(\tau). \quad (11)$$

Here, the sum is over all triplets of types $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ with $\mathbf{i} \leq \mathbf{j} \leq \mathbf{k}$. Moreover, ignoring terms of order N^{-1} , the coefficients $Q_{\mathbf{i}, \mathbf{j}, \mathbf{k}}$ are given by

$$\begin{aligned} Q_{\mathbf{1}, \mathbf{1}, \mathbf{1}} &= 0 & Q_{\mathbf{1}, \mathbf{1}, \mathbf{2}} &= 0 \\ Q_{\mathbf{1}, \mathbf{1}, \mathbf{3}} &= \frac{c_3 - c_1}{2} \sigma & Q_{\mathbf{1}, \mathbf{1}, \mathbf{4}} &= -\frac{A + 4c_1 - 4c_4}{8} \sigma \\ Q_{\mathbf{1}, \mathbf{2}, \mathbf{2}} &= 0 & Q_{\mathbf{1}, \mathbf{2}, \mathbf{3}} &= \frac{3A - 4c_1 - 4c_2 + 8c_3}{8} \sigma \\ Q_{\mathbf{1}, \mathbf{2}, \mathbf{4}} &= \frac{A - 4c_1 - 4c_2 + 8c_4}{8} \sigma & Q_{\mathbf{1}, \mathbf{3}, \mathbf{3}} &= \frac{c_3 - c_1}{2} \sigma \\ Q_{\mathbf{1}, \mathbf{3}, \mathbf{4}} &= \frac{3A - 8c_1 + 4c_3 + 4c_4}{8} \sigma & Q_{\mathbf{1}, \mathbf{4}, \mathbf{4}} &= \frac{3A - 4c_1 + 4c_4}{8} \sigma \\ Q_{\mathbf{2}, \mathbf{2}, \mathbf{2}} &= 0 & Q_{\mathbf{2}, \mathbf{2}, \mathbf{3}} &= \frac{3A - 4c_2 + 4c_3}{8} \sigma \\ Q_{\mathbf{2}, \mathbf{2}, \mathbf{4}} &= \frac{A - 2c_2 + 2c_4}{4} \sigma & Q_{\mathbf{2}, \mathbf{3}, \mathbf{3}} &= -\frac{A + 4c_2 - 4c_3}{8} \sigma \\ Q_{\mathbf{2}, \mathbf{3}, \mathbf{4}} &= \frac{A - 8c_2 + 4c_3 + 4c_4}{8} \sigma & Q_{\mathbf{2}, \mathbf{4}, \mathbf{4}} &= \frac{A - 2c_2 + 2c_4}{4} \sigma \\ Q_{\mathbf{3}, \mathbf{3}, \mathbf{3}} &= 0 & Q_{\mathbf{3}, \mathbf{3}, \mathbf{4}} &= 0 \\ Q_{\mathbf{3}, \mathbf{4}, \mathbf{4}} &= 0 & Q_{\mathbf{4}, \mathbf{4}, \mathbf{4}} &= 0 \end{aligned} \quad (12)$$

Note that these leading terms depend only on the selection parameters.

The Markov chain $(\mathbf{x}(\tau))_{\tau \geq 0}$ has four absorbing states, namely $\mathbf{e}_1 = (1, 0, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0, 0)$, $\mathbf{e}_3 = (0, 0, 1, 0)$ and $\mathbf{e}_4 = (0, 0, 0, 1)$, corresponding to fixation of types (T_1, P_1) , (T_1, P_2) , (T_2, P_1) and (T_2, P_2) , respectively. All other states which are in the form $\mathbf{y} = (y_1, y_2, y_3, y_4)$ with y_1, y_2, y_3, y_4 being multiples of N^{-1} that sum up to 1, are transient. An ergodic theorem (see, e.g., Karlin and Taylor, 1975) states that $(\mathbf{x}(\tau))_{\tau \geq 0}$ converges to the four absorbing states. Actually, the probability of transition from state \mathbf{x} at time step 0 to state \mathbf{y} at time step τ satisfies

$$P_{\mathbf{xy}}(\tau) = \mathbb{P}(\mathbf{x}(\tau) = \mathbf{y} \mid \mathbf{x}(0) = \mathbf{x}) \rightarrow P_{\mathbf{xy}}(\infty), \quad (13)$$

as $\tau \rightarrow \infty$, with $P_{\mathbf{xy}}(\infty) \neq 0$ only if \mathbf{y} is absorbing. Here, $P_{\mathbf{xy}}(\infty)$ represents the conditional probability of ultimate fixation in state \mathbf{y} , given an initial state \mathbf{x} . For the frequency of allele T_2 , this leads to

$$\begin{aligned} \mathbb{E}(x_{T_2}(\tau) \mid \mathbf{x}(0) = \mathbf{x}) &= \mathbb{E}(x_3(\tau) + x_4(\tau) \mid \mathbf{x}(0) = \mathbf{x}) \\ &= \sum_{\mathbf{y}=(y_1, y_2, y_3, y_4)} (y_3 + y_4) P_{\mathbf{xy}}(\tau) \\ &\rightarrow P_{\mathbf{x}e_3}(\infty) + P_{\mathbf{x}e_4}(\infty) = u_{T_2}(\mathbf{x}), \end{aligned} \quad (14)$$

as $\tau \rightarrow \infty$. Here, the limit $u_{T_2}(\mathbf{x})$ represents the conditional probability of ultimate fixation of allele T_2 , given an initial population state \mathbf{x} .

On the other hand, the frequency of allele T_2 at time step τ can be expressed as

$$x_{T_2}(\tau) = x_{T_2}(0) + \sum_{k=0}^{\tau-1} \Delta x_{T_2}(k). \quad (15)$$

Its conditional expected value given an initial population state $\mathbf{x}(0) = \mathbf{x}$ with

$$x_{T_2}(0) = x_3(0) + x_4(0) = \frac{1}{N} \quad (16)$$

takes the form

$$\mathbb{E}(x_{T_2}(\tau) \mid \mathbf{x}(0) = \mathbf{x}) = \frac{1}{N} + \sum_{k=0}^{\tau-1} \mathbb{E}(\Delta x_{T_2}(k) \mid \mathbf{x}(0) = \mathbf{x}). \quad (17)$$

Note that

$$\mathbb{E}(\Delta x_{T_2}(k) \mid \mathbf{x}(0) = \mathbf{x}) = \mathbb{E}(\mathbb{E}(\Delta x_{T_2}(k) \mid \mathbf{x}(k)) \mid \mathbf{x}(0) = \mathbf{x}). \quad (18)$$

Letting $\tau \rightarrow \infty$ yields

$$\begin{aligned} u_{T_2}(\mathbf{x}) &= \lim_{\tau \rightarrow \infty} \mathbb{E}(x_{T_2}(\tau) \mid \mathbf{x}(0) = \mathbf{x}) \\ &= \frac{1}{N} + \sum_{\tau=0}^{\infty} \mathbb{E}(\Delta x_{T_2}(\tau) \mid \mathbf{x}(0) = \mathbf{x}) \\ &= \frac{1}{N} + \sum_{\tau=0}^{\infty} \mathbb{E}(\mathbb{E}(\Delta x_{T_2}(\tau) \mid \mathbf{x}(\tau)) \mid \mathbf{x}(0) = \mathbf{x}) \\ &= \frac{1}{N} + \sum_{\mathbf{i}, \mathbf{j}, \mathbf{k}} Q_{\mathbf{i}, \mathbf{j}, \mathbf{k}} E_{\mathbf{i}, \mathbf{j}, \mathbf{k}}(\mathbf{x}), \end{aligned} \quad (19)$$

where

$$E_{\mathbf{i},\mathbf{j},\mathbf{k}}(\mathbf{x}) = \frac{2}{N^2} \sum_{\tau=0}^{\infty} \mathbb{E}(x_{\mathbf{i}}(\tau)x_{\mathbf{j}}(\tau)x_{\mathbf{k}}(\tau)|\mathbf{x}(0) = \mathbf{x}). \quad (20)$$

It remains to compute this quantity in order to obtain the conditional probability of ultimate fixation of allele T_2 . This will be done with the help of the ancestral recombination-selection graph that will provide a development of this quantity up to a given order with respect to the population-scaled intensity of selection and recombination.

4 Ancestral recombination-selection graph

4.1 General description

The ancestral recombination-selection graph is a stochastic process which combines the ancestral recombination graph (Griffiths, 1981, Hudson, 1983) and the ancestral selection graph (Krone and Neuhauser, 1997). It is a process that traces backwards in time the ancestral material of an ordered random sample of individuals at two or more loci under recombination in a population under selection (Donnelly and Kurtz, 1999, Fearnhead, 2003). It was used in Lessard and Kermany (2012) and Kermany and Lessard (2012) to compute the probability of fixation of new mutants in relation with the Hill-Robertson effect in favor of the evolution of recombination. Its topology is represented by events of different types: coalescence, recombination, selection. In this paper, we introduce intersexual selection besides viability selection. When only one event occurs at the same time step, it will be called pure. Simultaneous events can occur but their probability will be negligible in the limit of a large population size.

At each time step, there is no intersexual selection with probability $1 - As$ in which case, whatever the types of the individuals chosen to reproduce are, they produce an offspring with probability 1. In the case of intersexual selection (which occurs with probability As), reproduction is parent-dependent: the probability that an offspring is produced, denoted by $f_{\mathbf{i},\mathbf{j}}$, depends on the types \mathbf{i} and \mathbf{j} of the female and male, respectively, chosen to reproduce. In the situation where $\mathbf{i} = \mathbf{1}$ or $\mathbf{3}$, which means that the female carries allele P_1 , we have $f_{\mathbf{i},\mathbf{j}} = \frac{1}{2}$ whatever the type \mathbf{j} of the male is. Otherwise, we have

$$\begin{aligned} f_{\mathbf{2},\mathbf{1}} &= f_{\mathbf{2},\mathbf{2}} = f_{\mathbf{4},\mathbf{1}} = f_{\mathbf{4},\mathbf{2}} = 0, \\ f_{\mathbf{2},\mathbf{3}} &= f_{\mathbf{2},\mathbf{4}} = f_{\mathbf{4},\mathbf{3}} = f_{\mathbf{4},\mathbf{4}} = 1. \end{aligned} \quad (21)$$

This gives

$$q_1 = 1 - As \quad (22)$$

and

$$q_2 = (1 - As) + As = 1, \quad (23)$$

for the probabilities of a P₂-female to produce an offspring with a male carrying alleles T₁ and T₂, respectively. This has to be compared to

$$q_0 = (1 - As) + \frac{As}{2} = 1 - \frac{As}{2}, \quad (24)$$

for the probability of a P₁-female to produce an offspring whatever the type of the male is.

Similarly, an individual chosen to be replaced by an offspring is replaced whatever the type of the individual is with probability $1 - s$. On the other hand, there is viability selection in the form of type-dependent replacement with probability s . Actually in this case, we assume that replacement occurs with probability $1 - c_i$ if the type of the individual is \mathbf{i} . We set $c_1 = c_2 = 1$ et $c_3 = c_4 = 0$, so that a T₁-individual is replaced with probability $1 - s$ compared to

$$(1 - s) + s = 1, \quad (25)$$

for a T₂-individual.

In the following, we consider an ordered sample of n individuals numbered from 1 to n , taken at random without replacement in a population of size N at a given time step. The $N - n$ other individuals in the population are numbered from $n + 1$ to N . The individuals of the population can be assumed to occupy N distinct sites. The ordered sample is represented by a vector $\mathbf{z} = (z_1, \dots, z_n)$ where z_i is the type of individual $i = 1, \dots, n$. The corresponding vector for the ancestors (real or virtual, see below) of the sampled individuals one time step back is represented by \mathbf{z}' . All possible events from the latter to the former backwards in time are considered.

4.2 Pure coalescence events

There is pure coalescence of two individuals i and j for $1 \leq i < j \leq n$ if, one time step back, individual j is chosen at random (probability N^{-1}) to produce a non-recombinant offspring (probability $1 - R\sigma N^{-1}$) irrespective of the types of the parents (probability $1 - A\sigma N^{-1}$) and this offspring replaces the individual at site i (probability N^{-1}) irrespective of its type (probability $1 - \sigma N^{-1}$), or vice versa (factor 2). The probability of the whole event is

$$\frac{2}{N} \left(1 - \frac{A\sigma}{N}\right) \left(1 - \frac{R\sigma}{N}\right) \frac{1}{N} \left(1 - \frac{\sigma}{N}\right) = \frac{2}{N^2} \left(1 + O\left(\frac{1}{N}\right)\right), \quad (26)$$

in which case the sample size is reduced by 1. Then we have

$$z'_k = \begin{cases} z_k & \text{if } k < j, \\ z_{k+1} & \text{if } k \geq j, \end{cases} \quad (27)$$

and we write $\mathbf{z}' = C_{i,j}(\mathbf{z})$. Note that $z_i = z_j$.

4.3 Pure recombination events

There is pure recombination of individual i for $1 \leq i \leq n$ if, one time step back, two distinct individuals j and l with $N - n \leq j < l \leq N$ are chosen at random to mate and they produce parent-independently (that is, irrespective of their types) a recombinant offspring which replaces type-independently (that is, irrespective of the type of the individual to be replaced) the individual at site i . The probability of the whole event is

$$\frac{N-n}{N} \frac{N-n-1}{N} \left(1 - \frac{A\sigma}{N}\right) \frac{R\sigma}{N} \frac{1}{N} \left(1 - \frac{\sigma}{N}\right) = \frac{R\sigma}{N^2} \left(1 + O\left(\frac{1}{N}\right)\right), \quad (28)$$

in which case the sample size is increased by 1. Without loss of generality, it is assumed that the offspring carries the trait allele of individual j and the preference allele of individual l . Letting y_j and y_l be the respective types of individuals j and l , we have

$$z'_k = \begin{cases} z_k & \text{if } k < i, \\ y_j & \text{if } k = i, \\ y_l & \text{if } k = i + 1, \\ z_{k-1} & \text{if } k \geq i + 2, \end{cases} \quad (29)$$

and we write $\mathbf{z}' = R_i(\mathbf{z})$. Note that

$$z_i = (z_i(1), z_i(2)) = (z'_i(1), z'_{i+1}(2)), \quad (30)$$

where $z_i(1)$ and $z_i(2)$ refer to the trait and preference alleles, respectively, carried by individual i .

4.4 Pure viability selection events

There is pure viability selection of individual i for $1 \leq i \leq n$ if, one time step back, individual j with $N - n \leq j \leq N$ is chosen at random to produce parent-independently an exact copy of itself and the individual at site i is chosen to be replaced type-dependently by the offspring. If the type of the individual that was at site i is y_i , then the ancestral line of the offspring produced by individual j , called the incoming line, is real with probability $1 - c_{y_i}$, while it is the ancestral line of the individual that was at site i , called the continuous line, that is real with the complementary probability c_{y_i} . However, without knowing the type of the individual that was at site i , we do not know which

line is real and which line is not real, called virtual. Therefore, we keep track of both ancestral lines. Moreover, the probability of the whole event is

$$\frac{N-n}{N} \left(1 - \frac{A\sigma}{N}\right) \left(1 - \frac{R\sigma}{N}\right) \frac{1}{N} \frac{\sigma}{N} = \frac{\sigma}{N^2} \left(1 + O\left(\frac{1}{N}\right)\right), \quad (31)$$

in which case the sample size is increased by 1. Denoting the type of individual j that produced the offspring by y_j , we have

$$z'_k = \begin{cases} z_k & \text{if } k < i, \\ y_i & \text{if } k = i, \\ y_j & \text{if } k = i + 1, \\ z_{k-1} & \text{if } k \geq i + 2, \end{cases} \quad (32)$$

and we write $\mathbf{z}' = V_i(\mathbf{z})$. Note that

$$z_i = \begin{cases} z'_i & \text{with probability } c_{z'_i}, \\ z'_{i+1} & \text{with probability } 1 - c_{z'_i}. \end{cases} \quad (33)$$

4.5 Pure intersexual selection events

Finally, there is pure intersexual selection of individual at site i for $1 \leq i \leq n$ if, one time step back, a female j and a male l with $N-n \leq j, l \leq N$ are chosen to produce type-dependently a non-recombinant offspring and the individual at site i is chosen to be replaced type-independently by the offspring. The probability of this event is

$$\frac{N-n}{N} \frac{N-n}{N} \frac{A\sigma}{N} \left(1 - \frac{R\sigma}{N}\right) \frac{1}{N} \left(1 - \frac{\sigma}{N}\right) = \frac{A\sigma}{N^2} \left(1 + O\left(\frac{1}{N}\right)\right). \quad (34)$$

In the limit of a large population size, the possibility that the male and the female are the same individual can be neglected, in which case the sample size is increased by 2.

If y_j and y_l represent the respective types of the female j and the male l , then we have

$$z'_k = \begin{cases} z_k & \text{if } k \leq i, \\ y_j & \text{if } k = i + 1, \\ y_l & \text{if } k = i + 2, \\ z_{k-2} & \text{if } k \geq i + 3, \end{cases} \quad (35)$$

and we write $\mathbf{z}' = S_i(\mathbf{z})$. Note that

$$z_i = \begin{cases} z'_i & \text{with probability } 1 \text{ if } z'_{i+1} = \mathbf{2}, \mathbf{4} \text{ and } z'_{i+2} = \mathbf{1}, \mathbf{2}, \\ z'_{i+1} & \text{with probability } \frac{1}{2} \text{ if } z'_{i+1} = \mathbf{2}, \mathbf{4} \text{ and } z'_{i+2} = \mathbf{3}, \mathbf{4}, \\ z'_{i+2} & \text{with probability } \frac{1}{2} \text{ if } z'_{i+1} = \mathbf{2}, \mathbf{4} \text{ and } z'_{i+2} = \mathbf{3}, \mathbf{4}, \\ z'_i & \text{with probability } \frac{1}{2} \text{ if } z'_{i+1} = \mathbf{1}, \mathbf{3}, \\ z'_{i+1} & \text{with probability } \frac{1}{4} \text{ if } z'_{i+1} = \mathbf{1}, \mathbf{3}, \\ z'_{i+2} & \text{with probability } \frac{1}{4} \text{ if } z'_{i+1} = \mathbf{1}, \mathbf{3}. \end{cases} \quad (36)$$

4.6 Ancestral graph in the limit of a large population

Following the ancestry of a sample of size n in one time step back, there are $n(n-1)/2$ possible pure coalescence events and n possible pure events of recombination as well as n of viability selection and n of intersexual selection. All these events have probabilities of order $O(N^{-2})$. On the other hand, simultaneous events have probabilities of order $O(N^{-3})$. The probability that at least one event occurs is

$$p_n = \frac{2\lambda_n}{N^2} + O\left(\frac{1}{N^3}\right), \quad (37)$$

where

$$\lambda_n = \frac{n(n-1 + \sigma(1+R+A))}{2}. \quad (38)$$

Given that at least one event occurs and whatever the time before this event is, we have the conditional probabilities

$$\mathbb{P}(C_n) = \frac{1}{\lambda_n} + O\left(\frac{1}{N}\right), \quad (39)$$

$$\mathbb{P}(R_n) = \frac{R\sigma}{2\lambda_n} + O\left(\frac{1}{N}\right), \quad (40)$$

$$\mathbb{P}(V_n) = \frac{\sigma}{2\lambda_n} + O\left(\frac{1}{N}\right), \quad (41)$$

$$\mathbb{P}(S_n) = \frac{A\sigma}{2\lambda_n} + O\left(\frac{1}{N}\right), \quad (42)$$

where C_n , R_n , V_n et S_n designate pure events of coalescence, recombination, viability selection and intersexual selection, respectively.

Moreover, the time back in number of time steps before at least one event occurs, represented by τ_n , depends only on n and satisfies

$$\mathbb{P}(\tau_n > \tau) = (1 - p_n)^\tau, \quad (43)$$

for every integer $\tau \geq 0$, which corresponds to a geometric distribution of parameter p_n with

$$\mathbb{E}(\tau_n) = \sum_{\tau=0}^{\infty} (1 - p_n)^\tau = \frac{1}{p_n}. \quad (44)$$

The corresponding time in number of $N^2/2$ time steps in the limit of a large population size, represented by T_n , has a probability distribution given by

$$\mathbb{P}(T_n > t) = \lim_{N \rightarrow \infty} \mathbb{P}\left(\tau_n > \left\lfloor \frac{tN^2}{2} \right\rfloor\right), \quad (45)$$

for every real number $t \geq 0$, where $\lfloor \cdot \rfloor$ designates the floor value. Therefore, we have

$$\mathbb{P}(T_n > t) = \lim_{N \rightarrow \infty} (1 - p_n)^{\lfloor \frac{tN^2}{2} \rfloor} = e^{-\lambda_n t}, \quad (46)$$

for every real number $t \geq 0$, with

$$\mathbb{E}(T_n) = \int_0^{+\infty} e^{-\lambda_n t} dt = \frac{1}{\lambda_n} = \lim_{N \rightarrow \infty} \frac{2}{N^2 p_n} = \lim_{N \rightarrow \infty} \frac{2}{N^2} \mathbb{E}(\tau_n). \quad (47)$$

In other words, the time for a state change in the ancestry of the sample of size n in the limit of a large population size N follows an exponential distribution with parameter λ_n at the end of which a pure event that depends only on the state left occurs. In the same limit, the rates of pure coalescence, recombination, viability selection and intersexual selection are $n(n-1)/2$, $n\sigma R/2$, $n\sigma/2$ and $n\sigma A/2$, respectively, which sum up to λ_n . These parameters characterize the ancestral graph in the limit of a large population.

5 Application of the ancestral graph

We are now ready to compute the quantity $E_{\mathbf{i}, \mathbf{j}, \mathbf{k}}(\mathbf{x})$ for three types $\mathbf{i}, \mathbf{j}, \mathbf{k}$ and an initial population state $\mathbf{x}(0) = \mathbf{x}$. Let $\mathbf{z}(\tau)$ represent the ordered sample of three individuals chosen at random without replacement in the population of size N at time step τ . Define

$$\xi_{\mathbf{k}}^{(i)}(\tau) = \begin{cases} 1 & \text{if individual } i \text{ is of type } \mathbf{k}, \\ 0 & \text{otherwise,} \end{cases} \quad (48)$$

for $i = 1, 2, 3$. Given the type frequencies $x_{\mathbf{i}}(\tau)$, $x_{\mathbf{j}}(\tau)$ and $x_{\mathbf{k}}(\tau)$, the random variable $\xi_{\mathbf{i}}^{(1)}(\tau)\xi_{\mathbf{j}}^{(2)}(\tau)\xi_{\mathbf{k}}^{(3)}(\tau)$ takes the value 1 with probability

$$\frac{N^2}{(N-1)(N-2)} x_{\mathbf{i}}(\tau)x_{\mathbf{j}}(\tau)x_{\mathbf{k}}(\tau), \quad (49)$$

and 0 otherwise. Therefore, we have

$$\begin{aligned} \mathbb{P}(\mathbf{z}(\tau) = (\mathbf{i}, \mathbf{j}, \mathbf{k}) | \mathbf{x}(0) = \mathbf{x}) &= \mathbb{E}(\xi_{\mathbf{i}}^{(1)}(\tau)\xi_{\mathbf{j}}^{(2)}(\tau)\xi_{\mathbf{k}}^{(3)}(\tau) | \mathbf{x}(0) = \mathbf{x}) \\ &= \mathbb{E}(\mathbb{E}(\xi_{\mathbf{i}}^{(1)}(\tau)\xi_{\mathbf{j}}^{(2)}(\tau)\xi_{\mathbf{k}}^{(3)}(\tau) | x_{\mathbf{i}}(\tau), x_{\mathbf{j}}(\tau), x_{\mathbf{k}}(\tau)) | \mathbf{x}(0) = \mathbf{x}) \\ &= \frac{N^2}{(N-1)(N-2)} \mathbb{E}(x_{\mathbf{i}}(\tau)x_{\mathbf{j}}(\tau)x_{\mathbf{k}}(\tau) | \mathbf{x}(0) = \mathbf{x}). \end{aligned} \quad (50)$$

Plugging this into (20) yields

$$E_{\mathbf{i},\mathbf{j},\mathbf{k}}(\mathbf{x}) = \frac{2}{N^2} \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \sum_{\tau=0}^{\infty} \mathbb{P}(\mathbf{z}(\tau) = (\mathbf{i}, \mathbf{j}, \mathbf{k}) | \mathbf{x}(0) = \mathbf{x}). \quad (51)$$

Letting $\mathbf{z} = (\mathbf{i}, \mathbf{j}, \mathbf{k})$ and conditioning on the ancestral graph from time step τ to time step 0, we have

$$\mathbb{P}(\mathbf{z}(\tau) = \mathbf{z} | \mathbf{x}(0) = \mathbf{x}) = \sum_{G(\tau)} \mathbb{P}(\mathbf{z}(\tau) = \mathbf{z} | G(\tau), \mathbf{x}(0) = \mathbf{x}) \mathbb{P}(G(\tau)), \quad (52)$$

where $G(\tau)$ designates an ancestral graph for three individuals chosen at random without replacement in the population at time step τ .

The topology of the graph $G(\tau)$ is represented by a sequence of events backwards in time $G = (E_1, E_2, \dots, E_m)$ with E_i in the form R_n, S_n, \dots , for $i = 1, \dots, m$ for some finite number of events m . The number of ancestors following the last event backwards in time, which is the most ancient event, is denoted by n_G . Moreover, this last event occurs at time step back τ_G and the time spent with these n_G ancestors is denoted by τ_{n_G} . The time τ_G is a sum of independent geometric random variables and the time τ_{n_G} a geometric random variable independent of τ_G . Besides, we must have the inequalities $\tau_G \leq \tau < \tau_G + \tau_{n_G}$.

For instance, if $G = (R_3, V_4, C_5)$, then we have $n_G = 4$, $\tau_G = \tau_3 + \tau_4 + \tau_5$ where τ_3, τ_4 and τ_5 are independent geometric random variables of parameters p_3, p_4 and p_5 , respectively, and τ_{n_G} is a geometric random variable of parameter p_4 independent of τ_G .

The conditional probability that $\mathbf{z}(\tau) = \mathbf{z}$ given $G(\tau)$ and $\mathbf{x}(0) = \mathbf{x}$ depends in fact only on G and \mathbf{x} . Therefore, we introduce the notation

$$\mathbb{P}(\mathbf{z}(\tau) = \mathbf{z} | G(\tau), \mathbf{x}(0) = \mathbf{x}) = \mathbb{P}_G(\mathbf{z} | \mathbf{x}). \quad (53)$$

On the other hand, we have

$$\begin{aligned} \mathbb{P}(G(\tau)) &= \mathbb{P}(\tau_G \leq \tau < \tau_G + \tau_{n_G}) \mathbb{P}(G) \\ &= \mathbb{P}(\tau_G \leq \tau < \tau_G + \tau_{n_G}) \prod_{i=1}^m \mathbb{P}(E_i) \\ &= (\mathbb{P}(\tau_G + \tau_{n_G} > \tau) - \mathbb{P}(\tau_G > \tau)) \prod_{i=1}^m \mathbb{P}(E_i). \end{aligned} \quad (54)$$

Note that

$$\sum_{\tau=0}^{\infty} (\mathbb{P}(\tau_G + \tau_{n_G} > \tau) - \mathbb{P}(\tau_G > \tau)) = \mathbb{E}(\tau_G + \tau_{n_G}) - \mathbb{E}(\tau_G) = \mathbb{E}(\tau_{n_G}). \quad (55)$$

Then, we have

$$\sum_{\tau=0}^{\infty} \mathbb{P}(\mathbf{z}(\tau) = \mathbf{z} | \mathbf{x}(0) = \mathbf{x}) = \sum_G \mathbb{P}_G(\mathbf{z} | \mathbf{x}) \mathbb{P}(G) \mathbb{E}(\tau_{n_G}). \quad (56)$$

We conclude that

$$\begin{aligned} \lim_{N \rightarrow \infty} E_{\mathbf{i}, \mathbf{j}, \mathbf{k}}(\mathbf{x}) &= \lim_{N \rightarrow \infty} \frac{2}{N^2} \sum_{\tau=0}^{\infty} \mathbb{P}(\mathbf{z}(\tau) = \mathbf{z} | \mathbf{x}(0) = \mathbf{x}) \\ &= \sum_G \mathbb{P}_G(\mathbf{z} | \mathbf{x}) \mathbb{P}(G) \mathbb{E}(T_{n_G}). \end{aligned} \quad (57)$$

Here, $\mathbb{E}(T_{n_G})$ is the expected time with n_G ancestors with $N^2/2$ time steps as unit of time in the limit of a large population size, given by $\lambda_{n_G}^{-1}$. The calculation of the probability $\mathbb{P}_G(\mathbf{z} | \mathbf{x})$ in the limit of a large population size is presented in Appendix B. The summation is over all finite topologies G of the ancestral graph starting with three individuals and involving only pure events, the only ones with positive probabilities in the limit of a large population size. Such graph topologies are called pure. The number of pure graph topologies is actually infinite. However, only a finite number can be considered to get an approximation of any given order of the fixation probability $u_{T_2}(\mathbf{x})$ with respect to σ , the intensity of selection and recombination.

First, note that in all triplets $\mathbf{z} = (\mathbf{i}, \mathbf{j}, \mathbf{k})$ such that $Q_{\mathbf{i}, \mathbf{j}, \mathbf{k}} \neq 0$ in the fixation probability $u_{T_2}(\mathbf{x})$, there is at least one type where allele T_1 appears and also at least one type where allele T_2 appears. Therefore, $\mathbb{P}_G(\mathbf{z} | \mathbf{x}) = 0$ for every graph topology G such that the minimum number of ancestors at any point of the graph, denoted by $|G|$, is 1. On the other hand,

$$\mathbb{P}_G(\mathbf{z} | \mathbf{x}) \leq \frac{1}{N} \quad (58)$$

for every graph topology G such that $|G| \geq 2$, since at least one of the ancestors at time step 0 must carry allele T_2 .

For a random sample of i individuals in a population of size N , let W_i be the time back in number of $N^2/2$ time steps in the limit of a large population size before reaching a unique ancestor. The expected value is given by

$$\mathbb{E}(W_i) = \sum_{G \text{ for } i \text{ individuals} : |G| \geq 2} \mathbb{P}(G) \mathbb{E}(T_{n_G}). \quad (59)$$

It can be shown (see Appendix A) that $\mathbb{E}(W_i)$ is finite and increasing with i .

Let n_G^\dagger be the number of events in a pure graph topology G corresponding to an increase in the number of ancestors (actually pure events of recombination, viability selection or intersexual selection with increases of one or two ancestors for each). If G starts with three individuals and $n_G^\dagger > k$ for some integer $k \geq 1$, then we can write $G = (G_1, G_2)$ where G_1 and G_2 are pure graph topologies

such that $n_{G_1}^+ = k + 1$ and $n_{G_1} \leq 2(k + 1) + 3 = 2k + 5$. Note that this decomposition is not unique but that the number of possible decompositions is finite. Moreover, we have

$$\mathbb{P}(G) = \mathbb{P}(G_1)\mathbb{P}(G_2), \quad (60)$$

with

$$\mathbb{P}(G_1) = O(\sigma^{k+1}) \quad (61)$$

and

$$\sum_{G_2: |G_2| \geq 2} \mathbb{P}(G_2)\mathbb{E}(T_{n_{G_2}}) = \mathbb{E}(W_{n_{G_1}}) \leq \mathbb{E}(W_{2k+5}) < \infty. \quad (62)$$

We conclude that

$$\sum_{G: |G| \geq 2, n_G^+ > k} \mathbb{P}_G(\mathbf{z}|\mathbf{x})\mathbb{P}(G)\mathbb{E}(T_{n_G}) \leq \frac{1}{N}O(\sigma^{k+1}). \quad (63)$$

Finally, neglecting all terms of orders $N^{-1}O(\sigma^{k+1})$ and $O(N^{-2})$, we have the approximation

$$E_{i,j,k}(\mathbf{x}) \approx \sum_{G: |G| \geq 2, n_G^+ \leq k} \mathbb{P}_G(\mathbf{z}|\mathbf{x})\mathbb{P}(G)\mathbb{E}(T_{n_G}). \quad (64)$$

This approximation can be computed with MATHEMATICA using a development of $\mathbb{P}_G(\mathbf{z}|\mathbf{x})$ given in Appendix B and an algorithm presented in Appendix C which significantly shortens the calculation time.

6 Results and discussion

6.1 Fixation at the trait locus

Assume that the sexually preferred but viability-deleterious allele T_2 is introduced at random by a single mutation at the trait locus into a population of size N when the frequency of allele P_2 at the preference locus is x so that the population state is $\mathbf{x} = (1 - x, x, 0, 0)$. Then the probability of ultimate fixation of T_2 is given by the expression

$$u_{T_2} = xu_{T_2}(\mathbf{x} - N^{-1}\mathbf{e}_2 + N^{-1}\mathbf{e}_4) + (1 - x)u_{T_2}(\mathbf{x} - N^{-1}\mathbf{e}_1 + N^{-1}\mathbf{e}_3). \quad (65)$$

Under weak selection and weak recombination, we get the approximation

$$\begin{aligned}
u_{T_2} \approx & \frac{1}{N} + \sigma \left(-\frac{1}{2N} + \frac{Ax}{4N} \right) \\
& + \sigma^2 \left(\frac{1}{12N} - \frac{Ax}{12N} - \frac{7A^2x}{2160N} + \frac{13A^2x^2}{540N} \right) \\
& + \sigma^3 \left(\frac{283A^2x(1-x)}{86400N} + \frac{41A^2Rx(1-x)}{194400N} \right. \\
& \left. - \frac{7831A^3x}{54432000N} - \frac{5471A^3x^2}{4536000N} + \frac{73483A^3x^3}{54432000N} \right). \tag{66}
\end{aligned}$$

If the viabilities at the trait locus are interchanged so that the single mutant T_2 is viability-beneficial instead of being viability-deleterious, then only three signs in the above expression are changed to give the approximation

$$\begin{aligned}
u_{T_2} \approx & \frac{1}{N} + \sigma \left(\frac{1}{2N} + \frac{Ax}{4N} \right) \\
& + \sigma^2 \left(\frac{1}{12N} + \frac{Ax}{12N} - \frac{7A^2x}{2160N} + \frac{13A^2x^2}{540N} \right) \\
& + \sigma^3 \left(-\frac{283A^2x(1-x)}{86400N} + \frac{41A^2Rx(1-x)}{194400N} \right. \\
& \left. - \frac{7831A^3x}{54432000N} - \frac{5471A^3x^2}{4536000N} + \frac{73483A^3x^3}{54432000N} \right). \tag{67}
\end{aligned}$$

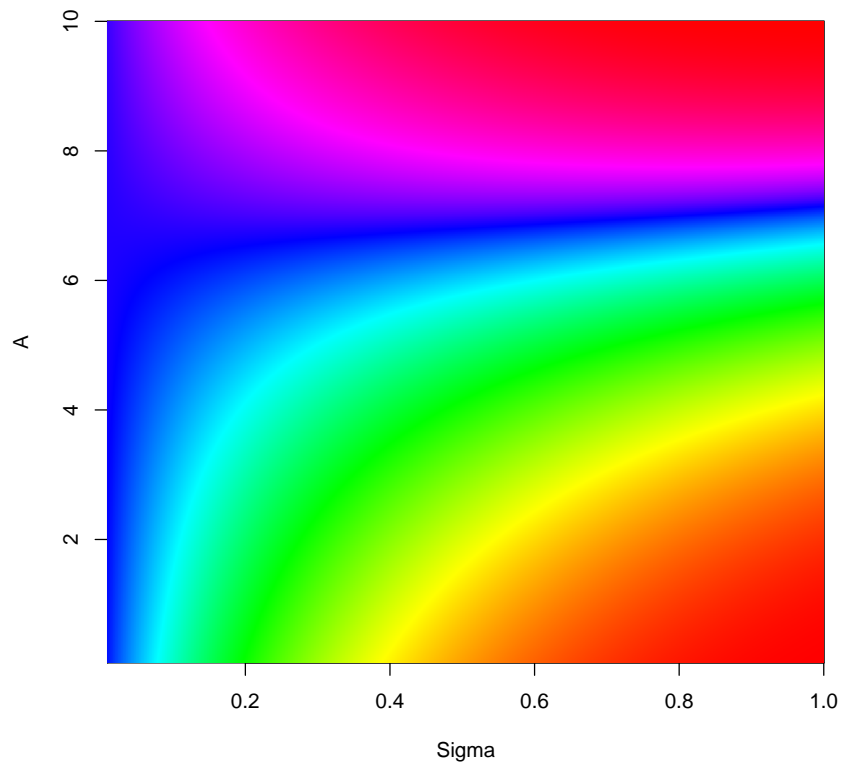
These approximations are valid for a population size N large enough and an intensity of selection and recombination σ small enough so that terms of order $\sigma^k N^{-l}$ for $k \geq 4$ or $l \geq 2$ can be ignored.

As expected, the leading effect of intersexual selection (term in σA) is to increase the probability of ultimate fixation of the sexually preferred allele T_2 . Moreover, if $Ax > 2$, then this effect is larger than the leading effect of viability selection so that it compensates this effect in the case of a viability-deleterious allele T_2 . Even in this case, the probability of ultimate fixation of T_2 under weak selection may exceed what it would be under neutrality, which is given by its initial frequency N^{-1} (see Fig. 1).

Surprisingly, the leading effect of recombination, though small (term in $\sigma^3 R$), is also to increase the probability of ultimate fixation of the sexually preferred allele T_2 . This appears to be the case for a viability-deleterious allele as well as a viability-beneficial allele. Random drift in a finite population is known to build up negative linkage disequilibrium (see, e.g. Barton and Otto, 2005, or Lessard and Kermany, 2012, and references therein). Recombination breaks down this disequilibrium and this increases the frequency of association between T_2 and P_2 . The effect of this increase is to promote the spread of T_2 up to fixation.

With free recombination or any fixed recombination rate in the limit of a large population size (see Appendix E), the probability of ultimate fixation of a viability-deleterious mutant T_2 that is sexually preferred by P_2 -females of initial frequency x is approximated under weak selection by

Figure 1: Fixation probability u_{T_2} for a viability-deleterious trait under intersexual selection as a function of the coefficient of selection σ and the coefficient of intersexual selection A for $x = 0.3$, $N = 1$ and $R = 1$. Note that the approximation is accurate only if $\sigma \ll 1$. The value increases from bottom (red) to top (violet).



$$\begin{aligned}
u_{T_2} \approx & \frac{1}{N} + \sigma \left(-\frac{1}{2N} + \frac{Ax}{4N} \right) \\
& + \sigma^2 \left(\frac{1}{12N} - \frac{Ax}{12N} - \frac{A^2x}{384N} + \frac{3A^2x^2}{128N} \right) \\
& + \sigma^3 \left(\frac{7A^2x(1-x)}{1536N} - \frac{A^3x}{2048N} - \frac{5A^3x^2}{6144N} + \frac{A^3x^3}{768N} \right). \quad (68)
\end{aligned}$$

The leading term of the difference between this approximation in the case of strong recombination and the previous approximation in the case of weak recombination is given by

$$\sigma^2 \left(-\frac{A^2x}{384N} + \frac{3A^2x^2}{128N} \right) - \sigma^2 \left(-\frac{7A^2x}{2160N} + \frac{13A^2x^2}{540N} \right) \geq 0, \quad (69)$$

with equality to 0 if and only if $x = 0$ or $x = 1$. This is also the case for a viability-beneficial mutant T_2 . This confirms that recombination favors the fixation of any mutant preferred allele.

The fact that recombination increases the probability of ultimate fixation of a sexually preferred mutant allele T_2 appears to be caused by an increase in linkage disequilibrium between this allele and the allele P_2 coding for the preference. Selection at the preference locus associated with a cost of being choosy or not as assumed in this paper may have a hitchhiking effect on the trait locus so that it may enhance or lessen the effect of recombination on the fixation probability of T_2 . However, this should not change the sign of this effect. Therefore, the same conclusion is expected in the case of a neutral preference locus.

6.2 Fixation at the preference locus

The probability of ultimate fixation of allele P_2 introduced at random by a single mutation at the preference locus into a population of size N when the frequency of the sexually preferred but viability-deleterious allele T_2 at the trait locus is y so that the population state is $\mathbf{y} = (1 - y, 0, y, 0)$ can be obtained in a similar way. Its expression is given by

$$u_{P_2} = yu_{P_2}(\mathbf{y} - N^{-1}\mathbf{e}_3 + N^{-1}\mathbf{e}_4) + (1 - y)u_{P_2}(\mathbf{y} - N^{-1}\mathbf{e}_1 + N^{-1}\mathbf{e}_2). \quad (70)$$

Using the coefficients $Q_{i,j,k}$ in Appendix D for the expected change in the frequency of P_2 and the same algorithm as before in calculating the expected values $E_{i,j,k}(\mathbf{y})$ under weak selection and weak recombination, we get the approximation

$$\begin{aligned}
u_{P_2} \approx & \frac{1}{N} + \sigma \left(-\frac{A}{8N} + \frac{Ay}{4N} \right) \\
& + \sigma^2 \left(-\frac{Ay(1-y)}{36N} + \frac{A^2}{192N} - \frac{37A^2y(1-y)}{2160N} \right) \\
& + \sigma^3 \left(\frac{Ay(1-y)(1-2y)}{1440N} + \frac{493A^2y(1-y)(1-2y)}{259200N} \right. \\
& + \frac{16993A^3y(1-y)(1-2y)}{108864000N} - \frac{101ARy(1-y)}{2592N} \\
& \left. + \frac{7739A^2Ry(1-y)}{777600N} \right). \tag{71}
\end{aligned}$$

If allele T_2 is viability-beneficial, then the approximation is given by

$$\begin{aligned}
u_{P_2} \approx & \frac{1}{N} + \sigma \left(-\frac{A}{8N} + \frac{Ay}{4N} \right) \\
& + \sigma^2 \left(\frac{Ay(1-y)}{36N} + \frac{A^2}{192N} - \frac{37A^2y(1-y)}{2160N} \right) \\
& + \sigma^3 \left(\frac{Ay(1-y)(1-2y)}{1440N} - \frac{493A^2y(1-y)(1-2y)}{259200N} \right. \\
& + \frac{16993A^3y(1-y)(1-2y)}{108864000N} + \frac{101ARy(1-y)}{2592N} \\
& \left. + \frac{7739A^2Ry(1-y)}{777600N} \right). \tag{72}
\end{aligned}$$

We see that the leading effect of intersexual selection (term in σA) increases the probability of ultimate fixation of P_2 only if the initial frequency of T_2 exceeds the initial frequency of T_1 (actually $y > 1/2$). Under this condition, the expected reproduction cost incurred by a female for not being choosy exceeds the expected reproduction cost incurred by a female for being choosy. This condition is unlikely to occur under the effects of random drift if T_2 is viability-deleterious.

In the critical case where the frequencies of T_1 and T_2 are equal ($y = 1/2$), there is no difference in the expected reproduction cost incurred by choosy females and non-choosy females. In this case, the leading effect of intersexual selection is to increase the probability of ultimate fixation of P_2 if T_2 is viability-beneficial and

$$\sigma^2 \left(\frac{A}{144N} + \frac{A^2}{192N} - \frac{37A^2}{8640N} \right) = \sigma^2 \left(\frac{A}{144N} + \frac{A^2}{1080N} \right) > 0, \tag{73}$$

which always holds, or if T_2 is viability-deleterious and the coefficient of intersexual selection is large enough so that

$$\sigma^2 \left(-\frac{A}{144N} + \frac{A^2}{192N} - \frac{37A^2}{8640N} \right) = \sigma^2 \left(-\frac{A}{144N} + \frac{A^2}{1080N} \right) > 0, \quad (74)$$

which holds when $A > 7.5$.

Finally, the leading effect of recombination (term in $\sigma^3 R$) is to increase the probability of ultimate fixation of P_2 unless T_2 is viability-deleterious and the coefficient of intersexual selection is not too large so that

$$-\frac{101ARy(1-y)}{2592N} + \frac{7739A^2Ry(1-y)}{777600N} \leq 0, \quad (75)$$

which occurs for $0 < y < 1$ when $A \leq 3.915$. By diminishing the magnitude of the negative linkage disequilibrium built up by random drift, it seems that recombination strengthens the hitchhiking effect of the trait locus on the preference locus. When the preferred trait is deleterious, the coefficient of intersexual selection must be high enough to compensate this effect.

6.3 Evolution of intersexual selection

In a finite population, conditions that increase the fixation probability of a new mutant may indicate the direction of evolution. We conclude from our findings that the best scenario for the evolution of a sexual preference for a viability-deleterious trait in a finite population is that the preference was neutral at least when introduced as a single mutant and increased in frequency either by random drift or by association with a viability-beneficial preferred trait before a single mutation on this trait rendered it more exaggerated and more preferred but viability-deleterious.

Our findings are in agreement with the phases in the history of a secondary sexual trait that were described by Fisher (1915, p. 187): *In the first it is favoured by natural selection, and being simple and easily apprehended as a "point," its advantage is slowly increased by the development of sexual selection in its favour; in the second phase it owes nothing to natural selection, which may even have turned against it, but it still increases in splendour and perfection, and the importance attached to it by the opposite sex still increases, so long as it retains a balance of advantage. Finally, perhaps, an equilibrium will be attained in which natural selection just balances sexual selection.*

Our results are based on fixation probabilities in the limit of a large finite population under very weak selection. While intersexual selection creates positive linkage disequilibrium, random drift does the opposite. Selection being weaker than drift (of order σN^{-1} with $\sigma \ll 1$ compared to N^{-1}), linkage disequilibrium tends to be negative. Recombination breaks down this disequilibrium and doing so reinforces intersexual selection so that it can hamper the effect of a viability-deleterious trait and even counterbalance its hitchhiking effect on the preference locus. This effect of recombination distinguishes intersexual selection models in finite populations from their infinite population versions (see, e.g., Kirkpatrick, 1982, Barton and Turelli, 1991).

Our analytical results have been obtained under simplifying assumptions such as a monoecious haploid population. The effects of intersexual selection and recombination may be affected by dioecy and diploidy among other more realistic assumptions. Numerical studies and simulations like those mentioned in the Introduction are more suitable to deal with such assumptions.

A Bound for $\mathbb{E}(W_i)$

Suppose that $i \geq 2$ individuals have $j \geq 2$ ancestors at a given time step. From this time step to the previous one, even allowing for simultaneous events, this number of ancestors can only go from j to $j-1$, j , $j+1$ and $j+2$ with positive probabilities. Moreover, these probabilities satisfy

$$\begin{aligned} p_{j,j-1} &\geq \frac{j(j-1)}{N^2} \left(1 - \frac{A\sigma}{N}\right) \left(1 - \frac{R\sigma}{N}\right) \left(1 - \frac{\sigma}{N}\right) \geq \frac{j(j-1)}{2N^2}, \\ p_{j,j+1} &\leq \frac{j}{N} \left(\frac{\sigma}{N} + \frac{R\sigma}{N} + \frac{A\sigma}{N}\right) \leq \frac{j}{12N^2}, \\ p_{j,j+2} &\leq \frac{j}{N} \left(\frac{R\sigma^2}{N^2} + \frac{A\sigma}{N}\right) \leq \frac{j}{12N^2}, \end{aligned} \quad (76)$$

if σ is small enough, with

$$p_{j,j} = 1 - p_{j,j-1} - p_{j,j+1} - p_{j,j+2}. \quad (77)$$

Then, the sojourn time back with j ancestors has expected value

$$\frac{1}{1 - p_{j,j}} \leq \frac{1}{p_{j,j-1}} \leq \frac{2N^2}{j(j-1)} \leq N^2. \quad (78)$$

In number of $N^2/2$ time steps, the expected sojourn time back with j ancestors is bounded by 2. Therefore, the time back in number of $N^2/2$ time steps in the limit of a large population size before reaching a single ancestor starting with i individuals, denoted by W_i , satisfies

$$\mathbb{E}(W_i) \leq 2\mathbb{E}(M_i), \quad (79)$$

where M_i represents the number of changes in the number of ancestors before reaching state 1 starting from state i .

Let $(U_n)_{n \geq 1}$ be a sequence of independent random variables having a uniform distribution on $[0, 1]$. Let $(X_n)_{n \geq 0}$ be a Markov chain on the positive integers with $X(0) = i \geq 2$ and 1 as absorbing state, defined by

$$\begin{aligned} X_n - X_{n-1} &= 3\mathbb{1}_{[0, q_{j,j+2}]}(U_n) + \mathbb{1}_{[1 - q_{j,j-1} - q_{j,j+1}, 1 - q_{j,j-1}]}(U_n) \\ &\quad - \mathbb{1}_{[1 - q_{j,j-1}, 1]}(U_n) \end{aligned} \quad (80)$$

if $X_{n-1} \neq 1$, where $\mathbb{1}_I$ designates the indicator function of the interval I . The transition probabilities of this Markov chain are given by

$$q_{1,1} = 1, \quad (81)$$

so that 1 is absorbing,

$$q_{j,j+k} = \frac{p_{j,j+k}}{1 - p_{j,j}} \leq \frac{2jN^2}{12j(j-1)N^2} \leq \frac{1}{6}, \quad (82)$$

for $j \geq 2$ and $k = 1, 2$, and

$$q_{j,j-1} = 1 - q_{j,j+1} - q_{j,j+2} \geq \frac{2}{3}. \quad (83)$$

Now, let $(Y_n)_{n \geq 0}$ be the Markov chain with $Y_0 = i$ and 1 as absorbing state, defined by

$$Y_n - Y_{n-1} = 2\mathbf{1}_{[0, \frac{1}{6}]}(U_n) + \mathbf{1}_{[\frac{1}{6}, \frac{1}{3}]}(U_n) - \mathbf{1}_{[\frac{1}{3}, 1]}(U_n), \quad (84)$$

if $Y_{n-1} \neq 1$. Then, we have

$$X_n = X_0 + \sum_{l=1}^n (X_l - X_{l-1}) \leq Y_0 + \sum_{l=1}^n (Y_l - Y_{l-1}) = Y_n, \quad (85)$$

so that

$$\mathbb{P}(X_n > 1 | X_0 = i) \leq \mathbb{P}(Y_n > 1 | Y_0 = i), \quad (86)$$

for every $n \geq 0$. On another hand, we have

$$\mathbb{E}(M_i) = \sum_{n=0}^{\infty} \mathbb{P}(M_i > n) = \sum_{n=0}^{\infty} \mathbb{P}(X_n > 1 | X_0 = i). \quad (87)$$

Therefore, we have

$$\mathbb{E}(M_i) = \sum_{n=0}^{\infty} \mathbb{P}(X_n > 1 | X_0 = i) \leq \sum_{n=0}^{\infty} \mathbb{P}(Y_n > 1 | Y_0 = i) = \mathbb{E}(N_i), \quad (88)$$

where N_i is defined as the number of transitions for the Markov chain $(Y_n)_{n \geq 0}$ to reach state 1 from state $i \geq 2$.

To conclude, we make a parallel between the Markov chain $(Y_n)_{n \geq 0}$ with initial state $i \geq 2$ and absorbing state 1 and a Bienaymé-Galton-Watson (BGW) branching process $(Z_n)_{n \geq 0}$ starting with $i - 1$ individuals. In this branching process for the size of a population with discrete, non-overlapping generations, every individual independently of all others produces 3, 2 or 0 offspring with respective probabilities $\frac{1}{6}$, $\frac{1}{6}$ and $\frac{2}{3}$. This corresponds to a transition in the process $(Y_n)_{n \geq 0}$. Therefore, the number of transitions in this process to reach state 1 from state $i \geq 2$ corresponds to the total number of individuals in the process $(Z_n)_{n \geq 0}$ starting with the $i - 1$ individuals of the initial generation and ending with the extinction of the population with none. This occurs with probability 1, since the expected number of offspring produced by an individual is $m = \frac{5}{6} < 1$. Finally, we have

$$\begin{aligned}
\mathbb{E}(N_i) &= \mathbb{E}\left(\sum_{n=0}^{\infty} Z_n\right) = \sum_{n=0}^{\infty} \mathbb{E}(Z_n) \\
&= \sum_{n=0}^{\infty} m^n \mathbb{E}(Z_0) = (i-1) \sum_{n=0}^{+\infty} \left(\frac{5}{6}\right)^n = 6(i-1). \tag{89}
\end{aligned}$$

We conclude that

$$\mathbb{E}(W_i) \leq 12(i-1), \tag{90}$$

for $i \geq 2$. Moreover, it is obvious that this expected value is increasing from $i \geq 2$ to $i+1$, since the time for reaching a single ancestor from $i+1$ individuals is always greater or equal to the time for reaching a single ancestor from any given subset of i of these individuals.

B Calculation of $\mathbb{P}_G(\mathbf{z}|\mathbf{x})$

We start with some additional notations. Let $\mathbf{x}(0) = \mathbf{x} = (x_i)_{i=1,\dots,4}$ be the vector of the initial type frequencies. Moreover, let the type configuration of an ordered sample \mathbf{z} be given by $\mathbf{n} = (n_i)_{i=1,\dots,4}$ where n_i is the number of individuals of type \mathbf{i} in the sample. Then, the sample size is $n = |\mathbf{n}| = \sum_{i=1}^4 n_i$.

For a graph topology $G = \emptyset$, we have

$$\mathbb{P}_G(\mathbf{z}|\mathbf{x}) = P_{\mathbf{x}}(\mathbf{n}) = \frac{1}{N^n} \prod_{i=1}^4 \frac{N_i!}{(N_i - n_i)!} = \prod_{i=1}^4 x_i^{n_i} + O\left(\frac{1}{N}\right), \tag{91}$$

where $(N_i)_{i=1,\dots,4} = N\mathbf{x}$.

Let us now consider a graph topology of a single pure event represented by $G = (E_1)$. Then we have

$$\mathbb{P}_G(\mathbf{z}|\mathbf{x}) = \sum_{\mathbf{z}'} \mathbb{P}_{E_1}(\mathbf{z}|\mathbf{z}') P_{\mathbf{x}}(\mathbf{n}'), \tag{92}$$

where

$$\mathbb{P}_{E_1}(\mathbf{z}|\mathbf{z}') = \begin{cases} \delta_{z_i, z_j} & \text{if } E_1 = C_n, \mathbf{z}' = C_{i,j}(\mathbf{z}), \\ (1 - c_{z'_i})\delta_{z_i, z'_{i+1}} + c_{z'_i}\delta_{z_i, z'_i} & \text{if } E_1 = V_n, \mathbf{z}' = V_i(\mathbf{z}), \\ \delta_{z_i(1), z'_i(1)}\delta_{z_i(2), z'_{i+1}(2)} & \text{if } E_1 = R_n, \mathbf{z}' = R_i(\mathbf{z}), \\ (1 - f_{z'_{i+1}, z'_{i+2}})\delta_{z_i, z'_i} \\ + f_{z'_{i+1}, z'_{i+2}}\left(\frac{1}{2}\delta_{z_i, z'_{i+1}} + \frac{1}{2}\delta_{z_i, z'_{i+2}}\right) & \text{if } E_1 = S_n, \mathbf{z}' = S_i(\mathbf{z}), \end{cases} \quad (93)$$

and 0 otherwise, with δ denoting Kronecker's delta. Now let

$$\Psi_{E_1}(\mathbf{n}, \mathbf{n}') = \sum_{\mathbf{z}' : \mathbf{n}'} \mathbb{P}_{E_1}(\mathbf{z}|\mathbf{z}'), \quad (94)$$

where the summation is over all ordered samples \mathbf{z}' having the type configuration \mathbf{n}' . In the following we always assume $\mathbf{e}_i \neq \mathbf{e}_j \neq \mathbf{e}_k \neq \mathbf{e}_i$. Then we have :

$$\Psi_{E_1}(\mathbf{n}, \mathbf{n}') = \begin{cases} \frac{n_i(n_i-1)}{2} & \text{if } E_1 = C_n \\ \text{and } \mathbf{n}' = \mathbf{n} - \mathbf{e}_i, \\ n(1 - c_i) + \sum_{\mathbf{j}} n_{\mathbf{j}} c_{\mathbf{j}} & \text{if } E_1 = V_n \\ \text{and } \mathbf{n}' = \mathbf{n} + \mathbf{e}_i, \\ n_i(\delta_{i, (j_1, k_2)} + \delta_{i, (k_1, j_2)}) & \text{if } E_1 = R_n \\ \text{and } \mathbf{n}' = \mathbf{n} - \mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k, \\ \sum_{\mathbf{j} \neq \mathbf{i}} n_{\mathbf{j}}(\delta_{\mathbf{j}, (j_1, i_2)} + \delta_{\mathbf{j}, (i_1, j_2)}) + n_i & \text{if } E_1 = R_n \\ \text{and } \mathbf{n}' = \mathbf{n} + \mathbf{e}_i, \\ A_{i, \mathbf{j}} & \text{if } E_1 = S_n \\ \text{and } \mathbf{n}' = \mathbf{n} + \mathbf{e}_i + \mathbf{e}_j, \\ B_i & \text{if } E_1 = S_n \\ \text{and } \mathbf{n}' = \mathbf{n} + 2\mathbf{e}_i, \end{cases} \quad (95)$$

and 0 otherwise, where

$$\begin{aligned}
A_{\mathbf{i},\mathbf{j}} &= \sum_{\mathbf{l} \neq \mathbf{i},\mathbf{j}} n_{\mathbf{l}} \left(2 - f_{\mathbf{i},\mathbf{j}} - f_{\mathbf{j},\mathbf{i}} + \frac{f_{\mathbf{l},\mathbf{i}}}{2} + \frac{f_{\mathbf{l},\mathbf{j}}}{2} + \frac{f_{\mathbf{i},\mathbf{l}}}{2} + \frac{f_{\mathbf{j},\mathbf{l}}}{2} \right) \\
&\quad + n_{\mathbf{j}} \left(f_{\mathbf{j},\mathbf{j}} + 2 - \frac{f_{\mathbf{i},\mathbf{j}}}{2} - \frac{f_{\mathbf{j},\mathbf{i}}}{2} \right) + n_{\mathbf{i}} \left(f_{\mathbf{i},\mathbf{i}} + 2 - \frac{f_{\mathbf{i},\mathbf{j}}}{2} - \frac{f_{\mathbf{j},\mathbf{i}}}{2} \right) \\
&= (n - n_{\mathbf{i}} - n_{\mathbf{j}}) (2 - f_{\mathbf{i},\mathbf{j}} - f_{\mathbf{j},\mathbf{i}}) + \sum_{\mathbf{l} \neq \mathbf{i},\mathbf{j}} n_{\mathbf{l}} \left(\frac{f_{\mathbf{l},\mathbf{i}}}{2} + \frac{f_{\mathbf{l},\mathbf{j}}}{2} + \frac{f_{\mathbf{i},\mathbf{l}}}{2} + \frac{f_{\mathbf{j},\mathbf{l}}}{2} \right) \\
&\quad + n_{\mathbf{j}} \left(f_{\mathbf{j},\mathbf{j}} + 2 - \frac{f_{\mathbf{i},\mathbf{j}}}{2} - \frac{f_{\mathbf{j},\mathbf{i}}}{2} \right) + n_{\mathbf{i}} \left(f_{\mathbf{i},\mathbf{i}} + 2 - \frac{f_{\mathbf{i},\mathbf{j}}}{2} - \frac{f_{\mathbf{j},\mathbf{i}}}{2} \right) \quad (96)
\end{aligned}$$

and

$$\begin{aligned}
B_{\mathbf{i}} &= n_{\mathbf{i}} + \sum_{\mathbf{l} \neq \mathbf{i}} n_{\mathbf{l}} \left(1 - f_{\mathbf{i},\mathbf{l}} + \frac{f_{\mathbf{l},\mathbf{i}}}{2} + \frac{f_{\mathbf{l},\mathbf{i}}}{2} \right) \\
&= n_{\mathbf{i}} + (n - n_{\mathbf{i}}) (1 - f_{\mathbf{i},\mathbf{i}}) + \sum_{\mathbf{l} \neq \mathbf{i}} n_{\mathbf{l}} \left(\frac{f_{\mathbf{l},\mathbf{i}}}{2} + \frac{f_{\mathbf{l},\mathbf{i}}}{2} \right). \quad (97)
\end{aligned}$$

Then we have

$$\mathbb{P}_G(\mathbf{z}|\mathbf{x}) = \sum_{\mathbf{n}'} \Psi_{E_1}(\mathbf{n}, \mathbf{n}') P_{\mathbf{x}}(\mathbf{n}'). \quad (98)$$

Next we consider $G = (E_1, E_2)$ where E_1, E_2 are two pure events. We have

$$\mathbb{P}_G(\mathbf{z}|\mathbf{x}) = \sum_{\mathbf{n}''} \Psi_G(\mathbf{n}, \mathbf{n}'') P_{\mathbf{x}}(\mathbf{n}''), \quad (99)$$

with

$$\begin{aligned}
\Psi_G(\mathbf{n}, \mathbf{n}'') &= \sum_{\mathbf{z}'' : \mathbf{n}''} \mathbb{P}_G(\mathbf{z}|\mathbf{z}'') \\
&= \sum_{\mathbf{z}'' : \mathbf{n}''} \sum_{\mathbf{z}'} \mathbb{P}_{E_1}(\mathbf{z}|\mathbf{z}') \mathbb{P}_{E_2}(\mathbf{z}'|\mathbf{z}'') \\
&= \sum_{\mathbf{z}'} \mathbb{P}_{E_1}(\mathbf{z}|\mathbf{z}') \Psi_{E_2}(\mathbf{n}', \mathbf{n}'') \\
&= \sum_{\mathbf{n}'} \Psi_{E_1}(\mathbf{n}, \mathbf{n}') \Psi_{E_2}(\mathbf{n}', \mathbf{n}''). \quad (100)
\end{aligned}$$

Recursively, for a graph topology of m pure events $G = (E_1, E_2, \dots, E_m)$, we obtain

$$\mathbb{P}_G(\mathbf{z}|\mathbf{x}) = \sum_{\mathbf{n}^{(m)}} \Psi_G(\mathbf{n}^{(0)}, \mathbf{n}^{(m)}) P_{\mathbf{x}}(\mathbf{n}^{(m)}), \quad (101)$$

with

$$\Psi_G(\mathbf{n}^{(0)}, \mathbf{n}^{(m)}) = \sum_{\mathbf{n}^{(1)}} \cdots \sum_{\mathbf{n}^{(m-1)}} \prod_{k=1}^m \Psi_{E_k}(\mathbf{n}^{(k-1)}, \mathbf{n}^{(k)}). \quad (102)$$

This provides a mean to compute $\mathbb{P}_G(\mathbf{z}|\mathbf{x})$ with MATHEMATICA.

C Algorithm for the computation of $E_{i,j,k}(\mathbf{x})$

In order to calculate $E_{i,j,k}(\mathbf{x})$ in Eq. (64) in the shortest possible time, we use a method inspired by Horner's rule for calculating a polynomial of degree n evaluated at some specific value x_0 . The rule consists in writing

$$P(x_0) = \sum_{k=0}^n a_k x_0^k \quad (103)$$

into the form

$$P(x_0) = a_0 + x_0(a_1 + x_0(a_2 + \cdots + x_0(a_{n-1} + x_0 a_n) \cdots)).$$

Here, we write

$$E_{i,j,k}(\mathbf{x}) = \sum_{G \text{ eligible}} \mathbb{P}_G(\mathbf{z}|\mathbf{x}) \mathbb{P}(G) \mathbb{E}(T_{n_G}) \quad (104)$$

into the form

$$\begin{aligned} E_{i,j,k}(\mathbf{x}) &= P_{\mathbf{x}}(\mathbf{n}) \mathbb{E}(T_3) + \sum_{E_1: (E_1) \text{ eligible}} \mathbb{P}(E_1) \sum_{\mathbf{n}'} \Psi_{E_1}(\mathbf{n}, \mathbf{n}') \\ &\times \left(P_{\mathbf{x}}(\mathbf{n}') \mathbb{E}(T_{n_{E_1}}) + \sum_{E_2: (E_1, E_2) \text{ eligible}} \mathbb{P}(E_2) \sum_{\mathbf{n}''} \Psi_{E_2}(\mathbf{n}', \mathbf{n}'') \right. \\ &\left. \times (P_{\mathbf{x}}(\mathbf{n}'') \mathbb{E}(T_{n_{E_2}}) + \cdots) \right). \end{aligned} \quad (105)$$

Moreover, a graph G is eligible if $|G| \geq 2$ and $n_G^+ \leq k$. This algorithm reduces the calculation time by a factor 1/2000.

D Coefficients Q for the probability of fixation of allele P_2

$$\begin{aligned}
Q_{1,1,1} &= 0, & Q_{1,1,2} &= \frac{-A-4c_1+4c_2}{8} \sigma, \\
Q_{1,1,3} &= 0, & Q_{1,1,4} &= -\frac{-A-4c_1+4c_4}{8} \sigma, \\
Q_{1,2,2} &= \frac{-A-4c_1+4c_2}{8} \sigma, & Q_{1,2,3} &= \frac{-c_1+2c_2-c_3}{2} \sigma, \\
Q_{1,2,4} &= \frac{A-4c_1+2c_2+2c_4}{4} \sigma, & Q_{1,3,3} &= 0, \\
Q_{1,3,4} &= \frac{-c_1-c_3+2c_4}{2} \sigma, & Q_{1,4,4} &= \frac{3A-4c_1+4c_4}{8} \sigma, \\
Q_{2,2,2} &= 0, & Q_{2,2,3} &= \frac{-3A+4c_2-4c_3}{8} \sigma, \\
Q_{2,2,4} &= 0, & Q_{2,3,3} &= \frac{A+4c_2-4c_3}{8} \sigma, \\
Q_{2,3,4} &= \frac{-A+2c_2-4c_3+2c_4}{4} \sigma, & Q_{2,4,4} &= 0, \\
Q_{3,3,3} &= 0, & Q_{3,3,4} &= \frac{A-4c_3+4c_4}{8} \sigma, \\
Q_{3,4,4} &= \frac{A-4c_3+4c_4}{8} \sigma, & Q_{4,4,4} &= 0.
\end{aligned} \tag{106}$$

E Strong recombination

With free recombination or any fixed recombination rate in the limit of a large population size, the ancestral material of a sample of size n is represented by a vector $\mathbf{n} = (n_1, n_2, n_3, n_4)$ where n_1 and n_2 are the numbers of lineages that are ancestral at the trait locus with T_1 and T_2 , respectively, at this locus, while n_3 and n_4 are the numbers of lineages that are ancestral at the preference locus with P_1 and P_2 , respectively, at this locus. With $N^2/2$ time steps as unit of time in a population of size $N \rightarrow \infty$, the rate of viability selection is $\sigma/2$ and the rate of sexual selection $A\sigma/2$ for each ancestral lineage, while the rate of coalescence is 1 for each pair of lineages that are ancestral at the same locus. Note that the rate of coalescence of each pair of lineages that are ancestral at different loci is 1, but the lineages are instantaneously separated anew in the limit of a large population size so that no change actually occurs. This leads to a function

$$\Psi_{E_1}(\mathbf{n}, \mathbf{n}') = \begin{cases} \frac{n_i(n_i-1)}{2} & \text{if } E_1 = C_n \\ & \text{and } \mathbf{n}' = \mathbf{n} - \mathbf{e}_i \\ & \text{with } \mathbf{i} \in \{\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}\}, \\ n(1 - c_{ij}) + n_1 c_{1j} + n_2 c_{2j} + n_3 c_{3j} + n_4 c_{4j} & \text{if } E_1 = V_n \\ & \text{and } \mathbf{n}' = \mathbf{n} + \mathbf{e}_i + \mathbf{e}_j \\ & \text{with } \mathbf{i} \in \{\mathbf{1}, \mathbf{2}\}, \mathbf{j} \in \{\mathbf{3}, \mathbf{4}\}, \\ A_{i,j} & \text{if } E_1 = S_n \\ & \text{and } \mathbf{n}' = \mathbf{n} + \mathbf{e}_i + \mathbf{e}_j \\ & \text{with } \mathbf{i} \in \{\mathbf{3}, \mathbf{4}\}, \mathbf{j} \in \{\mathbf{1}, \mathbf{2}\}, \end{cases} \quad (107)$$

and 0 otherwise, where

$$\begin{aligned} A_{i,j} &= n(1 - f_{i,j}) \\ &+ \frac{1}{2} (n_1(f_{i,j} + f_{i,1}) + n_2(f_{i,j} + f_{i,2}) + n_3(f_{i,j} + f_{3,j}) + n_4(f_{i,j} + f_{4,j})) \\ &= n \left(1 - \frac{1}{2} f_{i,j} \right) + \frac{1}{2} (n_1 f_{i,1} + n_2 f_{i,2} + n_3 f_{3,j} + n_4 f_{4,j}). \end{aligned} \quad (108)$$

Note that c_{ij} represents the probability for an individual of type \mathbf{i} at the trait locus and of type \mathbf{j} at the preference locus to be replaced in the event of viability selection, while $f_{i,j}$ represents the probability for a female of type \mathbf{i} at the preference locus mated to a male of type \mathbf{j} at the trait locus to produce an offspring in the event of sexual selection.

With T_2 -males being viability-deleterious but sexually preferred by P_2 -females, for instance, we have

$$c_{11} = 1, \quad c_{12} = 1, \quad c_{21} = 0, \quad c_{22} = 0,$$

and

$$f_{1,1} = \frac{1}{2}, \quad f_{1,2} = \frac{1}{2}, \quad f_{2,1} = 0, \quad f_{2,2} = 1.$$

Acknowledgements

This research was supported in part by NSERC of Canada, Grant no 8833.

References

- [1] BARTON, N. H. AND OTTO, S. P. (2005) Evolution of recombination due to random drift. *Genetics* **169**, 2353–2370.

- [2] BARTON, N. H. AND TURELLI, M. (1991) Natural and sexual selection on many loci. *Genetics* **127**, 229–255.
- [3] BERGEN, E. L., ROWELL, J. T., GOULD, F. AND SERVEDIO, M. R. (2012) Stochasticity in sexual selection enables divergence: Implications for moth pheromone evolution, *Evolutionary Biology* **39**, 271–281.
- [4] CARRIER, V. (1995) *Modèles de sélection sexuelle à deux espèces*. Ph.D. Thesis in Statistics, Université de Montréal, 225 pages.
- [5] CHARLESWORTH, D. AND CHARLESWORTH, B. (1981) Frequency dependence in a model of sexual selection. *Theoretical Population Biology* **19**, 163–168.
- [6] DARWIN, C. (1871) *The Descent of Man and Selection in Relation to Sex*. John Murray, London.
- [7] DE JONG, M.C.M. AND SABELIS, M.W. (1991). Limits to runaway sexual selection: The wallflower paradox. *J. evol. Biol.* **4**, 637–655.
- [8] DONNELLY, P. AND KURTZ, T. G. (1999) Genealogical processes for Fleming-Viot models with selection and recombination. *Annals of Applied Probability* **9**, 1091–1148.
- [9] EDWARD, D.A. AND CHAPMAN, T. (2011). The evolution and significance of male mate choice. *Trends in Ecology and Evolution* **26**, 647–654.
- [10] FAWCETT, T.W., KUIJPER, B., PEN, I. AND WEISSING, F.J. (2007). Should attractive males have more sons? *Behavioral Ecology* **18**, 71–80.
- [11] FEARNHEAD, P. (2003) Ancestral processes for non-neutral models of complex diseases. *Theoretical Population Biology* **63**, 115–130.
- [12] FISHER, R. A. (1930) *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- [13] FISHER, R. A. (1958) *The Genetical Theory of Natural Selection*. Dover, New York.
- [14] GOMULKIEWICZ, R. S. AND HASTINGS, A. (1990). Ploidy and evolution by sexual selection: A comparison of haploid and diploid female choice models near fixation equilibria. *Evolution* **44**, 757–770.
- [15] GREENSPOON, P. B. AND OTTO, S. P. (2009) Evolution by Fisherian selection in diploids. *Evolution* **63**, 1076–1083.
- [16] GRIFFITHS, R. C. (1981) Neutral two-locus multiple allele models with recombination. *Theoretical Population Biology* **19**, 169–186.
- [17] HEISLER, I. L. AND CURTSINGER, J. W. (1990). Dynamics of sexual selection in diploid populations. *Evolution* **44**, 1164–1176.

- [18] HUDSON, R. R. (1983) Properties of a neutral allele with intragenic recombination. *Theoretical Population Biology* **23**, 183–201.
- [19] KARLIN, S. AND TAYLOR, H. M. (1975) *A First Course in Stochastic Processes. Second Edition*. Academic Press, San Diego, California.
- [20] KERMANY, A. R. AND LESSARD, S. (2012) Effect of epistasis and linkage on fixation probability in three-locus models: An ancestral recombination-selection graph approach, *Theoretical Population Biology* **82**, 131–145.
- [21] KINGMAN, J. F. C. (1982) The coalescent. *Stochastic Processes and Their Applications* **13**, 235–248.
- [22] KIRKPATRICK, M. (1982) Sexual selection and the evolution of female choice. *Evolution* **36**, 1–12.
- [23] KRONE, S. M. AND NEUHAUSER, C. (1997) Ancestral processes with selection. *Theoretical Population Biology* **51**, 210–237.
- [24] KUIJPER, B., PEN, I. AND WEISSING, F. J. (2012) A Guide to Sexual Selection Theory. *Annual Review of Ecology, Evolution, and Systematics* **43**, 287–311.
- [25] LANDE, C.F. R. (1981). Models of speciation by sexual selection on polygenic traits. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 3721–3725.
- [26] LESSARD, S. AND KERMANY, A. R. (2012) Fixation probability in a two-locus model by the ancestral recombination-selection graph. *Genetics* **190**, 691–707.
- [27] MAYNARD SMITH, J. (1982) *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.
- [28] MAYNARD SMITH, J. (2000) Attitudes to animal behavior, in *Evolutionary Genetics: From Molecules to Morphology*, edited by R. S. Singh and C. B. Krimbas. Cambridge University Press, Cambridge, pp. 628–640.
- [29] MEAD, L.S. AND ARNOLD, S.J. (2004). Quantitative genetic models of sexual selection. *Trends Ecol. Evol.* **19**, 264–271.
- [30] MILLER, G. F. (2000) *The Mating Mind: How sexual choice shaped the evolution of human nature*. Heinemann, London.
- [31] MORAN, P. A. P. (1958) Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society* **54**, 60–71.
- [32] O’DONALD, P. (1962). The theory of sexual selection. *Heredity* **17**, 541–552.
- [33] O’DONALD, P. (1980). *Genetic Models of Sexual Selection*. Cambridge University Press, New York, NY.

- [34] OTTO, S. P. (1991). On evolution under sexual and viability selection: A two-locus diploid model. *Evolution* **45**, 1443–1457.
- [35] OTTO, S. P., SERVEDIO, M. R. AND NUISMER, S. L. (2008). Frequency-dependent selection and the evolution of assortative mating. *Genetics* **179**, 2091–2112.
- [36] PUTS, D.A. (2010) Beauty and the beast: mechanisms of sexual selection in humans. *Evolution and Human Behavior* **31**, 157–175.
- [37] SEGER, J. (1985) Unifying genetic models for the evolution of female choice. *Evolution* **39**, 1185–1193.
- [38] SERVEDIO, M. R. (2011). Limits to the evolution of assortative mating by female choice under restricted gene flow. *Proceedings of the Royal Society of London, B, Biological Sciences* **278**, 179–187.
- [39] TAZZYMAN, S.J. AND IWASA Y. (2010). Sexual selection can increase the effect of random genetic drift—a quantitative genetic model of polymorphism in *Oophaga pumilio*, the strawberry poison-dart frog. *Evolution* **64**, 1719–1728.
- [40] UYEDA, J. C., ARNOLD, S. J., HOHENLOHE, P. A. AND MEAD, L. S. (2009). Drift promotes speciation by sexual selection. *Evolution* **63**, 583–594.