

# The fate of recessive deleterious or overdominant mutations near mating-type loci under partial selfing

Emilie Tezenas<sup>1, 2, 3, \*</sup>, Tatiana Giraud<sup>1, +</sup>, Amandine Véber<sup>2, +</sup>, and Sylvain Billiard<sup>3, +</sup>

<sup>1</sup>Universite Paris-Saclay, CNRS, AgroParisTech, Ecologie Systematique et Evolution, 91190, Gif-sur-Yvette, France

<sup>2</sup>Université Paris Cité, CNRS, MAP 5, F-75006 Paris, France

<sup>3</sup>Univ. Lille, CNRS, UMR 8198 –Evo-Eco-Paleo, F-59000 Lille, France

<sup>+</sup>Co-senior author

\*Corresponding author : emilie.tezenas-du-montcel@ens-rennes.fr, Unité Evo-Eco-Paleo, Bât. SN2, Cité Scientifique, 59655 Villeneuve d’Ascq, France, Phone number : +33 3 62 26 85 31

December 21, 2022

## 1 Abstract

2 Large regions of suppressed recombination having extended **over** time occur in many organisms around genes  
3 involved in mating compatibility (sex-determining or mating-type genes). The sheltering of deleterious alleles has  
4 been proposed to be involved in such expansions. However, the dynamics of deleterious mutations partially linked to  
5 genes involved in mating compatibility are not well understood, especially in finite populations. In particular, under  
6 what conditions deleterious mutations are likely to be maintained for long enough near mating-compatibility genes  
7 remains to be evaluated, especially under selfing, which generally increases the purging rate of deleterious mutations.  
8 Using a branching process approximation, we studied the fate of a new deleterious or overdominant mutation in a  
9 diploid population, considering a locus carrying two permanently heterozygous mating-type alleles, and a partially  
10 linked locus at which the mutation appears. We obtained analytical and numerical results on the probability and  
11 purging time of the new mutation. We investigated the impact of recombination between the two loci and of the  
12 mating system (outcrossing, intra and inter-tetrad selfing) on the maintenance of the mutation. We found that the  
13 presence of a **fungal-like** mating-type locus (*i.e.* **not preventing diploid selfing**) always sheltered the mutation under  
14 selfing, *i.e.* it decreased the purging probability and increased the purging time of the mutations. The sheltering  
15 effect was higher in case of automixis (intra-tetrad selfing). This may contribute to explain why evolutionary strata  
16 of recombination suppression near the mating-type locus are found mostly in automictic (pseudo-homothallic) fungi.  
17 We also showed that rare events of deleterious mutation maintenance during strikingly long evolutionary times could  
18 occur, suggesting that deleterious mutations can indeed accumulate near the mating-type locus over evolutionary  
19 time scales. In conclusion, our results show that, although selfing purges deleterious mutations, these mutations  
20 can be maintained for very long times near a mating-type locus, which may contribute to promote the evolution of  
21 recombination suppression in sex-related chromosomes.

22 **Keywords:** Multitype branching process; Sheltering effect; Deleterious mutation; Mating-type locus; Selfing;  
23 Overdominance; Sex chromosome; Mating-type chromosome; Evolutionary strata; Automixis; Pseudo-homothallism;  
24 Extinction time.

## 25 1 Introduction

26 The evolution of sex chromosomes, and more generally of genomic regions lacking recombination, is widely studied  
27 in evolutionary biology as it raises multiple, unresolved questions (Ironsides, 2010, Yan et al., 2020, Hartmann,  
28 Ament-Velásquez, et al., 2021, Kratochvíl and Stöck, 2021, Jay et al., 2022). A striking feature of many sex  
29 and mating-type chromosomes is the absence of recombination in large regions around the sex-determining genes.  
30 Recombination suppression indeed evolved in various groups of plants and animals in several steps beyond the  
31 sex-determining genes, generating evolutionary strata of differentiation between sex chromosomes (Nicolas et al.,  
32 2004, Bergero and Charlesworth, 2009, Hartmann, Duhamel, et al., 2021, Kratochvíl and Stöck, 2021). The reasons  
33 for the gradual expansion of recombination cessation beyond sex-determining genes remain debated (Ironsides, 2010,  
34 Wright et al., 2016, Ponnikas et al., 2018, Hartmann, Duhamel, et al., 2021). Recombination suppression has  
35 extended progressively with time not only on many sex chromosomes but also on mating-type chromosomes in fungi  
36 (Hartmann, Duhamel, et al., 2021) and other supergenes (Yan et al., 2020, Jay et al., 2021).

37 The main hypothesis to explain such stepwise extension of recombination cessation on sex chromosomes has  
38 long been sexual antagonism (Charlesworth et al., 2005, Bergero and Charlesworth, 2009). Theoretical studies have  
39 indeed shown that the suppression of recombination may evolve to link alleles that are beneficial in only one sex to  
40 the sex-determining genes (Rice, 1987, Charlesworth et al., 2005, Ruzicka et al., 2020). However, this hypothesis  
41 has received little evidence from empirical studies despite decades of research (Ironsides, 2010, Dagilis et al., 2022).  
42 Moreover, the sexual antagonism hypothesis cannot explain the evolutionary strata found on fungal mating-type  
43 chromosomes. Indeed, in many fungi, two gametes can form a new individual only if they carry different mating  
44 types, but there is no sexual antagonism or other form of antagonistic selection between cells of opposite mating  
45 types; the cells of different mating types do not show contrasted phenotypes or footprints of diversifying selection  
46 (Bazzicalupo et al., 2019). Yet, evolutionary strata have been documented on the mating-type chromosomes of  
47 multiple fungi, with recombination suppression extending stepwise beyond mating-type determining genes (Fraser  
48 et al., 2004, Menkis et al., 2008, Branco et al., 2017, Branco et al., 2018, Hartmann, Duhamel, et al., 2021, Hartmann,  
49 Ament-Velásquez, et al., 2021, Vittorelli et al., 2022). Evolutionary strata have also been reported around other  
50 supergenes, *i.e.*, large genomic regions encompassing multiple genes linked by recombination suppression, such as  
51 in ants and butterflies (Yan et al., 2020, Jay et al., 2021). Several hypotheses alternative to sexual antagonism  
52 have been proposed and explored to explain the stepwise extension of recombination suppression on sex-related  
53 chromosomes (Ironsides, 2010, Hartmann, Duhamel, et al., 2021). Theoretical models suggested that recombination  
54 suppression could be induced by a divergence increase in regions in linkage disequilibrium with a sex-determining  
55 locus (Jeffries et al., 2021) or that inversions could be stabilized by dosage compensation on asymmetric XY-like  
56 sex chromosomes (Lenormand and Roze, 2022).

57 A promising, widely applicable hypothesis is the sheltering of deleterious alleles by inversions carrying a lower  
58 load than average in the population (Charlesworth and Wall, 1999, Antonovics and Abrams, 2004, Hartmann,  
59 Duhamel, et al., 2021, Jay et al., 2022). Inversions (or any suppressor of recombination in *cis*) can indeed behave as

60 overdominant: inversions with fewer recessive deleterious mutations than average are initially beneficial and increase  
61 in frequency, but can then occur in a homozygous state where they express their load, unless they are linked to a  
62 permanently heterozygous allele. In this case, they remain advantageous, and can reach fixation in the sex-related  
63 chromosome on which they appeared (Jay et al., 2022). The suppression of recombination is thereby selected  
64 for, and recessive deleterious mutations are permanently sheltered. The process can occur repeatedly, leading to  
65 evolutionary strata. Importantly, this is one of the few hypotheses able to explain the existence of evolutionary  
66 strata on fungal mating-type chromosomes and it can apply to any supergene with a permanently heterozygous  
67 allele (Llaurens et al., 2017, Jay et al., 2022).

68 A key point for the recombination suppressor to invade is that it must appear in populations where recessive  
69 deleterious mutations segregate near the mating-compatibility genes (Olito et al., 2022, Jay et al., 2022). We  
70 therefore need to understand whether such mutations can persist in the vicinity of permanently heterozygous alleles  
71 (such as those occurring at mating-type loci) and under what conditions. In particular, it is usually considered that  
72 selfing purges deleterious mutations (Glémin, 2007, Abu Awad and Billiard, 2017), while most evolutionary strata  
73 on fungal mating-type chromosomes have been reported in selfing (automictic) fungi (Branco et al., 2017, Branco  
74 et al., 2018, Hartmann, Ament-Velásquez, et al., 2021, Vittorelli et al., 2022). Indeed, because mating types are  
75 determined at the haploid stage in fungi, mating types do not prevent selfing when considering diploid individuals  
76 (Billiard et al., 2012). Some particular forms of selfing associated with a permanently heterozygous mating-type  
77 locus such as intra-tetrad mating (*i.e.* automixis, mating among gametes from the same meiosis) can however  
78 favor the maintenance of heterozygosity (Hood and Antonovics, 2000). Indeed, mating can only occur between  
79 haploid cells carrying different mating-type alleles, which maintains heterozygosity at the mating-type locus, and to  
80 some extent at flanking regions, thereby possibly sheltering deleterious alleles. We therefore need to study whether  
81 deleterious or overdominant mutations can be maintained near mating-type compatibility loci, even under selfing, to  
82 assess whether the mechanism of sheltering deleterious mutations can drive extensions of recombination suppression.

83 The dynamics of deleterious mutation frequencies in genomes have been extensively studied independently of the  
84 presence of a permanently heterozygous locus. Deterministic models and diffusion approximations have been used  
85 to study the dynamics of deleterious mutations in a one locus-two allele setup (Kimura, 1980, Ewens, 2004, Rice,  
86 2004), with the addition of sexual reproduction and in particular selfing (Ohta and Cockerham, 1974, Caballero  
87 and Hill, 1992, Abu Awad and Roze, 2018). Extensions of these models exist to cover the two locus-two allele  
88 case (Karlin, 1975) and multilocus systems (reviewed in Bürger, 2020), or to take stochastic fluctuations into  
89 account (Coron et al., 2013, Coron, 2014). However, the dynamics of deleterious mutations in genomic regions  
90 near a permanently heterozygous allele have been little studied. A deterministic model showed that a lethal allele  
91 can be sheltered in an outcrossing population only when it is completely linked to a self-incompatibility locus  
92 (Leach et al., 1986). Another deterministic model introduced selfing and showed with simulations that a lethal  
93 allele can be sheltered when it is completely linked to a mating-type allele, favored in a heterozygote state, and if  
94 there is intra-tetrad selfing (Antonovics et al., 1998). Assuming a variable recombination rate between the two loci,  
95 Antonovics and Abrams, 2004 showed that an overdominant allele lethal in a homozygous state could be maintained  
96 if recombination was twice as low as the selection for heterozygotes and mating occurred via intra-tetrad selfing.  
97 Stochastic simulations additionally showed that a recessive deleterious allele could be maintained completely linked  
98 to a self-incompatibility allele, especially when it is highly recessive, and when the number of self-incompatibility

99 alleles in the population is large (Llaurens et al., 2009), and that codominant weakly deleterious alleles could be  
100 maintained near loci under balancing selection in the major histocompatibility complex (MHC) in humans (Lenz  
101 et al., 2016).

102 Here, building on the work of Antonovics and Abrams, 2004, we use a similar though simplified two locus-two  
103 allele framework, taking into account the non-negligible reproductive stochasticity during the early stage of the  
104 dynamics of the mutant subpopulation, until it becomes extinct or reaches some appreciable fraction of the total  
105 population. More precisely, we consider a permanently heterozygous mating-type locus and a genetic load locus,  
106 and we assume that the recombination rate between the two loci is a fixed parameter. Individuals can reproduce  
107 via outcrossing, or via either one of two types of selfing, intra-tetrad mating or inter-tetrad mating. The two types  
108 of selfing depend on whether a given gamete mates with another gamete produced during the same meiosis event  
109 (within a tetrad) or with a gamete from a different meiosis (from another tetrad, App. 6). The distinction is  
110 important because intra-tetrad mating maintains more heterozygosity in some genomic regions than inter-tetrad  
111 mating (Hood and Antonovics, 2000). Starting with a continuous-time Moran process, we derive the rates at  
112 which individuals of each genotype are produced. Then, as a new mutation is carried by very few individuals at  
113 the beginning of its evolution, a branching process naturally arises. Indeed, in this initial phase two individuals  
114 carrying the mutant allele have an extremely low probability to mate with each other. Mutant-carrier individuals  
115 can thus be assumed to reproduce independently of each other, leading to an approximation of the dynamics of the  
116 subpopulation of mutant carriers by a branching process.

117 The use of branching processes has shown its utility to account for the dynamics of a newly arised mutant allele  
118 in a population. Many estimates of the fixation or purging time of mutants in stochastic models (Champagnat  
119 and Méléard, 2011, Collet et al., 2013) relied on the use of branching processes to approximate the dynamics of  
120 a newly appeared mutant allele and of a nearly-fixed one. A branching-process approximation was used to study  
121 a two locus-two allele model, with individual fitness depending on the allelic state at both loci (Ewens, 1967,  
122 Ewens, 1968). For the diploid case, the framework of a seven-type branching process that can be used to study the  
123 fate of a deleterious mutation has been described, without deriving any analytical result (Pollard, 1966, Pollard,  
124 1968). A similar branching process approximation was used to study the fate of a beneficial mutation with selfing  
125 (Pollak, 1987, Pollak and Sabran, 1992). Here, we use a similar framework but consider deleterious mutations  
126 and a permanently heterozygous locus. Modeling multiple loci suggests the use of multitype branching processes,  
127 which have been widely studied (Harris, 1964, Kesten and Stigum, 1966, Mode, 1971, Athreya and Ney, 1972,  
128 Sewastjanow, 1975, Pénisson, 2010). However, the multiplicity of types renders the derivation of analytical results  
129 on probabilities of extinction and on extinction times difficult (Heinzmann, 2009). We therefore use an analytical  
130 approach to study the probability that a new mutation is purged from the population, and a numerical approach to  
131 study the purging time (when purging occurs) to assess how long a deleterious or overdominant mutation remains in  
132 a population. We study in particular the impact of the mating system and of the level of linkage to a permanently  
133 heterozygous locus on the long-term maintenance of deleterious mutations near a fungal-like mating-type locus (*i.e.*  
134 not preventing diploid selfing).

## 135 2 Methods and Models

136 All parameters which will be needed below are listed in App. 5.

## 137 2.1 Population and stochastic dynamics

138 We consider diploid (or dikaryotic) individuals, represented by their mating-type chromosomes, that harbor two  
139 biallelic loci: one mating-type locus, with alleles  $A$  and  $a$ , and one load locus, with a wild allele  $B$  and a mutant allele  
140  $b$ . We model a fungal-like mating-type locus, so that mating is only possible between haploid cells carrying different  
141 alleles at the mating-type locus (this does not prevent diploid selfing as each diploid individual is heterozygous at  
142 the mating-type locus). Consequently, only four genotypes are admissible, denoted by  $G_1, \dots, G_4$  in Figure 1. We  
143 follow the evolution of  $(g(t))_{t \geq 0} = (g_1(t), \dots, g_4(t))_{t \geq 0}$ , where  $g_i(t)$  is the number of individuals of genotype  $G_i$  in  
144 the population at time  $t$ . We suppose that the reproduction dynamics is given by a biparental Moran model with  
145 selection. In this continuous-time model, a single individual is replaced successively and the total population size,  
146 denoted by  $N$ , remains constant. A change in the population state  $g$  occurs in three steps.

147 The first step is the production of an offspring. After a random time following an exponential law of parameter  
148  $N$ , an individual is chosen uniformly at random to reproduce. This means in particular that all individuals have  
149 the same probability to reproduce. Mathematically speaking, this formulation is equivalent to saying that each  
150 individual reproduces at rate 1. The chosen diploid individual produces haploid gametes, via meiosis, during which  
151 recombination takes place between the two loci with probability  $r$  (see Figure 1 (a)). The product of a meiosis  
152 is a tetrad that contains four haploid gametes (Figure 1 (b)). Mating can then occur through three modalities,  
153 illustrated in App. 6 (recall that two gametes can fuse only if they carry different mating-type alleles): (i) Intra-  
154 tetrad selfing, with probability  $f p_{in}$ : the two gametes are picked from the same tetrad, only one parent is involved;  
155 (ii) Inter-tetrad selfing, with probability  $f(1 - p_{in})$ : the two gametes are picked from two different tetrads produced  
156 by the same individual, only one parent is involved; (iii) Outcrossing, with probability  $1 - f$ : the two gametes  
157 are picked from tetrads produced by two different parents. In this case, the second parent is chosen uniformly at  
158 random in the remaining population, and produces haploid gametes via meiosis with the same recombination rate  
159  $r$ . An offspring is produced following the chosen mating system, its genotype thus depending on the genotypes of  
160 the parents involved and on the occurrence of a recombination event in the tetrads.

161 The second step is the offspring survival. We assume that the fitness of a genotype  $G_i$  is the probability that an  
162 offspring with that genotype survives, and we denote it by  $S_i$ . We consider two selection scenarios (Figure 1, left):  
163 (i) The partial dominance case, where the mutant allele  $b$  is always deleterious and recessive. Homozygotes  $bb$  and  
164 heterozygotes  $Bb$  at the load locus have fitness values (*i.e.* a probability of survival) of  $1 - s$  and  $1 - hs$ , respectively.  
165 Homozygotes  $BB$  have fitness 1; (ii) The overdominance case, where heterozygotes  $Bb$  are favored over  $BB$  and  $bb$   
166 individuals. In this case, the fitness of  $Bb$ ,  $bb$  and  $BB$  juveniles are respectively 1,  $1 - s_3$  and  $1 - s_4$ , with  $s_3 > s_4$   
167 so that the fitness of  $bb$  individuals is lower than the fitness of wild-type individuals  $BB$ . The mating-type locus is  
168 considered neutral regarding survival.

169 The third step occurs if the offspring survives, in which case an individual chosen uniformly at random in  
170 the extant population is chosen to die and to be replaced by the offspring. If the offspring does not survive, the  
171 population state  $(g_1, g_2, g_3, g_4)$  does not change.

172 A jump in the stochastic process is thus an increase by one of the number of genotype  $G_i$  individuals in the  
173 population, when an offspring of genotype  $G_i$  is produced and survives, and a concomitant decrease by one of the  
174 number of genotype  $G_j$  individuals in the population, when an adult of genotype  $G_j$  dies. If  $i = j$ , *i.e.* if the  
175 surviving offspring and the individual chosen to die have the same genotype, the composition of the population

176 does not change. We denote the jump rate from  $g$  to  $g + e_i - e_j$  by  $Q_{i,j}(g)$ , where  $e_i$  is the vector with a 1 in  
177 position  $i$  and zeros everywhere else.  $Q_{i,j}(g)$  is equal to the product of the rate at which an offspring of genotype  
178  $G_i$  is produced (first step), which we denote by  $T_g(+G_i)$ , of the probability that it survives ( $S_i$ , second step), and  
179 of the probability that the adult chosen to die is of genotype  $G_j$  (third step). Thus, we have

$$Q_{i,j}(g) = T_g(+G_i) \times S_i \times \frac{g_j}{N}.$$

180 The total rates at which individuals of different genotypes are produced are given in App. 7.1. For example, the  
181 rate at which an offspring of genotype  $G_1$  is produced when the current state of the population is  $g = (g_1, g_2, g_3, g_4)$   
182 is given by

$$\begin{aligned} T_g(+G_1) = & f g_1 \left( 1 - r \left( 1 - \frac{1}{4} p_{in} \right) + \frac{1}{4} (1 - p_{in}) r^2 \right) + f g_2 \frac{r}{4} (p_{in} + r (1 - p_{in})) \\ & + \frac{1-f}{N-1} \left[ g_1 \left( 1 - \frac{r}{2} \right) \left( (g_1 - 1) \left( 1 - \frac{r}{2} \right) + g_3 + g_4 \right) + g_2 r \left( g_2 \frac{r}{4} + \frac{1}{2} (g_3 + g_4) \right) \right. \\ & \left. + g_1 g_2 r \left( 1 - \frac{r}{2} \right) + g_3 g_4 \right]. \end{aligned}$$

183 The first two terms on the right-hand side, with a factor  $f$ , correspond to reproduction events by selfing.  
184 The third term, with a factor  $1 - f$ , corresponds to reproduction events by outcrossing. Each subterm then  
185 encompasses the rate at which each genotype is involved in the reproduction event, and the probability that the  
186 offspring produced is of genotype  $G_1$ , taking into account possible recombinations. For example, the subterm  
187  $(1 - f)/(N - 1) \times g_1(g_1 - 1)(1 - r/2)^2$  is the product of the total rate  $g_1 \times 1$  at which an individual of genotype 1  
188 reproduces, of the probability  $1 - f$  that reproduction happens by outcrossing, of the probability  $(g_1 - 1)/(N - 1)$   
189 that the second parent is chosen among the other individuals of genotype  $G_1$ , and of the probability  $(1 - r/2)^2$  that  
190 their offspring has genotype  $G_1$ .

## 191 2.2 Branching-Process approximation

192 Let us now consider that the population size  $N$  is very large. When a mutation appears at the load locus, it is carried  
193 by a single individual. Hence, during the initial phase of the dynamics of the mutation  $b$ , the number of individuals  
194 who carry the mutation remains small compared to the number of wild-type individuals. The number of wild-type  
195 individuals is of the same order of magnitude as the total population size  $N$ , and the number of mutation-carrier  
196 individuals is negligible. More precisely, we assume that, when  $N$  is large,

$$g_4 \approx N, \quad \text{and } g_i \ll N \quad \text{for } i = 1, 2, 3. \quad (1)$$

197 Under this assumption, the jump rates  $Q_{i,j}(g)$  of the process can be approximated by neglecting the terms of  
198 the form  $1/(N - 1) \times g_i \times g_j$ , with  $i, j \in \{1, 2, 3\}$ , as they are of order  $1/N$ . This means that mating by outcrossing  
199 between individuals carrying the mutation  $b$  can be neglected. As a consequence, the birth rates of the different  
200 genotypes are linear in  $g_i$ , and a reproduction law for each genotype that is independent of the number of individuals  
201 of all other mutant-carrier genotypes can be derived. The Moran process can then be approximated by a branching

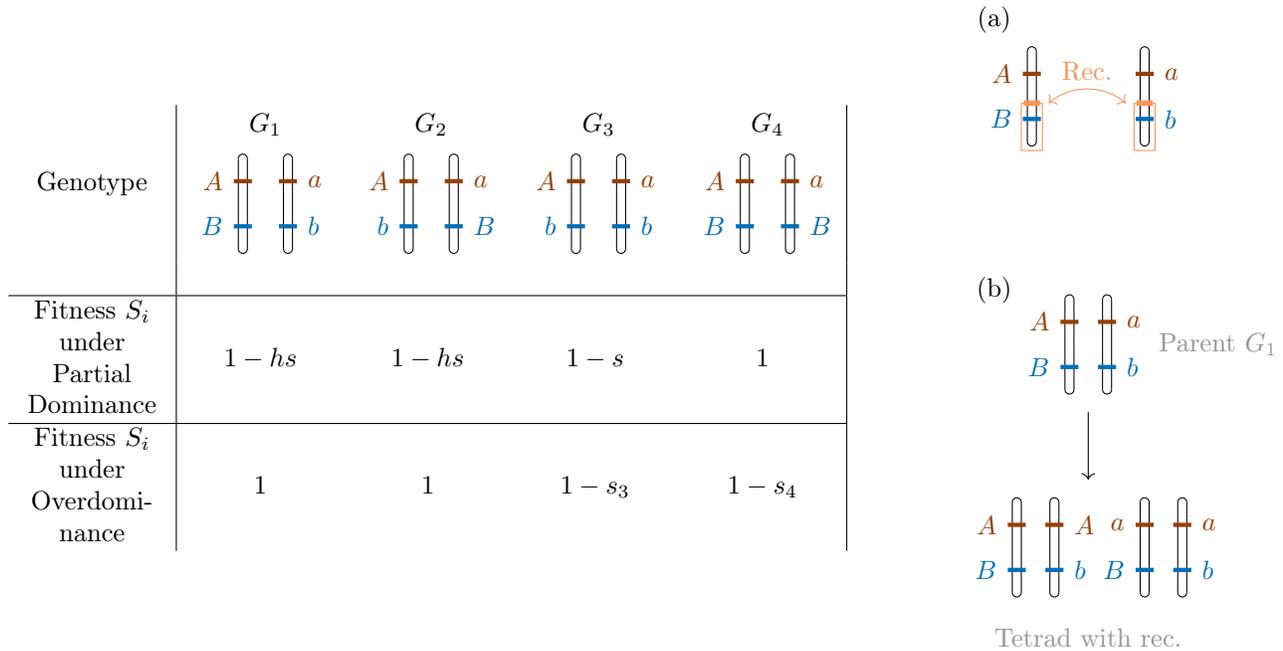


Figure 1: Schematic drawings of the genotypes considered and their parameters. (Left) Description of the possible genotypes in the population and their fitness  $S_i$  for the two selection scenarii considered (partial dominance and overdominance). (Right) (a) Position of a putative event of recombination between the mating-type locus and the load locus. (b) Example of a tetrad that can be obtained after a meiosis of an individual of genotype  $G_1$ , with recombination. Four gametes are produced, two of each mating type. In the second and third gamete from the left, combinations of alleles that did not exist in the parent are observed ( $A$  with  $b$  and  $a$  with  $B$ ).

202 process that follows the change in genotype counts for the mutation-carrier genotypes only.

203 We denote this branching process by  $(Z_t)_{t \geq 0}$ , where for each  $t \geq 0$ , we have  $Z_t = (Z_{t,1}, Z_{t,2}, Z_{t,3})$ , with  $Z_{t,i}$  the  
204 number of individuals of genotype  $G_i$  in the population at time  $t$ . To each genotype is associated a reproduction law,  
205 that is, a probability distribution on  $\mathbb{N}^3$  (vectors with three integer-valued coordinates) that gives the probability  
206 for an individual of that genotype to produce a given number of descendants of each genotype when it reproduces.  
207 Note that the rationale behind the branching process is different from the one for the Moran process. Indeed,  
208 each *replacement event* in the Moran model that involves an individual carrying the mutant allele  $b$  will be seen  
209 in the branching process as a *reproduction event*, in which the *offspring* is the mutant individual that is possibly  
210 produced during the first step of the Moran jump, and the *parent* is another mutant individual that is either one  
211 of the two actual parents in the replacement event, or the individual chosen to be replaced by the offspring in the  
212 Moran replacement event. A *reproduction event* of the branching process consists in the replacement of the parent  
213 by its *descendants*, which will be made of the mutant *offspring* when there is one, and of the mutant *parent* when  
214 it remains in the population. More precisely, we will encode three situations as follows: (i) when the replacement  
215 event in the Moran model corresponds to the reproduction of an individual of genotype  $G_i$ ,  $i \in \{1, 2, 3\}$  (via selfing  
216 or outcrossing with an individual of genotype  $G_4$ ), that this reproduction event generates a mutant offspring of  
217 genotype  $G_j$ ,  $j \in \{1, 2, 3\}$ , and the mutant parent is not chosen to die, we will see the *reproduction event* of the  
218 branching process as being an individual of genotype  $G_i$  having descendance vector  $e_i + e_j$ ; (ii) When the Moran  
219 replacement event leads to the reproduction of an individual of genotype  $G_i$ ,  $i \in \{1, 2, 3\}$  (via selfing or outcrossing  
220 with an individual of genotype  $G_4$ ), that this reproduction event generates an offspring of genotype  $G_4$ , and the  
221 mutant parent is not chosen to die, we will see the *reproduction event* as being an individual of genotype  $G_i$  having

222 descendance vector  $e_i$  (as non-mutant individuals are not accounted for in the branching process approximation).  
 223 Note that this reproduction event will imply no change in the population state, but for the sake of completeness  
 224 we indicate here all Moran replacement events that have non-vanishing rates as  $N$  tends to infinity; (iii) When  
 225 the Moran replacement event only involves non-mutant parents and an individual of genotype  $G_i$ ,  $i \in \{1, 2, 3\}$ , is  
 226 chosen to die, we will see the *reproduction event* as being an individual of genotype  $G_i$  having descendance vector 0  
 227 (corresponding to the *parent* being removed from the branching process and no mutant offspring being produced).  
 228 Other possible Moran replacement events occur at rates that vanish as  $N$  tends to infinity, and therefore do not  
 229 contribute to the *reproduction events* of the branching process. The rates at which *reproduction events* described  
 230 above occur are directly derived from the rates  $Q_{i,j}(g)$  of the Moran model, under the approximation stated in  
 231 Eq.(1). They are summarized in the matrices  $A$ ,  $T$ , and  $D$  defined as follows:

$$A = \begin{pmatrix} (fa(r) + (1-f)d(r))S_1 & (fc(r) + (1-f)\frac{r}{2})S_1 & (1-f)S_1 \\ (fc(r) + (1-f)\frac{r}{2})S_2 & (fa(r) + (1-f)d(r))S_2 & (1-f)S_2 \\ fb(r)S_3 & fb(r)S_3 & fS_3 \end{pmatrix},$$

$$T = \begin{pmatrix} (fb(r) + \frac{1-f}{2})S_4 & 0 & 0 \\ 0 & (fb(r) + \frac{1-f}{2})S_4 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$D = \begin{pmatrix} S_4 & 0 & 0 \\ 0 & S_4 & 0 \\ 0 & 0 & S_4 \end{pmatrix},$$

232 with

$$a(r) = 1 - r + \frac{r}{4}(1 - (1 - p_{in})(1 - r)), \quad b(r) = \frac{r}{4}(1 + (1 - p_{in})(1 - r)), \quad (2)$$

233

$$c(r) = \frac{r}{4}(1 - (1 - p_{in})(1 - r)), \quad \text{and} \quad d(r) = 1 - \frac{r}{2}. \quad (3)$$

234 The entries  $A_{ij}$  of matrix  $A$ ,  $T_{ij}$  of matrix  $T$  and  $D_{jj}$  of matrix  $D$  give the rates at which each individual of  
 235 genotype  $j$  reproduces and gives rise to a descendance vector respectively equal to  $e_i + e_j$  (situation (i)),  $e_i$  (situation  
 236 (ii)), and 0 (situation (iii)). An example of derivation of the matrix coefficients is given in App. 7.2.

### 237 2.3 Probability of purge and purging time

238 Under the assumption that the mutation is initially rare (after a mutation or migration event for example), we can  
 239 use the branching process approximation described in Section 2.2 to derive the probability and purging time of the  
 240 mutation from the population. In particular, our goal is to analyze the effect of the presence of a mating-type locus  
 241 near the load locus on the purge of the deleterious mutant  $b$ , *i.e.* on the extinction of the mutant-carrier population  
 242 described by the branching process.

#### 243 *Extinction Probability*

244 The probability of extinction of the branching process can be determined by looking at the eigenvalues of the  
 245 matrix  $C$  such that  $\mathbb{E}[Z_t|Z_0 = z_0] = z_0 e^{Ct}$  for  $t \geq 0$ , where  $z_0 \in \mathbb{N}^3$  is the initial state of the branching process  
 246  $(Z_t)_{t \geq 0}$  (Sewastjanow, 1975 in German, and Pénisson, 2010 for a statement of these results in English). Under the  
 247 assumption of irreducibility of the matrix  $C$ , results relying on the theory of Perron-Froebenius (see for example  
 248 Athreya and Ney, 1972) state that the process almost surely dies out (*i.e.* the mutation is purged with probability  
 249 1) if and only if  $\rho$ , the maximum eigenvalue of  $C$ , satisfies  $\rho \leq 0$ . When  $C$  is not irreducible, which occurs for  
 250 example if  $f = 0$  or  $f = 1$ , the result still holds but requires the use of the theory of final classes (Sewastjanow,  
 251 1975, cited in Pénisson, 2010). Details are given in App. 7.4.

252 We follow a method described in Bacaër, 2018, to compute the matrix  $C$  mentioned above and obtain

$$C_{ij} = \begin{cases} A_{ij} + T_{ij} & \text{if } i \neq j, \\ A_{jj} - \sum_{k \neq j} T_{kj} - D_{jj} & \text{if } i = j. \end{cases}$$

253 This gives

$$C = \begin{pmatrix} (fa(r) + (1-f)d(r))S_1 - S_4 & (fc(r) + (1-f)\frac{r}{2})S_1 & (1-f)S_1 \\ (fc(r) + (1-f)\frac{r}{2})S_2 & (fa(r) + (1-f)d(r))S_2 - S_4 & (1-f)S_2 \\ fb(r)S_3 & fb(r)S_3 & fS_3 - S_4 \end{pmatrix},$$

254 where the functions  $a, b, c, d$  were defined in Eqs. 2 and 3 (see details in App. 7.3).

255 We derived the dominant eigenvalue using Mathematica (Wolfram Research, 2015) and study its sign analytically  
 256 when possible, or numerically otherwise.

### 257 *Comparison with previous results*

258 Our results can be compared to the work of Ewens, 1967, who used a similar framework to study a random-mating  
 259 population with two biallelic loci under selection, one of which carried a new allele. Assuming that the frequency  
 260 of the gametes that carried a new allele was negligible compared to the frequencies of wild-type gametes, he used  
 261 a branching process approximation to study the probability that the new allele was purged from the population.  
 262 He considered a recombination rate  $R$  between the two loci, and fitnesses  $w_{ij}$  for each genotype (where  $i$  and  $j$   
 263 take the value 1 or 3 when loci are homozygous, and the value 2 when heterozygous). Setting  $w_{i1} = w_{i3} = 0$  for  
 264  $i = 1, 2, 3$  allows to force heterozygosity at the locus that does not carry the new allele in his model, and to compare  
 265 his findings with our results on the fate of a new allele appearing near a permanently heterozygous locus. The  
 266 dominant eigenvalue of the matrix driving the dynamics of the new allele in Ewens, 1967, is

$$\lambda_1 = \frac{w_{22}}{w_{32}} \quad (4)$$

267 with  $w_{22}$  being the fitness of individuals heterozygous for the new allele, and  $w_{32}$  the fitness of homozygous wild-  
 268 type individuals. As Ewens, 1967, considered a discrete-time branching process, this dominant eigenvalue must be  
 269 compared to one to deduce information on the new allele survival probability.

271 We investigate now to the potential effect of the presence of a mating-type locus on the maintenance of a mutant  
 272 allele in a population: as mating-type alleles are always heterozygous, any mutation appearing completely linked  
 273 to one mating-type allele is maintained in a heterozygous state as well. The load of the mutant allele is then less  
 274 expressed when the mutation is recessive, and the mutation is said to be "sheltered".

275 This potential *sheltering effect* can be explored by looking at the variation of the dominant eigenvalue  $\rho$  when  
 276 the recombination rate  $r$  is close to 0.5. Indeed, the quantity  $|\rho|$  can be seen as the rate of decay of the deleterious  
 277 mutant subpopulation (see the results on the probability of survival of a multitype branching process, Th. 3.1  
 278 of Heinzmann, 2009), and its value gives a rough approximation of the inverse of the mean time to extinction of  
 279 this subpopulation, *i.e.* of the mean purging time of the mutant allele  $b$ . Moreover, setting the recombination  
 280 rate to  $r = 0.5$  in our model allows us to consider a load locus completely unlinked to the mating-type locus, while  
 281 decreasing the value of  $r$  introduces some loose linkage between the two loci. We thus look at the derivative  $\frac{\partial \rho}{\partial r} |_{r=0.5}$   
 282 to obtain the variation of the dominant eigenvalue of  $C$  when departing from this unlinked state.

283 The sign of the derivative gives information on the existence of a sheltering effect due to the mating-type locus:  
 284 if  $\frac{\partial \rho}{\partial r} |_{r=0.5} < 0$ , then when  $r$  decreases from 0.5 to lower values, *i.e.* when linkage between the two loci appears,  
 285 the (negative) value of  $\rho$  increases, which means that the purging of the mutation becomes slower. In this case, the  
 286 mating-type locus has a sheltering effect. Otherwise, if  $\frac{\partial \rho}{\partial r} |_{r=0.5} > 0$ , the presence of a mating-type locus accelerates  
 287 the purging of a deleterious allele.

288 The absolute value of the derivative also gives information on the strength of the sheltering effect of the mating-  
 289 type locus. The closer to 0 the derivative is, the smaller the impact of the mating-type locus. We compute the  
 290 derivative and study its sign analytically. We then study the values of the derivative numerically in order to identify  
 291 the impact of each parameter on the sheltering effect of the mating-type locus.

292 We also look at the strength of the sheltering effect on mutations close to the mating-type locus, by studying  
 293 the eigenvalue variation around  $r = 0$ . Setting the recombination rate to  $r = 0$  models a situation where the load  
 294 locus is completely linked to the mating-type locus. Hence, the mutation is completely linked to one mating-type  
 295 allele, and maintained in a heterozygous state. Looking at the derivative  $\frac{\partial \rho}{\partial r} |_{r=0}$  allows us to quantify the impact  
 296 of departing from this situation by loosening the linkage between the two loci. We study the difference between the  
 297 derivative at  $r = 0.5$  and the derivative at  $r = 0$  to compare the effect of adding a small amount of linkage between  
 298 completely unlinked loci ( $r = 0.5$ ) and the effect of adding a small amount of recombination between completely  
 299 linked loci ( $r = 0$ ).

300 *Extinction time*

301 The mean time to extinction in a multitype branching process is finite for a subcritical process (that is, when  
 302 the principal eigenvalue  $\rho$  of  $C$  is less than 0), and infinite for a critical process (*i.e.* when  $\rho = 0$ , see Pötscher, 1985,  
 303 for the proof of existence and finiteness of extinction time moments). Previous work, in particular Theorem 4.2 in  
 304 Heinzmann, 2009, showed that a Gumbel law gives a good approximation of the law of the extinction time, provided  
 305 that the initial number of individuals in the branching process and the absolute value of the dominant eigenvalue are  
 306 both large. In our case, however, the mutation appears in a single individual, and the dominant eigenvalue is close

307 to zero, which prevents the use of the Gumbel law approximation. Therefore, we performed computer simulations  
308 to study the empirical distribution of the time to extinction of the process, *i.e.* the purging time of the  $b$  mutant  
309 allele.

310 The branching process was simulated with a Gillespie algorithm to obtain an empirical distribution for the time  
311 to extinction. More precisely, the Gillespie algorithm produces realizations of the stochastic process by iteratively  
312 updating the number of individuals of each genotype within the multitype branching process (Gillespie, 1976). To  
313 circumvent the problem of exponential increase of the population size in the supercritical case, the parameters were  
314 chosen so that the branching process was subcritical. The probability of extinction was thus equal to 1 and the  
315 mean time to extinction was finite. For each scenario, we looked at different values of the recombination rate  $r$ ,  
316 in order to study the impact of linkage between the load locus and the mating-type locus on the purging time of  
317 the mutant allele. We also chose different values for the selfing rate  $f$  in order to assess the impact of the mating  
318 system on the purging time of the mutant allele. For each set of parameters, 100,000 independent simulation runs  
319 were performed with the same initial condition (a single individual heterozygous at the load locus was introduced).

320 *Probability of a new mutation apparition before the first one is purged*

321 As a first step towards the study of the accumulation of deleterious mutations near a mating-type locus, we  
322 studied the probability that the deleterious mutation can be maintained long enough in the population so that a  
323 second mutation can appear before the first one is purged. We considered that a second mutation could appear  
324 during a reproduction event occurring in the population of mutation carriers (described by the branching process),  
325 on a region of a given length  $d = 10^6$  base pairs, at a rate of  $\mu = 10^{-8}$  mutations per base pairs per reproduction  
326 event. The mean number of reproduction events needed for a new mutation to appear in a region of length  $d$ ,  $\bar{n}_{ev}$ ,  
327 is the inverse of the mutation rate  $\mu$  multiplied by the length  $d$ :

$$\bar{n}_{ev} = \frac{1}{\mu \times d} = 10^2.$$

328 We then estimated the probability that a new mutation appears in such a genomic region before the first one is  
329 purged by counting the number of independent simulations in which the number of reproduction events exceeded  
330  $\bar{n}_{ev}$  before the branching process went extinct (*i.e.* before the purging of the first mutation), over 100,000 simulation  
331 runs. Note that we did not take into account the genotype of the individual on which the second mutation appears,  
332 and therefore we did not distinguish whether the second mutation appears on a chromosome that carries the first one  
333 or not. Our estimate thus does not exactly equals the probability to have two mutations on the same chromosome,  
334 but this gives an order of magnitude of the probability of deleterious mutation accumulation and of the impact of the  
335 mating system. The length of the genomic region on which a second mutation can appear was chosen arbitrarily, and  
336 changing it can also change the probability. However, the important point for the deleterious-mutation mechanism  
337 to work is that there exists a size for regions flanking mating-type loci that allows both inversions to appear and  
338 mutations to accumulate, so that inversions can trap several deleterious mutations when suppressing recombination.  
339 The value  $d = 10^6$  chosen here allows to cover such flanking regions.

340 We computed our estimate of the probability of deleterious mutations accumulation for  $r = 0.001$  (the two loci  
341 are close, strongly linked),  $r = 0.01$ ,  $r = 0.1$ , and  $r = 0.5$  (the two loci are distant, unlinked). We considered several

342 values of selfing and intra-tetrad mating rates  $f$  and  $p_{in}$  in order to assess the impact of the mating system on the  
343 probability of deleterious mutation accumulation near a mating-type locus.

### 344 3 Results

#### 345 3.1 Deleterious mutations are almost surely purged in the partial dominance case, 346 and can escape purge in the overdominance case

##### 347 *Partial Dominance scenario*

348 Under partial dominance, we find that the dominant eigenvalue  $\rho$  of the matrix  $C$  is always negative or null (see  
349 App. 8.2.1 and 8.2.2 for more details on the proof and computations). Previous theoretical results on branching  
350 processes state that, when  $\rho < 0$ , the probability that the deleterious mutation is purged from the population  
351 before it reaches a substantial frequency is one, and the mean time of purging is finite (see the Methods section).  
352 In particular, the probability of purging does not depend on the mating system ( $\rho < 0$  for any value of intratetrad,  
353 intertetrad and outcrossing rates), nor on the recombination probability, selection and dominance coefficients. The  
354 only exceptions are when the deleterious mutation is neutral ( $s = 0$ ) or behaves as neutral ( $h = 0$  and  $r = 0$ ,  
355 the mutation is neutral when heterozygous and completely linked to one mating-type allele), in which case the  
356 dominant eigenvalue is 0. The mutation is still purged from the population but previous theoretical results on  
357 branching processes state that this can take a much longer time compared to the case where  $\rho < 0$ , as the mean  
358 purging time would be infinite (see the Methods section).

359 Taking  $w_{22} = 1 - hs$  and  $w_{32} = 1$  in the model of Ewens, 1967, to mirror our partial dominance scenario, the  
360 dominant eigenvalue becomes  $1 - hs$ . It is always smaller than one, except when  $h = 0$  or  $s = 0$ , *i.e.* when the  
361 mutation is neutral in the heterozygous state. Except in those cases, the mutation is purged from the population  
362 with probability one. We therefore find the same results as Ewens, 1967, and we extend these results in the case  
363 where mating is not random among gametes. In particular, the mutation being neutral in the heterozygous case  
364 ( $h = 0$ ) is not sufficient to prevent the purging probability to be one when mating is not random: the mutation has  
365 to be completely linked to a permanently heterozygous locus ( $h = 0$  and  $r = 0$ ).

##### 366 *Overdominance scenario*

367 Under overdominance, the dominant eigenvalue  $\rho$  can take positive or negative values. When  $\rho$  is positive, the  
368 probability that the mutation escapes purging and that the number of mutation-carriers increases exponentially  
369 fast is strictly positive. The general conditions on the parameters for  $\rho$  to be positive in our model are given in  
370 App. 8.3.2, but they are difficult to interpret. Below, we describe a few simple cases in order to elucidate the role  
371 of each parameter, and then we complement the analysis with a numerical approach.

372 Similarly to the partial dominance case, the dominant eigenvalue is 0 when the mutation is neutral ( $s_3 = 0$ ,  
373 which implies  $s_4 = 0$  as well). The dynamics of the  $b$ -subpopulation (*i.e.* mutation-carriers) is then critical, which  
374 means that the mutant is purged with probability 1 but the mean purging time can be arbitrarily long (as the  
375 average extinction time of a critical branching process is infinite, see the Methods section).

376 When the mutation is not neutral ( $s_3 \neq 0$ ) but with no disadvantage to  $BB$  homozygotes ( $s_4 = 0$ ), we prove  
 377 that  $\rho < 0$  (see App. 8.3.2), which means that the dynamics of the  $b$  subpopulation is subcritical and that the  
 378 mutant allele is purged with probability 1. This shows that the overdominant mutant allele is not maintained in  
 379 the population when wild-type homozygotes are not disfavored compared to heterozygotes at the load locus. This  
 380 corresponds to a completely recessive mutation, and is in agreement with the results for the partial dominance case  
 381 with  $h = 0$ .

382 When the mutant allele is completely linked to a mating-type allele ( $r = 0$ ), or under complete outcrossing  
 383 ( $f = 0$ ), the dominant eigenvalue is equal to  $s_4$ , the selection coefficient for the fitness reduction of the  $BB$  wild-  
 384 type homozygotes. The dynamics of the  $b$  subpopulation is then supercritical, which means that there is a non-zero  
 385 probability that the mutant allele is not purged and, instead, reaches a significant number of carriers. Moreover,  
 386 the mutant allele is more favored in this case when selection against  $BB$  homozygotes is stronger as it induces a  
 387 stronger advantage of the  $Bb$  heterozygotes. A similar result can be derived from the work of Ewens, 1967. Taking  
 388  $w_{22} = 1$  and  $w_{32} = 1 - s_4$  in his model to mirror our overdominance scenario, the dominant eigenvalue of Eq. 4  
 389 becomes  $1/(1 - s_4)$ . As long as  $s_4 > 0$ , this eigenvalue is always greater than one, and its value increases as the  
 390 selection against wild-type homozygotes increases. This shows that the dynamics of an overdominant allele under  
 391 random gamete mating is similar as under complete outcrossing.

392 In the case of complete intra-tetrad selfing ( $f = 1, p_{in} = 1$ ), we find that  $\rho \geq 0$  if  $r \leq 2s_4$ , in agreement with the  
 393 results of Antonovics and Abrams, 2004. These results mean that the overdominant mutation can be maintained  
 394 under complete selfing if it is tightly linked to the mating-type locus ( $r$  small) or if the heterozygote advantage over  
 395 wild-type homozygotes is strong ( $s_4$  large).

396 In the case of complete selfing ( $f = 1$ ), we find that  $\rho = s_4 - s_3 \leq 0$  when  $r(2 - r - p_{in}(1 - r)) - 2s_3 \geq 0$ .  
 397 This shows that the dominant eigenvalue depends only on the selection coefficients when the recombination rate  $r$   
 398 exceeds a certain threshold (visible on the bottom panels of Figure 2). This means that, if the recombination rate is  
 399 larger than the strength of the selection against deleterious homozygotes, the mutation is purged with probability  
 400 one. Moreover, the purging time is shorter when the difference in fitness between the two homozygotes is larger.  
 401 The threshold on recombination increases as  $p_{in}$  increases, which means that the strength of the linkage between  
 402 the mating-type locus and the mutation has the highest effect under intra-tetrad selfing.

403 Figure 2 shows more generally that the mating system affects the purging of deleterious mutations. On Figure 2,  
 404 the probability of purging is one in blue areas (the dominant eigenvalue is negative), and positive but smaller than  
 405 one in red areas (the dominant eigenvalue is positive). The lines below which the mutation has a non-zero survival  
 406 probability under the framework of Antonovics and Abrams, 2004, *i.e.*  $r = 2s_4$  under complete intra-tetrad selfing,  
 407 are displayed as well. Comparing the panels for different values of intratetrad mating rate ( $p_{in}$ ) and selfing rate ( $f$ )  
 408 shows that selfing favors the purging of the mutant allele (the blue area becomes larger as  $f$  increases), whereas  
 409 intratetrad mating favors the maintenance of the deleterious allele (the blue area become smaller as  $p_{in}$  increases).  
 410 Indeed, selfing favors the creation of homozygous individuals, which are disfavored, and intra-tetrad selfing favors  
 411 the creation of heterozygous individuals, which are favored, compared to inter-tetrad selfing: the probability that  
 412 a heterozygous individual  $Bb$  produces a heterozygous offspring  $Bb$  is higher under intra-tetrad selfing (probability  
 413  $1 - r/2$ ) than under inter-tetrad selfing (probability  $1 - r + r^2/2$ ).

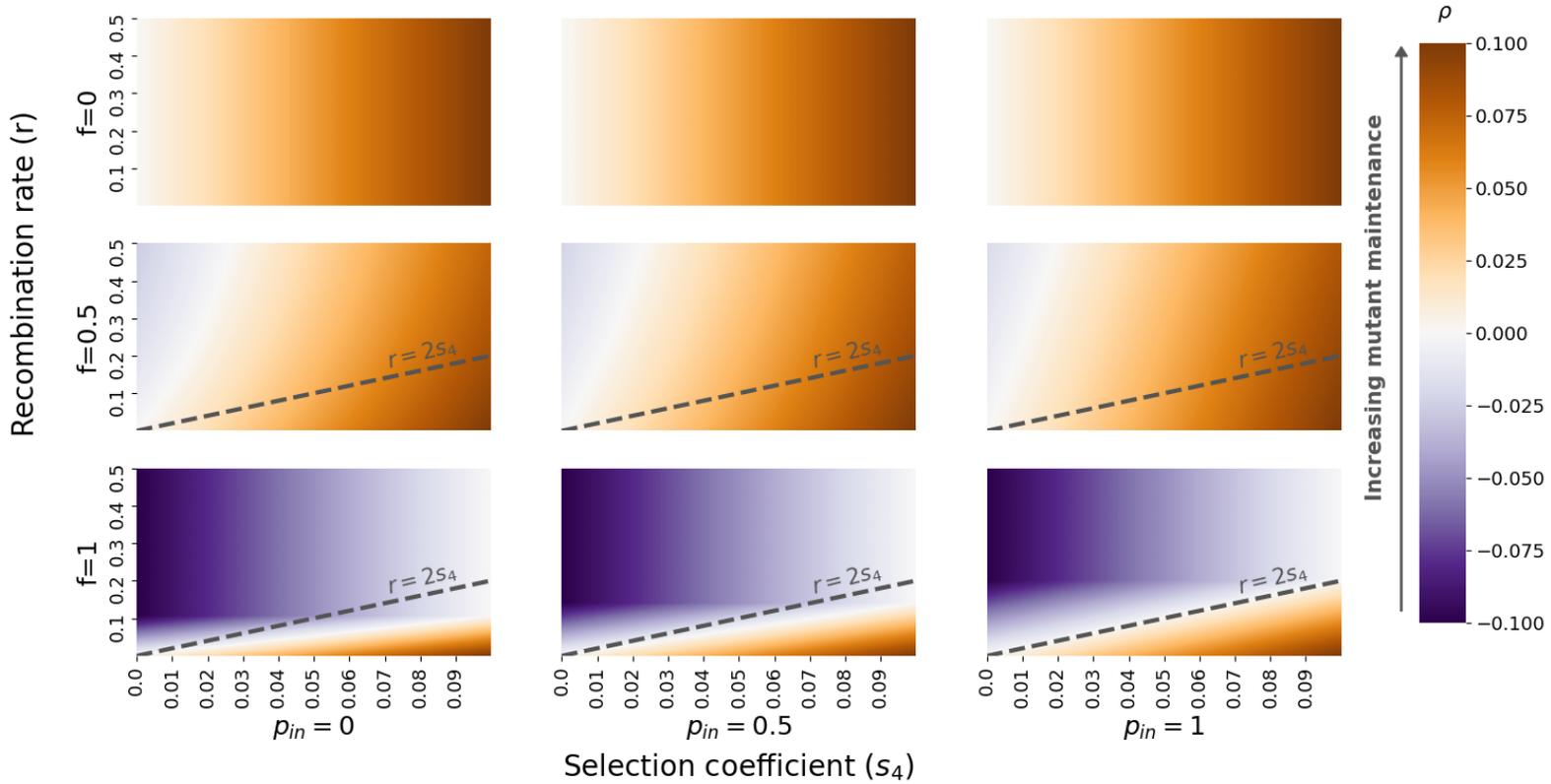


Figure 2: Dominant eigenvalue  $\rho$  for the overdominance scenario. When  $\rho \leq 0$  (blue areas), the mutation is purged with probability 1. When  $\rho > 0$  (red areas), the mutation has a non-zero probability to escape purging. The mutation is maintained longer in the population as  $\rho$  increases. All panels have the same axes. x-axis:  $s_4$ , selection coefficient for wild-type  $BB$  homozygotes. y-axis:  $r$ , recombination rate between the two loci. Each column corresponds to a value of  $p_{in}$  (intra-tetrad rate, 0, 0.5, 1), and each row to a value of  $f$  (selfing rate, 0, 0.5, 1). The selection coefficient for  $bb$  homozygotes is set to  $s_3 = 0.1$ . The line  $r = 2s_4$  is displayed for comparison with the findings in Antonovics and Abrams, 2004.

### 3.2 The presence of a mating-type locus has a sheltering effect under partial selfing

Looking at the derivative of the dominant eigenvalue at  $r = 0.5$ , we find that the presence of a mating-type locus near the mutation has a sheltering effect on the deleterious mutation, under partial selfing and in both selection scenarii. Indeed, the derivative  $\frac{\partial \rho}{\partial r} |_{r=0.5}$  is always negative, except when the mutation is neutral ( $s = 0$  under partial dominance or  $s_3 = 0$  under overdominance), when it is lethal ( $s = 1$ ) or dominant ( $h = 1$ ) under partial dominance, or under complete outcrossing ( $f = 0$ ) in both scenarii, in which cases the derivative is zero and there is no sheltering effect. Under complete selfing ( $f = 1$ ), the derivative is also null when the intratetrad coefficient  $p_{in}$  is below a certain threshold (see App 8.2.3 and 8.3.3 for the proof). As explained in the Methods section, this analysis shows that, in a wide range of situations, the rate of decay of the mutant subpopulation is lower when the mutation is linked to a mating-type locus, even loosely (*i.e.* as soon as  $r < 0.5$ ), than when recombination is free between the two loci. Hence, except for the particular cases cited above, the mating-type locus always has a sheltering effect on the deleterious mutation maintenance under partial selfing, independently of the mating system coefficients ( $f$  and  $p_{in}$ ) and of the selection and dominance coefficients ( $s$  and  $h$ , or  $s_3$  and  $s_4$ ).

Figure 3 shows that, under both partial dominance or overdominance, the variation of the derivative at  $r = 0.5$  is stronger when the selfing rate  $f$  (x-axis) or the intratetrad selfing probability  $p_{in}$  (y-axis) are high. This means that the sheltering effect of the mating-type locus is stronger under high selfing or high intratetrad mating. Two

430 forces oppose here: increasing selfing induces a greater production of homozygotes, which are disfavored, whereas  
 431 increasing intra-tetrad selfing rate or increasing the linkage with a mating-type locus favors the production of  
 432 heterozygotes, which are favored. The sheltering effect of the mating-type locus that counters the purging effect of  
 433 selfing is higher when selfing is higher, and this countering effect is reinforced by a high intra-tetrad mating rate.  
 434 Moreover, when approaching  $f = 1$ , the derivative decreases to 0. Indeed, the selection and dominance coefficients  
 435  $s$ ,  $s_3$  and  $h$  are here sufficiently small for the condition to have  $\frac{\partial \rho}{\partial r} \Big|_{r=0.5} = 0$  when  $f = 1$  to be met, for both  
 436 selection scenarii (see App. 8.2.3 and 8.3.3 for the derivation of this condition). This means that the dynamics of  
 437 the deleterious mutation is independent of the presence of a mating-type locus under complete selfing and weak  
 438 selection.

439 We explore the impact of other parameters in the Supplementary materials. Figure S1 shows that, under partial  
 440 dominance, the sheltering effect of a mating-type locus is stronger when the dominance coefficient  $h$  is lower ( $Bb$   
 441 heterozygotes, which are more prone to be created in the presence of a mating-type locus, are more favored) or  
 442 when the selection coefficient  $s$  is high (the differential in fitness between  $Bb$  heterozygotes and  $bb$  homozygotes  
 443 is higher). Similarly, Figure S2 shows that, under overdominance, the sheltering effect of the mating-type locus  
 444 is stronger when the selection against  $bb$  homozygotes is higher ( $s_3$  coefficient), whereas the selection against  $BB$   
 445 homozygotes does not impact the strength of the sheltering effect, suggesting that the dynamics of the deleterious  
 446 allele is mostly driven by the difference in fitness between the favored heterozygotes and the disfavored deleterious  
 447 homozygotes.

448 Looking at the derivative at  $r = 0$ , we show in App. 8.2.3 and App. 8.3.3 that it is also negative in both selection  
 449 scenarii. This means that the eigenvalue decreases, *i.e.* that the mutation is less maintained in the population as soon  
 450 as the two loci are no longer completely linked. Figure S3 shows that the difference  $\Delta \left( \frac{\partial \rho}{\partial r} \right) = \frac{\partial \rho}{\partial r} \Big|_{r=0.5} - \frac{\partial \rho}{\partial r} \Big|_{r=0}$   
 451 is always positive, which means that the absolute value of the derivative at  $r = 0$  is larger than the absolute value  
 452 of the derivative at  $r = 0.5$ . This shows that the sheltering effect is stronger on mutations closely linked to the  
 453 mating-type locus : adding a small chance of recombination on previously completely linked loci ( $r = 0$ ) has a  
 454 greater impact on the maintenance of deleterious mutations than adding a small amount of linkage between two  
 455 previously completely unlinked loci ( $r = 0.5$ ). The largest difference between the two derivatives occurs for selfing  
 456 rates close to one, the derivative being then zero at  $r = 0.5$ , while the derivative at  $r = 0$  approaches  $-1$ . This  
 457 shows that the linkage to the mating-type locus particularly impacts the strength of its sheltering effect under high  
 458 selfing.

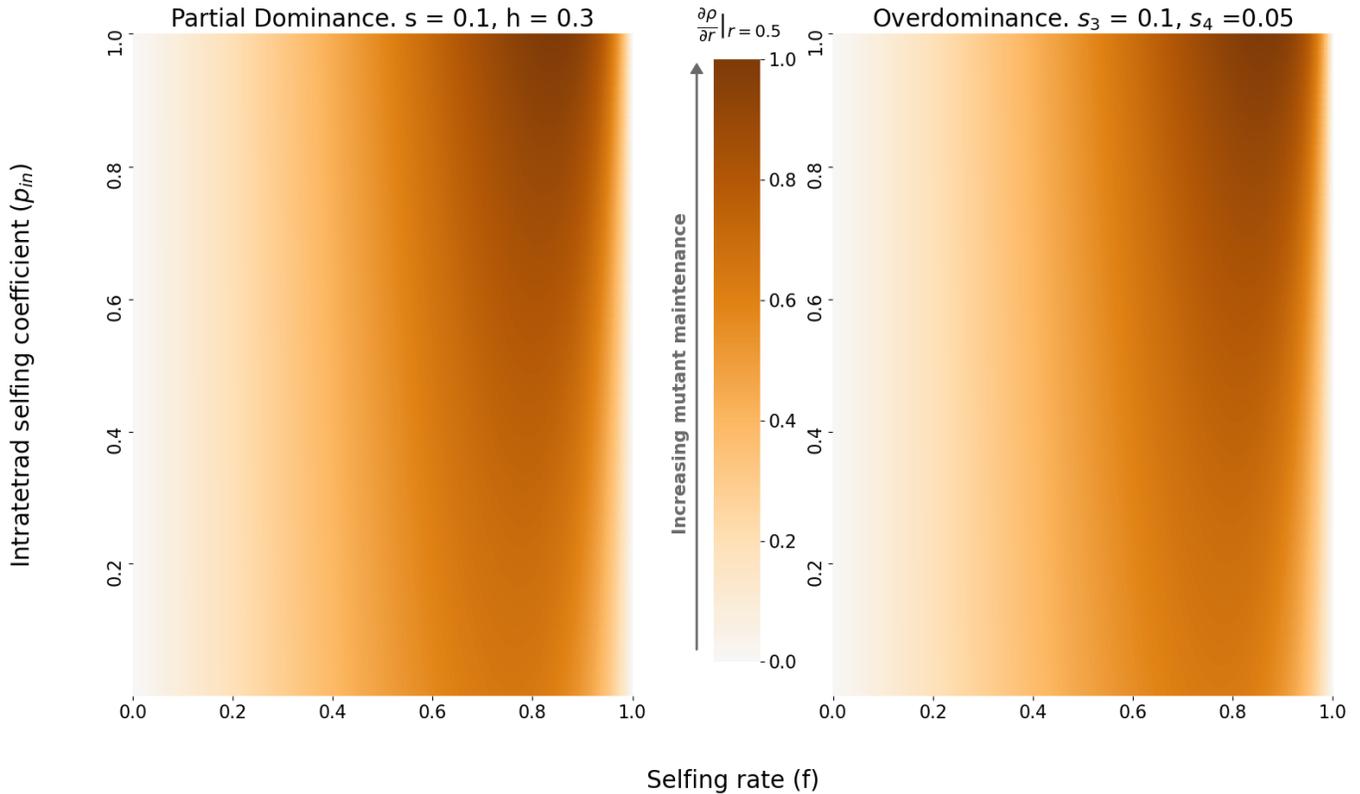


Figure 3: Relative variation of the derivative of the dominant eigenvalue in the partial dominance case (left) and the overdominance case (right). For each panel, the values of  $\frac{\partial \rho}{\partial r}|_{r=0.5}$  range from a minimal value, which is negative, to zero. We divided each value of the derivative by this minimum in order to plot values between 0 and 1 for every panel. This enables us to compare the effect of the presence of a mating-type locus on the same scale for both selection scenarii. x-axis: selfing rate  $f$ . y-axis: intratetrad selfing rate  $p_{in}$ . The darker the color, the more the mating-type locus shelters the mutation, thus promoting its maintenance.

### 3.3 Rare events of maintenance of the deleterious mutation occur in both selection scenarii, paving the way for an accumulation of mutations

The empirical distribution of the purging time of the deleterious mutation in the partial dominance case is shown on Figure 4: for ca. 75% of the independent runs, the mutation was rapidly purged, while in some rare cases (ca. 1%), the purge took very long (several orders of magnitude longer than the 75% percentile empirically obtained from the 100,000 runs). Note that the approximation of the distribution of the time to extinction by a Gumbel law (Th. 4.1 of Heinzmann, 2009) falls short here, because the initial number of individuals (one) and the absolute value of  $\rho$  (given in the caption) are too small.

Consistently with our results that  $\frac{\partial \rho}{\partial r} < 0$ , the sheltering effect of the mating-type locus implies that the purging time increases when the recombination rate decreases (Figure 4, and Figure S5 for the overdominant case). We also consistently find that increasing selfing decreases the purging time (Figures S4 and S6). In each case, the closer  $\rho$  is to zero, the more extreme the rare events are : the distribution of the 1% longest purging times is stretched towards higher values when  $\rho$  gets closer to zero, while the distributions of the 75% shortest remain similar.

Figure 5 displays the probability that the mutation can be maintained long enough in the population for another mutation to appear in a region of  $10^6$  bp near the mating-type locus. This probability is nonnegligible (of the order of 1% to 10%), which shows that accumulation events are rare but still occur near mating-type loci. This is true even

475 under selfing as the sheltering effect of the mating-type locus can counter the purging effect of selfing. Indeed, when  
 476 the recombination rate between the first mutation and the mating-type locus is high ( $r = 0.5$  or  $r = 0.1$ ), modeling  
 477 a situation where the distance between the two loci is large, the probability that a second mutation appears before  
 478 the first one is purged decreases with increasing selfing, even with high intra-tetrad selfing rates. However, when the  
 479 first mutation is closer to the mating-type locus (lower recombination rates), the probability that a second mutation  
 480 appears before the first one is purged under selfing is similar to the probability under complete outcrossing. The  
 481 presence of a mating-type locus can thus facilitate the accumulation of deleterious mutations in its flanking regions,  
 482 especially in highly selfing populations.

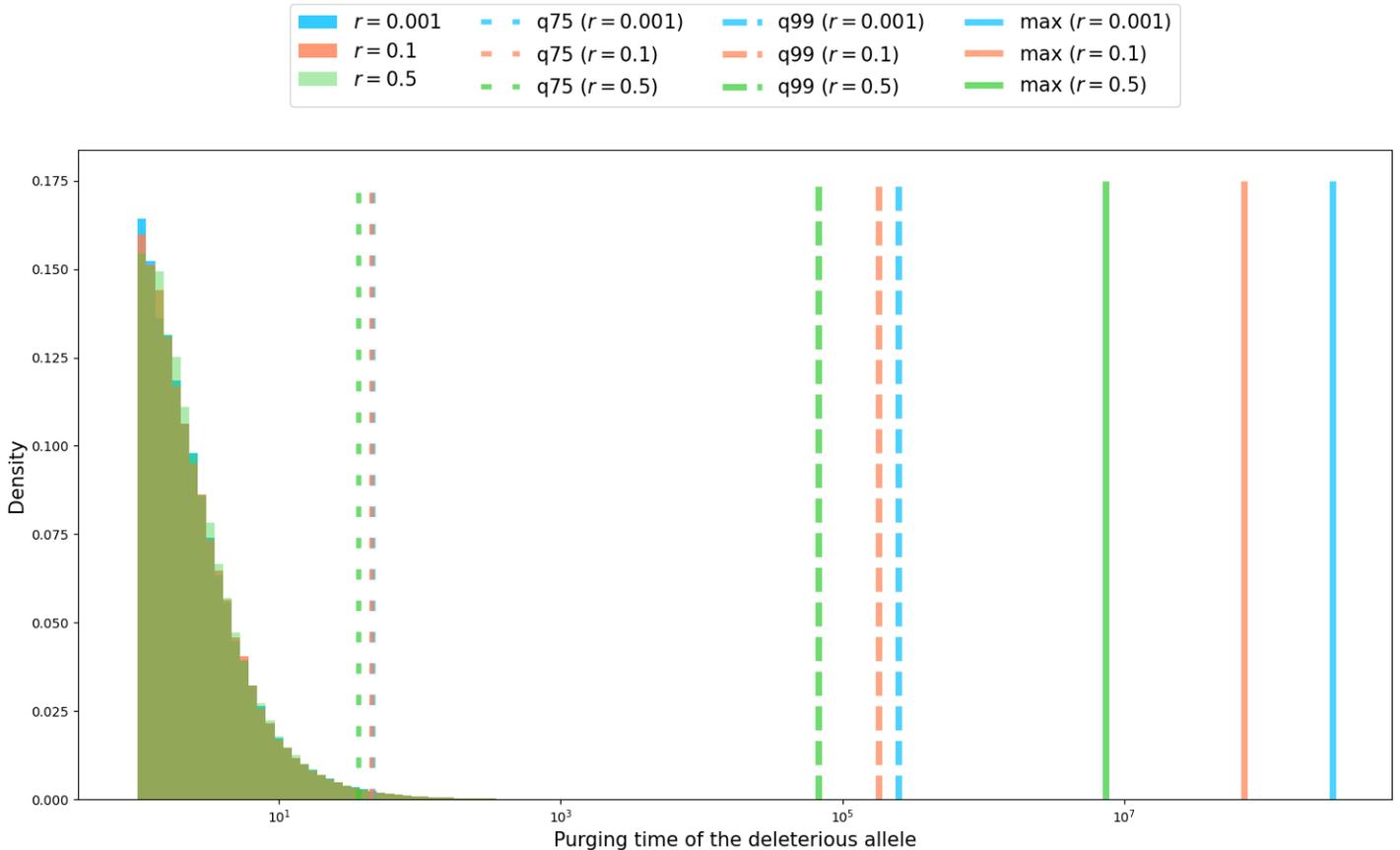


Figure 4: Empirical distribution of the deleterious allele purging time for the partial dominance scenario. A total of 100,000 simulations were run, with  $s = 0.1$ ,  $h = 0.1$ ,  $f = 0.5$ ,  $p_{in} = 0.5$ , starting from one heterozygous individual ( $X_0 = (1, 0, 0)$ ), and for three values of the recombination rate ( $r = 0.001$  in blue,  $r = 0.1$  in red and  $r = 0.5$  in green). The respective values for  $\rho$  are  $\rho = -0.0101$ ,  $\rho = -0.0106$  and  $\rho = -0.0307$ . The x-axis is log-scaled. The large-dotted lines represent the 75<sup>th</sup> percentile ( $q75$ ), the dashed lines indicate the 99<sup>th</sup> percentile ( $q99$ ), and solid lines the maximum value ( $max$ ) of the purging time. Maximum values are several order of magnitudes higher than the 75th percentile of the empirical distribution of the purging time.

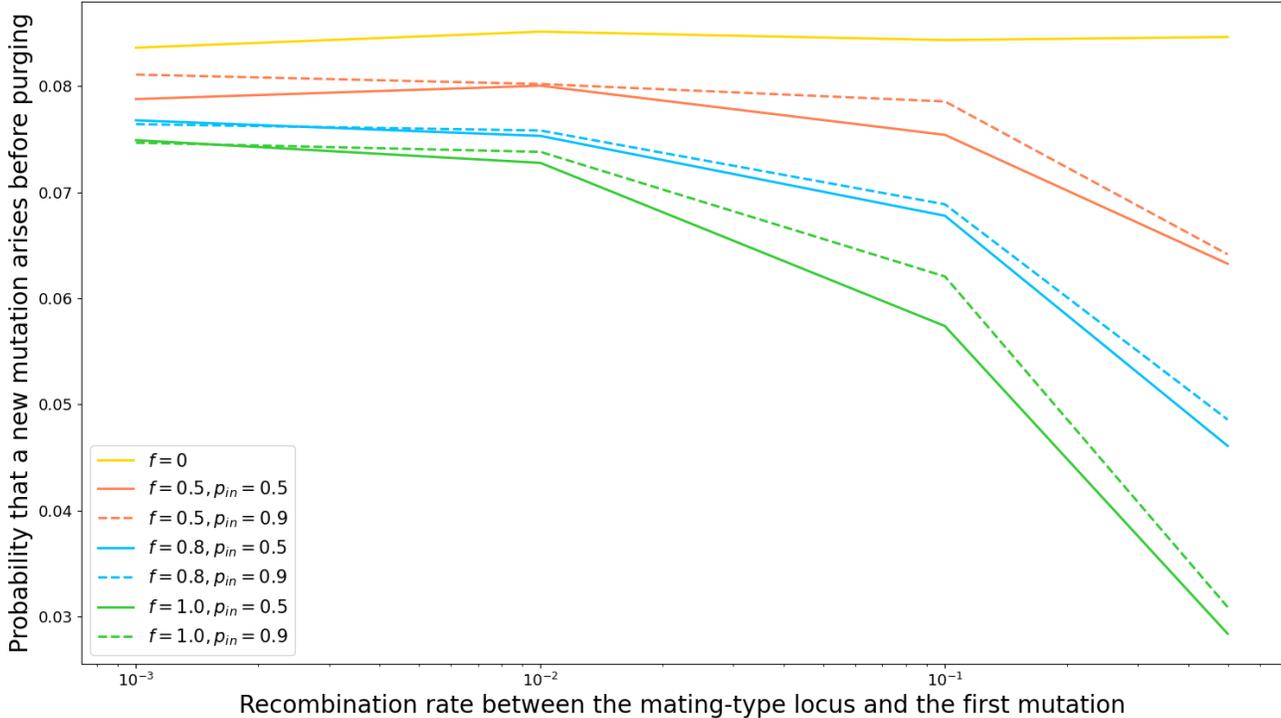


Figure 5: Probability that a new mutation appears in a region of length  $10^6$  bp before the first mutation is purged from the population, under the partial dominance scenario, depending on the recombination rate between the first mutation and the mating-type locus. We considered a mutation rate per base pair per reproduction event of  $10^{-8}$ . Here, the reproduction events are those of the branching process, that change the composition of the mutant-carriers subpopulation. The probability that a new mutation appears before the purge of the first one is approximated by the proportion of simulation runs for which the number of reproduction events exceeds the expected number of events needed for a new mutation to appear (see text). For each set of parameters  $(r, f, p_{in})$ , 100,000 independent simulations were run. Colors correspond to different values of the selfing rate  $f$ , and line styles to different values of the intra-tetrad selfing rate  $p_{in}$ . When  $f = 0$ , a single curve is displayed, as the value of  $p_{in}$  has no impact under complete outcrossing. For all simulations, we set  $s = 0.1$  and  $h = 0.1$ .

## 483 4 Discussion

484 *Partially recessive deleterious mutations are almost surely purged in finite time while overdominant mutations can*  
 485 *persist*

486 We have shown that partially recessive deleterious mutations close to a fungal-like mating-type locus (*i.e.* that  
 487 does not prevent diploid selfing) are almost surely purged in finite time, except when they are neutral or behave as  
 488 neutral. In the overdominance case, the probability of purge depends on parameter values. Low selfing rates, high  
 489 intra-tetrad selfing rates or tight linkage to the mating-type locus increases both the maintenance probability and  
 490 persistence of the overdominant allele, whereas a high selfing rate favors its purge.

491 In particular, if linkage is complete (corresponding to  $r = 0$  here, or to the case where the inversion encompasses  
 492 a permanently heterozygous locus in Jay et al., 2022), an overdominant allele may be maintained in a population  
 493 and even sweep to fixation with non-zero probability, which confirms previous findings (Antonovics et al., 1998,  
 494 Antonovics and Abrams, 2004, Jay et al., 2022). This means that, although selfing purges deleterious mutations, a

495 mating-type locus can have a sheltering effect in its flanking regions.

496 In general, the overdominant allele is maintained longer and with a higher probability in the population when the  
497 fitness advantage of heterozygotes over homozygotes is higher, in line with previous simulation results (Antonovics  
498 and Abrams, 2004). This conclusion is sensible: if the mutant is strongly favored in a heterozygous state, it can be  
499 maintained in this state in the population.

500 *The presence of the mating-type locus has a sheltering effect under selfing*

501 We found that, in both selection scenarii, the presence of the mating-type locus had no effect on the maintenance  
502 of deleterious mutations under outcrossing, but always had a sheltering effect under selfing, which strengthened as  
503 the selfing rate increased. Indeed, selfing increases homozygosity and thus accelerates the purge of a deleterious  
504 allele, whereas the presence of a permanently heterozygous mating-type locus induces more heterozygosity in its  
505 flanking regions, that counters the purging effect of selfing. The sheltering effect of a mating-type locus is thus all the  
506 more tangible as it counters the strong purging effect induced by selfing. Increasing intra-tetrad selfing also induces  
507 more heterozygosity and thus slightly reinforces the sheltering effect of the mating-type locus. This is consistent with  
508 the findings that, in fungi, ascomycetes that reproduce via outcrossing and live as haploids do not show evolutionary  
509 strata (Skinner et al., 1993, Zhong et al., 2002, Phan et al., 2003, Kuhn et al., 2006, Jin et al., 2007, Malkus et al.,  
510 2009) whereas pseudo-homothallic ascomycete fungi, living as dikaryotic and undergoing mostly intra-tetrad selfing,  
511 are those with evolutionary strata around their mating-type locus (Menkis et al., 2008, Hartmann, Duhamel, et al.,  
512 2021, Hartmann, Ament-Velásquez, et al., 2021, Vittorelli et al., 2022). In basidiomycetes also, the species with  
513 evolutionary strata are dikaryotic and automictic, e.g. *Microbotryum fungi* and *Agaricus bisporus* var. *bisporus*  
514 (Branco et al., 2017, Branco et al., 2018, Foulongne-Oriol et al., 2021). This may be explained by the fact that  
515 intra-tetrad selfing favors the accumulation of deleterious alleles near the mating-type locus, which in turn can  
516 promote selection for recombination suppression because there will be more variability in the number of mutations  
517 present in a genomic region close to the mating-type locus, and therefore more fragments having a much lower  
518 number of deleterious mutations than average in the population (Jay et al., 2022).

519 Additionally, we found that the sheltering effect of a mating-type locus was stronger when the mutation was  
520 more strongly recessive. Indeed, the purging effect of selfing on partially recessive mutations is stronger for more  
521 recessive mutations (Charlesworth and Charlesworth, 1987, Caballero and Hill, 1992, Arunkumar et al., 2015), in  
522 which case the opposite force of the sheltering effect of a mating-type locus is strengthened. This is in agreement  
523 with the results of studies on the sheltered load linked to a self-incompatibility locus, showing that completely  
524 recessive deleterious mutations are more easily fixed than partially recessive ones (Llaurens et al., 2009). This  
525 also confirms results on the fixation of inversions encompassing recessive deleterious mutations and linked to a  
526 permanently heterozygous locus (Olito et al., 2022, Jay et al., 2022). These results showed that inversions became  
527 fixed with a higher probability when segregating deleterious mutations were more strongly recessive.

528 *Rare events of long maintenance of deleterious mutations in the population can occur*

529 We further found that rare events of long maintenance of deleterious mutations in the population occurred  
530 under both selection scenarii. This shows that some deleterious mutations can persist in the population for an  
531 extended period of time before being purged, especially near the mating-type locus: in approximately 1% of our  
532 simulations, the purge of the deleterious mutation took several orders of magnitude longer than the 75% percentile

533 empirically obtained from the 100,000. These surprisingly long purging times are likely to be due to the dynamics  
534 of the mutant being almost critical (the dominant eigenvalue in the branching process approximation is negative,  
535 but close to zero). However, from a modeling perspective very little is currently known about these trajectories,  
536 and more generally about the extinction time of multitype branching processes. Studying the extinction time of a  
537 deleterious allele in a one locus-two allele setting with a unitype branching process approximation and a diffusion  
538 approximation showed that the standard deviation of the mean extinction time was higher than the mean itself (Nei,  
539 1971), which is a feature that was also found in our simulations of multitype branching processes. These results  
540 show that the extinction time of deleterious alleles is highly variable, producing long-lasting mutations that may  
541 induce an accumulation of deleterious alleles near a mating-type locus, which is a prerequisite for recombination  
542 suppression to extend away from this locus (Jay et al., 2022).

#### 543 *The dynamics of deleterious mutations heavily relies on the mating system*

544 Our results show that the mating system, and selfing in particular, is a prevailing force impacting the dynamics  
545 of deleterious mutations. Indeed, we found that a mating-type locus shelters mutations and thus favors their  
546 maintenance, but increasing selfing reduces the maintenance of mutations with a stronger effect. This result is  
547 congruent with previous studies showing that an increase in the selfing rate induces i) a reduction of the mutational  
548 load at a given locus or at multiple non-interacting loci far from mating-type compatibility loci (Charlesworth et al.,  
549 1990, for a deterministic model, Abu Awad and Roze, 2018, for diffusion approximation), and ii) a reduction of the  
550 purging time of deleterious mutations (Caballero and Hill, 1992).

551 However, we observed a particular behavior when the population reproduced only via selfing. Under complete  
552 selfing in our setting, the existence of a sheltering effect of a mating-type locus strongly depended on the values  
553 of the intra-tetrad selfing rate: the sheltering effect of the mating-type locus was detectable only when the intra-  
554 tetrad selfing coefficient exceeded a certain threshold, that depended on the dominance and selection coefficients.  
555 This strong effect of departing from complete selfing had previously been noted: introducing a small amount of  
556 outcrossing in a selfing population can lead to sharp changes in the dynamics of a deleterious mutation, whereas  
557 adding a small amount of selfing in an outcrossing population induces a smoother change (Holsinger and Feldman,  
558 1985).

#### 559 *Limits of the methods*

560 Our results are limited to the case of a single load locus, in interaction with a heterozygous mating-type  
561 locus, and may not apply when considering different frameworks, such as multiple epistatic loci or with additional  
562 beneficial mutations, especially regarding the impact of the mating system. Indeed, selfing has a non-monotonous  
563 effect depending on the tightness of linkage between multiple interacting loci (Abu Awad and Roze, 2018): at low  
564 selfing rates, increasing linkage between loci increases the mutation load, whereas the opposite effect is observed at  
565 high selfing rates. Selfing also has a non-monotonous effect on genetic variation in populations under stabilizing  
566 selection (Lande and Porcher, 2015, Clo and Opedal, 2021). In addition, selfing can enhance the fixation chances  
567 of a deleterious allele when it hitchhikes during a selective sweep (Hartfield and Otto, 2011, Hartfield and Glémin,  
568 2014). Moreover, the impact of the mating system on the maintenance of deleterious mutations may be different  
569 if the number of individuals carrying the mutant allele exceeds a certain threshold. In this case, the branching  
570 process approximation does not hold anymore, and a deterministic model in large population may be used to

571 further describe the dynamics of the deleterious allele (Durrett and Schweinsberg, 2004, Durrett, 2008 Section  
572 6.1.3). The impact of the mating system then remains unclear: in large populations, selfing reduces the effective  
573 population size, which impairs the efficiency of selection and increases the mutational load of the population, but it  
574 also bolsters homozygosity, which favors the purge of deleterious mutations (Pollak, 1987, Caballero and Hill, 1992,  
575 Charlesworth and Wright, 2001, Wright et al., 2008).

576 Another limitation of our approach is that we considered a fixed recombination rate for simplicity, but allowing  
577 this rate to vary would allow us to test whether recombination suppression could evolve. Such an outcome may  
578 depend on the strength of selection against the deleterious mutation, as well as on the mating system (Antonovics  
579 and Abrams, 2004, Abu Awad and Roze, 2018). In some previous models, the impact of a modifier of recombination  
580 in the form of a multi-allelic locus was studied by simulations, but no analytical results were obtained (Feldman,  
581 1972, Palsson, 2002, Antonovics and Abrams, 2004, Lenormand and Roze, 2022). The multitype branching process  
582 framework developed here would also be an interesting approach to obtain numerical results on this more complex  
583 situation, but analytical results would probably be out of reach because of the increase in complexity of the model.

### 584 *Conclusion and Perspectives*

585 In conclusion, our findings show that a mating-type locus has a sheltering effect on nearby deleterious mutations,  
586 especially in case of selfing and automixis, which can then play a role in the evolution of recombination suppression  
587 near mating-compatibility loci (Antonovics and Abrams, 2004, Jay et al., 2022). This may contribute to explain  
588 why evolutionary strata of recombination suppression near the mating-type locus are found mostly in automictic  
589 (pseudo-homothalic) fungi (Menkis et al., 2008, Branco et al., 2017, Branco et al., 2018, Hartmann et al., 2020,  
590 Hartmann, Ament-Velásquez, et al., 2021, Foulongne-Oriol et al., 2021, Vittorelli et al., 2022).

591 The results obtained here on the accumulation of deleterious mutations should apply, beyond fungal-like mating-  
592 type loci, to other permanently heterozygous loci, such as supergenes (Llaurens et al., 2017). In contrast, sporophytic  
593 or gametophytic plant self-incompatibility loci prevent diploid selfing, leading to a completely different evolutionary  
594 scenario in their flanking regions as imposed by complete outcrossing. The diversity of observed patterns regarding  
595 the presence or absence, length and number of evolutionary strata around these regions (Uyenoyama, 2005) may  
596 be explained, in addition to the mating system, by other factors controlling the long-term behavior of deleterious  
597 mutations which are not studied here, such as the number of alleles at supergenes, the length of the haploid  
598 phase (Jay et al., 2022), or the presence of multiple load loci that are possibly physically linked and with epistatic  
599 interactions (Abu Awad and Roze, 2018, Lenormand and Roze, 2022). The questions of the genome-wide impact of  
600 a mating-type locus, and of the interaction between a permanently heterozygous locus and background mutations,  
601 are currently debated (Abu Awad and Waller, 2021). The branching process framework developed here could be  
602 applied to diploid individuals carrying a load locus with two alleles, undergoing selfing or outcrossing, in order to  
603 investigate the dynamics of a new deleterious mutation in a population with or without a mating-type locus.

604 Our results showing the long maintenance of deleterious mutations in the vicinity of permanently heterozygous  
605 loci pave the way for future investigations on the accumulation of deleterious mutations. Previous studies (Coron  
606 et al., 2013, Coron, 2014) on mutational meltdown, showing that deleterious mutations accumulate faster when  
607 other mutations are already fixed, also encourage future work in this direction.

## 608 Acknowledgments

609 We thank Paul Jay and Denis Roze for insightful discussions, Aurélien Tellier for handling the recommendation  
610 process, and three anonymous reviewers for their useful and very constructive comments.

## 611 Funding and Conflict of Interest

612 This work was supported by the European Research Council (ERC) EvolSexChrom (832352) grant to TG. ET,  
613 SB and AV acknowledge support from the chaire program « Mathematical modeling and biodiversity » (Ecole  
614 Polytechnique, Museum National d'Histoire Naturelle, Veolia Environnement, Fondation X). The authors of this  
615 preprint declare that they have no financial conflict of interest with the content of this article. TG and SB are  
616 recommenders at PCIEvolBiol.

## 617 References

- 618 Abu Awad, D., & Billiard, S. (2017). The double edged sword: The demographic consequences of the evolution of  
619 self-fertilization: THE DOUBLE EDGED SWORD. *Evolution*, 71(5), 1178–1190. [https://doi.org/10.1111/  
620 evo.13222](https://doi.org/10.1111/evo.13222)
- 621 Abu Awad, D., & Roze, D. (2018). Effects of partial selfing on the equilibrium genetic variance, mutation load, and  
622 inbreeding depression under stabilizing selection. *Evolution*, 72(4), 751–769. [https://doi.org/10.1111/evo.  
623 13449](https://doi.org/10.1111/evo.13449)
- 624 Abu Awad, D., & Waller, D. (2021). Conditions for maintaining and eroding pseudo-overdominance and its contri-  
625 bution to inbreeding depression. <https://doi.org/10.1101/2021.12.16.473022>
- 626 Antonovics, J., O’Keefe, K., & Hood, M. E. (1998). Theoretical population genetics of mating-type linked haplo-  
627 lethal alleles. *International Journal of Plant Sciences*, 159(2), 192–198. <https://doi.org/10.1086/297538>
- 628 Antonovics, J., & Abrams, J. Y. (2004). Intratetrad mating and the evolution of linkage relationships. *Evolution*,  
629 58(4), 702–709. <https://doi.org/10.1111/j.0014-3820.2004.tb00403.x>
- 630 Arunkumar, R., Ness, R. W., Wright, S. I., & Barrett, S. C. H. (2015). The evolution of selfing is accompanied  
631 by reduced efficacy of selection and purging of deleterious mutations. *Genetics*, 199(3), 817–829. [https:  
632 //doi.org/10.1534/genetics.114.172809](https://doi.org/10.1534/genetics.114.172809)
- 633 Athreya, K. B., & Ney, P. E. (1972). *Branching processes*. Springer Berlin Heidelberg. [https://doi.org/10.1007/978-  
634 3-642-65371-1](https://doi.org/10.1007/978-3-642-65371-1)
- 635 Bacaër, N. (2018). Sur l’extinction des populations avec plusieurs types dans un environnement aléatoire. *Comptes  
636 Rendus Biologies*, 341(3), 145–151. <https://doi.org/10.1016/j.crv.2018.01.009>
- 637 Bazzicalupo, A. L., Carpentier, F., Otto, S. P., & Giraud, T. (2019). Little evidence of antagonistic selection in the  
638 evolutionary strata of fungal mating-type chromosomes (*Microbotryum lychnidis-dioicae*). *G3 Genes/Genomes/Genetics*,  
639 9(6), 1987–1998. <https://doi.org/10.1534/g3.119.400242>
- 640 Bergero, R., & Charlesworth, D. (2009). The evolution of restricted recombination in sex chromosomes. *Trends in  
641 Ecology & Evolution*, 24(2), 94–102. <https://doi.org/10.1016/j.tree.2008.09.010>

- 642 Billiard, S., López-Villavicencio, M., Hood, M. E., & Giraud, T. (2012). Sex, outcrossing and mating types: Unsolved  
643 questions in fungi and beyond: Sexy fungi. *Journal of Evolutionary Biology*, *25*(6), 1020–1038. <https://doi.org/10.1111/j.1420-9101.2012.02495.x>  
644
- 645 Branco, S., Badouin, H., Rodríguez de la Vega, R. C., Gouzy, J., Carpentier, F., Aguilera, G., Siguenza, S., Branden-  
646 burg, J.-T., Coelho, M. A., Hood, M. E., & Giraud, T. (2017). Evolutionary strata on young mating-type  
647 chromosomes despite the lack of sexual antagonism. *Proceedings of the National Academy of Sciences*,  
648 *114*(27), 7067–7072. <https://doi.org/10.1073/pnas.1701658114>
- 649 Branco, S., Carpentier, F., Rodríguez de la Vega, R. C., Badouin, H., Snirc, A., Le Prieur, S., Coelho, M. A.,  
650 de Vienne, D. M., Hartmann, F. E., Begerow, D., Hood, M. E., & Giraud, T. (2018). Multiple convergent  
651 supergene evolution events in mating-type chromosomes. *Nature Communications*, *9*(1), 2000. <https://doi.org/10.1038/s41467-018-04380-9>  
652
- 653 Bürger, R. (2020). Multilocus population-genetic theory. *Theoretical Population Biology*, *9*. <https://doi.org/https://doi.org/10.1016/j.tpb.2019.09.004>  
654
- 655 Caballero, A., & Hill, W. G. (1992). Effects of partial inbreeding on fixation rates and variation of mutant genes.  
656 *Genetics*, *131*(2), 493–507. <https://doi.org/10.1093/genetics/131.2.493>
- 657 Champagnat, N., & Méléard, S. (2011). Polymorphic evolution sequence and evolutionary branching. *Probability*  
658 *Theory and Related Fields*, *151*(1), 45–94. <https://doi.org/10.1007/s00440-010-0292-9>
- 659 Charlesworth, B., & Wall, J. D. (1999). Inbreeding, heterozygote advantage and the evolution of neo-X and neo-Y  
660 sex chromosomes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1414),  
661 51–56. <https://doi.org/10.1098/rspb.1999.0603>
- 662 Charlesworth, D., & Charlesworth, B. (1987). Inbreeding depression and its evolutionary consequences. *Annual*  
663 *Review of Ecology and Systematics*, *18*, 237–268.
- 664 Charlesworth, D., Morgan, M. T., & Charlesworth, B. (1990). Inbreeding depression, genetic load, and the evolution  
665 of outcrossing rates in a multilocus system with no linkage. *Evolution*, *44*(6), 1469–1489. <https://doi.org/10.1111/j.1558-5646.1990.tb03839.x>  
666
- 667 Charlesworth, D., & Wright, S. I. (2001). Breeding systems and genome evolution. *Current Opinion in Genetics &*  
668 *Development*, *11*(6), 685–690. [https://doi.org/10.1016/S0959-437X\(00\)00254-9](https://doi.org/10.1016/S0959-437X(00)00254-9)
- 669 Charlesworth, D., Charlesworth, B., & Marais, G. (2005). Steps in the evolution of heteromorphic sex chromosomes.  
670 *Heredity*, *95*(2), 118–128. <https://doi.org/10.1038/sj.hdy.6800697>
- 671 Clo, J., & Opedal, Ø. H. (2021). Genetics of quantitative traits with dominance under stabilizing and directional  
672 selection in partially selfing species. *Evolution*, *75*(8), 1920–1935. <https://doi.org/10.1111/evo.14304>
- 673 Collet, P., Méléard, S., & Metz, J. A. J. (2013). A rigorous model study of the adaptive dynamics of Mendelian  
674 diploids. *Journal of Mathematical Biology*, *67*(3), 569–607. <https://doi.org/10.1007/s00285-012-0562-5>
- 675 Coron, C., Méléard, S., Porcher, E., & Robert, A. (2013). Quantifying the mutational meltdown in diploid popula-  
676 tions. *The American Naturalist*, *181*(5), 623–636. <https://doi.org/10.1086/670022>
- 677 Coron, C. (2014). Stochastic modeling of density-dependent diploid populations and the extinction vortex. *Advances*  
678 *in Applied Probability*, *46*(2), 446–477. <https://doi.org/10.1239/aap/1401369702>

- 679 Dagilis, A. J., Sardell, J. M., Josephson, M. P., Su, Y., Kirkpatrick, M., & Peichel, C. L. (2022). Searching for  
680 signatures of sexually antagonistic selection on stickleback sex chromosomes. *Philosophical Transactions of*  
681 *the Royal Society B: Biological Sciences*, 377(1856), 20210205. <https://doi.org/10.1098/rstb.2021.0205>
- 682 Durrett, R., & Schweinsberg, J. (2004). Approximating selective sweeps. *Theoretical Population Biology*, 66(2),  
683 129–138. <https://doi.org/10.1016/j.tpb.2004.04.002>
- 684 Durrett, R. (2008). *Probability models for DNA sequence evolution*. Springer New York. [https://doi.org/10.1007/](https://doi.org/10.1007/978-0-387-78168-6)  
685 [978-0-387-78168-6](https://doi.org/10.1007/978-0-387-78168-6)
- 686 Ewens, W. J. (1967). The probability of fixation of a mutant : Two-locus case. *Evolution*, 21(3), 532–540. <https://doi.org/10.1111/j.1558-5646.1967.tb03409.x>
- 687 [//doi.org/10.1111/j.1558-5646.1967.tb03409.x](https://doi.org/10.1111/j.1558-5646.1967.tb03409.x)
- 688 Ewens, W. J. (1968). Some applications of multiple-type branching processes in population genetics. *Journal of*  
689 *the Royal Statistical Society: Series B (Methodological)*, 30(1), 164–175. [https://doi.org/10.1111/j.2517-](https://doi.org/10.1111/j.2517-6161.1968.tb01515.x)  
690 [6161.1968.tb01515.x](https://doi.org/10.1111/j.2517-6161.1968.tb01515.x)
- 691 Ewens, W. J. (2004). *Mathematical population genetics* (Vol. 27). Springer New York. [https://doi.org/10.1007/978-](https://doi.org/10.1007/978-0-387-21822-9)  
692 [0-387-21822-9](https://doi.org/10.1007/978-0-387-21822-9)
- 693 Feldman, M. W. (1972). Selection for linkage modification: I. Random mating populations. *Theoretical Population*  
694 *Biology*, 3(3), 324–346. [https://doi.org/10.1016/0040-5809\(72\)90007-X](https://doi.org/10.1016/0040-5809(72)90007-X)
- 695 Foulongne-Oriol, M., Taskent, O., Kües, U., Sonnenberg, A. S. M., van Peer, A. F., & Giraud, T. (2021). Mating-  
696 type locus organization and mating-type chromosome differentiation in the bipolar edible button mushroom  
697 *Agaricus bisporus*. *Genes*, 12(7), 1079. <https://doi.org/10.3390/genes12071079>
- 698 Fraser, J. A., Diezmann, S., Subaran, R. L., Allen, A., Lengeler, K. B., Dietrich, F. S., & Heitman, J. (2004).  
699 Convergent evolution of chromosomal sex-determining regions in the animal and fungal kingdoms (David  
700 Page, Ed.). *PLoS Biology*, 2(12), e384. <https://doi.org/10.1371/journal.pbio.0020384>
- 701 Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled  
702 chemical reactions. *Journal of Computational Physics*, 22(4), 403–434. [https://doi.org/10.1016/0021-](https://doi.org/10.1016/0021-9991(76)90041-3)  
703 [9991\(76\)90041-3](https://doi.org/10.1016/0021-9991(76)90041-3)
- 704 Glémin, S. (2007). Mating systems and the efficacy of selection at the molecular level. *Genetics*, 177(2), 905–916.  
705 <https://doi.org/10.1534/genetics.107.073601>
- 706 Harris, T. E. (1964). *The theory of branching processes* [OCLC: 931719544]. Springer-Verlag.
- 707 Hartfield, M., & Otto, S. P. (2011). Recombination and hitchhiking of deleterious alleles: Recombination and the  
708 undesirable hitchhiker. *Evolution*, 65(9), 2421–2434. <https://doi.org/10.1111/j.1558-5646.2011.01311.x>
- 709 Hartfield, M., & Glémin, S. (2014). Hitchhiking of deleterious alleles and the cost of adaptation in partially selfing  
710 species. *Genetics*, 196(1), 281–293. <https://doi.org/10.1534/genetics.113.158196>
- 711 Hartmann, F. E., Rodríguez de la Vega, R. C., Gladioux, P., Ma, W.-J., Hood, M. E., & Giraud, T. (2020). Higher  
712 gene flow in sex-related chromosomes than in autosomes during fungal divergence. *Molecular Biology and*  
713 *Evolution*, 37(3), 668–682. <https://doi.org/10.1093/molbev/msz252>
- 714 Hartmann, F. E., Duhamel, M., Carpentier, F., Hood, M. E., Foulongne-Oriol, M., Silar, P., Malagnac, F., Grognet,  
715 P., & Giraud, T. (2021). Recombination suppression and evolutionary strata around mating-type loci in  
716 fungi: Documenting patterns and understanding evolutionary and mechanistic causes. *New Phytologist*,  
717 229(5), 2470–2491. <https://doi.org/10.1111/nph.17039>

- 718 Hartmann, F. E., Ament-Velásquez, S. L., Vogan, A. A., Gautier, V., Le Prieur, S., Berramdane, M., Snirc, A.,  
719 Johannesson, H., Grognet, P., Malagnac, F., Silar, P., & Giraud, T. (2021). Size variation of the nonrecom-  
720 bining region on the mating-type chromosomes in the fungal *Podospora anserina* species complex (J. Parsch,  
721 Ed.). *Molecular Biology and Evolution*, *38*(6), 2475–2492. <https://doi.org/10.1093/molbev/msab040>
- 722 Heinzmann, D. (2009). Extinction times in multitype markov branching processes. *Journal of Applied Probability*,  
723 *46*(1), 296–307. <https://doi.org/10.1239/jap/1238592131>
- 724 Holsinger, K. E., & Feldman, M. W. (1985). Selection in complex genetic systems. VI. Equilibrium properties  
725 of two locus selection models with partial selfing. *Theoretical Population Biology*, *28*(1), 117–132. [https://doi.org/10.1016/0040-5809\(85\)90024-3](https://doi.org/10.1016/0040-5809(85)90024-3)
- 726
- 727 Hood, M. E., & Antonovics, J. (2000). Intratetrad mating, heterozygosity, and the maintenance of deleterious alleles  
728 in *Microbotryum violaceum* (= *Ustilago violacea*). *Heredity*, *85*(3), 231–241. <https://doi.org/10.1046/j.1365-2540.2000.00748.x>
- 729
- 730 Ironside, J. E. (2010). No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome  
731 evolution. *BioEssays*, *32*(8), 718–726. <https://doi.org/10.1002/bies.200900124>
- 732 Jay, P., Chouteau, M., Whibley, A., Bastide, H., Parrinello, H., Llaurens, V., & Joron, M. (2021). Mutation load at  
733 a mimicry supergene sheds new light on the evolution of inversion polymorphisms. *Nature Genetics*, *53*(3),  
734 288–293. <https://doi.org/10.1038/s41588-020-00771-1>
- 735 Jay, P., Tezenas, E., Véber, A., & Giraud, T. (2022). Sheltering of deleterious mutations explains the stepwise  
736 extension of recombination suppression on sex chromosomes and other supergenes (L. D. Hurst, Ed.).  
737 *PLOS Biology*, *20*(7), e3001698. <https://doi.org/10.1371/journal.pbio.3001698>
- 738 Jeffries, D. L., Gerchen, J. F., Scharmann, M., & Pannell, J. R. (2021). A neutral model for the loss of recombination  
739 on sex chromosomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1832),  
740 20200096. <https://doi.org/10.1098/rstb.2020.0096>
- 741 Jin, Y., Allan, S., Baber, L., Bhattarai, E. K., Lamb, T. M., & Versaw, W. K. (2007). Rapid genetic mapping in  
742 *Neurospora crassa*. *Fungal Genetics and Biology*, *44*(6), 455–465. <https://doi.org/10.1016/j.fgb.2006.09.002>
- 743 Karlin, S. (1975). General two-locus selection models: Some objectives, results and interpretations. *Theoretical*  
744 *Population Biology*, *7*(3), 364–398. [https://doi.org/10.1016/0040-5809\(75\)90025-8](https://doi.org/10.1016/0040-5809(75)90025-8)
- 745 Kesten, H., & Stigum, B. P. (1966). A limit theorem for multidimensional Galton-Watson processes. *The Annals*  
746 *of Mathematical Statistics*, *37*(5), 1211–1223. <https://doi.org/10.1214/aoms/1177699266>
- 747 Kimura, M. (1980). Average time until fixation of a mutant allele in a finite population under continued muta-  
748 tion pressure: Studies by analytical, numerical, and pseudo-sampling methods. *Proceedings of the National*  
749 *Academy of Sciences*, *77*(1), 522–526. <https://doi.org/10.1073/pnas.77.1.522>
- 750 Kratochvíl, L., & Stöck, M. (2021). Preface. *Philosophical Transactions of the Royal Society B: Biological Sciences*,  
751 *376*(1832), 20200088. <https://doi.org/10.1098/rstb.2020.0088>
- 752 Kuhn, M.-L., Gout, L., Howlett, B. J., Melayah, D., Meyer, M., Balesdent, M.-H., & Rouxel, T. (2006). Genetic  
753 linkage maps and genomic organization in *Leptosphaeria maculans*. *European Journal of Plant Pathology*,  
754 *114*(1), 17–31. <https://doi.org/10.1007/s10658-005-3168-6>
- 755 Lande, R., & Porcher, E. (2015). Maintenance of quantitative genetic variance under partial self-fertilization, with  
756 implications for evolution of selfing. *Genetics*, *200*(3), 891–906. <https://doi.org/10.1534/genetics.115.176693>

- 757 Leach, C. R., Mayo, O., & Morris, M. M. (1986). Linkage disequilibrium and gametophytic self-incompatibility.  
758 *Theoretical and Applied Genetics*, 73(1), 102–112. <https://doi.org/10.1007/BF00273726>
- 759 Lenormand, T., & Roze, D. (2022). Y recombination arrest and degeneration in the absence of sexual dimorphism.  
760 *Science*, 375(6581), 663–666. <https://doi.org/10.1126/science.abj1813>
- 761 Lenz, T. L., Spirin, V., Jordan, D. M., & Sunyaev, S. R. (2016). Excess of deleterious mutations around HLA  
762 genes reveals evolutionary cost of balancing selection. *Molecular Biology and Evolution*, 33(10), 2555–2564.  
763 <https://doi.org/10.1093/molbev/msw127>
- 764 Llaurens, V., Gonthier, L., & Billiard, S. (2009). The sheltered genetic load linked to the *S* locus in plants: New  
765 insights from theoretical and empirical approaches in sporophytic self-incompatibility. *Genetics*, 183(3),  
766 1105–1118. <https://doi.org/10.1534/genetics.109.102707>
- 767 Llaurens, V., Whibley, A., & Joron, M. (2017). Genetic architecture and balancing selection: The life and death of  
768 differentiated variants. *Molecular Ecology*, 26(9), 2430–2448. <https://doi.org/10.1111/mec.14051>
- 769 Malkus, A., Song, Q., Cregan, P., Arseniuk, E., & Ueng, P. P. (2009). Genetic linkage map of *Phaeosphaeria*  
770 *nodorum*, the causal agent of stagonospora nodorum blotch disease of wheat. *European Journal of Plant*  
771 *Pathology*, 124(4), 681–690. <https://doi.org/10.1007/s10658-009-9454-y>
- 772 Menkis, A., Jacobson, D. J., Gustafsson, T., & Johannesson, H. (2008). The mating-type chromosome in the  
773 filamentous ascomycete *Neurospora tetrasperma* represents a model for early evolution of sex chromosomes.  
774 *PLoS Genetics*, 4(3), e1000030. <https://doi.org/10.1371/journal.pgen.1000030>
- 775 Mode, C. J. (1971). *Multitype branching processes: Theory and applications*. American Elsevier Pub. Co.
- 776 Nei, M. (1971). Extinction time of deleterious mutant genes in large populations. *Theoretical Population Biology*,  
777 2(4), 419–425. [https://doi.org/10.1016/0040-5809\(71\)90030-X](https://doi.org/10.1016/0040-5809(71)90030-X)
- 778 Nicolas, M., Marais, G., Hykelova, V., Janousek, B., Laporte, V., Vyskot, B., Mouchiroud, D., Negrutiu, I.,  
779 Charlesworth, D., & Monéger, F. (2004). A gradual process of recombination restriction in the evolu-  
780 tionary history of the sex chromosomes in dioecious plants. *PLoS Biology*, 3(1), e4. <https://doi.org/10.1371/journal.pbio.0030004>
- 781
- 782 Ohta, T., & Cockerham, C. C. (1974). Detrimental genes with partial selfing and effects on a neutral locus. *Genetical*  
783 *Research*, 23(2), 191–200. <https://doi.org/10.1017/S0016672300014816>
- 784 Olito, C., Ponnikas, S., Hansson, B., & Abbott, J. K. (2022). Consequences of partially recessive deleterious genetic  
785 variation for the evolution of inversions suppressing recombination between sex chromosomes. *Evolution*,  
786 76(6), 1320–1330. <https://doi.org/10.1111/evo.14496>
- 787 Palsson, S. (2002). Selection on a modifier of recombination rate due to linked deleterious mutations. *Journal of*  
788 *Heredity*, 93(1), 22–26. <https://doi.org/10.1093/jhered/93.1.22>
- 789 Pénisson, S. (2010). *Conditional limit theorems for multitype branching processes and illustration in epidemiological*  
790 *risk analysis* (Doctoral dissertation). Universität Potsdam, Université Paris-Sud.
- 791 Phan, H. T. T., Ford, R., & Taylor, P. W. J. (2003). Mapping the mating type locus of *Ascochyta rabiei*, the causal  
792 agent of ascochyta blight of chickpea. *Molecular Plant Pathology*, 4(5), 373–381. <https://doi.org/10.1046/j.1364-3703.2003.00185.x>
- 793
- 794 Pollak, E. (1987). On the theory of partially inbreeding finite populations. I. Partial selfing. *Genetics*, 117(2), 353–  
795 360. <https://doi.org/10.1093/genetics/117.2.353>

- 796 Pollak, E., & Sabran, M. (1992). On the theory of partially inbreeding finite populations. III. Fixation probabilities  
797 under partial selfing when heterozygotes are intermediate in viability. *Genetics*, *131*(4), 979–985. <https://doi.org/10.1093/genetics/131.4.979>  
798
- 799 Pollard, J. H. (1966). On the use of the direct matrix product in analysing certain stochastic population models.  
800 *Biometrika*, *53*(3), 397.
- 801 Pollard, J. H. (1968). The multi-type galton-watson process in a genetical context. *Biometrics*, *24*(1), 147. <https://doi.org/10.2307/2528466>  
802
- 803 Ponnikas, S., Sigeman, H., Abbott, J. K., & Hansson, B. (2018). Why do sex chromosomes stop recombining? *Trends*  
804 *in Genetics*, *34*(7), 492–503. <https://doi.org/10.1016/j.tig.2018.04.001>
- 805 Pötscher, B. M. (1985). Moments and order statistics of extinction times in multitype branching processes and their  
806 relation to random selection models. *Bulletin of Mathematical Biology*, *47*(2), 263–272.
- 807 Rice, S. H. (2004). *Evolutionary theory: Mathematical and conceptual foundations*. Sinauer Associates.
- 808 Rice, W. R. (1987). The accumulation of sexually antagonistic genes as a selective agent promoting the evolution  
809 of reduced recombination between primitive sex chromosomes. *Evolution*, *41*(4), 911–914. <https://doi.org/10.1111/j.1558-5646.1987.tb05864.x>  
810
- 811 Ruzicka, F., Dutoit, L., Czuppon, P., Jordan, C. Y., Li, X.-Y., Olito, C., Runemark, A., Svensson, E. I., Yazdi, H. P.,  
812 & Connallon, T. (2020). The search for sexually antagonistic genes: Practical insights from studies of local  
813 adaptation and statistical genomics. *Evolution Letters*, *4*(5), 398–415. <https://doi.org/10.1002/evl3.192>
- 814 Sewastjanow, B. (1975). *Verzweigungsprozesse* (R. Oldenbourg Verlag).
- 815 Skinner, D. Z., Budde, A. D., Farman, M. L., Smith, J. R., Leung, H., & Leong, S. A. (1993). Genome organization of  
816 *Magnaporthe grisea*: Genetic map, electrophoretic karyotype, and occurrence of repeated DNAs. *Theoretical*  
817 *and Applied Genetics*, *87*(5), 545–557. <https://doi.org/10.1007/BF00221877>
- 818 Uyenoyama, M. K. (2005). Evolution under tight linkage to mating type. *New Phytologist*, *165*(1), 63–70. <https://doi.org/10.1111/j.1469-8137.2004.01246.x>  
819
- 820 Vittorelli, N., Snirc, A., Levert, E., Gautier, V., Lalanne, C., De Filippo, E., Rodríguez de la Vega, R. C., Gladieux,  
821 P., Guillou, S., Zhang, Y., Tejomurthula, S., Grigoriev, I. V., Debuchy, R., Silar, P., Giraud, T., & Hartmann,  
822 F. E. (2022). Stepwise recombination suppression around the mating-type locus in the fungus *Schizothecium*  
823 *tetrasporum* (ascomycota, sordariales). <https://doi.org/10.1101/2022.07.20.500756>
- 824 Wolfram Research, I. (2015). *Mathematica* (Version 10.1). Champaign, IL, 2015.
- 825 Wright, A. E., Dean, R., Zimmer, F., & Mank, J. E. (2016). How to make a sex chromosome. *Nature Communica-*  
826 *tions*, *7*(1), 12087. <https://doi.org/10.1038/ncomms12087>
- 827 Wright, S. I., Ness, R. W., Foxe, J. P., & Barrett, S. C. H. (2008). Genomic consequences of outcrossing and selfing  
828 in plants. *International Journal of Plant Sciences*, *169*(1), 105–118. <https://doi.org/10.1086/523366>
- 829 Yan, Z., Martin, S. H., Gotzek, D., Arsenault, S. V., Duchon, P., Helleu, Q., Riba-Grognuz, O., Hunt, B. G., Salamin,  
830 N., Shoemaker, D., Ross, K. G., & Keller, L. (2020). Evolution of a supergene that regulates a trans-species  
831 social polymorphism. *Nature Ecology & Evolution*, *4*(2), 240–249. <https://doi.org/10.1038/s41559-019-1081-1>  
832

833 Zhong, S., Steffenson, B. J., Martinez, J. P., & Ciuffetti, L. M. (2002). A molecular genetic map and electrophoretic  
 834 karyotype of the plant pathogenic fungus *Cochliobolus sativus*. *Molecular Plant-Microbe Interactions*, 15(5),  
 835 481–492. <https://doi.org/10.1094/MPMI.2002.15.5.481>

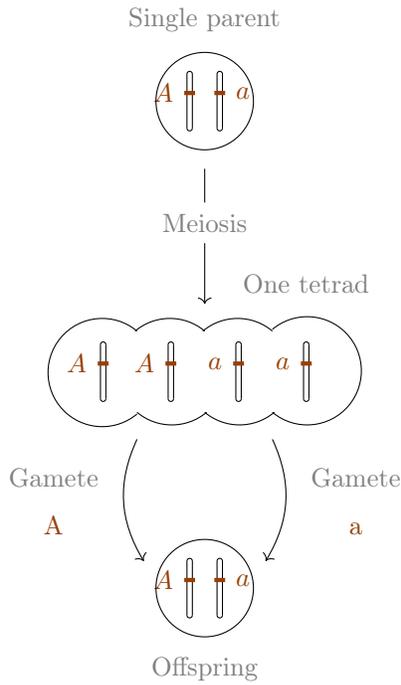
## 836 5 Appendix: Table of notation

$N$	Population size
$G_1, \dots, G_4$	Genotypes
$(g_1, \dots, g_4)$	Number of individuals of each genotype
$f$	Selfing probability
$p_{in}$ and $p_{out} = 1 - p_{in}$	Intra- and Inter-tetrad selfing probabilities
$r$	Recombination rate
$S_i$	Probability of survival of an offspring of genotype $i \in \{1, 2, 3, 4\}$ (see Figure 1)
$s$	Selection coefficient in the partial dominance case
$h$	Dominance coefficient in the partial dominance case
$s_3, s_4$	Selection coefficients in the overdominance case

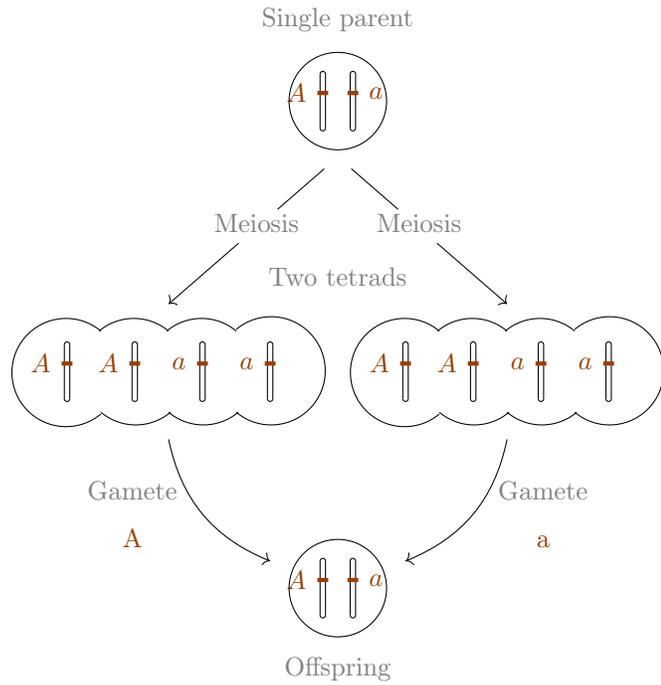
Table 1



(A) Intra-tetrad selfing



(B) Inter-tetrad selfing



(C) Outcrossing

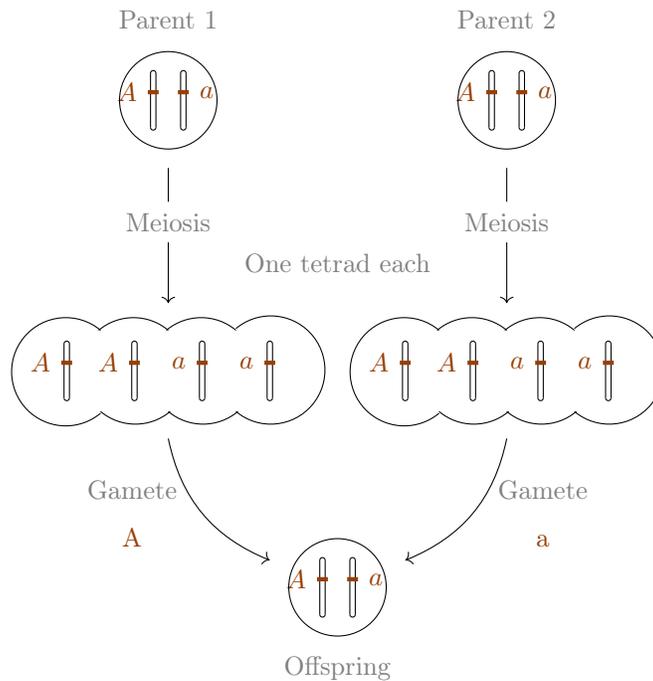


Figure 6: Schematic representation of the three mating systems considered in the model. Individuals are represented by a pair of mating-type chromosomes, with the mating-type locus displayed. A diploid offspring is generated by the fusion of two gametes carrying different mating-type alleles (A and a). (A) Under intra-tetrad selfing, both gametes are picked from the same tetrad; only one parent is involved. (B) Under inter-tetrad selfing, the two gametes are picked from two different tetrads (meioses) produced by the same diploid parent; only one parent is involved. (C) Under outcrossing, the two gametes are picked in tetrads produced by different parents.

838 **7 Appendices for the Method section**

839 **7.1 Rates of creation of offspring with given genotypes (Moran process)**

Parental Genotypes	Intra/Inter - Tetrad	Recombination	Genotype of offspring			
			$G_1$	$G_2$	$G_3$	$G_4$
$G_1$ $fg_1$	Intra $p_{in}$	(A) $(1-r)$	$fg_1p_{in}(1-r)$	0	0	0
		(A) $r$	$\frac{1}{4}fg_1p_{in}r$	$\frac{1}{4}fg_1p_{in}r$	$\frac{1}{4}fg_1p_{in}r$	$\frac{1}{4}fg_1p_{in}r$
	Inter $p_{out}$	(AA) $(1-r)^2$	$fg_1p_{out}(1-r)^2$	0	0	0
		(AP) $2(1-r)r$	$\frac{1}{2}2fg_1p_{out}(1-r)r$	0	$\frac{1}{4}2fg_1p_{out}(1-r)r$	$\frac{1}{4}2fg_1p_{out}(1-r)r$
		(PP) $r^2$	$\frac{1}{4}fg_1p_{out}r^2$	$\frac{1}{4}fg_1p_{out}r^2$	$\frac{1}{4}fg_1p_{out}r^2$	$\frac{1}{4}fg_1p_{out}r^2$
$G_2$ $fg_2$	Intra $p_{in}$	(A) $(1-r)$	0	$fg_2p_{in}(1-r)$	0	0
		(P) $r$	$\frac{1}{4}fg_2p_{in}r$	$\frac{1}{4}fg_2p_{in}r$	$\frac{1}{4}fg_2p_{in}r$	$\frac{1}{4}fg_2p_{in}r$
	Inter $p_{out}$	(AA) $(1-r)^2$	0	$fg_2p_{out}(1-r)^2$	0	0
		(AP) $2(1-r)r$	0	$\frac{1}{2}2fg_2p_{out}(1-r)r$	$\frac{1}{4}2fg_2p_{out}(1-r)r$	$\frac{1}{4}2fg_2p_{out}(1-r)r$
		(PP) $r^2$	$\frac{1}{4}fg_2p_{out}r^2$	$\frac{1}{4}fg_2p_{out}r^2$	$\frac{1}{4}fg_2p_{out}r^2$	$\frac{1}{4}fg_2p_{out}r^2$
$G_3 fg_3$	Same tetrad (homozyg.)	0	0	$fg_3$	0	
$G_4 fg_4$	Same tetrad (homozyg.)	0	0	0	$fg_4$	

Table 2: Table summarizing the rates of production of an offspring of each genotype (last four columns) in case of **selfing**. *Parental Genotype*: The genotype of the individual involved in the mating event; *Intra/Inter-tetrad*: Mating through intra- of inter-tetrad selfing (see section 2.1 for definitions); *Recombination*: Occurrence of a recombination event in the tetrads from which gametes are picked. "A" stands for "Absence" in one tetrad, "P" stands for "Presence" in one tetrad. We use only one letter when the two gametes come from the same tetrad or when one of the genotypes involved is homozygous at the load locus. For example, (AP) indicates that recombination occurred in one tetrad but not in the other.  $G_i$ : the rate at which an offspring of genotype  $G_i$  is produced, due to the scenario of parental genotype, intra/inter tetrad selfing and presence/absence of recombination considered. The total rate  $T_g(+G_i)$  at which a new offspring of genotype  $G_i$  is created when the population state is  $g = (g_1, g_2, g_3, g_4)$  is then the sum of all the rates appearing in column  $G_i$  in this Table, Table 3 and Table 4.

Parental Genotypes	Recombination	Genotype of offspring			
		$G_1$	$G_2$	$G_3$	$G_4$
$G_1G_1$ $(1-f) \times \frac{g_1-1}{N-1}$	(AA) $(1-r)^2$	$(1-f)g_1 \frac{g_1-1}{N-1} (1-r)^2$	0	0	0
	(AP) $2(1-r)r$	$\frac{1}{2}2(1-f)g_1 \frac{g_1-1}{N-1} (1-r)r$	0	$\frac{1}{4}2(1-f)g_1 \frac{g_1-1}{N-1} (1-r)r$	$\frac{1}{4}2(1-f)g_1 \frac{g_1-1}{N-1} (1-r)r$
	(PP) $r^2$	$\frac{1}{4}(1-f)g_1 \frac{g_1-1}{N-1} r^2$	$\frac{1}{4}(1-f)g_1 \frac{g_1-1}{N-1} r^2$	$\frac{1}{4}(1-f)g_1 \frac{g_1-1}{N-1} r^2$	$\frac{1}{4}(1-f)g_1 \frac{g_1-1}{N-1} r^2$
$G_1G_2$ $2(1-f) \times \frac{g_2}{N-1}$	(AA) $(1-r)^2$	0	0	$\frac{1}{2}2(1-f)g_1 \frac{g_2}{N-1} (1-r)^2$	$\frac{1}{2}2(1-f)g_1 \frac{g_2}{N-1} (1-r)^2$
	(AP) $(1-r)r$	$\frac{1}{2}2(1-f)g_1 \frac{g_2}{N-1} (1-r)r$	0	$\frac{1}{4}2(1-f)g_1 \frac{g_2}{N-1} (1-r)r$	$\frac{1}{4}2(1-f)g_1 \frac{g_2}{N-1} (1-r)r$
	(PA) $r(1-r)$	0	$\frac{1}{2}2(1-f)g_1 \frac{g_2}{N-1} r(1-r)$	$\frac{1}{4}2(1-f)g_1 \frac{g_2}{N-1} r(1-r)$	$\frac{1}{4}2(1-f)g_1 \frac{g_2}{N-1} r(1-r)$
	(PP) $r^2$	$\frac{1}{4}2(1-f)g_1 \frac{g_2}{N-1} r^2$	$\frac{1}{4}2(1-f)g_1 \frac{g_2}{N-1} r^2$	$\frac{1}{4}2(1-f)g_1 \frac{g_2}{N-1} r^2$	$\frac{1}{4}2(1-f)g_1 \frac{g_2}{N-1} r^2$
$G_1G_3$ $2(1-f) \times \frac{g_3}{N-1}$	(A) $1-r$	$\frac{1}{2}2(1-f)g_1 \frac{g_3}{N-1} (1-r)$	0	$\frac{1}{2}2(1-f)g_1 \frac{g_3}{N-1} (1-r)$	0
	(P) $r$	$\frac{1}{4}2(1-f)g_1 \frac{g_3}{N-1} r$	$\frac{1}{4}2(1-f)g_1 \frac{g_3}{N-1} r$	$\frac{1}{2}2(1-f)g_1 \frac{g_3}{N-1} r$	0
$G_1G_4$ $2(1-f) \times \frac{g_4}{N-1}$	(A) $1-r$	$\frac{1}{2}2(1-f)g_1 \frac{g_4}{N-1} (1-r)$	0	0	$\frac{1}{2}2(1-f)g_1 \frac{g_4}{N-1} (1-r)$
	(P) $r$	$\frac{1}{4}2(1-f)g_1 \frac{g_4}{N-1} r$	$\frac{1}{4}2(1-f)g_1 \frac{g_4}{N-1} r$	0	$\frac{1}{2}2(1-f)g_1 \frac{g_4}{N-1} r$

Table 3: Part 1 of the table summarizing the rates of production of an offspring of each genotype (last four columns) in case of **outcrossing**. *Parental Genotype*: The genotype of the individuals involved in the mating event; *Recombination*: Occurrence of a recombination event in the tetrads from which gametes are picked. "A" stands for "Absence" in one tetrad, "P" stands for "Presence" in one tetrad. We use only one letter when the two gametes come from the same tetrad or when one of the genotypes involved is homozygous at the load locus. For example, (AP) indicates that recombination occurred in one tetrad but not the other.  $G_i$ : the rate at which an offspring of genotype  $G_i$  is produced, due to the scenario of parental genotype, intra/inter tetrad selfing and presence/absence of recombination considered. The total rate  $T_g(+G_i)$  at which a new offspring of genotype  $G_i$  is created when the population state is  $g = (g_1, g_2, g_3, g_4)$  is then the sum of all the rates appearing in column  $G_i$  in this Table, Table 2 and Table 4.

Parental Genotypes	Recombination	Genotype of offspring			
		$G_1$	$G_2$	$G_3$	$G_4$
$G_2G_2$ $(1-f) \times \frac{g_2-1}{N-1}$	(AA) $(1-r)^2$	0	$(1-f)g_2 \frac{g_2-1}{N-1} (1-r)^2$	0	0
	(AP) $2(1-r)r$	0	$\frac{1}{2}2(1-f)g_2 \frac{g_2-1}{N-1} (1-r)r$	$\frac{1}{4}2(1-f)g_2 \frac{g_2-1}{N-1} (1-r)r$	$\frac{1}{4}2(1-f)g_2 \frac{g_2-1}{N-1} (1-r)r$
	(PP) $r^2$	$\frac{1}{4}(1-f)g_2 \frac{g_2-1}{N-1} r^2$	$\frac{1}{4}(1-f)g_2 \frac{g_2-1}{N-1} r^2$	$\frac{1}{4}(1-f)g_2 \frac{g_2-1}{N-1} r^2$	$\frac{1}{4}(1-f)g_2 \frac{g_2-1}{N-1} r^2$
$G_2G_3$ $2(1-f) \times \frac{g_3}{N-1}$	(A) $1-r$	0	$\frac{1}{2}(1-f)g_2 \frac{g_3}{N-1} (1-r)$	$\frac{1}{2}2(1-f)g_2 \frac{g_3}{N-1} (1-r)$	0
	(P) $r$	$\frac{1}{4}2(1-f)g_2 \frac{g_3}{N-1} r$	$\frac{1}{4}2(1-f)g_2 \frac{g_3}{N-1} r$	$\frac{1}{2}2(1-f)g_2 \frac{g_3}{N-1} r$	0
$G_2G_4$ $2(1-f) \times \frac{g_4}{N-1}$	(A) $1-r$	0	$\frac{1}{2}2(1-f)g_2 \frac{g_4}{N-1} (1-r)$	0	$\frac{1}{2}2(1-f)g_2 \frac{g_4}{N-1} (1-r)$
	(P) $r$	$\frac{1}{4}2(1-f)g_2 \frac{g_4}{N-1} r$	$\frac{1}{4}2(1-f)g_2 \frac{g_4}{N-1} r$	0	$\frac{1}{2}2(1-f)g_2 \frac{g_4}{N-1} r$
$G_3G_3$ $(1-f) \times \frac{g_3-1}{N-1}$	Same tetrads	0	0	$(1-f)g_3 \frac{g_3-1}{N-1}$	0
$G_3G_4$ $2(1-f) \times \frac{g_4}{N-1}$	Same tetrads	$\frac{1}{2}2(1-f)g_3 \frac{g_4}{N-1}$	$\frac{1}{2}2(1-f)g_3 \frac{g_4}{N-1}$	0	0
$G_4G_4$ $(1-f) \times \frac{g_4-1}{N-1}$	Same tetrads	0	0	0	$(1-f)g_4 \frac{g_4-1}{N-1}$

Table 4: Part 2 of the table summarizing the rates of production of an offspring of each genotype (last four columns) in case of **outcrossing**. *Parental Genotype*: The genotype of the individuals involved in the mating; *Recombination*: Occurrence of a recombination event in the tetrads from which gametes are picked. "A" stands for "Absence" in one tetrad, "P" stands for "Presence" in one tetrad. We use only one letter when the two gametes come from the same tetrad or when one of the genotypes involved is homozygous at the load locus. For example, (AP) indicates that recombination occurred in one tetrad but not the other.  $G_i$ : the rate at which an offspring of genotype  $G_i$  is produced, due to the scenario of parental genotype, intra/inter tetrad selfing and presence/absence of recombination considered. The total rate  $T_g(+G_i)$  at which a new offspring of genotype  $G_i$  is created when the population state is  $g = (g_1, g_2, g_3, g_4)$  is then the sum of all the rates appearing in column  $G_i$  in this Table, Table 2 and Table 3.

841 The total rate at which an offspring of a given genotype is produced is then obtained by summing the rates  
 842 along each column  $G_i$  in Tables 2, 3 and 4. This gives:

$$T_g(+G_1) = fg_1 \left(1 - r + \frac{r}{4} \left(1 - (1 - p_{in})(1 - r)\right)\right) + fg_2 \frac{r}{4} \left(1 - (1 - p_{in})(1 - r)\right) \\ + \frac{1-f}{N-1} \left[ g_1 \left(1 - \frac{r}{2}\right) \left((g_1 - 1) \left(1 - \frac{r}{2}\right) + g_3 + g_4\right) + g_2 r \left((g_2 - 1) \frac{r}{4} + \frac{1}{2}(g_3 + g_4)\right) + g_1 g_2 r \left(1 - \frac{r}{2}\right) + g_3 g_4 \right],$$

$$T_g(+G_2) = fg_1 \frac{r}{4} \left(1 - (1 - p_{in})(1 - r)\right) + fg_2 \left(1 - r + \frac{r}{4} \left(1 - (1 - p_{in})(1 - r)\right)\right) \\ + \frac{1-f}{N-1} \left[ g_2 \left(1 - \frac{r}{2}\right) \left((g_2 - 1) \left(1 - \frac{r}{2}\right) + g_3 + g_4\right) + g_1 r \left((g_1 - 1) \frac{r}{4} + \frac{1}{2}(g_3 + g_4)\right) + g_1 g_2 r \left(1 - \frac{r}{2}\right) + g_3 g_4 \right],$$

$$T_g(+G_3) = fg_1 \frac{r}{4} \left(1 + (1 - p_{in})(1 - r)\right) + fg_2 \frac{r}{4} \left(1 + (1 - p_{in})(1 - r)\right) + fg_3 \\ + \frac{1-f}{N-1} \left[ g_1 (g_1 - 1) \frac{r}{2} \left(1 - \frac{r}{2}\right) + g_2 (g_2 - 1) \frac{r}{2} \left(1 - \frac{r}{2}\right) + g_1 g_2 \left(1 - r + \frac{r^2}{2}\right) + g_3 (g_1 + g_2 + (g_3 - 1)) \right],$$

$$T_g(+G_4) = fg_1 \frac{r}{4} \left(1 + (1 - p_{in})(1 - r)\right) + fg_2 \frac{r}{4} \left(1 + (1 - p_{in})(1 - r)\right) + fg_4 \\ + \frac{1-f}{N-1} \left[ g_1 (g_1 - 1) \frac{r}{2} \left(1 - \frac{r}{2}\right) + g_2 (g_2 - 1) \frac{r}{2} \left(1 - \frac{r}{2}\right) + g_1 g_2 \left(1 - r + \frac{r^2}{2}\right) + g_4 (g_1 + g_2 + (g_4 - 1)) \right].$$

## 843 7.2 Reproduction law for the branching process

844 We give here an example of how the reproduction laws for the branching process are derived from the rates of the  
 845 Moran process, using the approximate regime (1).

846 Let us derive the coefficient  $A_{12}$  of the matrix  $A$ , which is the rate at which an individual of genotype  $G_2$   
 847 generates an offspring of genotype  $G_1$  and survives. Equivalently, this is the rate at which an individual of genotype  
 848  $G_2$  generates a **descendance vector** equal to  $e_1 + e_2$ .

849 Using the rates obtained for the Moran model, the rate at which an individual of genotype  $G_2$  produces an  
 850 offspring of genotype  $G_1$  is:

$$f \frac{r}{4} \left(1 - (1 - p_{in})(1 - r)\right) + (1 - f) \left[ r \left( \frac{g_2 - 1}{N - 1} \times \frac{r}{4} + \frac{1}{2} \left( \frac{g_3}{N - 1} + \frac{g_4}{N - 1} \right) \right) + \frac{g_1}{N - 1} r \left(1 - \frac{r}{2}\right) \right]. \quad (5)$$

851 The first term, with a factor  $f$ , is the rate at which an individual of genotype  $G_2$  produces an offspring of  
 852 genotype  $G_1$  by selfing. The second term, with a factor  $1 - f$ , is the rate at which an individual of genotype  $G_2$   
 853 produces an offspring of genotype  $G_1$  by outcrossing. In this term, the fractions of the form  $\frac{g_i}{N-1}$  represent the  
 854 probabilities that an individual of genotype  $G_i$  is chosen to mate with the  $G_2$  parent.

855 Using the approximation (1), *i.e.* assuming that  $g_4 \approx N$  and  $g_i \ll N$  for  $i = 1, 2, 3$ , we obtain that the quantity  
 856 in Eq. (5) can be approximated by:

$$f \frac{r}{4} \left( 1 - (1 - p_{in})(1 - r) \right) + (1 - f) \frac{r}{2}.$$

857 To obtain  $A_{12}$ , it remains to multiply this rate by the probability that the offspring survives,  $S_1$ , and the  
858 probability that the parent  $G_2$  is not chosen to die,  $\frac{N-1}{N}$ . As the population size  $N$  is considered large, the latter  
859 probability is approximately equal to 1.

860 This gives:

$$A_{12} = \left[ f \frac{r}{4} \left( 1 - (1 - p_{in})(1 - r) \right) + (1 - f) \frac{r}{2} \right] \times S_1.$$

### 861 7.3 Equation for the expected value of the size of the mutant population

862 This appendix gives the details of the derivation of the coefficients of the matrix  $C$  defined by

$$\frac{d}{dt} \mathbb{E}_{Z_0}[Z_t] = \mathbb{E}_{Z_0}[Z_t]C,$$

863 following Bacaër, 2018. Note that this is the same matrix defined in Athreya and Ney, 1972, Eq. 9, part. V.7.2., or  
864 in Pénisson, 2010, Eq. 1.1.16, but here we use the methodology described by Bacaër, 2018 to derive its coefficients.

865 In the following, type  $j$  refers to the genotype  $G_j$ . We will use the standard notation  $s^z := s_1^{z_1} s_2^{z_2} \dots s_d^{z_d}$  for  $s$   
866 and  $z$  two vectors of the same dimension  $d$ .

#### 867 7.3.1 Notation

868 For all  $t \geq 0$ , let us denote the expected value of the process at time  $t$  by  $E(t)$ :

$$E(t) = \begin{pmatrix} E_1(t) \\ E_2(t) \\ E_3(t) \end{pmatrix} = \begin{pmatrix} \mathbb{E}[Z_{t,1}] \\ \mathbb{E}[Z_{t,2}] \\ \mathbb{E}[Z_{t,3}] \end{pmatrix}.$$

For  $z \in \mathbb{N}^3$  and  $t \geq 0$ , we let  $p(t, z) = \mathbb{P}(Z_t = z)$  be the probability that the system is found in state  $z$  at time  
 $t$ . Let  $f(t, \cdot)$  be the generating function of the variable  $Z_t$ : for all  $s \in [0, 1]^3$ ,

$$f(t, s) := \sum_{z \in \mathbb{N}^3} p(t, z) s^z = \mathbb{E} \left[ s_1^{Z_{t,1}} s_2^{Z_{t,2}} s_3^{Z_{t,3}} \right].$$

869 Recalling that  $Y^j$  stands for the random vector of number of descendants of each type generated by the re-  
870 production of a type  $j$  individual, we also define  $\pi_j(z) = \mathbb{P}(Y^j = (z_1, z_2, z_3))$ . As indicated in the main text,  
871 the rates at which an individual of type  $j$  reproduces and gives rise to a **descendance vector**  $e_i + e_j$ ,  $e_i$  or 0 are  
872 respectively  $A_{ij}$ ,  $T_{ij}$  and  $D_{jj}$ . We denote the total rate at which a reproduction event occurs for a parent of type  
873  $j$  by  $c_j := \sum_i A_{ij} + \sum_i T_{ij} + D_{jj}$ .

874 The reproduction law of type  $j$  individuals is then given by, for every  $i \in \{1, 2, 3\}$ ,

$$\mathbb{P}(Y^j = e_i + e_j) = \frac{A_{ij}}{c_j}, \quad \mathbb{P}(Y^j = e_i) = \frac{T_{ij}}{c_j}, \quad \mathbb{P}(Y^j = 0) = \frac{D_{jj}}{c_j}.$$

875 Finally, let  $h_j$  be the generating function of the reproduction law of type  $j$  individuals, for  $j \in \{1, 2, 3\}$ . That  
876 is, for  $s \in [0, 1]^3$ ,

$$h_j(s) = \sum_{z \in \mathbb{N}^3} \pi_j(z) s^z = \mathbb{E} \left[ s_1^{Y_1^j} s_2^{Y_2^j} s_3^{Y_3^j} \right].$$

#### 877 7.3.2 Ordinary differential edecompoquation (ODE) satisfied by $(E(t))_{t \geq 0}$

The reproduction law of each type has finite moments of all order, because the number of descendants produced can  
not exceed 2. That garantees that there is no explosion of the population in finite time. Hence, standard results  
on multi-dimensional random variables (see for example Athreya and Ney, 1972) give us that, for all types  $j$  and

all  $t \geq 0$ ,

$$E_j(t) = \frac{\partial f}{\partial s_j}(t, \mathbb{1}),$$

878 with  $\mathbb{1} = (1, 1, 1)$ , which gives

$$\frac{dE_j(t)}{dt} = \frac{\partial^2 f}{\partial s_j \partial t}(t, \mathbb{1}) = \frac{\partial}{\partial s_j} \left( \sum_{z \in \mathbb{N}^3} \frac{\partial p}{\partial t}(t, z) s^z \right) \Big|_{s=\mathbb{1}}.$$

879 The variation of  $p$  over time  $\frac{\partial p(t, z)}{\partial t}$  can be decomposed into two terms. For  $z \in \mathbb{N}^3$ ,

$$\frac{\partial p}{\partial t}(t, z) = - \sum_{j=1}^3 z_j c_j p(t, z) + \sum_{j=1}^3 \sum_{\substack{u, v \in \mathbb{N}^3 \\ u+v=z}} (u_j + 1) c_j p(t, u + e_j) \pi_j(v).$$

880 The first term is the rate at which the population departs from state  $z$ , and is given by the sum over all types  
 881  $j$  of the rate at which individuals of type  $j$  reproduce. The second term is the rate at which the population arrives  
 882 in state  $z$  from another state, and can be decomposed according to the individual type whose reproduction changes  
 883 the population state. Note that the **descendance vector** generated during the reproduction event ( $v$ ) counts the  
 884 parent when it does not die, implying that the population is formally decreased by one individual of type  $j$  and  
 885 increased by a vector  $v$  during the reproduction event. In other words, if the population starts from a state  $u + e_j$   
 886 and an individual of type  $j$  reproduces by creating a vector  $v$  of descendants, the final state of the population is  
 887  $u + e_j - e_j + v = u + v$ .

888 Back to the derivative of  $f$  with respect to  $t$ , we use the fact that the rates  $c_j$  are independent of the current  
 889 state of the population to re-arrange the sums and obtain:

$$\begin{aligned} \frac{\partial f}{\partial t}(t, s) &= \sum_{z \in \mathbb{N}^3} \frac{\partial p}{\partial t}(t, z) s^z \\ &= \sum_{z \in \mathbb{N}^3} \left( - \sum_{j=1}^3 z_j c_j p(t, z) s^z + \sum_{j=1}^3 \sum_{\substack{u, v \in \mathbb{N}^3 \\ u+v=z}} (u_j + 1) c_j p(t, u + e_j) \pi_j(v) s^z \right) \\ &= \sum_j c_j \left( - \sum_{z \in \mathbb{N}^3} z_j p(t, z) s^z + \sum_{z \in \mathbb{N}^3} \sum_{\substack{u, v \in \mathbb{N}^3 \\ u+v=z}} (u_j + 1) p(t, u + e_j) \pi_j(v) s^z \right) \\ &= \sum_j c_j \left( -s_j \sum_{z \in \mathbb{N}^3} z_j p(t, z) s^{z-e_j} + \sum_{v \in \mathbb{N}^3} \sum_{u \in \mathbb{N}^3} (u_j + 1) p(t, u + e_j) \pi_j(v) s^{u+v} \right) \\ &= \sum_j c_j \left( -s_j \sum_{z \in \mathbb{N}^3} z_j p(t, z) s^{z-e_j} + \sum_{v \in \mathbb{N}^3} \pi_j(v) s^v \sum_{u \in \mathbb{N}^3} (u_j + 1) p(t, u + e_j) s^u \right) \\ &= \sum_j c_j \left( -s_j \frac{\partial f}{\partial s_j}(t, s) + h_j(s) \frac{\partial f}{\partial s_j}(t, s) \right) \\ &= \sum_j c_j (h_j(s) - s_j) \frac{\partial f}{\partial s_j}(t, s). \end{aligned}$$

890 Writing  $\delta_{i,j} = 1$  if  $i = j$  and  $\delta_{i,j} = 0$  otherwise, we then obtain for the expected value:

$$\begin{aligned}
\frac{dE_i}{dt} &= \frac{\partial}{\partial s_i} \frac{\partial f}{\partial t}(t, s) \Big|_{s=\mathbb{1}} \\
&= \sum_j c_j \left( \frac{\partial h_j}{\partial s_i}(\mathbb{1}) - \delta_{i,j} \right) \frac{\partial f}{\partial s_j}(t, \mathbb{1}) + \sum_j c_j (h_j(\mathbb{1}) - 1) \frac{\partial^2 f}{\partial s_i \partial s_j}(t, \mathbb{1}) \\
&= \sum_j c_j \left( \frac{\partial h_j}{\partial s_i}(\mathbb{1}) - \delta_{i,j} \right) E_j(t) + \sum_j c_j (h_j(\mathbb{1}) - 1) \frac{\partial^2 f}{\partial s_i \partial s_j}(t, \mathbb{1}) \\
&= \sum_j c_j \left( \frac{\partial h_j}{\partial s_i}(\mathbb{1}) - \delta_{i,j} \right) E_j(t),
\end{aligned}$$

where the last equality arises from the fact that, because  $h_j$  is a generating function,

$$h_j(\mathbb{1}) - 1 = \sum_{z \in \mathbb{N}^3} \pi_j(z) - 1 = 0.$$

891 The matrix  $C$  we are looking for is thus defined by  $C_{ij} = c_j \left( \frac{\partial h_j}{\partial s_i}(\mathbb{1}) - \delta_{ij} \right)$  for  $1 \leq i, j \leq 3$ .

892 Furthermore, we have, for all  $j$ ,

$$h_j(s) = \frac{1}{c_j} \left( \sum_i A_{ij} s_i s_j + \sum_i T_{ij} s_i + D_{jj} \right).$$

893 Combining the above, we arrive at

$$C_{ij} = \begin{cases} A_{ij} + T_{ij} & \text{if } i \neq j, \\ A_{jj} - \sum_{k \neq j} T_{kj} - D_{jj} & \text{if } i = j. \end{cases}$$

894 In conclusion, the matrix  $C$  is given by

$$C = \begin{pmatrix} (fa(r) + (1-f)d(r))S_1 - S_4 & (fc(r) + (1-f)\frac{r}{2})S_1 & (1-f)S_1 \\ (fc(r) + (1-f)\frac{r}{2})S_2 & (fa(r) + (1-f)d(r))S_2 - S_4 & (1-f)S_2 \\ fb(r)S_3 & fb(r)S_3 & fS_3 - S_4 \end{pmatrix}, \quad (6)$$

895 with

$$\begin{aligned}
a(r) &= 1 - r + \frac{r}{4} \left( 1 - (1 - p_{in})(1 - r) \right), & b(r) &= \frac{r}{4} \left( 1 + (1 - p_{in})(1 - r) \right), \\
c(r) &= \frac{r}{4} \left( 1 - (1 - p_{in})(1 - r) \right), & \text{and } d(r) &= 1 - \frac{1}{2}r.
\end{aligned}$$

## 7.4 Reducibility of the matrix $C$ and probability of extinction of the branching process

We will use the standard notation  $s^z := s_1^{z_1} s_2^{z_2} \dots s_d^{z_d}$  for  $s$  and  $z$  two vectors of the same dimension  $d$ .

Assessing the type of branching process at hand (super-, sub-, or critical) relies on the study of the eigenvalues of the matrix  $C$ . We use results of Sewastjanow, 1975 detailed in Pénisson, 2010 to obtain conditions on the almost-sure extinction of the process. When the matrix  $C$  is irreducible, the Perron-Frobenius theory of positive matrices states that it has a unique dominant eigenvalue. The branching process is then super-, sub-, or critical when this dominant eigenvalue is respectively positive, negative, or zero (Athreya and Ney, 1972, V.7.2.). In particular, the probability of extinction is equal to 1 when  $\rho \leq 0$ .

In our case, the matrix  $C$  can be reducible (for example, when  $f = 0$ ). In order to obtain a result on the probability of extinction in the subcritical case, we use the theory of sub-processes and of final classes. We recall below useful definitions and the principal result used (Sewastjanow, 1975).

Let  $(Z_t)_{t>0}$  be a multitype branching process, with types in a finite set  $K$ . The equivalence relation of *communication* is defined by: for all states  $k_i, k_j \in K$ , we say that  $k_i$  and  $k_j$  *communicate*, if and only if there exist  $s, t > 0$  such that

$$\mathbb{P}_{e_{k_i}}(Z_{s,k_j} > 0) > 0 \quad \text{and} \quad \mathbb{P}_{e_{k_j}}(Z_{t,k_i} > 0) > 0.$$

This means that there exists a time at which the probability that the population described by a branching process initiated with a single individual of type  $k_i$  contains an individual of type  $k_j$  is positive, and a time at which the probability that the population described by a branching process initiated with a single individual of type  $k_j$  contains an individual of type  $k_i$  is positive as well. If a subset  $\tilde{K} = \{k_1, \dots, k_p\}$  is a class for the communication equivalence relation (meaning that each state of  $\tilde{K}$  communicates with all the others but communicates with none of the states in  $\tilde{K}^c$ ), the  $\tilde{K}$ -subprocess is the process defined for all  $t > 0$  by

$$\tilde{Z}_t := (Z_{t,k_1}, \dots, Z_{t,k_p}),$$

which is the vector  $Z_t$  from which only the coordinates of the types in the class  $\tilde{K}$  are kept.  $(\tilde{Z}_t)_{t \geq 0}$  is still a branching process, and is by definition irreducible.

Let  $F_{t,k_i} : s \in [0, 1]^d \mapsto \mathbb{E}_{e_{k_i}} [s^{Z_t}]$  be the generating function of the process  $(Z_t)_{t \geq 0}$  at time  $t$ , starting with one individual of type  $k_i$ .  $\tilde{K} = \{k_1, \dots, k_p\}$  is then said to be a *final class* if it is non-empty, and satisfies the property that there exists  $t > 0$  such that for all  $k_i \in \tilde{K}$  and  $s \in [0, 1]^d$ ,  $F_{t,k_i}(s)$  is of the form

$$F_{t,k_i}(s) = \alpha_{k_i,1}(t, s) s_{k_1} + \dots + \alpha_{k_i,p}(t, s) s_{k_p},$$

where the coefficients  $\alpha_{k_i,j}$  can be expressed using the coordinates  $s_k$  of  $s$  such that  $k \notin \tilde{K}$ . In other words,  $F_{t,k_i}(s)$  is linear in  $s_k$  for all  $k \in \tilde{K}$ . The interpretation of this property is that whenever the population starts from a single individual of type  $k_i \in \tilde{K}$ , at any time  $t \geq 0$  there is one, and only one, individual of a type  $k_j \in \tilde{K}$  (and potentially other individuals with types in  $\tilde{K}^c$ ). The following result gives a condition for the almost sure extinction of the process  $(Z_t)_{t \geq 0}$  in the general case where the matrix  $C$  is not necessarily irreducible. Recall that the Perron's root

918  $\rho$  of a process, when it exists, is a real eigenvalue of the matrix associated with the process such that all real parts  
 919 of other eigenvalues are smaller than  $\rho$  (see Pénisson, 2010 Th. 1.1.7 and the following ones for a more detailed  
 920 definition).

921 **Proposition 1** (Prop. 1.1.22 in Pénisson, 2010)

922 Let  $(Z_t)_{t>0}$  be a continuous time Galton-Watson process, and let  $\rho = \max_{\tilde{K}} \rho_{\tilde{K}}$  be the maximal value of the Perron's  
 923 roots of all the possible  $\tilde{K}$ -subprocesses. Then the process  $(Z_t)_{t>0}$  almost surely dies out if and only if there are no  
 924 final classes and  $\rho \leq 0$ .

925 Let us verify that our branching process does not contain a final class. For that, we show that the generating  
 926 function of the process starting from any state has a non-zero coefficient of degree zero, and thus cannot be linear.

927 For any  $t > 0, r \in [0, 1]^3$  and any  $j \in \{1, 2, 3\}$ , we can decompose the generating function into

$$F_{t,j}(r) = \mathbb{E}_{e_j}(r^{Z_t}) = \mathbb{P}_{e_j}(Z_t = \mathbf{0}) + \sum_{z \in \mathbb{N}^3 \setminus \{\mathbf{0}\}} \mathbb{P}_{e_j}(Z_t = z) r^z.$$

928 Let us prove that  $\mathbb{P}_{e_j}(Z_t = \mathbf{0}) > 0$  for every  $j \in \{1, 2, 3\}$ . This will prove that the generating function cannot be  
 929 linear for any initial type, and thus that the process does not contain any final classes.

930 Let  $j \in \{1, 2, 3\}$ ,  $\tau_1$  be the time of the first reproduction event, and  $Y_1^j$  be the **descendance vector** created at  
 931 that time. We have

$$\begin{aligned} \mathbb{P}_{e_j}(Z_t = \mathbf{0}) &\geq \mathbb{P}_{e_j}(\{\tau_1 \leq t\} \cap \{Y_1^j = \mathbf{0}\}) \\ &= \mathbb{P}_{e_j}(Y_1^j = \mathbf{0} | \tau_1 \leq t) \mathbb{P}_{e_j}(\tau_1 \leq t) \\ &= \mathbb{P}_{e_j}(\tau_1 \leq t) \times \mathbb{P}(Y^j = \mathbf{0}) \\ &= (1 - e^{-c_j t}) \frac{D_{jj}}{c_j}, \end{aligned}$$

932 where  $c_j$  is the total rate of reproduction of an individual of type  $j$ , and  $D_{jj}$  is the rate at which an individual of  
 933 type  $j$  reproduces and gives rise to a null vector of descendants. Hence,  $\mathbb{P}_{e_j}(Z_t = \mathbf{0}) > 0$  when  $D_{jj} > 0$ .

934 For both selection scenarii,  $D_{11} = D_{22} = D_{33} = S_4$ . In the partial dominance selection scenario,  $S_4 = 1$ , and  
 935 in the overdominant selection scenario,  $S_4 = 1 - s_4$ . Having  $D_{jj} = 0$  for any  $j$  is impossible in the first scenario  
 936 and requires  $s_4 = 1$  in the second scenario, which means that the wild allele is lethal, which is not a reasonable  
 937 assumption. We thus take  $s_4 < 1$ . As a consequence, the generating function cannot be linear, and the process  
 938 does not contain any final class.

939 The result of Proposition 1 then applies here, and the sign of the dominant eigenvalue of matrix  $C$  gives a  
 940 condition on the almost-sure extinction of the process.

## 941 8 Appendices for the Results section

### 942 8.1 Supplementary figures

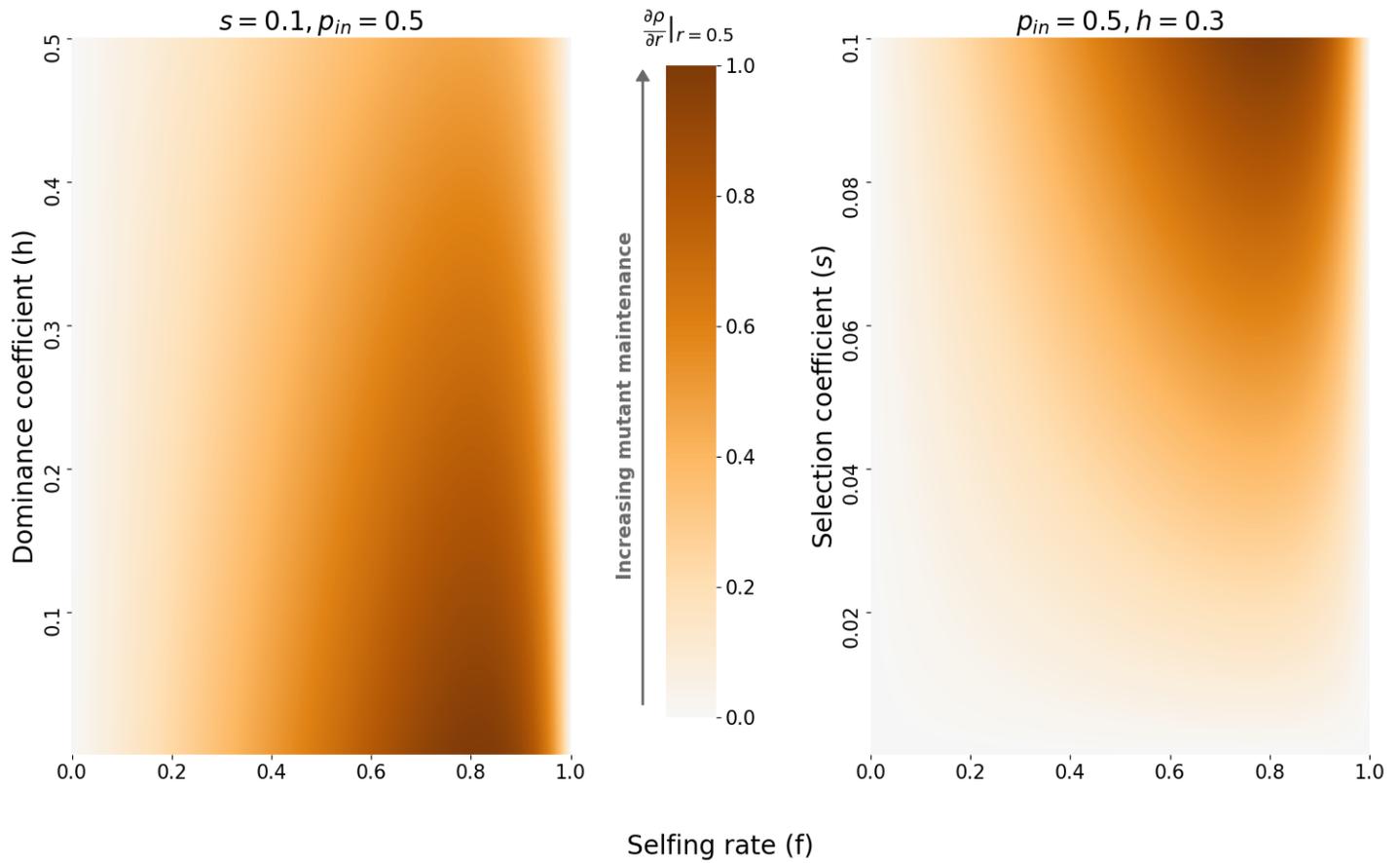


Figure S1: Relative variation of the derivative of the eigenvalue in the partial dominance case, for varying selfing rate  $f$  (x-axis), dominance coefficient  $h$  (y-axis, left) and selection coefficient  $s$  (y-axis, right). For each panel, the values of  $\frac{\partial \rho}{\partial r} |_{r=0.5}$  range from a minimal value, which is negative, to zero. We divided each value of the derivative by this minimum in order to plot values between 0 and 1 for every panel. This enables us to compare the impact of different parameters ( $h$ ,  $s$  and  $f$ ) on the sheltering effect of the mating-type locus. The darker the color, the more the mating-type locus shelters the mutation, thus promoting its maintenance.

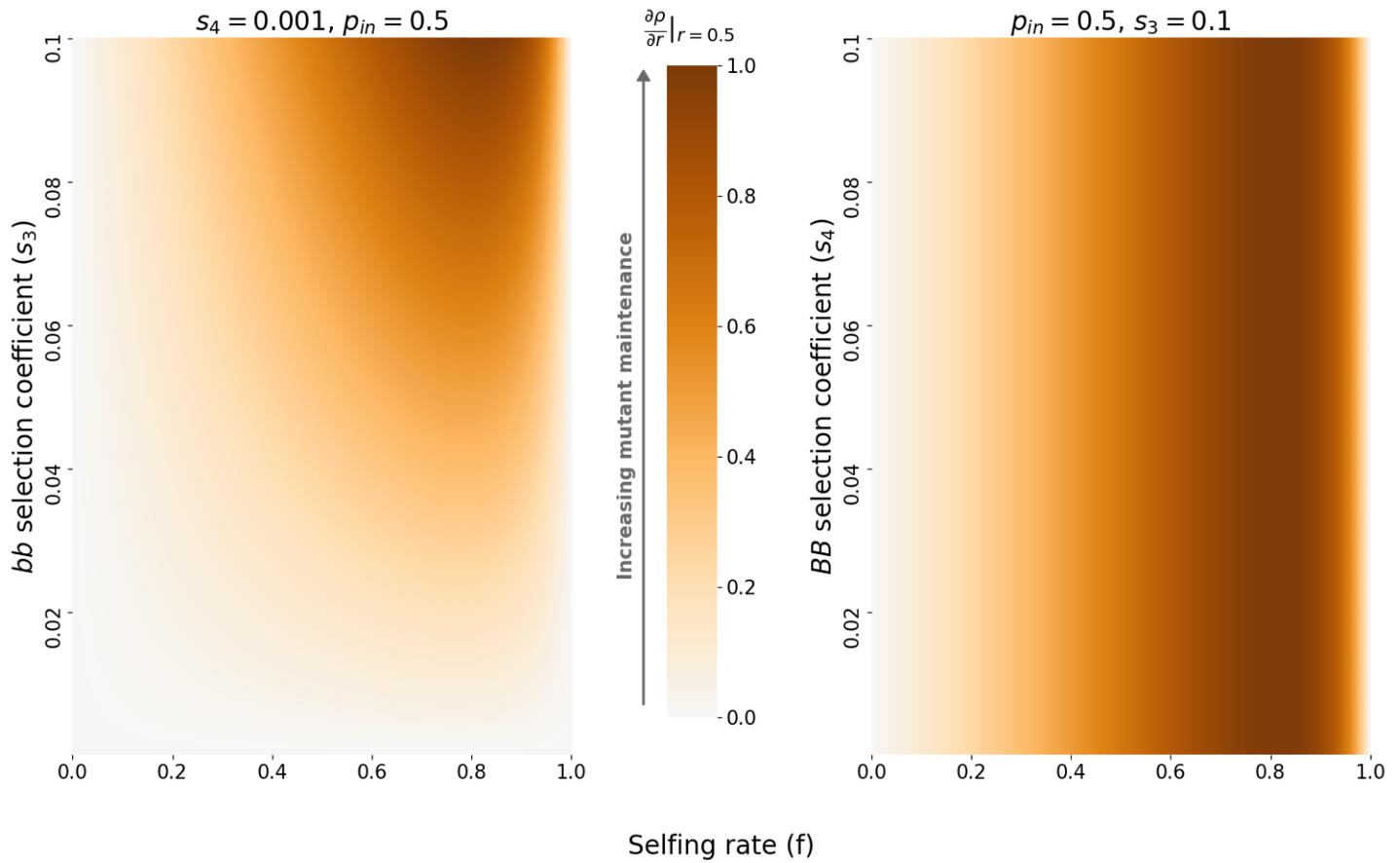


Figure S2: Relative variation of the derivative of the eigenvalue in the overdominance case, for varying selfing rate  $f$  (x-axis), and selection coefficients  $s_3$  (y-axis, left) and  $s_4$  (y-axis, right), with  $s_3 > s_4$ . For each panel, the values of  $\frac{\partial \rho}{\partial r} |_{r=0.5}$  range from a minimal value, which is negative, to zero. We divided each value of the derivative by this minimum in order to plot values between 0 and 1 for every panel. This enables us to compare the impact of different parameters ( $s_3$ ,  $s_4$  and  $f$ ) on the sheltering effect of the mating-type locus. The darker the color, the more the mating-type locus shelters the mutation, thus promoting its maintenance.

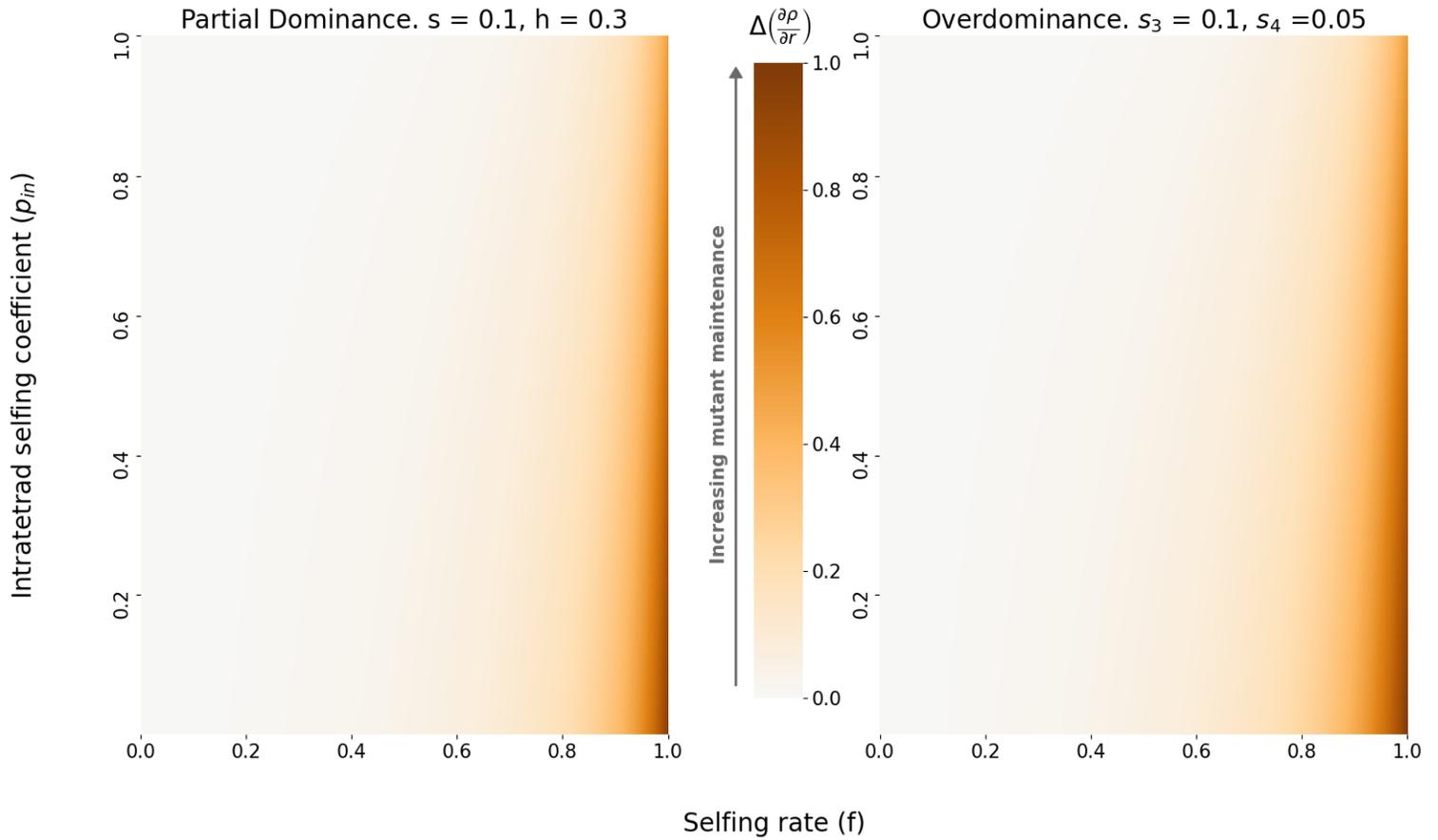


Figure S3: Difference between the dominant eigenvalue derivative at  $r = 0.5$  and at  $r = 0$ ,  $\Delta\left(\frac{\partial \rho}{\partial r}\right) = \left.\frac{\partial \rho}{\partial r}\right|_{r=0.5} - \left.\frac{\partial \rho}{\partial r}\right|_{r=0}$ . The left panel shows the partial dominance case, the right panel shows the overdominance case, for varying selfing rate  $f$  (x-axis), and intra-tetrad selfing rate (y-axis). The difference is always positive, with both derivative being negative (see App. 8.2.3 and App. 8.3.3). This means that the absolute value of the derivative at  $r = 0$  is always greater than the absolute value of the derivative at  $r = 0.5$ . The darker the color, the larger the difference between the two derivatives.

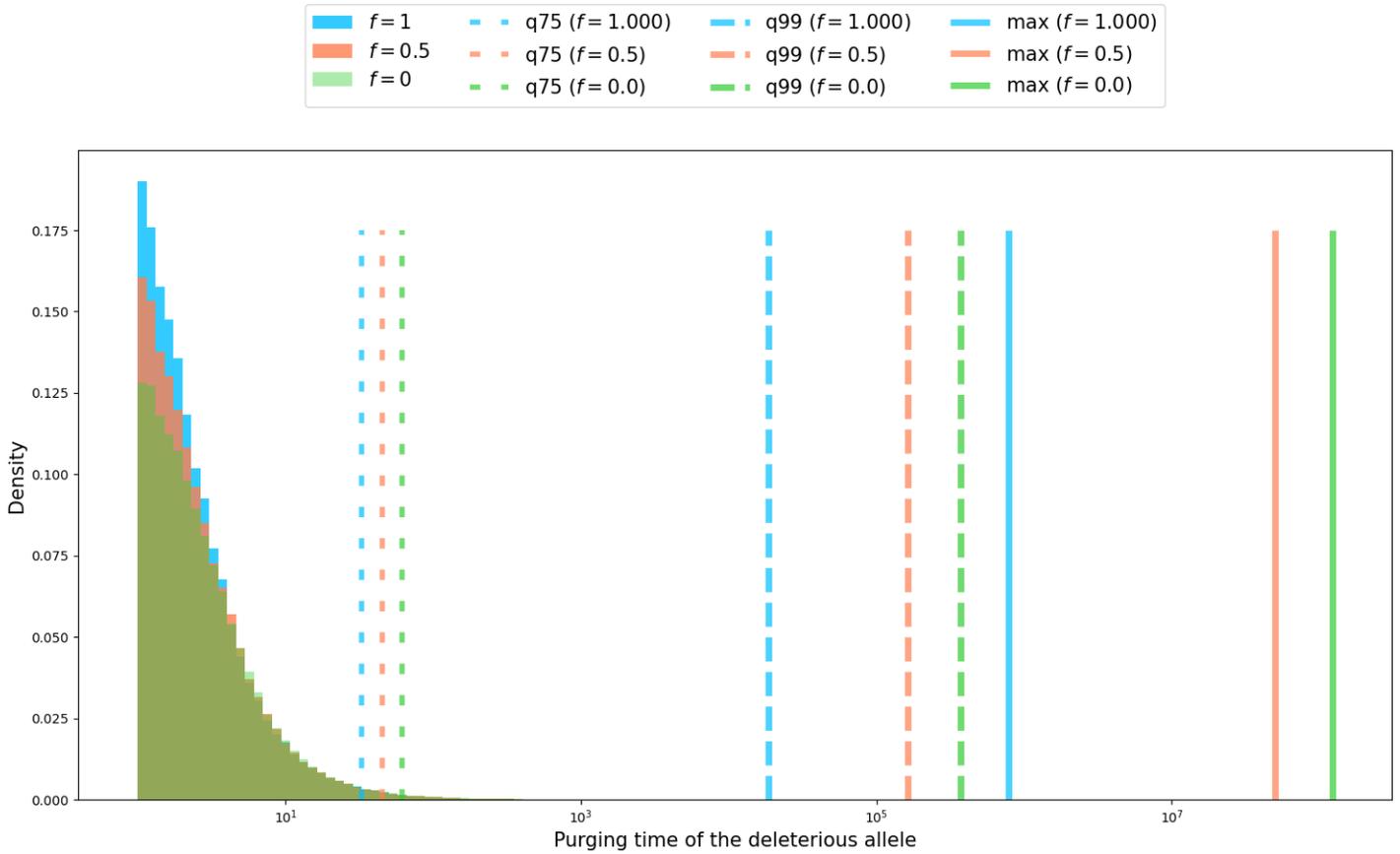


Figure S4: Empirical distribution of the deleterious allele purging time for the partial dominance scenario. A total of 100,000 simulations were run, with  $s = 0.1$ ,  $h = 0.1$ ,  $r = 0.1$ ,  $p_{in} = 0.5$ , starting from one heterozygous individual ( $X_0 = (1, 0, 0)$ ), and for three values of the selfing rate ( $f = 0$  in green,  $f = 0.5$  in red and  $f = 1$  in blue). The respective values for  $\rho$  are  $\rho = -0.0100$ ,  $\rho = -0.0157$  and  $\rho = -0.0818$ . The x-axis is log-scaled. The large-dotted lines represent the 75<sup>th</sup> percentile ( $q75$ ), the dashed lines indicate the 99<sup>th</sup> percentile ( $q99$ ), and solid lines the maximum value ( $max$ ) of the purging time. Maximum values are several order of magnitudes higher than the 75th percentile of the empirical distribution of the purging time.

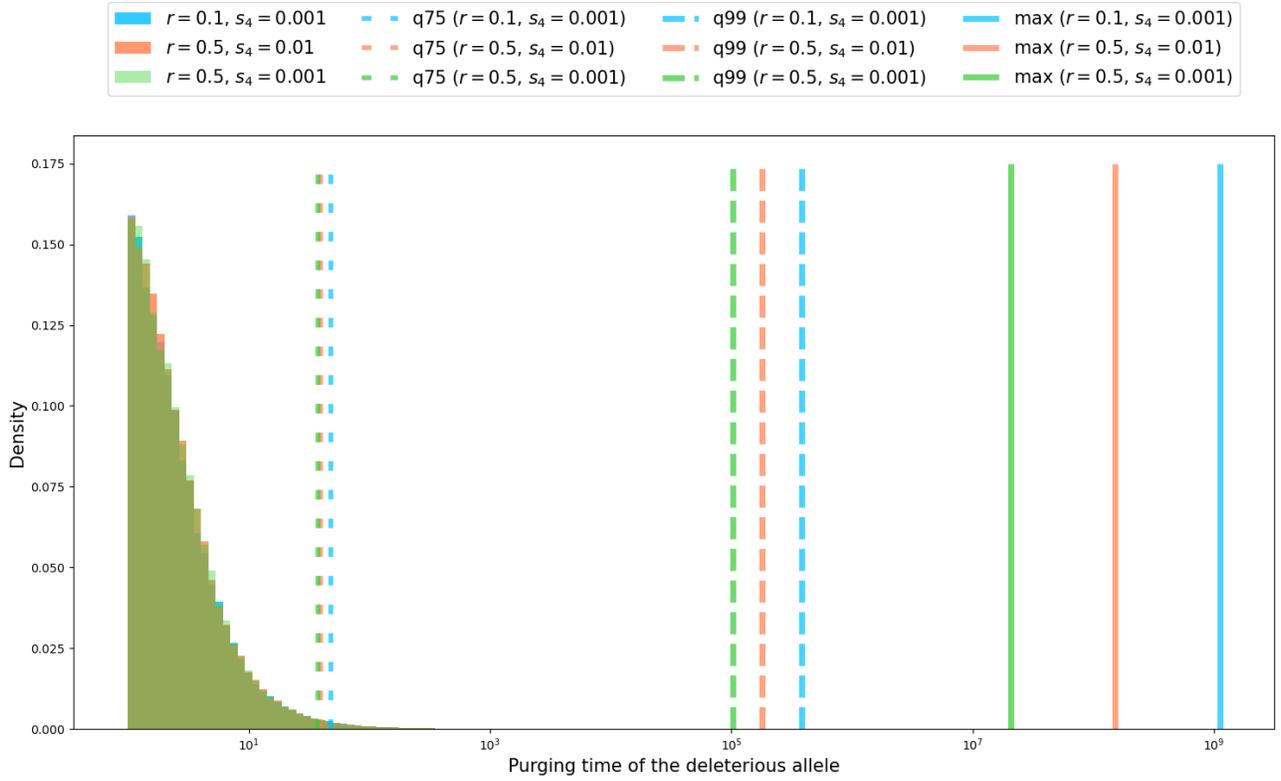


Figure S5: Empirical distribution of the deleterious allele purging time for the overdominance scenario. A total of 100,000 simulations were run, with  $s_3 = 0.1$ ,  $f = 0.5$ ,  $p_{in} = 0.5$ , starting from one heterozygous individual ( $X_0 = (1, 0, 0)$ ), for several values of the recombination rate and of the selection coefficient  $s_4$  ( $r = 0.1, s_4 = 0.001$  in blue,  $r = 0.5, s_4 = 0.01$  in red, and  $r = 0.5, s_4 = 0.001$  in green). The respective values for  $\rho$  are  $\rho = -0.0052$ ,  $\rho = -0.0129$  and  $\rho = -0.0219$ . The parameters were chosen so that the process is sub-critical and thus the purging time is almost surely finite. The x-axis is log-scaled. The large-dotted lines represent the 75<sup>th</sup> percentile (q75), the dashed lines indicate the 99<sup>th</sup> percentile (q99), and solid lines the maximum value (max) of the purging time. Maximum values are several order of magnitudes higher than the 75th percentile of the empirical distribution of the purging time.

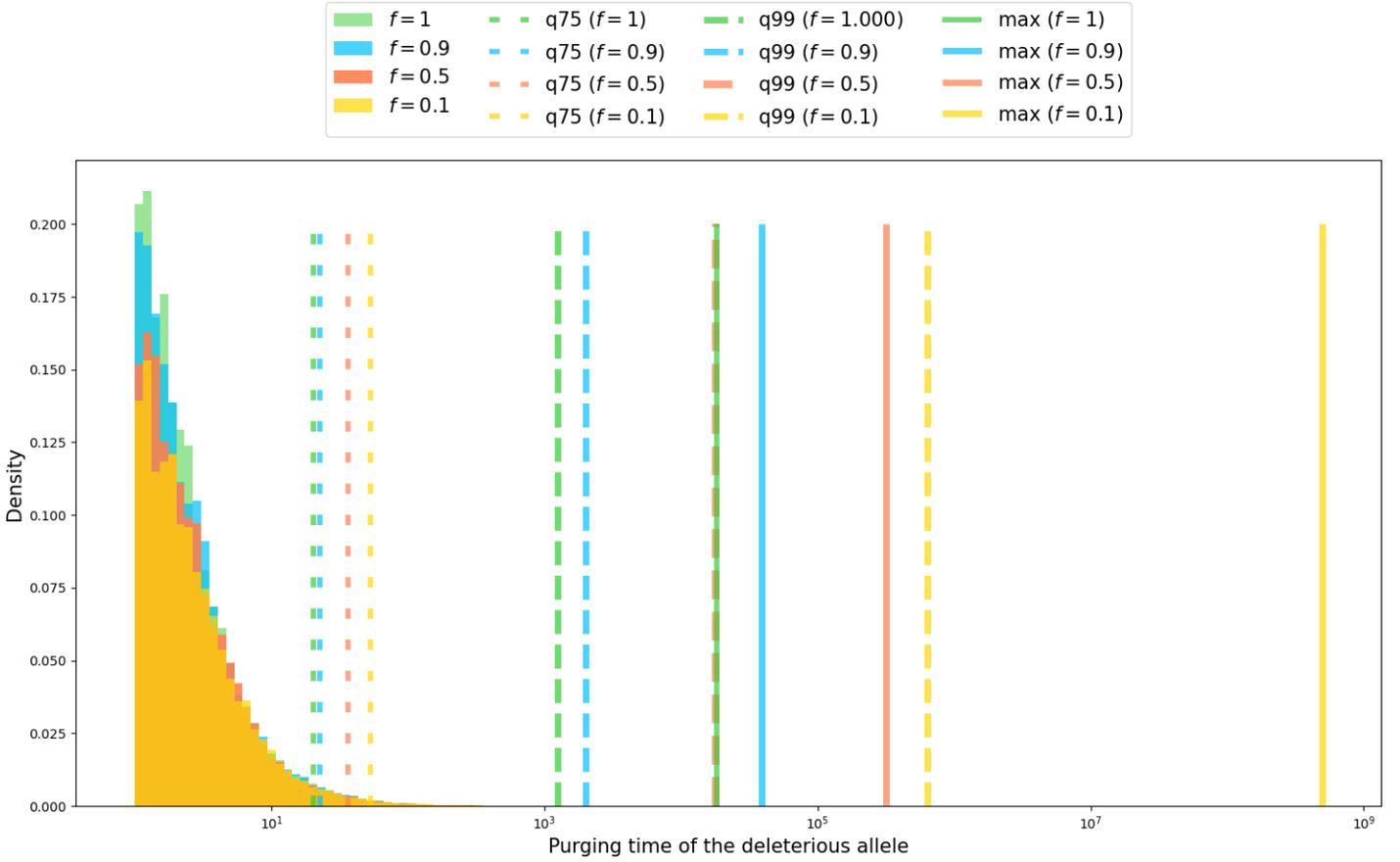


Figure S6: Empirical distribution of the deleterious allele purging time for the overdominant scenario. A total of 100,000 simulations were run, with  $s_3 = 0.5$ ,  $p_{in} = 0.5$ ,  $s_4 = 0.01$ ,  $r = 0.4$ , starting from one heterozygous individual ( $X_0 = (1, 0, 0)$ ), for four values of the selfing rate ( $f = 0.1$  in yellow,  $f = 0.5$  in red,  $f = 0.9$  in blue,  $f = 1$  in green). The respective values for  $\rho$  are  $\rho = -0.0035$ ,  $\rho = -0.0713$ ,  $\rho = -0.1905$  and  $\rho = -0.25$ . Parameters were chosen so that the process is sub-critical and thus the purging time is almost surely finite. The x-axis is log-scaled. The large-dotted lines represent the 75<sup>th</sup> percentile (*q75*), dashed lines indicate the 99<sup>th</sup> percentile (*q99*), and solid lines the maximum value (*max*) of the purging time. Maximum values are several order of magnitudes higher than the 75th percentile of the empirical distribution of the purging time. Note that the selection coefficient for *bb* homozygotes is high ( $s_3 = 0.5$ ).

## 943 8.2 The dominant eigenvalue, its sign and its derivative: partial dominance scenario

### 944 8.2.1 Determination of the dominant eigenvalue

945 The eigenvalues computed with Mathematica (Wolfram Research, 2015) are, for the partial dominance case,

$$\lambda_0 = -r - hs(1 - r), \quad \lambda_+ = \frac{1}{4}(\beta + \sqrt{\Delta}), \quad \lambda_- = \frac{1}{4}(\beta - \sqrt{\Delta}),$$

946 where

$$\beta = f(-r(1 - hs)\alpha + 2(1 - s)) - 2(1 + hs), \quad \alpha = 2 - r - p_{in}(1 - r), \quad (7)$$

947 and

$$\Delta = (\beta + 4hs)^2 - 8fsr\alpha(1-h)(1-hs). \quad (8)$$

948 It is straightforward to see that  $\lambda_+ > \lambda_-$ .

949

950 Let us prove that we also have  $\lambda_+ > \lambda_0$ . We used Geogebra to assist us in the calculations.

$$\lambda_+ > \lambda_0 \quad \text{if and only if (iff)} \quad \frac{1}{4}(\beta + \sqrt{\Delta}) \geq \lambda_0 \quad \text{iff} \quad \sqrt{\Delta} \geq 4\lambda_0 - \beta.$$

951 If  $4\lambda_0 - \beta \leq 0$ , the last inequality is straightforward, as  $\sqrt{\Delta} \geq 0$ .

952 Let us study the sign of  $4\lambda_0 - \beta$ . We define  $P(r) := 4\lambda_0 - \beta = a_2r^2 + a_1r + a_0$ , with

$$a_2 = -f(1-hs)(1-p_{in}) < 0, \quad a_1 = f(1-hs)(2-p_{in}) - 4(1-hs) < 0, \quad \text{and} \quad a_0 = -2f(1-s) + 2(1-hs) > 0.$$

953  $P$  is a second-order polynomial, with negative quadratic coefficient and positive coefficient of order zero (because  
954  $1-hs > 1-s > f(1-s)$ ). Thus,  $P$  admits two roots, one negative and one positive. We denote the positive root  
955 by  $r_P$ . For  $r \in [0, r_P]$ , we have  $P(r) \geq 0$ , and for  $r > r_P$ , we have  $P(r) < 0$ . Consequently, we readily obtain that  
956 when  $r > r_P$ ,  $\lambda_+ > \lambda_0$ .

957 Let us now consider the case  $r \in [0, r_P]$ . For such an  $r$ , using that  $4\lambda_0 - \beta \geq 0$ , we can write that

$$\sqrt{\Delta} \geq 4\lambda_0 - \beta \quad \text{iff} \quad \Delta \geq (4\lambda_0 - \beta)^2.$$

958 Let us write  $Q(r) := (4\lambda_0 - \beta)^2 - \Delta = b_2r^2 + b_1r + b_0$ , with

$$b_2 = f(1-hs)(1-p_{in}) > 0, \quad b_1 = 2(1-hs) - f(1-hs)(2-p_{in}) - fs(1-h)(1-p_{in}),$$

$$\text{and} \quad b_0 = -2(1-hs)(1-f) - p_{in}fs(1-h) < 0.$$

959  $Q$  is a second-order polynomial, with positive quadratic coefficient, and negative coefficient of order 0. Hence,  
960  $Q$  admits two roots, one negative, and one positive. We denote the positive root by  $r_Q$ . In order to prove that  
961  $\lambda_+ > \lambda_0$ , we have to prove that  $Q(r) \leq 0$  when  $P(r) > 0$ , *i.e.* when  $r \in [0, r_P]$ . As  $Q(0) < 0$  and  $Q$  has only  
962 one positive root, proving that  $Q(r_P) < 0$  will imply that  $Q(r) \leq 0$  for  $r \in [0, r_P]$ . Let us prove that  $Q(r_P) < 0$ .  
963 Noting that the quadratic coefficients of  $P$  and  $Q$  are the opposites of one another, we use the equation  $P(r_P) = 0$   
964 to obtain

$$Q(r_P) = r_P \left( -2(1-hs) - fs(1-h)(1-p_{in}) \right) + fs(1-h)(2-p_{in}).$$

965 Seeing  $Q(r_P)$  as an affine function of  $r_P$ , we obtain that the function  $r_P \mapsto Q(r_P)$  admits a unique root, which  
 966 is positive, and that we will denote by  $r_P^0$  :

$$r_P^0 = \frac{fs(1-h)(2-p_{in})}{2(1-hs) + fs(1-h)(1-p_{in})}.$$

967 We wish to prove that  $r_P \geq r_P^0$ , as it implies that  $Q(r_P) \leq 0$ . Having  $r_P \geq r_P^0$  is equivalent to having  $P(r_P^0) \geq 0$ ,  
 968 as  $r_P$  is the unique positive root of  $P$  and  $P(0) \geq 0$ . Consequently, it only remains to prove that  $P(r_P^0) \geq 0$ , which  
 969 is equivalent to

$$\left(2(1-hs) + fs(1-h)(1-p_{in})\right)^2 P(r_P^0) \geq 0.$$

970 To obtain this result, an efficient way is to consider the left-hand term as a polynomial in  $p_{in}$ . Let us write  
 971  $K(p_{in}) = (2(1-hs) + fs(1-h)(1-p_{in}))^2 P(r_0) = c_2 p_{in}^2 + c_1 p_{in} + c_0$ , with

$$c_2 = f^2 s(1-s)(1-fs)(1-h) > 0, \quad c_1 = 2(1-s)f^3 s^2(1-h)^2 > 0$$

and  $c_0 = (1-h)^2 f^2 s^2(1-hs + f(s-1)) + 4(1-f)(1-hs)^2 > 0$ .

972  $K$  is thus a second-degree polynomial in  $p_{in}$ , with a positive quadratic coefficient, and a minimum reached for a  
 973 negative value (minimum reached at  $-c_1/(2c_2) < 0$ ).  $K$  is thus monotonic for positive abscissa, and the coefficient  
 974 of order zero is positive. Consequently, for all  $p_{in} \geq 0$ , we have  $K(p_{in}) \geq 0$ . We have then  $P(r_0) \geq 0$ , which  
 975 concludes the proof that  $\lambda_+ \geq \lambda_0$ .

976 Based on the result we just obtained, from now on we write  $\rho = \lambda_+$ .

977

### 978 8.2.2 Sign of the dominant eigenvalue

979 We prove that  $\rho < 0$ , except when  $s = 0$ , or when  $h = 0$  and  $r = 0$ , in which cases  $\rho = 0$ . Recall the notation  $\alpha, \beta$   
 980 from (7) and  $\Delta$  from (8).

981 First, considering that  $r \in [0, 1]$  and  $p_{in} \in [0, 1]$ , we have  $0 < \alpha < 2$ , which leads to  $\beta < 0$ .

982 When  $s = 0$  or  $(r, h) = (0, 0)$ ,  $\Delta = \beta^2$ , which gives, as  $\beta < 0$ ,  $\sqrt{\Delta} = \sqrt{\beta^2} = |\beta| = -\beta$ . We then have  
 983  $\rho = \frac{1}{4}(\beta - \beta) = 0$ .

984 Let us now consider the case where  $s \neq 0$  and  $(r, h) \neq (0, 0)$ . We have

$$\beta + \sqrt{\Delta} > 0 \quad \text{iff} \quad 0 > \beta > -\sqrt{\Delta} \quad \text{iff} \quad 0 < \beta^2 < \Delta \quad \text{iff} \quad \beta^2 - \Delta < 0.$$

985 Moreover,  $\beta^2 - \Delta = -16f(1-s)hs + 8fr\alpha(1-hs)s + 16hs$ . The sign of  $\rho$  is thus the sign of  $fh(2-r\alpha)s +$   
 986  $2h(1-f) + r\alpha f$ , which is an affine function of  $s$ . The slope and intercept of this function are both non-positive  
 987 when  $(r, h) \neq 0$  or  $s \neq 0$ , which gives  $\rho < 0$  in those cases.

988 **8.2.3 Derivative of the dominant eigenvalue**

989 The derivative of  $\rho$  is

$$\frac{\partial \rho}{\partial r} = \frac{1}{4} \left( \beta'(r) + \frac{1}{2} \frac{\Delta'(r)}{\sqrt{\Delta}} \right),$$

990 Evaluating this derivative at  $r = 0.5$ , and using that  $\beta'(0.5) = -f(1 - hs)$ , we obtain

$$\frac{\partial \rho}{\partial r} \Big|_{r=0.5} = -\frac{1}{4} f(1 - hs) \left( 1 + \frac{\beta(0.5) + 4s}{\sqrt{\Delta}} \right).$$

991 Simple calculations lead to  $\frac{\partial \rho}{\partial r} \Big|_{r=0.5} = 0$  when  $f = 0$ , or  $s = 0$ , or  $s = 1$ , or  $h = 1$ , or  $f = 1$  and  $p_{in} \leq 3 - \frac{8s(1-h)}{1-hs}$ .  
 992 In the latter case, whether the inequality is verified or not determines the sign of  $\Delta$  and therefore the value of  $\sqrt{\Delta}$ ,  
 993 which is either equal to  $\beta(0.5) + 4s$  or  $-(\beta(0.5) + 4s)$ . The derivative is then either equal to zero or strictly negative.

994 For the rest of this paragraph, we study the sign of the derivative when none of the above cases is met.

995 Let us write  $\gamma = \beta(0.5) + 4s$ . If  $\gamma \geq 0$ , we readily obtain  $\frac{\partial \rho}{\partial r} \Big|_{r=0.5} < 0$ . Let us then assume that  $\gamma < 0$ . In this  
 996 case, we have

$$\frac{\partial \rho}{\partial r} \Big|_{r=0.5} < 0 \quad \text{iff} \quad 1 + \frac{\gamma}{\sqrt{\Delta}} > 0 \quad \text{iff} \quad \sqrt{\Delta} > -\gamma.$$

997 As  $-\gamma > 0$ , this comes down to

$$\frac{\partial \rho}{\partial r} \Big|_{r=0.5} < 0 \quad \text{iff} \quad \Delta > \gamma^2 \quad \text{iff} \quad (1 - s)(f - 1) < 0,$$

998 which is indeed satisfied.

In conclusion, we have shown that, in the general case,

$$\frac{\partial \rho}{\partial r} \Big|_{r=0.5} < 0.$$

999 We also compute the derivative at  $r = 0$ . We have  $\beta(0) = 2f(1 - s) - 2(1 + hs)$ ,  $\beta'(0) = -f(1 - hs)(2 - p_{in})$ ,  
 1000  $\Delta(0) = (2f(1 - s) - 2(1 - hs))^2$ , and  $\Delta'(0) = 2\beta'(0)(\beta(0) + 4hs) - 8fs(1 - h)(1 - hs)(2 - p_{in})$ .

1001 After simplification, this gives

$$\frac{\partial \rho}{\partial r} \Big|_{r=0} = -\frac{f(1 - hs)(2 - p_{in})}{4} \left( 1 + \text{sgn}(f(1 - s) - (1 - hs)) + \frac{2s(1 - h)}{|f(1 - s) - (1 - hs)|} \right),$$

1002 with  $\text{sgn}(f(1 - s) - (1 - hs))$  is equal to 1 (respectively to  $-1$ ) when  $f(1 - s) - (1 - hs)$  is positive (resp.  
 1003 negative).

1004 We obtain immediately that

$$\frac{\partial \rho}{\partial r} \Big|_{r=0} \leq 0.$$

## 1005 8.3 The dominant eigenvalue, its sign and its derivative: overdominance scenario

### 1006 8.3.1 Determination of the dominant eigenvalue

The eigenvalues computed with Mathematica (Wolfram Research, 2015) for the overdominant case are

$$\lambda_0 = s_4 - r, \quad \lambda_+ = \frac{1}{4}(\beta + \sqrt{\Delta}), \quad \lambda_- = \frac{1}{4}(\beta - \sqrt{\Delta})$$

1007 with

$$\beta = f(r[p_{in}(1-r) + r - 2] - 2s_3 + 2) + 4s_4 - 2, \quad (9)$$

1008 and

$$\Delta = (\beta - 4s_4)^2 + 8frs_3(p_{in}(1-r) + r - 2) \quad (10)$$

1009 Here again, we obviously have  $\lambda_+ > \lambda_-$ .

1010 We follow the same method as in the partial dominance case to prove that  $\lambda_+ > \lambda_0$ .

We have

$$\lambda_+ > \lambda_0 \quad \text{if and only if (iff)} \quad \frac{1}{4}(\beta + \sqrt{\Delta}) \geq \lambda_0 \quad \text{iff} \quad \sqrt{\Delta} \geq 4\lambda_0 - \beta.$$

1011 If  $4\lambda_0 - \beta \leq 0$ , the last inequality is straightforward, as  $\sqrt{\Delta} \geq 0$ . Let us thus study the sign of  $4\lambda_0 - \beta$ . Let us  
1012 define the function  $P$  by  $P(r) := 4\lambda_0 - \beta = a_2r^2 + a_1r + a_0$ , with

$$a_2 = -f(1 - p_{in}) < 0, \quad a_1 = f(2 - p_{in}) - 4 < 0, \quad \text{and} \quad a_0 = 2(1 - f(1 - s_3)) > 0.$$

1013  $P$  is a second-order polynomial, with a negative quadratic coefficient and a positive coefficient of order 0. Hence  
1014  $P$  admits two roots, one which is negative and one which is positive. We denote the positive root by  $r_P$ . For  
1015  $r \in [0, r_P]$ , we have  $P(r) \geq 0$ , and for  $r > r_P$ , we have  $P(r) < 0$ . Consequently, we readily obtain that when  $r > r_P$ ,  
1016 the conclusion follows.

1017 Let us now consider  $r \in [0, r_P]$ . For such an  $r$ , as  $4\lambda_0 - \beta \geq 0$ , again we have

$$\sqrt{\Delta} \geq 4\lambda_0 - \beta \quad \text{iff} \quad \Delta \geq (4\lambda_0 - \beta)^2.$$

1018 Let us define the function  $Q$  by  $Q(r) := (4\lambda_0 - \beta)^2 - \Delta = b_2r^2 + b_1r + b_0$ , with

$$b_2 = f(1 - p_{in}) > 0, \quad b_1 = f(1 + s_3)(p_{in} - 1) - f + 2, \quad \text{and} \quad b_0 = -fs_3p_{in} - 2(1 - f) < 0.$$

1019  $Q$  is a second-order polynomial, with positive quadratic coefficient, and negative coefficient of order 0. Hence,  
1020  $Q$  admits two roots, one negative and one positive. We denote the positive root by  $r_Q$ . In order to prove that  
1021  $\lambda_+ > \lambda_0$ , we have to prove that  $Q(r) \leq 0$  when  $P(r) > 0$ , *i.e.* when  $r \in [0, r_P]$ . As  $Q(0) < 0$  and  $Q$  has only  
1022 one positive root, proving that  $Q(r_P) < 0$  will imply that  $Q(r) \leq 0$  for  $r \in [0, r_P]$ . Let us prove that  $Q(r_P) < 0$ .  
1023 Noting that the quadratic coefficients of  $P$  and  $Q$  are the opposites of one another, we use the equation  $P(r_P) = 0$   
1024 to obtain

$$Q(r_P) = r_P \left( -2 - fs_3(1 - p_{in}) \right) + fs_3(2 - p_{in}).$$

1025 Seeing  $Q(r_P)$  as an affine function of  $r_P$ , we obtain that the function  $r_P \mapsto Q(r_P)$  admits a unique root, which  
1026 is positive, and that we will denote by  $r_P^0$ :

$$r_P^0 = \frac{fs_3(2 - p_{in})}{2 + fs_3(1 - p_{in})}.$$

1027 We wish to prove that  $r_P \geq r_P^0$ , as it implies that  $Q(r_P) \leq 0$ . Having  $r_P \geq r_P^0$  is equivalent to having  $P(r_0) \geq 0$ ,  
1028 as  $r_P$  is the unique positive root of  $P$  and  $P(0) \geq 0$ . There is thus left to prove that  $P(r_P^0) \geq 0$ , which is equivalent  
1029 to

$$\left( 2 + fs_3(1 - p_{in}) \right)^2 P(r_P^0) \geq 0.$$

1030 To obtain this result, an efficient way is to consider the left-hand term as a polynomial in  $p_{in}$ .

1031 Let us write  $K(p_{in}) = (2 + fs_3(1 - p_{in}))^2 P(r_P^0) = c_2 p_{in}^2 + c_1 p_{in} + c_0$ , with

$$c_2 = f^2 s_3(1 - s_3)(1 - fs_3) > 0, \quad c_1 = 2(1 - s_3)f^3 s_3^2 > 0 \quad \text{and} \quad c_0 = f^2 s_3^2((1 - f)(1 - s_3) + s_3) + 4(1 - f) > 0.$$

1032  $K$  is thus a second-degree polynomial in  $p_{in}$ , with a positive quadratic coefficient and positive coefficient of  
1033 order 0, that reaches its minimum for a negative value (minimum reached at  $-c_1/(2c_2) < 0$ ). Consequently, for all  
1034  $p_{in} \geq 0$ , we have  $K(p_{in}) \geq 0$ . We have then  $P(r_P^0) \geq 0$ , which concludes the proof that  $\lambda_+ \geq \lambda_0$ .

1035 Based on the result we just obtained, from now on we write  $\rho = \lambda_+$ .

### 1036 8.3.2 Sign of the dominant eigenvalue

1037 In this selection scenario,  $\rho$  is not of constant sign.

1038 The condition for  $\rho \geq 0$  is

$$\begin{aligned} & \sqrt{f^2 \left[ r(p_{in}(1 - r) + r - 2) - 2s_3 + 2 \right]^2 + 4f(2(s_3 - 1) - r(2s_3 - 1))((p_{in} - 1)r - p_{in} + 2)} + 4 \\ & + f \left[ r(p_{in}(1 - r) + r - 2) - 2s_3 + 2 \right] + 4s_4 \geq 2. \end{aligned}$$

1039 We compute the dominant eigenvalue and study its sign for simple cases, and then use a numerical approach  
1040 to complete the analysis (Figure 2). Under complete intra-tetrad selfing ( $f = 1, p_{in} = 1$ ), we have  $\rho = s_4 - r/2$  if

1041  $s_3 \geq r/2$  and  $\rho = s_4 - s_3$  if  $s_3 < r/2$ . As  $s_4 \leq s_3$ , the condition to have  $\rho \geq 0$  reduces to  $r \leq 2s_4$ . This is consistent  
1042 with the results of Antonovics and Abrams, 2004, as the authors set  $s_3 = 1$  and thus obtain  $\rho = s_4 - r/2$ . Under  
1043 complete selfing ( $f = 1$ ), if  $r(2 - r - p_{in}(1 - r)) - 2s_3 \geq 0$ , then  $\rho = s_4 - s_3 \leq 0$ . This shows that the value of  
1044 the dominant eigenvalue, and thus the dynamics of the process, depends only on the selection strength when the  
1045 recombination rate  $r$  exceeds a certain threshold. Moreover, this threshold depends only on the selection coefficient  
1046 for homozygous deleterious ( $s_3$ ), and on the probability of intra-tetrad mating ( $p_{in}$ ). This threshold appears on  
1047 the bottom panels of Figure 2. Under complete outcrossing ( $f = 0$ ), we have  $\rho = s_4 \geq 0$ . When the mutation is  
1048 completely linked to a mating-type allele ( $r = 0$ ), we have  $\rho = s_4 \geq 0$ . When the mutation is neutral ( $s_3 = 0$ ,  
1049 implying  $s_4 = 0$  as well), we have  $\rho = 0$ . Finally, when  $BB$  homozygotes are not disfavored ( $s_4 = 0$ ), we have  $\rho < 0$ .  
1050 Indeed, in this case,  $\beta = fr[p_{in}(1 - r) + r - 2] - 2fs_3 + 2(f - 1) < 0$ . We thus have

$$\rho \geq 0 \quad \text{iff} \quad \sqrt{\Delta} \geq -\beta \geq 0 \quad \text{iff} \quad \Delta \geq \beta^2 \quad \text{iff} \quad 8frs_3(p_{in}(1 - r) + r - 2) \geq 0.$$

1051 But we trivially have  $p_{in}(1 - r) + r - 2 \leq 0$ , and so the condition is not met and  $\rho < 0$ .

### 1052 8.3.3 Derivative of the dominant eigenvalue

1053 The derivative of the largest eigenvalue  $\rho$  is

$$\frac{\partial \rho}{\partial r} = \frac{1}{4} \left( \beta'(r) + \frac{1}{2} \frac{\Delta'(r)}{\sqrt{\Delta}} \right).$$

1054 Moreover, we have

$$\beta'(r) = 2f(1 - p_{in})r + f(p_{in} - 2)$$

1055 and

$$\Delta'(r) = 2\beta'(r)(\beta(r) - 4s_4) + 8fs_3(2(1 - p_{in})r + p_{in} - 2).$$

1056 Evaluating these quantities at  $r = 0.5$ , we obtain  $\beta'(0.5) = -f$ , and so

$$\frac{\partial \rho}{\partial r} \Big|_{r=0.5} = -\frac{1}{4}f \left( 1 + \frac{\beta(0.5) - 4s_4 + 4s_3}{\sqrt{\Delta}} \right).$$

1057 Simple calculations lead to  $\frac{\partial \rho}{\partial r} \Big|_{r=0.5} = 0$  when  $f = 0$ , or  $s_3 = 0$ , or  $f = 1$  and  $p_{in} \leq 3 - 8s_3$ .

1058 For the rest of this paragraph, we study the sign of the derivative when none of the above cases is met.

1059 Let us write  $\gamma = \beta(r = 0.5) - 4s_4 + 4s_3$ . If  $\gamma \geq 0$ , we readily obtain that  $\frac{\partial \rho}{\partial r} \Big|_{r=0.5} < 0$ . Let us then assume that  
1060  $\gamma < 0$ . In this case, we have

$$\frac{\partial \rho}{\partial r} \Big|_{r=0.5} < 0 \quad \text{iff} \quad 1 + \frac{\gamma}{\sqrt{\Delta}} > 0 \quad \text{iff} \quad \sqrt{\Delta} > -\gamma.$$

1061 As  $-\gamma > 0$ ,

$$\frac{\partial \rho}{\partial r} \Big|_{r=0.5} < 0 \quad \text{iff} \quad \Delta > \gamma^2 \quad \text{iff} \quad (\beta - 4s_4)^2 + 8frs_3(p_{in}(1-r) + r - 2) > (\beta - 4s_4)^2 + 8s_3(\beta - 4s_4) + 16s_3^2.$$

After some simplifications, we obtain

$$\frac{\partial \rho}{\partial r} \Big|_{r=0.5} < 0 \quad \text{iff} \quad f < \frac{8}{3},$$

1062 which is always satisfied as  $f \in [0, 1]$ .

In conclusion, we have shown that

$$\frac{\partial \rho}{\partial r} \Big|_{r=0.5} < 0.$$

1063 We also compute the derivative at  $r = 0$ . We have  $\beta(0) = 2f(1 - s_3) + 4s_4 - 2$ ,  $\beta'(0) = -f(2 - p_{in})$ ,

1064  $\Delta(0) = (2f(1 - s_3) - 2)^2$ , and  $\Delta'(0) = 2\beta'(0)(\beta(0) - 4s_4) - 8fs_3(2 - p_{in})$ .

1065 After simplification, this gives

$$\frac{\partial \rho}{\partial r} \Big|_{r=0} = -\frac{f(2 - p_{in})}{4} \left( 1 + \text{sgn}(2f(1 - s_3) - 2) + \frac{4s_4}{|2f(1 - s_3) - 4s_4|} \right),$$

1066 with  $\text{sgn}(2f(1 - s_3) - 2)$  is equal to 1 (respectively to  $-1$ ) when  $2f(1 - s_3) - 2$  is positive (resp. negative).

1067 We obtain immediately that

$$\frac{\partial \rho}{\partial r} \Big|_{r=0} \leq 0.$$