



HAL
open science

The evolution of recombination in self-fertilizing organisms

Roman Stetsenko, Denis Roze

► **To cite this version:**

Roman Stetsenko, Denis Roze. The evolution of recombination in self-fertilizing organisms. *Genetics*, 2022, 10.1093/genetics/iyac114 . hal-03768504

HAL Id: hal-03768504

<https://hal.science/hal-03768504>

Submitted on 3 Sep 2022

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

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

The evolution of recombination in self-fertilizing organisms

Roman Stetsenko^{*,†} and Denis Roze^{*,†}

* CNRS, IRL 3614 Evolutionary Biology and Ecology of Algae, 29688 Roscoff,
France

† Sorbonne Université, Station Biologique de Roscoff, 29688 Roscoff, France

Running title: Selfing and recombination

Keywords: epistasis, genetic interference, meiosis, multilocus model, mating systems

Address for correspondence:

Denis Roze

Station Biologique de Roscoff

Place Georges Teissier, CS90074

29688 Roscoff Cedex

France

Phone: (+33) 2 56 45 21 39

Fax: (+33) 2 98 29 23 24

email: roze@sb-roscoff.fr

ABSTRACT

Cytological data from flowering plants suggest that the evolution of recombination rates is affected by the mating system of organisms, as higher chiasma frequencies are often observed in self-fertilizing species compared with their outcrossing relatives.

5 Understanding the evolutionary cause of this effect is of particular interest, as it may shed light on the selective forces favoring recombination in natural populations. While previous models showed that inbreeding may have important effects on selection for recombination, existing analytical treatments are restricted to the case of loosely linked loci and weak selfing rates, and ignore the stochastic effect of genetic interference (Hill-
10 Robertson effect), known to be an important component of selection for recombination in randomly mating populations. In this article, we derive general expressions quantifying the stochastic and deterministic components of selection acting on a mutation affecting the genetic map length of a whole chromosome along which deleterious mutations occur, valid for arbitrary selfing rates. The results show that selfing generally
15 increases selection for recombination caused by interference among mutations as long as selection against deleterious alleles is sufficiently weak. While interference is often the main driver of selection for recombination under tight linkage or high selfing rates, deterministic effects can play a stronger role under intermediate selfing rates and high recombination, selecting against recombination in the absence of epistasis, but
20 favoring recombination when epistasis is negative. Individual-based simulation results indicate that our analytical model often provides accurate predictions for the strength of selection on recombination under partial selfing.

INTRODUCTION

Genetic recombination lies at the heart of the sexual life cycle, and is often considered as one of the main evolutionary benefits of sexual reproduction (Otto, 2021). However, considerable variation for the rate and position of meiotic crossovers along chromosomes exists within and between species (Kong et al., 2010; Johnston et al., 2016; Stapley et al., 2017; Ritz et al., 2017; Brand et al., 2018; Samuk et al., 2020), showing that recombination can evolve over short timescales. Rapid changes in recombination rates have also been observed during artificial selection experiments, in response to selection on recombination itself or on other traits (reviewed in Otto and Barton, 2001), or following major genomic rearrangements such as whole genome duplications (e.g., Wright et al., 2015). Although substantial progress has been achieved over recent years in our understanding of the molecular mechanisms governing crossover formation (Gray and Cohen, 2016; Zelkowski et al., 2019), the evolutionary forces acting on recombination in natural populations remain elusive. In most species, at least one crossover per bivalent seems required to ensure the proper segregation of chromosomes during meiosis I, while data on human trisomies suggest that homologs may fail to separate when they are entangled by too many crossovers (Koehler et al., 1996). These mechanistic constraints probably set lower and upper bounds to the genetic map length of chromosomes, but may leave some space for evolutionary change to occur. The evolution of recombination may also be affected by indirect selective forces, stemming from the effect of recombination on genetic variation. In particular, higher recombination rates may be favored when negative linkage disequilibria (LD) between selected loci exist within populations (i.e., when beneficial alleles at some loci tend to be associated with deleterious alleles at other loci), as recombination then increases the variance in fitness among offspring and the efficiency of natural selection (Otto and Lenormand, 2002; Agrawal, 2006). Different possible sources of negative LD have been identified, including epistatic interactions among loci (Charlesworth, 1990; Barton, 1995) and the Hill-Robertson effect, that tends to generate negative LD between selected loci from the random fluctuations of genotype frequencies occurring in finite populations (Hill and Robertson, 1966; Felsenstein, 1974; Otto and Barton, 1997; Barton and Otto, 2005; Roze and Barton, 2006). Analytical and simulation models

have shown that the stochastic component of selection for recombination (due to the
55 Hill-Robertson effect) may be stronger than deterministic components generated by
epistasis even when population size is rather large, especially when linkage is tight
(Otto and Barton, 2001; Keightley and Otto, 2006; Roze, 2021).

An interesting pattern observed in several genera of flowering plants is that
self-fertilizing species tend to have higher chiasma frequencies than their outcrossing
60 relatives (Roze and Lenormand, 2005; Ross-Ibarra, 2007). Detailed comparisons be-
tween the genetic maps of the selfing *Arabidopsis thaliana* and its outcrossing relative
Arabidopsis lyrata also point to higher recombination rates in *A. thaliana* (Kuittinen
et al., 2004; Hansson et al., 2006; Kawabe et al., 2006). A possible explanation for
higher recombination rates in selfers could be that polymorphism between homologs
65 hinders recombination (Borts and Haber, 1987). However, this hypothesis does not
stand up to closer scrutiny, as available data suggest that substantial levels of diver-
gence are needed to prevent meiotic recombination (Chen and Jinks-Robertson, 1999),
while data from tetraploid rye (Benavente and Sybenga, 2004) and from crosses be-
tween *A. thaliana* strains with different levels of divergence show that crossovers may
70 occur preferentially in heterozygous regions (Ziolkowski et al., 2015; Blackwell et al.,
2020), possibly due to a positive effect of mismatches among homologs on crossover
initiation (Blackwell et al., 2020). Because recombination between homozygous loci
has no genetic effect, selfing reduces the efficiency of recombination in breaking LD
(Nordborg, 2000; Wright et al., 2008), and one may thus expect that increased rates of
75 recombination could evolve to compensate for this effect. However, indirect selection
for recombination should vanish under complete selfing (as heterozygosity should then
be extremely rare), and the effect of selfing on selection for recombination may thus
be non-monotonic. Furthermore, it is not immediately obvious that models for the
evolution of recombination under random mating can be directly transposed to the
80 case of partial selfers. Indeed, simulation models have shown that recombination may
be favored under different conditions in partially selfing than in outcrossing popula-
tions, though the exact mechanisms remained unclear (Charlesworth et al., 1977, 1979;
Holsinger and Feldman, 1983).

A three-locus model of the effect of partial selfing on the evolution of recom-
85 bination was analyzed by Roze and Lenormand (2005). The results showed that cor-

relations in homozygosity across loci caused by partial selfing generate a selective force on recombination that is absent under random mating. By breaking correlations in homozygosity, recombination is favored when dominance-by-dominance epistasis is negative (meaning that double homozygotes have a lower fitness than expected based on the fitness of single homozygotes), as recombination increases the mean fitness of offspring produced by selfing. In the absence of dominance-by-dominance epistasis (or when it is positive), recombination is generally disfavored. The analysis also showed that even a small selfing rate has important consequences, with the effect of breaking correlations in homozygosity quickly becoming the dominant source of indirect selection on recombination. However, the method used by Roze and Lenormand (2005) only holds when effective recombination rates are large, and thus breaks down when the selfing rate is not small, or when loci are tightly linked — yet tightly linked loci should be the ones contributing most to indirect selection for recombination, unless selfing is strong. Furthermore, this model and the previous simulation models mentioned above are deterministic, considering infinite populations. From previous results on selection for recombination in finite, randomly mating populations (Otto and Barton, 2001; Keightley and Otto, 2006; Roze, 2021), and from the fact that the effective size of highly selfing populations may be strongly reduced by interference effects among loci (Glémin and Ronfort, 2013; Roze, 2016), it seems likely that the Hill-Robertson effect should be an important component of selection for recombination in selfing organisms, but this has not been quantified.

In this article, we provide a general analysis of selection for recombination caused by interactions among deleterious alleles, in populations with arbitrary selfing rates. In a first step, we revisit Roze and Lenormand’s (2005) deterministic three locus model, and show that linkage disequilibrium is the main source of selection for recombination in the case of tightly linked loci or under strong selfing, while the effect of correlations in homozygosity stays negligible. In a second step, we explore how the stochastic component of selection for recombination (Hill-Robertson effect) is affected by the mating system, by extending Roze’s (2021) finite population model to partial selfing. Last, we extrapolate the results from the stochastic and deterministic three-locus models in order to quantify the overall strength of selection acting on a modifier allele increasing the map length of a whole chromosome, and compare the predictions

obtained with results from individual-based simulations. The results confirm that the Hill-Robertson effect is often the main component of selection for recombination
120 when the selfing rate is high, usually generating stronger benefits of recombination as selfing increases, but not always. Deterministic effects may become important under intermediate selfing rates and when the chromosomal map length is sufficiently high, and may either increase or decrease selection for recombination depending on the sign and magnitude of the different components of epistasis.

125

METHODS

Our baseline model is the three-locus deterministic recombination modifier model with partial selfing considered by Roze and Lenormand (2005), that we re-analyze in order to obtain more accurate results for arbitrary selfing and recombination rates. In a second step, this model is extended to include the effect of random drift in
130 finite populations. Finally, extrapolations are used to predict the overall strength of selection acting on a modifier affecting the genetic map length of a whole chromosome.

The three-locus model

Genetic architecture

The model considers two selected loci (each with two alleles, A, a and B, b re-
135 spectively) and a recombination modifier locus (with two alleles M and m). Throughout the following, a, b and m will also be used to refer to the three loci, in order to keep the notation simple. All notations used are summarized in Table 1. Alleles a and b are deleterious and affect fitness by a factor $1 - s$ in homozygotes and $1 - sh$ in heterozygotes (s and h thus correspond to the selection and dominance coefficients
140 of alleles a and b). Deleterious alleles are generated by mutation at a rate u per generation; back mutation is ignored. Throughout the paper, we assume that h is significantly greater than zero and that $u \ll s$, so that the equilibrium frequency of deleterious alleles remains small. As in Roze and Lenormand (2005), epistasis between alleles a and b is decomposed into three components (see Table 2): additive-by-additive

145 epistasis $e_{a \times a}$ represents the effect of the interaction between two deleterious alleles, one at each selected locus (either on the same or on different chromosomes), while additive-by-dominance epistasis $e_{a \times d}$ represents the effect of the interaction between three deleterious alleles, and dominance-by-dominance epistasis $e_{d \times d}$ the effect of the interaction between four deleterious alleles. The recombination modifier locus affects
150 the baseline recombination rate r_{ab} between the two selected loci, so that individuals with genotypes MM , Mm and mm at the modifier locus have recombination rates r_{ab} , $r_{ab} + \delta r_{ab} h_m$ and $r_{ab} + \delta r_{ab}$, respectively, with δr_{ab} denoting the effect of the modifier and h_m the dominance coefficient of allele m . Note that because recombination only affects the genotype of meiotic products when it occurs between heterozygous loci,
155 any effect of the modifier on recombination rates between itself and the selected loci (r_{ma} , r_{mb}) will not generate any indirect selection at the modifier locus (e.g., Barton, 1995; Otto and Barton, 1997). The results given throughout the paper are valid for any ordering of the three loci along the chromosome (i.e., either $m - a - b$ or $a - m - b$).

160 *Genetic associations*

The change in frequency at the modifier locus involves different forms of genetic associations between alleles among loci, which are defined as follows. As in Barton and Turelli (1991) and Kirkpatrick et al. (2002), we define indicator variables $X_{j,\emptyset}$ and $X_{\emptyset,j}$ that equal 1 if the lowercase allele j (where j may be m , a or b) is present (or 0 if absent)
165 on the maternally ($X_{j,\emptyset}$) or paternally ($X_{\emptyset,j}$) inherited gene of an individual (note that the terms “maternally” and “paternally inherited” are simply used to differentiate the two haplotypes of an individual). The average of these indicator variables over all individuals in the population gives the frequency of allele j for maternally and paternally inherited genes, $p_{j,\emptyset}$ and $p_{\emptyset,j}$, respectively. The frequency of allele j in the
170 population, p_j , is thus given by $(p_{j,\emptyset} + p_{\emptyset,j})/2$. Centered variables $\zeta_{j,\emptyset}$ and $\zeta_{\emptyset,j}$ are defined as:

$$\zeta_{j,\emptyset} = X_{j,\emptyset} - p_{j,\emptyset}, \quad \zeta_{\emptyset,j} = X_{\emptyset,j} - p_{\emptyset,j}. \quad (1)$$

The genetic association between the sets of alleles \mathbb{U} and \mathbb{V} (that may be \emptyset , a , b , m , ab , ma , mb , mab) present on the maternally and paternally inherited haplotypes of

the same individual is given by:

$$D_{\mathbb{U},\mathbb{V}} = \mathbb{E}[\zeta_{\mathbb{U},\mathbb{V}}], \quad (2)$$

175 with \mathbb{E} the average over all individuals in the population, while

$$\zeta_{\mathbb{U},\mathbb{V}} = \left(\prod_{j \in \mathbb{U}} \zeta_{j,\emptyset} \right) \left(\prod_{k \in \mathbb{V}} \zeta_{\emptyset,k} \right). \quad (3)$$

Because our model does not include any sex-of-origin effect, we always have $D_{\mathbb{U},\mathbb{V}} = D_{\mathbb{V},\mathbb{U}}$. Associations between alleles on the same haplotype $D_{\mathbb{U},\emptyset} = D_{\emptyset,\mathbb{U}}$ will be denoted $D_{\mathbb{U}}$ for simplicity. For example, $D_{a,a}$ measures the excess of homozygotes for allele a (departure from Hardy-Weinberg equilibrium), while D_{ab} is the linkage disequilibrium
180 between alleles a and b .

Life cycle

Each generation starts by selection between newly formed diploid individuals, followed by meiosis and syngamy. The effect of selection on allele frequencies and genetic associations can be computed using the multilocus genetics framework of Kirkpatrick et al. (2002). For this, the fitness of an individual relative to the mean fitness
185 of the population is written as:

$$\frac{W}{\overline{W}} = 1 + \sum_{\mathbb{U},\mathbb{V}} \alpha_{\mathbb{U},\mathbb{V}} (\zeta_{\mathbb{U},\mathbb{V}} - D_{\mathbb{U},\mathbb{V}}), \quad (4)$$

where the “selection coefficients” $\alpha_{\mathbb{U},\mathbb{V}}$ represent the effect of selection acting on the sets of loci \mathbb{U} and \mathbb{V} present on the maternally and paternally inherited haplotypes of an individual. Since we do not assume any sex-of-origin effect we have $\alpha_{\mathbb{U},\mathbb{V}} = \alpha_{\mathbb{V},\mathbb{U}}$,
190 while $\alpha_{\mathbb{U},\emptyset}$ will be denoted $\alpha_{\mathbb{U}}$ for simplicity. The combined effect of selection at loci a and b can thus be represented by 9 coefficients: α_a, α_b represent the effective strength of selection against the deleterious alleles a and b , $\alpha_{a,a}, \alpha_{b,b}$ the effect of dominance at the two loci, while $\alpha_{ab}, \alpha_{a,b}, \alpha_{ab,a}, \alpha_{ab,b}$ and $\alpha_{ab,ab}$ represent epistatic interactions, measured as deviations from additivity. Throughout the paper, we will assume that
195 selection is weak (s is of order ϵ , where ϵ is a small term) while epistasis is weaker ($e_{a \times a}, e_{a \times b}, e_{b \times b}$ of order ϵ^2). Assuming that deleterious alleles stay at low frequency,

and under the fitness matrix given by Table 2, $\alpha_{\mathbb{U},\mathbb{V}}$ coefficients are, to leading order:

$$\begin{aligned}\alpha_a &= \alpha_b \approx -sh, & \alpha_{a,a} &= \alpha_{b,b} \approx -s(1-2h), \\ \alpha_{ab} &= \alpha_{a,b} \approx e_{a \times a} + (sh)^2, & \alpha_{ab,a} &= \alpha_{ab,b} \approx e_{a \times d} + s^2h(1-2h), \\ \alpha_{ab,ab} &\approx e_{d \times d} + s^2(1-2h)^2.\end{aligned}\tag{5}$$

The effect of selection on genetic associations (in terms of $\alpha_{\mathbb{U},\mathbb{V}}$ coefficients) is given by equations 9 and 15 in Kirkpatrick et al. (2002), while the change in frequency of the
200 modifier is given by:

$$\Delta p_m = \sum_{\mathbb{U},\mathbb{V}} \alpha_{\mathbb{U},\mathbb{V}} D_{m\mathbb{U},\mathbb{V}}.\tag{6}$$

After selection, individuals produce gametes to form the zygotes of the next generation. During syngamy, a proportion σ of fertilizations involves two gametes produced by the same parent (selfing), while a proportion $1 - \sigma$ involves gametes sampled at random from the whole population (outcrossing). Recombination and syngamy do not
205 change allele frequencies, but do change genetic associations, their effect being given by equations 13 and 14 in Roze and Lenormand (2005). These equations, together with the equations describing the effect of selection on allele frequencies and genetic associations, have been implemented in a *Mathematica* notebook (available as Supplementary Material) that can be used to automatically generate recursions on these
210 variables.

Approximations

Throughout the paper we assume that the effect of the recombination modifier (δr_{ab}) is weak and compute all results to the first order in δr_{ab} . Our three-locus diploid model can be described by 36 genotype frequencies (leading to 35 independent variables), or alternatively by 3 allele frequencies (p_m, p_a, p_b) and 32 genetic associations.
215 A separation of timescales argument can be used to reduce this large number of variables: in particular, when selection is weak relative to recombination, allele frequencies change slowly while genetic associations are rapidly eroded by recombination. In this case, one can show that genetic associations quickly reach a quasi-equilibrium value, which can be computed by assuming that allele frequencies remain constant (Barton and Turelli, 1991; Nagylaki, 1993; Kirkpatrick et al., 2002). In this quasi-linkage equilibrium (QLE) state, associations can be expressed in terms of allele frequencies and of
220

the parameters of the model, and these expressions can then be plugged into equation 6 to obtain the change in frequency of the modifier. This is the approach used in 225 Roze and Lenormand (2005) to quantify the strength of selection for recombination under partial selfing, assuming that recombination rates are sufficiently large (relative to the strength of selection) and that the selfing rate is not too large (as selfing reduces the effect of recombination). However, more accurate expressions can in principle be obtained in situations where deleterious alleles are maintained at mutation – selection 230 balance: indeed, in this case changes in allele frequencies depend solely on the effect of the recombination modifier, and the QLE approximation thus requires only that δr_{ab} is sufficiently small relative to effective recombination rates (Roze, 2014; Gervais and Roze, 2017; Roze, 2021). This is the approach used in the present paper to obtain approximations that remain valid under low effective recombination (i.e., when 235 recombination rates r_{ij} or when the outcrossing rate $1 - \sigma$ are of order ϵ).

The Hill-Robertson effect

An expression for the strength of selection for recombination due to the Hill-Robertson effect between two deleterious alleles in a randomly mating, diploid population was derived in Roze (2021). Generalizing this analysis to partial selfing would 240 be extremely tedious, as a very large number of stochastic moments of genetic associations and allele frequencies would need to be computed. However, in the case of tightly linked loci (r_{ij} of order ϵ), which should be the ones contributing most to selection for recombination unless the selfing rate is very high, separation-of-timescales arguments can be used to show that the effects of selection against deleterious alleles, recom- 245 bination and drift under partial selfing can be predicted by replacing the dominance coefficient h , recombination rates r_{ij} and the population size N by the effective coefficients $h(1 - F) + F$, $r_{ij}(1 - F)$ and $N/(1 + F)$ in the expressions obtained under random mating, where $F = \sigma/(2 - \sigma)$ is the inbreeding coefficient (e.g., Nordborg, 1997; Glémin and Ronfort, 2013; Roze, 2016). We thus introduced these effective 250 coefficients into the expression derived in Roze (2021) in order to explore how self-fertilization affects selection for recombination generated by the Hill-Robertson effect. As we will see, comparisons with multilocus simulations indicate that this approach often yields correct results.

Multilocus extrapolation

255 Following Roze (2021), the three-locus analysis can be extrapolated to predict the overall strength of selection on a modifier affecting the genetic map length R of a whole chromosome (by an amount δR). For simplicity, we assume that the modifier is located at the mid-point of a linear chromosome and that the position of each crossover is sampled in a uniform distribution along the chromosome (no interference).
 260 Deleterious mutations occur at a rate U per haploid chromosome per generation, and we assume that all mutations have the same selection and dominance coefficients s and h , the position of each new mutation being sampled in a uniform distribution along the chromosome (infinite site model). We make the simplifying assumption that epistatic effects are identical between all pairs of loci across the genome. Neglecting the
 265 effect of interactions between more than two mutations (for simplicity), the strength of indirect selection for recombination can be obtained by integrating the result from the three-locus analysis over the genetic map (see Supplementary Material). When the mean number of deleterious alleles per chromosome (n) is large and with epistasis, more accurate expressions for selection coefficients $\alpha_{\mathbb{U},\mathbb{V}}$ at loci j and k segregating for
 270 deleterious alleles are given by:

$$\alpha_j \approx -sh + 2ne_{a \times a}, \quad \alpha_{j,j} \approx -s(1 - 2h) + 2ne_{a \times d}, \quad (7)$$

$$\alpha_{jk} = \alpha_{j,k} \approx e_{a \times a} + \alpha_j \alpha_k, \quad \alpha_{jk,k} \approx e_{a \times d} + \alpha_j \alpha_{k,k}, \quad \alpha_{jk,jk} \approx e_{a \times d} + \alpha_{j,j} \alpha_{k,k}$$

(see Supplementary Material). An approximate expression for the mean number of deleterious alleles per chromosome n (taking into account the effects of epistasis and selfing) can be obtained from equation 27 in Abu Awad and Roze (2020). We also computed the overall strength of selection for recombination under the diploid model
 275 of synergistic epistasis considered by Charlesworth et al. (1991), in which the fitness of individuals is given by $W = \exp[-(\alpha \tilde{n} + \beta \tilde{n}^2/2)]$, with $\tilde{n} = hn_{\text{he}} + n_{\text{ho}}$ and where n_{he} and n_{ho} are the numbers of heterozygous and homozygous mutations present in the genome of the individual. As shown in Abu Awad and Roze (2020), this is equivalent to setting $e_{a \times a} = -\beta h^2$, $e_{a \times d} = -\beta h(1 - 2h)$ and $e_{d \times d} = -\beta(1 - 2h)^2$ in the present
 280 model.

Individual-based simulations

Our simulation program (written in C++ and available from Zenodo) is equivalent to the program used in Roze (2021), with the addition of partial selfing and the different forms of epistasis. It represents a population of N diploids, each carrying a pair of linear chromosomes. At each generation, the number of new deleterious mutations per chromosome is drawn from a Poisson distribution with parameter U , and their position along the chromosome is drawn from a uniform distribution between 0 and 1. The fitness of each individual is computed as:

$$W = W_c (1 - sh)^{n_{he}} (1 - s)^{n_{ho}} (1 + e_{a \times a})^{n_2} (1 + e_{a \times d})^{n_3} (1 + e_{d \times d})^{n_4} \quad (8)$$

where n_{he} and n_{ho} are the number of heterozygous and homozygous deleterious mutations in the genome of the individual, and where n_2 , n_3 and n_4 are the number of interactions between 2, 3 and 4 deleterious alleles at two loci, given by:

$$\begin{aligned} n_2 &= \frac{1}{2} n_{he} (n_{he} - 1) + 2n_{he} n_{ho} + 2n_{ho} (n_{ho} - 1), \\ n_3 &= n_{he} n_{ho} + 2n_{ho} (n_{ho} - 1), \\ n_4 &= \frac{1}{2} n_{ho} (n_{ho} - 1) \end{aligned} \quad (9)$$

(e.g., Abu Awad and Roze, 2020). The term W_c corresponds to a direct fitness effect of the chromosome map length R ; as in Roze (2021), it is set to $W_c = \exp(-cR)$, so that c measures a direct fitness cost per crossover. This parameter reflects the fertility cost of having too many crossovers (see Introduction); it also ensures that R does not evolve towards very large values, and allows simple comparisons with analytical predictions.

To produce individuals of the next generation, parents are sampled according to their fitness, each new individual being produced by selfing with probability σ . The recombination modifier locus is located at the midpoint of the chromosome, with an infinite number of possible alleles coding for different values of R (the map length of the individual being given by the average of its two modifier alleles). One replicate was performed for each parameter set. During the first 20,000 generations R is fixed to 1 in order to reach mutation–selection balance for deleterious alleles. Then, for an extra 5×10^6 generations (increased up to 5×10^7 generations for high values of σ), mutations

are introduced at a rate 10^{-4} at the modifier locus, each mutation multiplying the value of R by a random number drawn from a Gaussian distribution with mean 1 and variance 0.04. To allow for large effect mutations, a proportion 0.05 of mutations have an additive effect on R drawn from a uniform distribution between -1 and 1 (the new value being set to zero if it is negative). The average map length \bar{R} , average fitness \bar{W} , average number of deleterious mutations per chromosome n and number of fixed mutations are recorded every 500 generations (fixed mutations are removed from the population in order to reduce execution time). The equilibrium value of map length is computed as the time average of \bar{R} , after removing the first 5×10^5 generations to allow \bar{R} to equilibrate. In order to study the effect of variable selection coefficients of deleterious mutations, we modified the above baseline simulation program so that each new mutation is associated with a value of s drawn from a log-normal distribution (ensuring $s > 0$): the value of $\ln(s)$ is drawn from a Gaussian distribution with variance sd^2 and mean $\ln(\bar{s}) - sd^2/2$ (so that the average selection coefficient \bar{s} stays constant for different values of sd).

RESULTS

The deterministic three-locus model

We first reiterate the result obtained by Roze and Lenormand (2005) under high effective recombination in a more general version, and then provide a new analysis for the case of weak effective recombination.

High effective recombination

As found by Roze and Lenormand (2005), the joint effects of selfing and the recombination modifier generate an association $D_{mab,ab}$ even in the absence of selection. This stems from the fact that recombination tends to break correlations in homozygosity between loci generated by partial selfing (see also Roze, 2009). The association $D_{mab,ab}$ has the sign of $-\delta r_{ab}$, reflecting the fact that the modifier allele increasing recombination tends to be found on genetic backgrounds in which the correlation in

homozygosity between loci a and b is relatively weaker. At QLE and to the first order in δr_{ab} , it is given by:

$$D_{mab,ab}^0 \approx -\frac{2\delta r_{ab}(1-2r_{ab})F(1-F)[H_m(1+2F\gamma_{ab})+F(\gamma_{ma}+\gamma_{mb}-\gamma_{ab})]}{[1+F(\gamma_{ma}+\gamma_{mb}+\gamma_{ab})](1+2F\gamma_{ab})^2} pq_{mab}, \quad (10)$$

335 where the superscript “0” indicates that the effect of selection at loci a and b is neglected in this expression, and where $H_m = h_m + p_m(1 - 2h_m)$, $\gamma_{ij} = r_{ij}(1 - r_{ij})$ and $pq_{mab} = p_m q_m p_a q_a p_b q_b$ (with $q_j = 1 - p_j$). Equation 10 generalizes equation 32 in Roze and Lenormand (2005) to arbitrary h_m and any ordering of the three loci along the chromosome. It shows that $D_{mab,ab}^0$ vanishes under random mating ($F = 0$) and
 340 full selfing ($F = 1$), because in these cases the mating system does not generate any correlation in homozygosity among loci.

The other associations in equation 6 are generated by selection against the deleterious alleles (of order ϵ) and by the modifier effect. At QLE, associations $D_{mi,i}$, $D_{mij,i}$ and $D_{mi,ij}$ (where i, j are either a or b) are of order ϵ , while associations D_{mi} ,
 345 $D_{m,i}$, D_{mij} , $D_{m,ij}$ and $D_{mi,j}$ are of order ϵ^2 (expressions for these associations are given in Appendix A). As a result, one obtains from equations 5 and 6 that the change in frequency of the modifier is, to leading order:

$$\Delta p_m \approx \alpha_{a,a} D_{ma,a} + \alpha_{b,b} D_{mb,b} + \alpha_{ab,ab} D_{mab,ab}^0, \quad (11)$$

the association $D_{ma,a}$ being given by:

$$D_{ma,a} \approx \alpha_{b,b} F D_{mab,ab}^0, \quad (12)$$

(and symmetrically for $D_{mb,b}$). $D_{ma,a}$ is positive when $\delta r_{ab} > 0$ and $h < 1/2$ (since $\alpha_{b,b}$
 350 and $D_{mab,ab}^0$ are both negative in this case), reflecting the fact that the modifier allele m increasing recombination tends to be found on more homozygous backgrounds at locus a . Indeed when $h < 1/2$, homozygous genotypes at loci a and b have, on average, a lower fitness than heterozygous genotypes, and the correlation in homozygosity generated by partial selfing increases the efficiency of selection, lowering the frequency
 355 of homozygous genotypes. As the modifier allele increasing recombination tends to break this correlation, it is associated to a relative excess in homozygosity at each selected locus. This effect disfavors recombination (since homozygotes have a lower

fitness than heterozygotes on average), which is reflected by the first two terms of equation 11. Equations 11 and 12 give:

$$\Delta p_m \approx (2\alpha_{a,a}\alpha_{b,b}F + \alpha_{ab,ab}) D_{mab,ab}^0, \quad (13)$$

360 which, using equation 5, becomes:

$$\Delta p_m \approx [s^2(1-2h)^2(1+2F) + e_{d \times d}] D_{mab,ab}^0, \quad (14)$$

yielding Equation 36 in Roze and Lenormand (2005). Equations 10 and 14 show that increased recombination is favored when dominance-by-dominance epistasis ($e_{d \times d}$) is sufficiently negative. Indeed, in this case, breaking correlations in homozygosity (thus increasing the frequency of genotypes that are homozygous at one selected locus and
365 heterozygous at the other) tends to increase the mean fitness of offspring.

Low effective recombination

In the previous analysis, terms in $r_{mi}(1-F)$ appear in the denominators of the expressions for D_{mi} and $D_{m,i}$ at QLE (see Appendix A), causing these expressions to diverge (i.e., tend to infinity) as effective recombination rates $r_{mi}(1-F)$ tends to
370 zero. Similarly, terms in $r_{mab}(1-F)$ appear in the denominators of D_{mab} , $D_{m,ab}$, $D_{ma,b}$ and $D_{mb,a}$, where $r_{mab} = (r_{ma} + r_{mb} + r_{ab})/2$ is the probability that at least one recombination event occurs between the three loci. This indicates that these associations should play a more important role in the case of tightly linked loci. In order to explore this regime, we re-analyzed the model in the case where all recombination
375 rates are of order ϵ (see Supplementary Material). This analysis shows that the effect of the linkage disequilibrium D_{ab} between deleterious alleles (which was negligible under high effective recombination) becomes predominant when loci are tightly linked. A general expression for D_{ab} at equilibrium in terms of $\alpha_{\mathbb{U},\mathbb{V}}$ coefficients is given by:

$$D_{ab} \approx \frac{\tilde{\alpha}_{ab} - \tilde{\alpha}_a\tilde{\alpha}_b + 2FG_{ab}(\alpha_a + \alpha_{a,a})(\alpha_b + \alpha_{b,b})}{\tilde{r}_{ab} - \tilde{\alpha}_a - \tilde{\alpha}_b} pq_{ab}, \quad (15)$$

where the tilde denotes “effective coefficients”: $\tilde{r}_{ab} = r_{ab}(1-F)$, $\tilde{\alpha}_a = \alpha_a(1+F) +$
380 $\alpha_{a,a}F$ (and similarly for $\tilde{\alpha}_b$), while $\tilde{\alpha}_{ab} = \alpha_{ab}(1+\phi_{ab}) + 2\alpha_{a,b}F + (\alpha_{ab,a} + \alpha_{ab,b})(F + \phi_{ab}) +$
 $\alpha_{ab,ab}\phi_{ab}$, where ϕ_{ab} is the probability of joint identity-by-descent at the two loci (which is approximately F when r_{ab} is small, see Appendix A). Finally, G_{ab} in equation 15 refers to the identity disequilibrium between the two loci, defined as $G_{ab} = \phi_{ab} - F^2$

(Weir and Cockerham, 1973), and thus approximately equal to $F(1 - F)$ under tight
 385 linkage. Equation 15 simplifies to $(\alpha_{ab} - \alpha_a\alpha_b)pq_{ab}/(r_{ab} - \alpha_a - \alpha_b)$ in the absence of
 selfing; this is equivalent to the result obtained by Barton (1995) under strong re-
 combination (equation 9b in Barton, 1995), except that α_a and α_b now appear in the
 denominator, due to our assumption that r_{ab} is small (of order ϵ). With selfing, two
 important differences appear: (i): the recombination rate and selection coefficients are
 390 replaced by effective coefficients \tilde{r}_{ab} , $\tilde{\alpha}_a$, $\tilde{\alpha}_b$, $\tilde{\alpha}_{ab}$ (since increased homozygosity affects
 both the effects of recombination and of selection, as explained in Appendix B), and
 (ii): an extra term, involving the identity disequilibrium G_{ab} , appears in the numerator
 and tends to generate positive linkage disequilibrium between deleterious alleles.

Using the expressions for $\alpha_{U,V}$ coefficients given by equation 5, equation 15
 395 simplifies to:

$$D_{ab} \approx \frac{s^2(1-h)^2(1+2F)G_{ab} + \tilde{e}}{\tilde{r}_{ab} + 2s\tilde{h}}pq_{ab}, \quad (16)$$

with $\tilde{e} = e_{a \times a}(1 + 2F + \phi_{ab}) + 2e_{a \times d}(F + \phi_{ab}) + e_{d \times d}\phi_{ab}$ and $\tilde{h} = h(1 - F) + F$. The
 fact that partial selfing generates positive linkage disequilibrium between deleterious
 alleles in the absence of epistasis has been noticed in previous analytical and simulation
 studies (Roze and Lenormand, 2005; Kamran-Disfani and Agrawal, 2014), and may
 400 be understood as follows. Lineages with different histories of inbreeding coexist in
 a partially selfing population: lineages that have been inbred for many generations
 tend to be very homozygous (at all loci), while lineages that have been inbred for
 fewer generations tend to be less homozygous. This is the basis of correlations in
 homozygosity among loci, represented by the identity disequilibrium G_{ab} . Because
 405 homozygosity increases the efficiency of selection, the frequency of deleterious alleles
 tends to be lower within lineages that have been inbred for longer (purging), and higher
 in less inbred lineages, resulting in positive linkage disequilibrium between deleterious
 alleles. As shown by equation 16, D_{ab} may become negative when epistasis is negative
 and sufficiently strong, the relative importance of $e_{a \times d}$ and $e_{d \times d}$ increasing as the
 410 selfing rate increases.

The fact that the allele associated with higher recombination tends to erode
 D_{ab} more rapidly in turn generates genetic associations between the modifier and the

selected loci. In particular, we have:

$$D_{mab} \approx -\frac{\delta\tilde{r}_{ab}H_m D_{ab}}{\tilde{r}_{mab} - \tilde{\alpha}_a - \tilde{\alpha}_b} pq_m, \quad D_{ma} \approx \frac{\tilde{\alpha}_b D_{mab}}{\tilde{r}_{ma} - \tilde{\alpha}_a} \quad (17)$$

with $\delta\tilde{r}_{ab} = \delta r_{ab}(1-F)$, $\tilde{r}_{mab} = r_{mab}(1-F)$, and again $\tilde{\alpha}_a = \tilde{\alpha}_b \approx -s\tilde{h}$, $H_m =$
 415 $h_m + p_m(1-2h_m)$ (D_{mb} is given by a symmetric expression). These equations take
 the same form as under random mating (e.g., equations 9c and 11 in Barton, 1995)
 and are interpreted in the same way: when $D_{ab} > 0$, AB and ab haplotypes are in
 excess in the population, and the allele increasing recombination (allele m if $\delta r_{ab} > 0$)
 tends to reduce this excess, and thus becomes more associated with Ab , aB haplotypes
 420 (which is reflected by a negative value of D_{mab}). By contrast, when $D_{ab} < 0$, the allele
 increasing recombination becomes more associated with ab , AB haplotypes (which is
 reflected by a positive value of D_{mab}). When $D_{mab} < 0$, selection is less efficient in the
 m background (because the frequency of extreme haplotypes AB and ab is lower in
 this background), leading to positive values of D_{ma} , D_{mb} (deleterious alleles are more
 425 frequent in the m background). When $D_{mab} > 0$, selection is more efficient in the
 m background (because the frequency of extreme haplotypes AB and ab is higher in
 this background), and m thus becomes better purged from deleterious alleles (which
 is reflected by negative values of D_{ma} , D_{mb}).

Under weak recombination, separation of timescale arguments can be used to
 430 express the other associations that appear in equation 6 in terms of D_{mab} , D_{ma} and
 D_{mb} (e.g., Roze, 2016); one obtains:

$$\begin{aligned} D_{mab,ab} &\approx D_{mab,a} \approx D_{ma,ab} \approx D_{m,ab} \approx D_{ma,b} \approx F D_{mab}, \\ D_{m,a} &\approx D_{ma,a} \approx F D_{ma}, \quad D_{m,b} \approx D_{mb,b} \approx F D_{mb}. \end{aligned} \quad (18)$$

Plugging these into equation 6 yields:

$$\Delta p_m \approx \tilde{\alpha}_a D_{ma} + \tilde{\alpha}_b D_{mb} + \tilde{\alpha}_{ab} D_{mab}, \quad (19)$$

which again takes the same form as under random mating (equation 8 in Barton, 1995).

From equations 17 and 19, one obtains:

$$\Delta p_m \approx -\frac{\delta\tilde{r}_{ab}H_m D_{ab} pq_m}{\tilde{r}_{mab} - \tilde{\alpha}_a - \tilde{\alpha}_b} \left[\tilde{\alpha}_{ab} + \tilde{\alpha}_a \tilde{\alpha}_b \left(\frac{1}{\tilde{r}_{ma} - \tilde{\alpha}_a} + \frac{1}{\tilde{r}_{mb} - \tilde{\alpha}_b} \right) \right], \quad (20)$$

435 equivalent to equation 12 in Barton (1995). Given that the equilibrium frequency of
 deleterious alleles is approximately $u/(s\tilde{h})$, thus of order u/ϵ , equations 16 and 20

show that under weak effective recombination (i.e., when \tilde{r} coefficients are of order ϵ), selection for recombination generated by the linkage disequilibrium D_{ab} is of order $\delta r_{ab} u^2 / \epsilon$, thus stronger than selection for recombination generated by the term in $D_{mab,ab}^0$ seen in the previous subsection (of order $\delta r_{ab} u^2$, see equation 14). When \tilde{r} coefficients are not small (of order 1), however, indirect selection generated by D_{ab} becomes of order $\delta r_{ab} u^2 \epsilon^2$ (and thus negligible). A general expression for the change in frequency of the modifier, valid under both weak and strong effective recombination, can thus be obtained by summing equations 14 and 20. We noticed that in some cases, more accurate expressions can be obtained by taking into account indirect effects of $D_{mab,ab}^0$ on other genetic associations between the modifier and selected loci (these terms should be negligible under both weak and strong effective recombination, but improve the precision of our approximations for intermediate values of recombination rates); these expressions are given in Appendix C. Two further improvements to our approximations are made in Appendix C, leading to more accurate expressions when epistasis is of the same order of magnitude as the strength of selection ($e_{a \times a}$, $e_{a \times d}$, $e_{d \times d}$ of order ϵ) and when the selfing rate is high (see Supplementary Material for derivations).

Figures 1 and S1 show that our approximations provide correct predictions for the change in frequency at the modifier locus, for all values of the selfing rate between 0 and 1. The dots in these figures correspond to the results of deterministic simulations, obtained by iterating exact recurrence equations for the 36 genotype frequencies: allele M is fixed during the first 3000 generations to reach mutation–selection balance at the selected loci, then allele m is introduced in frequency 0.01 and the population is allowed to evolve for an extra 1000 generations, $\Delta p_m / pq_{mab}$ being averaged over the last 500 generations (see Supplementary Material). As can be seen in Figure 1, indirect selection is mostly driven by D_{ab} when recombination rates are small (left figures) or when the selfing rate approaches 1. In the absence of epistasis, the identity disequilibrium generates positive D_{ab} which disfavors recombination (as breaking positive D_{ab} reduces the variance in fitness, see Figures 1A, 1B). Negative epistasis generates negative D_{ab} , which favors increased recombination when effective recombination rates are sufficiently small (equation 20, Figures 1C-F). The relative effect of the term in $D_{mab,ab}^0$ becomes more important in the case of loosely linked loci and for intermediate selfing

rates, either when epistasis is absent (in which case it disfavors recombination, see
 470 equation 14, Figure 1B), or when the dominance-by-dominance component of epistasis
 is important (in which case it favors increased recombination when $e_{d \times d} < 0$, Figure
 1F). Finally, selection on recombination vanishes under complete selfing ($\sigma = 1$), as
 recombination is ineffective in that case.

The Hill-Robertson effect

475 An expression for the strength of selection for recombination caused by the
 Hill-Robertson effect between two deleterious alleles in a diploid, randomly mating
 population was derived by Roze (2021). This expression consists of a sum of terms
 corresponding to different mechanisms generating selection for recombination, that all
 take the form $\delta r_{ab} (sh)^2 u^2 / [N_e \prod_{i=1}^5 (r_{\mathbb{U}_i} + x_i sh)]$ where N_e is the effective population
 480 size, \mathbb{U}_i is either mab , ma , mb or ab , and x_i equals 1, 2, 3 or 4. In the case of loosely
 linked loci ($r_{\mathbb{U}_i} \gg sh$), the terms in sh in the denominator may be neglected and
 the strength of selection for recombination increases as sh increases — being roughly
 proportional to $(sh)^2$. In the case of tightly linked loci ($r_{\mathbb{U}_i} \ll sh$), however, recombi-
 nation rates $r_{\mathbb{U}_i}$ may be neglected in the denominator and selection for recombination
 485 now decreases as sh increases — being roughly proportional to $1/(sh)^3$.

In order to extend these results to partially selfing populations, we replaced
 δr_{ab} , $r_{\mathbb{U}_i}$ and h by the effective parameters $\delta \tilde{r}_{ab} = \delta r_{ab} (1 - F)$, $\tilde{r}_{\mathbb{U}_i} = r_{\mathbb{U}_i} (1 - F)$ and
 $\tilde{h} = h(1 - F) + F$ in the expression derived in Roze (2021). The effect of increas-
 ing selfing (thus increasing F) on the strength of selection for recombination due to
 490 the Hill-Robertson effect can be understood using the same reasoning as above. When
 effective recombination rates are large relative to the strength of selection against dele-
 terious alleles (selfing rate not too large, $r_{\mathbb{U}_i}$ not too small so that $\tilde{r}_{\mathbb{U}_i} \gg s\tilde{h}$), selection
 for recombination becomes approximately proportional to $\delta \tilde{r}_{ab} (s\tilde{h})^2 / [N_e \prod_{i=1}^5 \tilde{r}_{\mathbb{U}_i}] =$
 $\delta r_{ab} (s\tilde{h})^2 / [N_e (1 - F)^4 \prod_{i=1}^5 r_{\mathbb{U}_i}]$, which increases as F increases (mostly due to the
 495 decreased effective recombination rates $\tilde{r}_{\mathbb{U}_i}$ at the denominator, but also to the increase
 in \tilde{h}). When effective recombination rates are small relative to the strength of selection
 (high selfing and/or tightly linked loci, so that $\tilde{r}_{\mathbb{U}_i} \ll s\tilde{h}$), selection for recombination
 is approximately proportional to $\delta \tilde{r}_{ab} / [N_e (s\tilde{h})^3] = \delta r_{ab} (1 - F) / [N_e (s\tilde{h})^3]$, which de-

creases as F increases (mostly due to the decreased effect of the modifier $\delta\tilde{r}_{ab}$, but
 500 also to the increase in \tilde{h}). One thus predicts that increasing selfing should generally
 have a non-monotonic effect on selection for recombination due to the Hill-Robertson
 effect (selection for recombination first increasing, and then decreasing towards zero
 as F increases from 0 to 1), the position of the maximum depending on sh and on
 the values of recombination rates (however, selfing may always decrease selection for
 505 recombination when sh is sufficiently strong and/or linkage sufficiently tight). These
 verbal predictions are illustrated by Figure 2. Note that selfing also has the additional
 effect of reducing N_e by a factor $1/(1+F)$ (Pollak, 1987; Nordborg, 2000), thus in-
 creasing the strength of the Hill-Robertson effect, but this effect stays minor relative
 to the effect of selfing on effective recombination rates. Selfing may cause stronger
 510 reductions in N_e when deleterious alleles are segregating at many linked loci, however,
 through background selection effects (Glémin and Ronfort, 2013; Roze, 2016).

These results also provide us with some understanding of the relative impor-
 tance of stochastic and deterministic sources of selection for recombination: when
 effective recombination rates are large, indirect selection due to the Hill-Robertson ef-
 515 fect is of order $\delta r_{ab} u^2 \epsilon^2 / N_e$, and should thus be negligible relative to the deterministic
 component (of order $\delta r_{ab} u^2$). When effective recombination rates are small (of order
 ϵ), however, selection due to the Hill-Robertson effect is now of order $\delta r_{ab} u^2 / (N_e \epsilon^3)$,
 and may thus become of the same order of magnitude or stronger than deterministic
 terms (which are then of order $\delta r_{ab} u^2 / \epsilon$, as shown in the previous subsection).

520 **Multilocus extrapolation**

The three-locus analysis can be extended to compute the overall strength of
 indirect selection acting on a modifier affecting the genetic map length of a whole
 chromosome. Neglecting the effect of interactions between deleterious alleles at more
 than two loci, this can be done by integrating the result from the three-locus model
 525 over all possible positions of alleles a and b . As in Roze (2021), we consider a linear
 chromosome with map length R (in Morgans), along which deleterious mutations occur
 at a rate U per chromosome per generation. The modifier is located at the mid-point of
 the chromosome, allele m increasing map length by an amount $\delta R/2$ when heterozygous

and δR when homozygous (see Methods). The strength of selection for allele m , defined
 530 as $s_m = \Delta p_m / (\frac{1}{2} p_m q_m)$, can be decomposed into three terms:

$$s_m = s_{\text{direct}} + s_{\text{det}} + s_{\text{HR}} \quad (21)$$

where s_{direct} corresponds to the effect of direct selective pressures acting on R (due to
 any direct effect of R on the fitness of individuals), s_{det} to indirect selection caused
 by deterministic interactions between deleterious alleles, and s_{HR} to indirect selection
 caused by the Hill-Robertson effect. Assuming a direct fitness cost of crossovers so
 535 that fitness decreases as e^{-cR} as R increases, the direct selection term is given by
 $s_{\text{direct}} = -c \delta R (1 + F)$ to the first order in δR (e.g., Gervais and Roze, 2017). From
 the analysis above, the term s_{det} can be further decomposed into $s_{D_{ab}} + s_{D_{mab,ab}^0}$, where
 $s_{D_{ab}}$ corresponds to the overall effect of deterministically generated linkage disequilibria
 (D_{ab}) between deleterious alleles (either by epistasis or by identity disequilibria),
 540 given by equations 15 and 20, while $s_{D_{mab,ab}^0}$ corresponds to the overall effect of as-
 sociations of the form $D_{mab,ab}$ generated by correlations in homozygosity and by the
 modifier effect, given by equations 10 and 13. Because $s_{D_{ab}}$ should be mostly driven by
 tightly linked loci, recombination rates are approximated by genetic distances in equa-
 tions 15 and 20 (before integrating over the genetic map), while δr_{ab} is approximated
 545 by $\delta R r_{ab} / R$ (Roze, 2021). By contrast, loosely linked loci can make a stronger contri-
 bution to $s_{D_{mab,ab}^0}$, and the recombination rate between loci i and j in equation 10 is
 thus expressed in terms of the genetic distance d_{ij} between these loci using Haldane's
 mapping function $r_{ij} = [1 - e^{-2d_{ij}}] / 2$ (Haldane, 1919), while δr_{ab} is approximated
 by $\delta R e^{-2d_{ab}} d_{ab} / R$ (see Supplementary Material). Note that when U is not small (so
 550 that the mean number of deleterious alleles per chromosome n may be large), and with
 epistasis, the expressions for $\alpha_{\mathbb{U},\mathbb{V}}$ coefficients given by equation 7 must be used instead
 of equation 5.

Concerning the integration of the term generated by the Hill-Robertson effect,
 Roze (2021) showed that using scaled recombination rates $\rho_{\mathbb{U}} = r_{\mathbb{U}} / (sh)$ eliminates sh
 555 from the integrand, sh only appearing in the integration limits, given by $R / (2sh)$. The
 same method can be used when $r_{\mathbb{U}}$, δr_{ab} and h are changed to effective parameters in
 order to incorporate the effect of selfing. In particular, defining scaled recombination
 rates $\tilde{\rho}_{\mathbb{U}} = \tilde{r}_{\mathbb{U}} / (s\tilde{h})$, one finds that the overall strength of indirect selection generated

by the Hill-Robertson effect is given by the same expression as under random mating
 560 (equation 3 in Roze, 2021), except that the integration limits become $R(1-F)/(2s\tilde{h})$
 and that the whole expression is multiplied by a factor $1/(1-F)^2$. When $R(1-F) \gg$
 $s\tilde{h}$ (which implies that the selfing rate is not too large and s is sufficiently small), the
 integral can be approximated by the same integral taken between 0 and infinity, which
 is ≈ 1.8 (Roze, 2021), yielding:

$$s_{\text{HR}} \approx \frac{1.8 \delta R U^2}{N_e R^3 (1-F)^2}. \quad (22)$$

Equation 22 is equivalent to equation 1 in Roze (2021), replacing R by $\tilde{R} = R(1-F)$
 and δR by $\delta\tilde{R} = \delta R(1-F)$. Note that this approximation is not valid when F ap-
 proaches 1, however, as $R(1-F) \gg s\tilde{h}$ cannot hold in this case. Finally, the effective
 population size N_e may be significantly lowered by background selection effects when
 U is not small. Assuming that the reduction in N_e is mostly due to tightly linked loci,
 570 one obtains that $N_e \approx [N/(1+F)] \exp[-2U/(\tilde{R} + 2s\tilde{h})]$ (e.g., Hudson and Kaplan,
 1995; Roze, 2016), where N is the census population size.

The evolutionarily stable map length (R_{ES}) corresponds to the value of R for
 which direct and indirect selection balance each other, that is, $s_m = 0$. Figure 3 shows
 R_{ES} as a function of the selfing rate σ , for different values of the direct cost of re-
 575 combination c . The curves correspond to the analytical predictions, obtained using
Mathematica by numerically integrating the results of the three-locus model over the
 genetic map to obtain s_{det} and s_{HR} for a range of values of R , and finding the value
 of R for which $s_m = 0$ by interpolation (see Supplementary Material). Note that the
 more precise approximations given in Appendix C have been used to compute s_{det} ,
 580 but using equations 10, 13, 15 and 20 often yields similar results. Figure 3 shows that
 the ES map length generally increases as the selfing rate increases, due to stronger
 indirect selection caused by the Hill-Robertson effect. While indirect selection van-
 ishes under full selfing ($\sigma = 1$), the analytical model predicts that the maximum map
 length is reached for values of σ very close to 1, in particular when $c = 10^{-3}$ and
 585 $c = 10^{-4}$. In the absence of epistasis, the deterministic component of indirect selec-
 tion selects against recombination ($s_{\text{det}} < 0$). Figure 3 shows that this deterministic
 component stays negligible relative to the Hill-Robertson effect for parameter values
 leading to low R_{ES} ($c = 0.01$), while its relative effect becomes more important for

parameter values leading to higher R_{ES} ($c = 10^{-4}$). This agrees with the prediction
 590 that the Hill-Robertson effect becomes stronger than deterministic effects when effective recombination rates are sufficiently small. In the absence of epistasis, we generally found that s_{det} is mainly driven by $s_{D_{ab}}$, the effect of $s_{D_{mab,ab}^0}$ staying negligible (not shown). Figure 3 also shows that extrapolations from our three-locus model often provide accurate predictions of the evolutionarily stable map length observed in the
 595 simulations, with discrepancies appearing for high values of the selfing rate. These discrepancies may be due to the fact that using effective recombination coefficients to transpose the result obtained under random mating to the case of a partially selfing population (as we did to compute s_{HR}) should strictly only hold when loci are sufficiently tightly linked (e.g., Padhukasahasram et al., 2008; Roze, 2016), while loosely
 600 linked loci may significantly contribute to selection for recombination when the selfing rate is high. They may also be caused by higher-order genetic associations (involving more than two selected loci) that are not taken into account in the analysis.

Figure 4 shows the effects of the deleterious mutation rate U , population size N , selection and dominance coefficients of deleterious alleles (s , h) on the evolutionarily
 605 stable map length (Figure S2 shows the same results with R on a log scale). As predicted, increasing U and/or decreasing N leads to stronger effects of Hill-Robertson interference between deleterious alleles, favoring higher values of R (see also Roze, 2021). As can be seen in Figure 4A, our analytical approximations overestimate the ES map length when U is high: this is probably due to the effect of higher-order genetic
 610 associations. Furthermore, in some simulations with $U = 0.5$, R fell to a very small value at some point during the simulation, which led to a quick accumulation of heterozygous mutations, the only surviving individuals being heterozygous for haplotypes carrying different deleterious alleles in repulsion. This accumulation of slightly deleterious mutations at the heterozygous state in a low recombining region has previously
 615 been described as pseudo-overdominance (e.g., Charlesworth and Charlesworth, 1997; Pálsson and Pamilo, 1999; Waller, 2021). This occurred for $0.5 \leq \sigma \leq 0.7$, in which case the simulation was terminated, as the number of segregating mutations quickly became very large. Figure 4C shows that selection for recombination is stronger when deleterious alleles are more weakly selected, as already found by Roze (2021); further-
 620 more, above a given value of s , the equilibrium map length decreases as the selfing rate

increases (which can be understood from the reasoning given in the previous subsection and Figure 2). By contrast, the dominance coefficient h of deleterious alleles has only little effect on the ES map length (Figure 4D). Figures S3 and S4 show that the results are not significantly affected by introducing variability in the selection coefficients of deleterious alleles into the simulation program, nor by limiting to 100 or 1000 the number of loci at which deleterious mutations may occur. Figure S5 shows that the strength of selection for recombination due to the Hill-Robertson effect is accurately predicted by equation 22 when selection against deleterious alleles is sufficiently weak, as long as the selfing rate is not too high.

Figure 5 shows the effect of each form of diploid epistasis ($e_{a \times a}$, $e_{a \times d}$, $e_{d \times d}$) and of Charlesworth et al.'s (1991) synergistic epistasis coefficient β (that combines the three forms of epistasis, see Methods) on the ES map length. Note that only negative values of epistasis were considered, as combinations of deleterious mutations quickly become beneficial when epistasis is positive and U is not very small. Negative epistasis tends to lower the strength of the Hill-Robertson effect (dotted curves in Figure 5), by increasing the effective strength of selection against deleterious alleles. Nevertheless, the deterministic effects generated by negative epistasis increase the overall strength of selection for recombination. Figure S6 shows that this increase is driven by the term in D_{ab} in the case of negative additive-by-additive ($e_{a \times a}$) and additive-by-dominance ($e_{a \times d}$) epistasis, the term in $D_{mab,ab}^0$ being negligible when $e_{a \times a} < 0$, while it disfavors recombination when $e_{a \times d} < 0$ (one can show that this last effect is generated by the first term of equation 13, since effective dominance coefficients $a_{j,j}$ are increased by $e_{a \times d}$, as shown by equation 7). In the case of dominance-by-dominance epistasis ($e_{d \times d}$), the large increase in recombination observed for moderate values of the selfing rate is generated by the term in $D_{mab,ab}^0$, and thus corresponds to the benefit of recombination previously described by Roze and Lenormand (2005). One can note that the model overestimates the strength of selection for recombination in this case. This is possibly due to the fact that, while the effect of $e_{d \times d}$ on effective dominance coefficients $a_{j,j}$ is negligible as long as epistasis is sufficiently weak, one can show that $a_{j,j}$ increases with $e_{d \times d}$ when dominance-by-dominance epistasis becomes the main source of selection against deleterious alleles (see Appendix G in Roze, 2009), reducing selection for recombination through the first term of equation 13.

DISCUSSION

Digging into the causes of general empirical patterns such as the positive correlation between selfing rate and chiasma frequency observed within several families of flowering plants may help us to gain a better understanding of the selective forces affecting the evolution of recombination rates. While there is no obvious reason why the mechanistic constraints associated with chromosomal segregation during meiosis should differ between outcrossing and selfing species, the mating system of organisms does affect the benefit of recombination through its effect on genetic variation. Although indirect selective forces acting on recombination are expected to vanish under complete selfing (as heterozygosity should then be very rare), the results presented in this article show that intermediate selfing rates may either increase or decrease selection for recombination caused by interference (Hill-Robertson effect) among deleterious mutations, depending on parameter values. Roughly, selfing leads to stronger selection for increased chromosomal map length as long as $sh \ll R(1 - F)$ (mostly due to the fact that selfing reduces effective recombination rates, thus increasing the strength of genetic associations), while selfing decreases selection for recombination when $sh \gg R(1 - F)$, due to the fact that changes in recombination rates have little effect on genetic associations in this regime. Given that most deleterious mutations seem to have weak fitness effects (e.g., Charlesworth, 2015), selection for recombination should thus be increased by selfing, and be maximized for selfing rates slightly below 1 (as illustrated by Figures 3 and 4).

In agreement with previous results (Otto and Barton, 2001; Keightley and Otto, 2006; Roze, 2021), we found that interference is often the main driver of selection for recombination when linkage is tight (or when the selfing rate is strong), while deterministic effects tend to become more important when recombination is frequent. In the absence of epistasis, the variance in the degree of inbreeding among individuals caused by partial selfing (associated with a more efficient purging of deleterious alleles in more inbred lineages) generates positive associations among deleterious alleles, selecting against recombination in infinite populations. As a result, the equilibrium map length may be lower under moderate selfing rates than under random mating in large, highly recombining populations, as can be seen in Figure 3 with $c = 10^{-4}$.

As in standard models of infinite, randomly-mating populations (Charlesworth, 1990; 685 Barton, 1995), negative epistasis between mutations tends to favor recombination by generating negative linkage disequilibria between deleterious alleles. This effect is often maximized at high selfing rates, again due to the fact that selfing reduces effective recombination rates, thus increasing the magnitude of linkage disequilibria. Furthermore, components of epistasis involving dominance ($e_{a \times d}$, $e_{d \times d}$) also contribute to 690 generating linkage disequilibrium (through an effective epistasis parameter) when the selfing rate is greater than zero.

Under high effective recombination, Roze and Lenormand (2005) showed that correlations in homozygosity among loci generated by partial selfing (identity disequilibria) are the main drivers of indirect selection on recombination rates, favoring 695 higher recombination when dominance-by-dominance epistasis is negative (as recombination then benefits from a short-term advantage, by increasing the mean fitness of offspring). The results of the present article temper the relative importance of this effect, by showing that it becomes negligible relative to the effect of deterministically generated linkage disequilibria and Hill-Robertson interference under high selfing 700 and in the case of tightly linked loci. Nevertheless, moderate selfing rates may favor a significant increase in chromosomal map length through this short-term benefit of breaking identity disequilibria, provided that dominance-by-dominance interactions are negative and represent a major component of epistasis (Figure 5C). In principle, the average sign and overall importance of dominance-by-dominance effects can be 705 inferred from the shape of the relation between the degree of inbreeding of individuals and their fitness (Crow and Kimura, 1970, p. 80), negative $e_{d \times d}$ causing a faster than linear decline in fitness with inbreeding. This method was used on several plant species and did not yield any clear evidence for negative $e_{d \times d}$ (Willis, 1993; Falconer and Mackay, 1996); however, the methodology used (involving experimental crosses to 710 increase the degree of inbreeding of individuals) generates a bias against finding negative $e_{d \times d}$, as deleterious alleles may have been purged from the more highly inbred lines. Using an experimental protocol that avoids this bias, Sharp and Agrawal (2016) found evidence for negative $e_{d \times d}$ (on viability) between EMS-induced mutations in *Drosophila melanogaster*. While more work is needed to assess the generality of this 715 result, previous experimental studies showed that epistasis is typically quite variable

among pairs of loci (de Visser and Elena, 2007; Kouyos et al., 2007; Martin et al., 2007). As shown by Otto and Feldman (1997), recombination tends to be less favored when epistasis is variable, and it would thus be of interest to extend our model to more realistic fitness landscapes including distributions of epistasis.

720 While the indirect benefits of increased crossover rates may be strong when recombination is rare (e.g., Keightley and Otto, 2006), they typically become rather weak under frequent recombination (Roze, 2021), and one may thus wonder to what extent our model can explain a positive effect of selfing on chiasma frequency when at least one crossover per chromosome occurs during meiosis. When our model is modified
725 so that 1 crossover per bivalent necessarily occurs (leading to a minimum map length of 50cM, that is, $R = 0.5$) and letting R evolve above 0.5 using a similar model as before, one indeed observes very limited effects of indirect selection on the evolutionarily stable map length (except for high selfing rates) in the absence of epistasis, for a deleterious mutation rate per chromosome of $U = 0.1$ (Figure 6). Substantial increases
730 in recombination under partial selfing can be favored in the presence of negative $e_{d \times d}$, however, chromosomal map length being maximized for moderate selfing rates in this case (Figure 6A). Furthermore, higher chromosomal mutation rates lead to important increases in R at high selfing rates due to stronger Hill-Robertson effects, as can be seen on Figure 6B for $U = 0.5$. More generally, equilibrium values of R with one
735 obligate crossover per bivalent should be approximately the same as in Figures 3 – 5 when setting to 0.5 the points falling below $R = 0.5$ (with some slight differences caused by the fact that the distribution of the number of crossovers per chromosome is not Poisson anymore in this modified model).

Are these results consistent with empirical data on the effect of selfing on recom-
740 bination? Figure S7 shows chiasma count data from several families of Angiosperms (comparing closely related species with contrasted mating systems) that were used to generate Figure 1 in Roze and Lenormand (2005). As can be seen on Figure S7, the increase in chromosome map length in selfing species compared with their out-crossing relatives is generally of the order 20 to 30cM, which seems consistent with
745 the results obtained here for moderate chromosomal mutation rates, either due to the Hill-Robertson effect or to negative values of $e_{a \times a}$, $e_{a \times d}$ or Charlesworth et al.’s (1991) synergistic epistasis coefficient β . The results obtained under negative $e_{d \times d}$ seem less

consistent with the empirical patterns, as they show that R should be maximized for moderate selfing rates (Figures 5, 6). As the data do not include any precise selfing rate estimate, it is difficult to assess whether the observed patterns are more consistent with a gradual increase in R as selfing increases (as predicted for example under negative $e_{a \times a}$, $e_{a \times d}$ or β) or with a sharper increase at higher selfing rates (as predicted under the action of the Hill-Robertson effect alone). Further insights could be gained by comparing the genetic maps of closely related species for which estimates of selfing rates are available.

As already noted in Roze (2021), the fact that chromosomal map length may evolve towards values greater than 0.5 even in the absence of selfing (as can be seen on Figure S7) may seem at odds with the predictions of our model. Several explanations can be proposed to explain this discrepancy. First, the direct fitness cost of increasing R above 0.5 may be very low, as suggested by recent experiments on *A. thaliana* mutants in which map length is increased up to 7.8 fold without any clear effect on fertility (Fernandes et al., 2018). Second, relaxing the (unrealistic) assumption that crossovers occur anywhere along the chromosome with the same probability may generate stronger effects of indirect selection on the evolutionary stable map length. In particular, data from *Caenorhabditis elegans* and humans suggest that structural constraints may impose restrictions on the possible localization of obligate crossovers along chromosomes (Koehler et al., 1996; Ottolini et al., 2015; Altendorfer et al., 2020); the fact that crossovers are often more frequent in subtelomeric regions (in plants and animals) may possibly reflect such constraints (e.g., Haenel et al., 2018). Restricting the position of the obligate crossover in a given chromosomal region should increase the strength of indirect selection acting on a modifier allele increasing recombination in other regions, particularly if the modifier is located in regions with lower recombination. Effects of non-uniform positions of crossovers along chromosomes and crossover interference will be explored in a future work. Third, sweeps of beneficial mutations may also increase the strength of indirect selection for recombination (Hartfield et al., 2010; Roze, 2021). Note that this may also enhance the effect of selfing on recombination, since transitions towards predominant selfing are often followed by a number of phenotypic changes such as reduced flower size (“selfing syndrome”, e.g., Cutter, 2019) that may be seen as adaptations to the new mating system, thus involving the spread

780 of newly beneficial alleles that may (at least transiently) favor higher recombination rates. Recombination rates may also be further increased after transitions to selfing due to stronger Hill-Robertson effects caused by population bottlenecks and extinction/recolonization dynamics that often characterize selfing populations (e.g., Guo et al., 2009; Willi et al., 2018; Orsucci et al., 2020).

785 Interestingly, predominantly selfing species often maintain low rates of outcrossing in nature (e.g., Bonnin et al., 2001; Bomblies et al., 2010), which may reflect the effect of selective forces favoring the maintenance of recombination. Indeed, deterministic and stochastic multilocus models showed that individuals outcrossing at a low rate are often selectively favored over complete selfers in conditions where recombination
790 is advantageous (negative epistasis or finite population size, Charlesworth et al., 1991; Kamran-Disfani and Agrawal, 2014). However, higher effective recombination rates could in principle be achieved by increasing either the outcrossing rate of individuals or the number of crossovers at meiosis, and one could imagine that one solution or the other may be favored depending on the different types of direct and indirect se-
795 lective forces that may act on outcrossing vs. recombination modifiers. Exploring the joint evolution of outcrossing and recombination rates in predominantly selfing species would thus be of interest, both from a theoretical and empirical perspective.

Data availability. All derivations are provided in the *Mathematica* notebook available as Supplementary Material at doi.org/10.5281/zenodo.6783728, along with the
800 C++ codes used to run the simulations.

Acknowledgements. We thank four anonymous reviewers for helpful comments, and the Bioinformatics and Computing Service of Roscoff’s Biological Station (Abims
805 platform) for computing time.

Funding. This work was funded by the Agence Nationale pour la Recherche (SelfRecomb project: ANR-18-CE02-0017-02, and GenAsex project: ANR-17-CE02-0016-01).

810 **Competing interests.** We declare no competing interest.

LITERATURE CITED

- Abu Awad, D. and D. Roze. 2020. Epistasis, inbreeding depression and the evolution of self-fertilization. *Evolution* 74:1301–1320.
- Agrawal, A. F. 2006. Evolution of sex: why do organisms shuffle their genotypes?
815 *Curr. Biol.* 16:R696–R704.
- Altendorfer, E., L. I. Láscares-Lagunas, S. Nadarajan, and M. P. Colaiácovo. 2020. Crossover position drives chromosome remodeling for accurate meiotic chromosome segregation. *Curr. Biol.* 30:1329–1338.
- Barton, N. H. 1995. A general model for the evolution of recombination. *Genet. Res.*
820 65:123–144.
- Barton, N. H. and S. P. Otto. 2005. Evolution of recombination due to random drift. *Genetics* 169:2353–2370.
- Barton, N. H. and M. Turelli. 1991. Natural and sexual selection on many loci. *Genetics* 127:229–255.
- 825 Benavente, E. and J. Sybenga. 2004. The relation between pairing preference and chiasma frequency in tetrasomics of rye. *Genome* 47:122–133.
- Blackwell, A. R., J. Dłuzewska, M. Szymanska-Lejman, S. Desjardins, A. J. Tock, N. Kbir, C. Lambing, E. J. Lawrence, T. Bieluszewski, B. Rowan, J. D. Higgins, P. A. Ziolkowski, and I. R. Henderson. 2020. MSH2 shapes the meiotic crossover
830 landscape in relation to interhomolog polymorphism in *Arabidopsis*. *EMBO J.* 39:e104858.
- Bomblies, K., L. Yant, R. A. Laitinen, S.-T. Kim, J. D. Hollister, N. Warthmann, J. Fitz, and D. Weigel. 2010. Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of *Arabidopsis thaliana*. *PLoS Genetics*
835 6:e1000890.
- Bonnin, I., J. Ronfort, F. Wozniak, and I. Olivieri. 2001. Spatial effects and rare outcrossing events in *Medicago truncatula* (Fabaceae). *Mol. Ecol.* 10:1371–1383.

- Borts, R. H. and J. E. Haber. 1987. Meiotic recombination in yeast: alteration by multiple heterozygosities. *Science* 237:1459–1465.
- 840 Brand, C. L., M. V. Cattani, S. B. Kingan, E. L. Landeen, and D. C. Presgraves. 2018. Molecular evolution at a meiosis gene mediates species differences in the rate and patterning of recombination. *Curr. Biol.* 28:1289–1295.
- Charlesworth, B. 1990. Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genet. Res.* 55:199–221.
- 845 ———. 2015. Causes of natural variation in fitness: Evidence from studies of *Drosophila* populations. *Proc. Natl. Acad. Sci. U. S. A.* 112:1662–1669.
- Charlesworth, B. and D. Charlesworth. 1997. Rapid fixation of deleterious alleles can be caused by Muller’s ratchet. *Genet. Res.* 70:63–73.
- Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1991. Multilocus models of
850 inbreeding depression with synergistic selection and partial self-fertilization. *Genet. Res.* 57:177–194.
- Charlesworth, D., B. Charlesworth, and C. Strobeck. 1977. Effects of selfing on selection for recombination. *Genetics* 86:213–226.
- . 1979. Selection for recombination in partially self-fertilizing populations. *Genetics* 93:237–244.
855
- Chen, W. and S. Jinks-Robertson. 1999. The role of the mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast. *Genetics* 151:1299–1313.
- Crow, J. F. and M. Kimura. 1970. *An Introduction to Population Genetics Theory*.
860 Harper and Row, New York.
- Cutter, A. D. 2019. Reproductive transitions in plants and animals: selfing syndrome, sexual selection and speciation. *New Phytol.* 224:1080–1094.
- de Visser, J. A. G. M. and S. F. Elena. 2007. The evolution of sex: empirical insights into the roles of epistasis and drift. *Nat. Rev. Genet.* 8:139–149.

- 865 Falconer, D. S. and T. F. C. Mackay. 1996. Introduction to Quantitative Genetics. Addison Wesley Longman, Harlow.
- Felsenstein, J. 1974. The evolutionary advantage of recombination. *Genetics* 78:737–756.
- Fernandes, J. B., M. Séguéla-Arnaud, C. Larchevêque, and R. Mercier. 2018. Unleashing meiotic crossovers in hybrid plants. *Proc. Natl. Acad. Sci. U. S. A.* 115:2431–2436.
- 870 Gervais, C. and D. Roze. 2017. Mutation rate evolution in partially selfing and partially asexual organisms. *Genetics* 207:1561–1575.
- Glémin, S. and J. Ronfort. 2013. Adaptation and maladaptation in selfing in outcrossing species: new mutations versus standing variation. *Evolution* 67:225–240.
- 875 Gray, D. and P. E. Cohen. 2016. Control of meiotic crossovers: from double-strand break formation to designation. *Ann. Rev. Gen.* 50:175–210.
- Guo, Y. L., J. S. Bechsgaard, T. Slotte, B. Neuffer, M. Lascoux, D. Weigel, and M. H. Schierup. 2009. Recent speciation of *Capsella rubella* from *Capsella grandiflora*, associated with loss of self-incompatibility and an extreme bottleneck. *Proc. Natl. Acad. Sci. U. S. A.* 106:5246–5251.
- 880 Haenel, Q., T. G. Laurentino, M. Roesti, and D. Berner. 2018. Meta-analysis of chromosome-scale crossover rate variation in eukaryotes and its significance to evolutionary genomics. *Mol. Ecol.* 27:2477–2497.
- 885 Haldane, J. B. S. 1919. The combination of linkage values and the calculation of distances between the loci of linked factors. *J. Genet.* 8:299–309.
- Hansson, B., A. Kawabe, S. Preuss, H. Kuittinen, and D. Charlesworth. 2006. Comparative gene mapping in *Arabidopsis lyrata* chromosomes 1 and 2 and the corresponding *A. thaliana* chromosome 1: recombination rates, rearrangements and centromere location. *Genet. Res.* 87:75–85.
- 890

- Hartfield, M., S. P. Otto, and P. D. Keightley. 2010. The role of advantageous mutations in enhancing the evolution of a recombination modifier. *Genetics* 184:1153–1164.
- Hill, W. G. and A. Robertson. 1966. The effect of linkage on limits to artificial selection. 895 *Genet. Res.* 8:269–294.
- Holsinger, K. E. and M. W. Feldman. 1983. Linkage modification with mixed random mating and selfing: a numerical study. *Genetics* 103:323–333.
- Hudson, R. R. and N. L. Kaplan. 1995. Deleterious background selection with recombination. *Genetics* 141:1605–1617.
- 900 Johnston, S. E., C. Bérénos, J. Slate, and J. M. Pemberton. 2016. Conserved genetic architecture underlying individual recombination rate variation in a wild population of Soay sheep (*Ovis aries*). *Genetics* 203:583–598.
- Kamran-Disfani, A. and A. F. Agrawal. 2014. Selfing, adaptation and background selection in finite populations. *J. Evol. Biol.* 27:1360–1371.
- 905 Kawabe, A., B. Hansson, A. Forrest, J. Hagenblad, and D. Charlesworth. 2006. Comparative gene mapping in *Arabidopsis lyrata* chromosomes 6 and 7 and *A. thaliana* chromosome IV: evolutionary history, rearrangements and local recombination rates. *Genet. Res.* 88:45–56.
- Keightley, P. D. and S. P. Otto. 2006. Interference among deleterious mutations favours 910 sex and recombination in finite populations. *Nature* 443:89–92.
- Kirkpatrick, M., T. Johnson, and N. H. Barton. 2002. General models of multilocus evolution. *Genetics* 161:1727–1750.
- Koehler, K. E., R. Scott Hawley, S. Sherman, and T. Hassold. 1996. Recombination and nondisjunction in human and flies. *Hum. Mol. Genet.* 5:1495–1504.
- 915 Kong, A., G. Thorleifsson, D. F. Gudbjartsson, G. Masson, A. Sigurdsson, A. Jonasdottir, G. Bragi Walters, A. Jonasdottir, A. Gylfason, K. T. Kristinsson, S. A. Gudjonsson, M. L. Frigge, A. Helgason, U. Thorsteinsdottir, and K. Stefansson.

2010. Fine-scale recombination rate differences between sexes, populations and individuals. *Nature* 467:1099–1103.
- 920 Kouyos, R. D., O. K. Silander, and S. Bonhoeffer. 2007. Epistasis between deleterious mutations and the evolution of recombination. *Trends Ecol. Evol.* 22:308–315.
- Kuittinen, H., A. A. de Haan, C. Vogl, S. Oikarinen, J. Leppälä, M. Koch, T. Mitchell-Olds, C. H. Langley, and O. Savolainen. 2004. Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *A. thaliana*. *Genetics* 168:1575–1584.
- 925 Martin, G., S. F. Elena, and T. Lenormand. 2007. Distributions of epistasis in microbes fit predictions from a fitness landscape model. *Nat. Genet.* 39:555–560.
- Nagylaki, T. 1993. The evolution of multilocus systems under weak selection. *Genetics* 134:627–647.
- Nordborg, M. 1997. Structured coalescent processes on different time scales. *Genetics* 930 146:1501–1514.
- . 2000. Linkage disequilibrium, gene trees and selfing: and ancestral recombination graph with partial self-fertilization. *Genetics* 154:923–929.
- Orsucci, M., P. Milesi, J. Hansen, J. Girodolle, S. Glémin, and M. Lascoux. 2020. Shift in ecological strategy helps marginal populations of shepherd’s purse (*Capsella bursa-pastoris*) to overcome a high genetic load: competition avoidance and colonization. *Proc. Roy. Soc. (Lond.) B* 287:20200463.
- 935
- Otto, S. P. 2021. Selective interference and the evolution of sex. *J. Hered.* 112:9–18.
- Otto, S. P. and N. H. Barton. 1997. The evolution of recombination: removing the limits to natural selection. *Genetics* 147:879–906.
- 940 ———. 2001. Selection for recombination in small populations. *Evolution* 55:1921–1931.
- Otto, S. P. and M. W. Feldman. 1997. Deleterious mutations, variable epistatic interactions, and the evolution of recombination. *Theor. Popul. Biol.* 51:134–47.

- Otto, S. P. and T. Lenormand. 2002. Resolving the paradox of sex and recombination.
945 Nat. Rev. Genet. 3:252–261.
- Ottolini, C. S., L. J. Newnham, A. Capalbo, S. A. Natesan, H. A. Joshi, D. Cimadomo,
D. K. Griffin, K. Sage, M. C. Summers, A. R. Thornhill, E. Housworth, A. D.
Herbert, L. Rienzi, F. M. Ubaldi, A. H. Handyside, and E. R. Hoffmann. 2015.
950 Genome-wide maps of recombination and chromosome segregation in human oocytes
and embryos show selection for maternal recombination rates. Nat. Genet. 47:727–
737.
- Padhukasahasram, B., P. Marjoram, J. D. Wall, C. D. Bustamante, and M. Nord-
borg. 2008. Exploring population genetics models with recombination using efficient
forward-time simulations. Genetics 178:2417–2427.
- 955 Pálsson, S. and P. Pamilo. 1999. The effects of deleterious mutations on linked, neutral
variation in small populations. Genetics 153:475–483.
- Pollak, E. 1987. On the theory of partially inbreeding finite populations. I. Partial
selfing. Genetics 117:353–360.
- Ritz, K. R., M. A. F. Noor, and N. D. Singh. 2017. Variation in recombination rate:
960 adaptive or not? Trends Genet. 33:364–374.
- Ross-Ibarra, J. 2007. Genome size and recombination in angiosperms: a second look.
J. Evol. Biol. 20:800–806.
- Roze, D. 2009. Diploidy, population structure and the evolution of recombination.
Am. Nat. 174:S79–S94.
- 965 ———. 2014. Selection for sex in finite populations. J. Evol. Biol. 27:1304–1322.
- . 2016. Background selection in partially selfing populations. Genetics 203:937–
957.
- . 2021. A simple expression for the strength of selection on recombination
generated by interference among mutations. Proc. Natl. Acad. Sci. U. S. A.
970 118:e2022805118.

- Roze, D. and N. H. Barton. 2006. The Hill-Robertson effect and the evolution of recombination. *Genetics* 173:1793–1811.
- Roze, D. and T. Lenormand. 2005. Self-fertilization and the evolution of recombination. *Genetics* 170:841–857.
- 975 Samuk, K., B. Manzano-Winkler, K. R. Ritz, and M. A. F. Noor. 2020. Natural selection shapes variation in genome-wide recombination rate in *Drosophila pseudoobscura*. *Curr. Biol.* 30:1517–1528.
- Sharp, N. P. and A. F. Agrawal. 2016. The decline in fitness with inbreeding: evidence for negative dominance-by-dominance epistasis in *Drosophila melanogaster*. *J. Evol.*
980 *Biol.* 29:857–864.
- Stapley, J., P. G. D. Feulner, S. E. Johnston, A. W. Santure, and C. M. Smadja. 2017. Variation in recombination frequency and distribution across eukaryotes: patterns and processes. *Phil. Trans. Roy. Soc. (Lond.) B* 372:20160455.
- Waller, D. M. 2021. Addressing Darwin’s dilemma: can pseudo-overdominance explain
985 persistent inbreeding depression and load? *Evolution* 75:779–793.
- Weir, B. S. and C. C. Cockerham. 1973. Mixed self and random mating at two loci. *Genet. Res.* 21:247–262.
- Willi, Y., M. Fracassetti, S. Zoller, and J. Van Buskirk. 2018. Accumulation of mutational load at the edges of a species range. *Mol. Biol. Evol.* 35:781–791.
- 990 Willis, J. H. 1993. Effects of different levels of inbreeding on fitness components in *Mimulus guttatus*. *Evolution* 47:864–876.
- Wright, K. M., B. Arnold, K. Xue, M. Surinova, J. O’Connell, and K. Bomblies. 2015. Selection on meiosis genes in diploid and tetraploid *Arabidopsis arenosa*. *Mol. Biol. Evol.* 32:944–955.
- 995 Wright, S. I., R. W. Ness, J. P. Foxe, and S. C. H. Barrett. 2008. Genomic consequences of outcrossing and selfing in plants. *Int. J. Plant Sci.* 169:105–118.

Zelkowski, M., M. A. Olson, M. Wang, and W. Pawłowski. 2019. Diversity and determinants of meiotic recombination landscapes. *Trends Genet.* 35:359–370.

1000 Ziolkowski, P. A., L. E. Berchowitz, C. Lambing, N. E. Yelina, X. Zhao, K. A. Kelly, K. Choi, L. Ziolkowska, V. June, E. Sanchez-Moran, C. Franklin, G. P. Copenhaver, and I. R. Henderson. 2015. Juxtaposition of heterozygous and homozygous regions causes reciprocal crossover remodelling via interference during *Arabidopsis* meiosis. *eLife* 4:e03708.

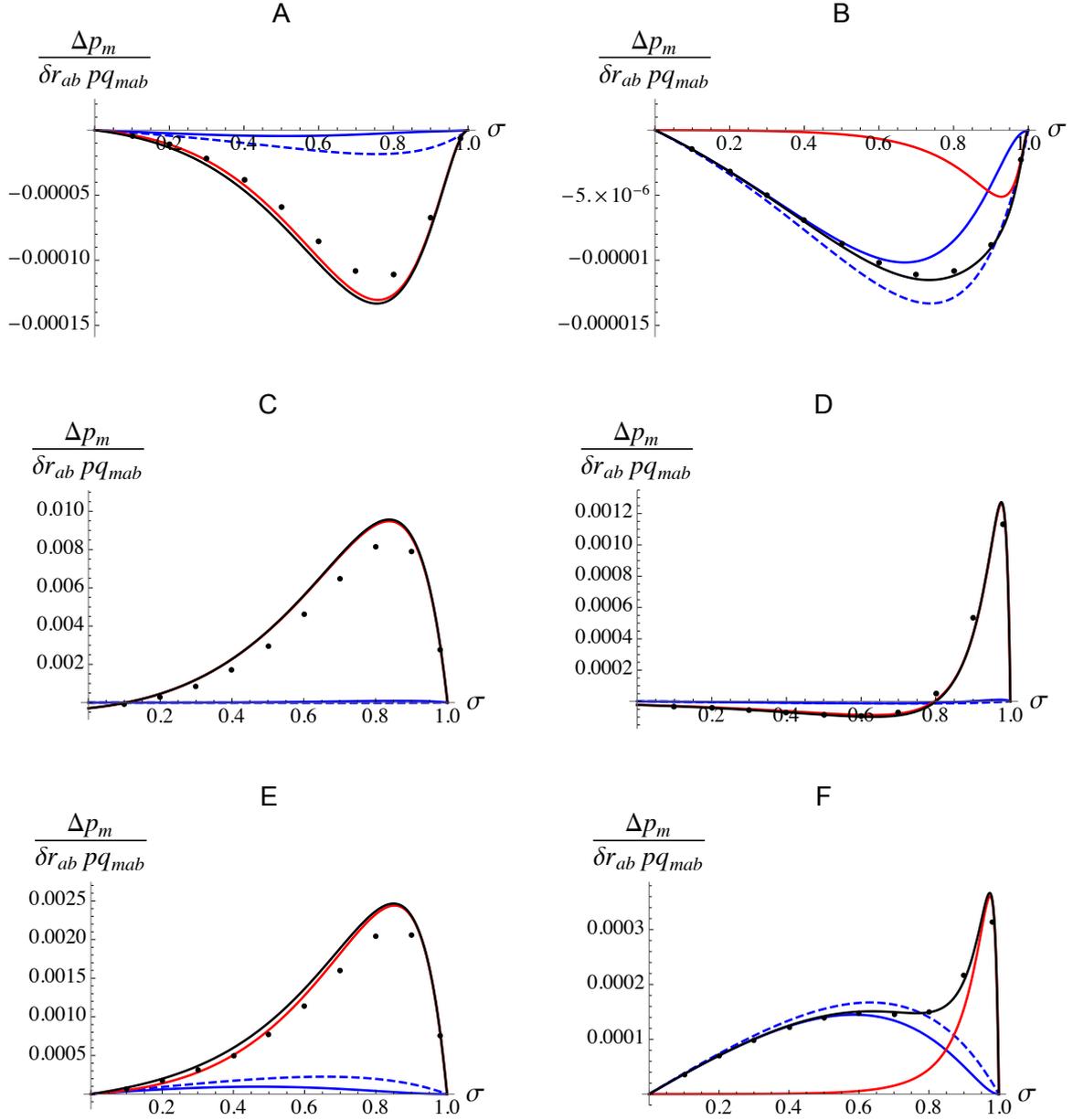
Table 1: Parameters and variables of the model.

1005

σ	Selfing rate
s, h, \tilde{h}	Selection, dominance and effective dominance coefficients of deleterious alleles
$e_{a \times a}, e_{a \times d}, e_{d \times d}, \tilde{e}$	Additive-by-additive, additive-by-dominance, dominance-by-dominance epistasis and effective epistasis between deleterious alleles
β	Charlesworth et al.'s (1991) synergistic epistasis coefficient
ϵ	Strength of selection (scaling parameter)
u, U	Deleterious mutation rate per locus, and per chromosome
r_{ij}, \tilde{r}_{ij}	Recombination rate and effective recombination rate between loci i and j
$\delta r_{ab}, \tilde{\delta} r_{ab}$	Effect and effective effect of the modifier on r_{ab}
h_m	Dominance coefficient of allele m at the modifier locus
N, N_e	Census and effective population size
R, R_{ES}	Chromosome map length and evolutionarily stable map length
δR	Effect of the modifier on R
$D_{\mathbb{U}, \mathbb{V}}$	Genetic association between the sets \mathbb{U} and \mathbb{V} of loci present on the maternally and paternally inherited haplotypes of an individual
$\alpha_{\mathbb{U}, \mathbb{V}}$	Effect of selection on the sets \mathbb{U} and \mathbb{V} of loci present on the maternally and paternally inherited haplotypes of an individual
W, \bar{W}	Fitness of an individual and average fitness
$F = \sigma / (2 - \sigma)$	Inbreeding coefficient (probability of identity by descent at one locus)
ϕ_{ab}	Joint probability of identity by descent at loci a and b
$G_{ab} = \phi_{ab} - F^2$	Identity disequilibrium between loci a and b
p_j, q_j	Frequencies of the lower- and uppercase allele at locus j
n	Mean number of deleterious alleles per chromosome

Table 2: Fitness matrix in the three-locus model.

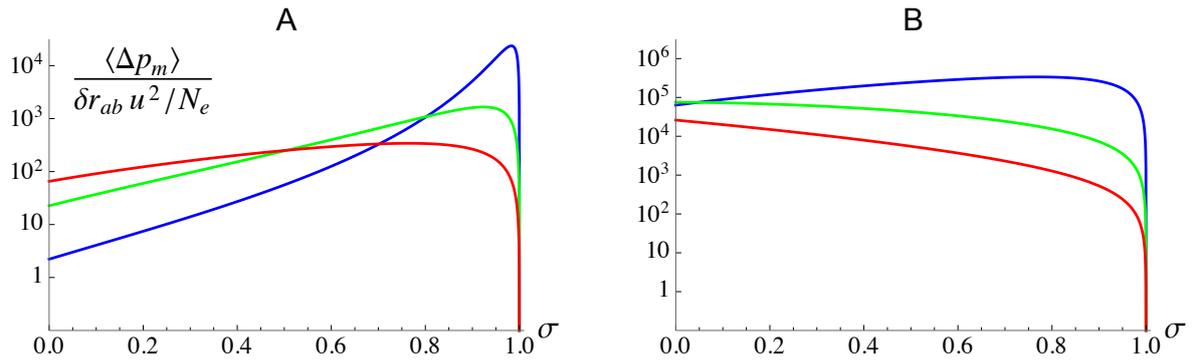
	AA	Aa	aa
BB	1	$1 - hs$	$1 - s$
Bb	$1 - hs$	$(1 - hs)^2 + e_{a \times a}$	$(1 - hs)(1 - s) + 2e_{a \times a} + e_{a \times d}$
bb	$1 - s$	$(1 - hs)(1 - s) + 2e_{a \times a} + e_{a \times d}$	$(1 - s)^2 + 4e_{a \times a} + 4e_{a \times d} + e_{d \times d}$



1010

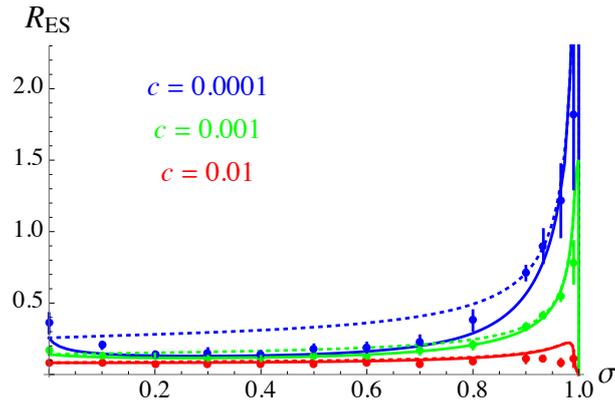
Figure 1. Effect of the selfing rate σ on the change in frequency of allele m in the deterministic three-locus model (scaled by $\delta r_{ab} p q_{mab}$) with no epistasis (A and B), negative additive-by-additive epistasis ($e_{a \times a} = -0.001$; C and D) and negative dominance-by-dominance epistasis ($e_{d \times d} = -0.001$; E and F), and for different recombination rates (A, C, E: $r_{ma} = r_{ab} = 0.01$; B, D, F: $r_{ma} = r_{ab} = 0.1$). Dots correspond to deterministic simulation results (see text for details) and black curves to the result obtained from equations 6 and C1 – C11, which is the sum of a term generated by $D_{ab} - D_{a,b}$ (red curves) and a term generated by D_{mab}^0 (blue curves);

dashed blue curves correspond to the strong recombination approximation (Equation
1020 11). Note that the red and black curves are nearly indistinguishable in C, D. The effect
of additive-by-dominance epistasis ($e_{a \times d}$) is similar to the effect of additive-by-additive
epistasis ($e_{a \times a}$) for these parameter values, and is shown in Figure S1. Loci are in the
order $m - a - b$, parameters values are: $s = 0.01$, $h = 0.2$, $h_m = 0.5$, while $\delta r_{ab} = 0.01$
and $u = 10^{-5}$ in the simulations.



1025

Figure 2. Effect of the selfing rate σ on the expected change in frequency of allele m (scaled by $\delta r_{ab} u^2 / N_e$ and on a log scale) generated by the Hill-Robertson effect between alleles a and b , for different values of the strength of selection against deleterious alleles (blue: $s = 0.005$; green: $s = 0.02$; red: $s = 0.05$) and recombination rates (A: $r_{ma} = r_{ab} = 0.05$; B: $r_{ma} = r_{ab} = 0.005$). The value of σ maximizing selection for recombination decreases as s increases and as recombination decreases. When selection is sufficiently strong and linkage sufficiently tight, selection for recombination decreases monotonously as σ increases (as can be seen in B for $s = 0.02$, $s = 0.05$). Loci are in the order $m - a - b$, the dominance coefficient of deleterious alleles is set to $h = 0.2$.



1035

Figure 3. Effect of the selfing rate σ on the evolutionarily stable map length R_{ES} , for different values of the cost of recombination c . Solid curves correspond to the analytical predictions obtained by solving $s_{\text{direct}} + s_{\text{det}} + s_{\text{HR}} = 0$ for R , where s_{det} and s_{HR} are obtained by integrating the expressions for the strength of indirect selection generated
1040 by deterministic effects (for s_{det}) and by the Hill-Robertson effect (for s_{HR}) over the genetic map (see Supplementary Material). Dotted curves correspond to the predictions obtained when ignoring indirect selection caused by deterministic effects (i.e., solving $s_{\text{direct}} + s_{\text{HR}} = 0$ for R). Dots correspond to individual-based simulation results; in this and the following figures, error bars are obtained by dividing the simulation output
1045 (after removing the first 5×10^5 generations) into 10 batches and calculating the variance of the average map length per batch, error bars measuring ± 1.96 SE. Parameter values are $N = 20,000$, $U = 0.1$, $s = 0.02$, $h = 0.2$, $e_{a \times a} = e_{a \times d} = e_{d \times d} = 0$.

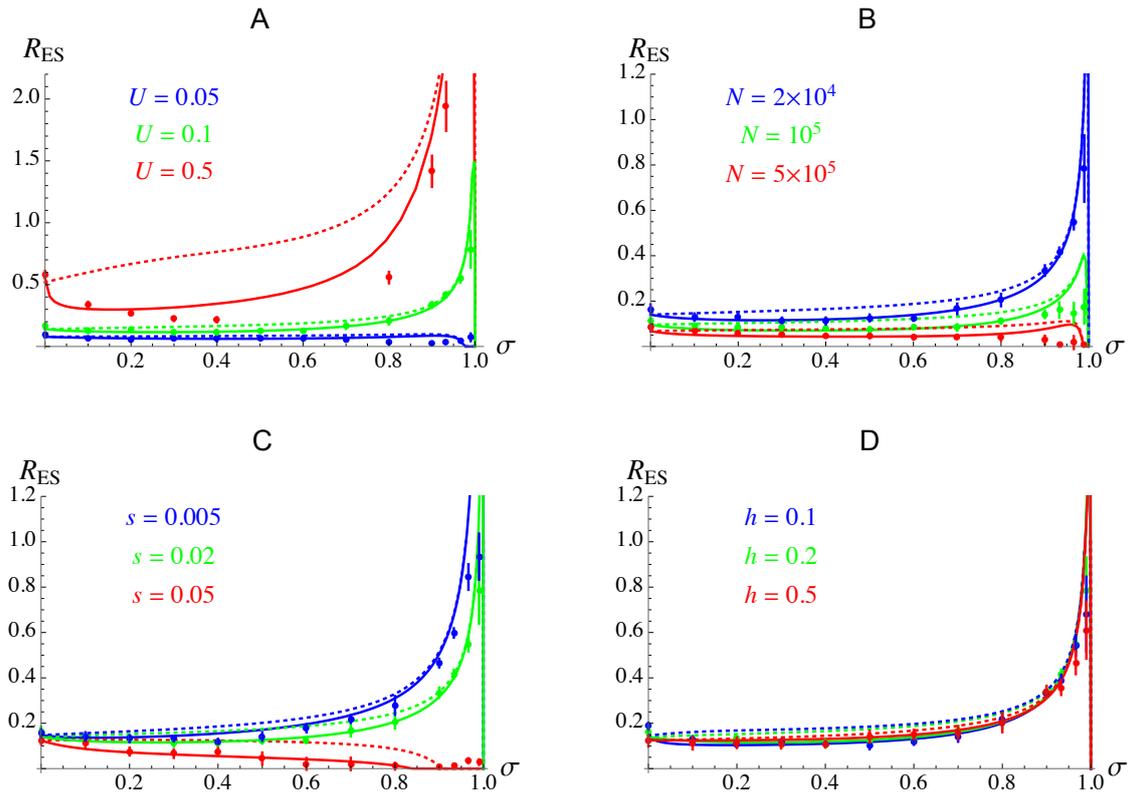
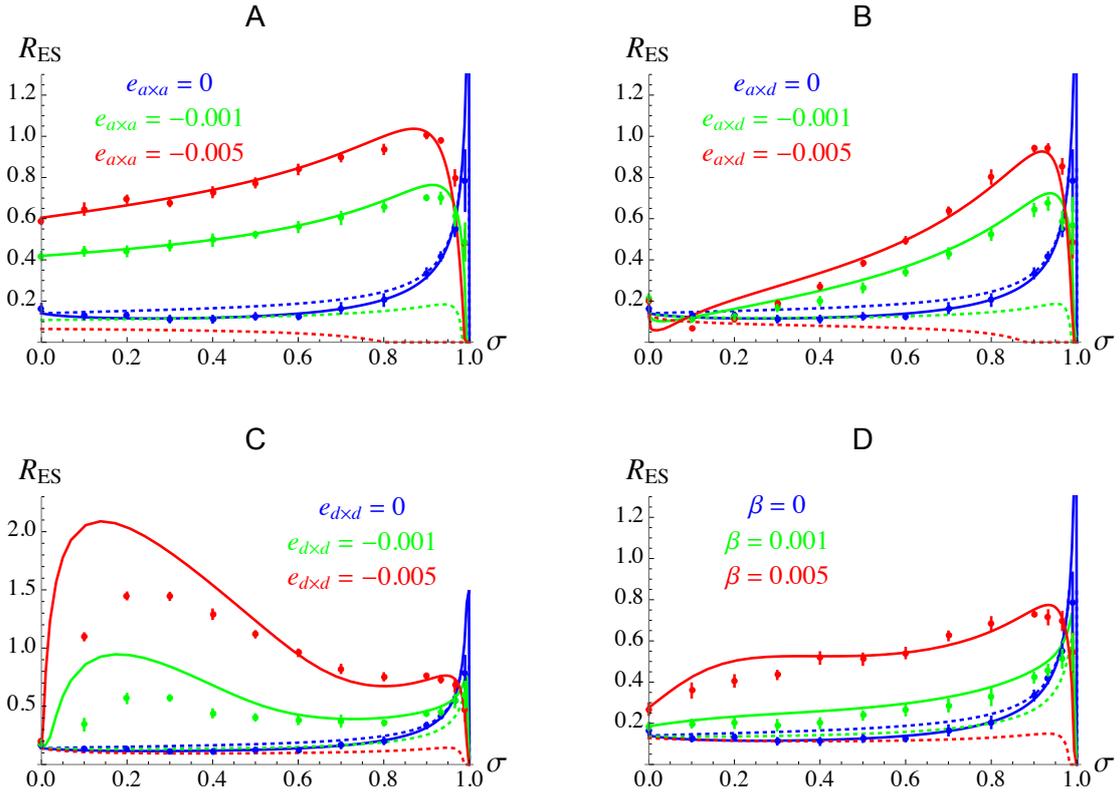


Figure 4. Effect of the selfing rate σ on the evolutionarily stable map length R_{ES} ,
 1050 for different values of the mutation rate U , population size N , strength of selection
 and dominance coefficient of deleterious alleles (s , h). Dots correspond to simulation
 results, curves have the same meaning as in Figure 3, and default parameter values
 are as in Figure 3 with $c = 0.001$ (no epistasis).



1055 **Figure 5.** Effect of the selfing rate σ on the evolutionarily stable map length R_{ES} , for
 different values of additive-by-additive (A), additive-by-dominance (B) and dominance-
 by-dominance (C) epistasis, and different values of Charlesworth et al.'s (1991) syn-
 1060 synergistic epistasis coefficient β (D). Note the different scale of the y-axis in C. Dots
 correspond to simulation results, curves have the same meaning as in Figure 3, and
 default parameter values are as in Figure 3 with $c = 0.001$.

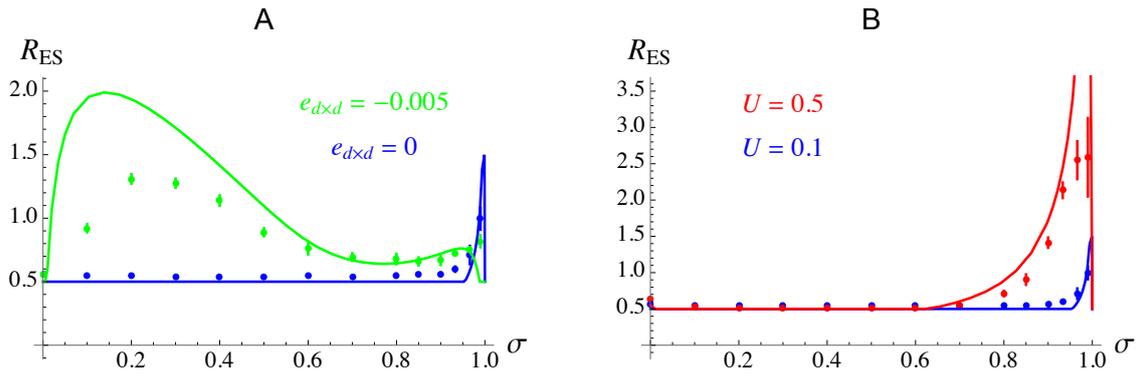


Figure 6. Effect of the selfing rate σ on the evolutionarily stable map length R_{ES} , when at least one crossover per chromosome occurs during meiosis (leading to a minimal map length of $R = 0.5$). Dots correspond to simulation results, curves correspond to analytical predictions; default parameter values are as in Figure 3 with $c = 0.001$.
 1065 A: Blue: no epistasis, green: $e_{d \times d} = -0.005$. B: Blue: $U = 0.1$, red: $U = 0.5$.

APPENDIX A: HIGH EFFECTIVE RECOMBINATION

Under high effective recombination, the associations between the modifier and the selected loci involved in equation 6 are given by (see *Mathematica* notebook for derivation):

$$D_{mab,a} \approx D_{ma,ab} \approx F (\alpha_b + \alpha_{b,b}) D_{mab,ab}^0 \quad (\text{A1})$$

$$D_{ma} \approx \frac{(1 + 2Fr_{ma}) [2F (\alpha_a + \alpha_{a,a}) \alpha_{b,b} + \alpha_{ab,b} + \alpha_{ab,ab}]}{r_{ma} (1 - F)} D_{mab,ab}^0 \quad (\text{A2})$$

$$D_{m,a} \approx \frac{F (1 + 2r_{ma})}{1 + 2Fr_{ma}} D_{ma} \quad (\text{A3})$$

$D_{mab,b}$, $D_{mb,ab}$, D_{mb} and $D_{m,b}$ being given by symmetric expressions. D_{mab} , $D_{m,ab}$ and $D_{ma,b}$ are given by:

$$D_{mab} \approx \frac{(1 + 2Fr_{ma}) (1 + 2Fr_{mb})}{(1 - F) [r_{mab} + 2F (r_{ma}r_{mb} + r_{ma}r_{ab} + r_{mb}r_{ab}) + 6F^2 r_{ma}r_{mb}r_{ab}]} \times \left[(1 + 2Fr_{ab}) T_{ab} D_{mab,ab}^0 - \delta r_{ab} H_m (D_{ab} - D_{a,b}) pq_m - \delta r_{ab} (1 - H_m) (D_{mab,m} - D_{ma,mb}) \right] \quad (\text{A4})$$

$$D_{m,ab} \approx \frac{F [(1 + 2Fr_{ma}) (1 + 2Fr_{mb}) + (1 - F) (r_{ma} + r_{mb} + 4Fr_{ma}r_{mb})]}{(1 + 2Fr_{ma}) (1 + 2Fr_{mb})} D_{mab} \quad (\text{A5})$$

$$D_{ma,b} \approx \frac{F}{(1 - F) [r_{mab} + 2F (r_{ma}r_{mb} + r_{ma}r_{ab} + r_{mb}r_{ab}) + 6F^2 r_{ma}r_{mb}r_{ab}]} \times \left[(1 + 2Fr_{ma}) \left[(1 + 2Fr_{mb}) (1 + 2Fr_{ab}) + (1 - F) (r_{mb} + r_{ab} + 4Fr_{mb}r_{ab}) \right] T_{ab} D_{mab,ab}^0 - (1 + 2Fr_{mb}) [1 - (1 - 3F) r_{ma}] \left[\delta r_{ab} H_m (D_{ab} - D_{a,b}) pq_m + \delta r_{ab} (1 - H_m) (D_{mab,m} - D_{ma,mb}) \right] \right] \quad (\text{A6})$$

($D_{mb,a}$ being given by a symmetric expression), with $H_m = h_m + (1 - 2h_m) p_m$, and:

$$T_{ab} = \alpha_{ab} + \alpha_{ab,a} + \alpha_{ab,b} + \alpha_{ab,ab} + 2F (\alpha_a + \alpha_{a,a}) (\alpha_b + \alpha_{b,b}) \quad (\text{A7})$$

$$D_{ab} - D_{a,b} \approx \frac{(1 - F) [\tilde{\alpha}_{ab} - \tilde{\alpha}_a \tilde{\alpha}_b + 2F G_{ab} (\alpha_a + \alpha_{a,a}) (\alpha_b + \alpha_{b,b})]}{r_{ab} (1 - F)} pq_{ab} \quad (\text{A8})$$

$$D_{mab,m} - D_{ma,mb} \approx \frac{F (r_{ma} + r_{mb} - r_{ab} - 2r_{ma}r_{mb})}{1 + F (r_{ma} + r_{mb} + r_{ab} - 2r_{ma}r_{mb})} (D_{ab} - D_{a,b}) pq_m. \quad (\text{A9})$$

The identity disequilibrium G_{ab} is given by $\phi_{ab} - F^2$, with:

$$\phi_{ab} = F \left[1 - \frac{2(1 - F) r_{ab} (1 - r_{ab})}{1 + 2Fr_{ab} (1 - r_{ab})} \right]. \quad (\text{A10})$$

APPENDIX B: INTERPRETING EFFECTIVE SELECTION COEFFICIENTS

The form of the coefficients $\tilde{\alpha}_a$ and $\tilde{\alpha}_{ab}$ representing the effective strength of selection against deleterious alleles and the effective epistasis between pairs of deleterious alleles in a partially selfing population can be understood as follows. Assuming that deleterious alleles stay rare in the population ($u \ll s$), we can neglect homozygotes for deleterious alleles produced by outcrossing, and consider that homozygosity for deleterious alleles necessarily implies identity-by-descent among these alleles. Then, if an individual carries allele a on one of its haplotypes, with probability F it also carries allele a on its second haplotype. In that case, the fitness effect of the deleterious allele (α_a) is increased by α_a due to the presence of a on the second haplotype, and by $\alpha_{a,a}$ due to the interaction among those two deleterious alleles: therefore, the effective strength of selection experienced by deleterious alleles is $\tilde{\alpha}_a = (1 + F) \alpha_a + F \alpha_{a,a}$ (which, using equation 5, yields $\tilde{\alpha}_a \approx -s\tilde{h}$). A similar reasoning can be used to compute the effective epistasis coefficient. With probability ϕ_{ab} , an individual is identical-by-descent at both loci a and b ; with probability $F - \phi_{ab}$ it is identical-by-descent at the first locus and not at the second (or at the second and not at the first), while with the complementary probability $1 - 2F + \phi_{ab}$ it is identical-by-descent at neither locus. Therefore, an individual carrying alleles a and b on one of its haplotypes is $aabb$ with probability ϕ_{ab} , $aaBb$ (or $Aabb$) with probability $F - \phi_{ab}$, and $AaBb$ with probability $1 + 2F - \phi_{ab}$. In the first case ($aabb$), the fitness effect of the interaction between a and b (α_{ab}) is increased by the interaction between a and b present on the other haplotype (α_{ab}), twice the interaction between deleterious alleles in trans ($\alpha_{a,b}$), twice the additive-by-dominance interactions ($\alpha_{ab,a}$, $\alpha_{ab,b}$), and by the dominance-by-dominance interaction ($\alpha_{ab,ab}$). In the second case ($aaAb$ or $Aabb$), it is increased by the interaction between deleterious alleles in trans ($\alpha_{a,b}$) and by the additive-by-dominance interaction ($\alpha_{ab,a}$ or $\alpha_{ab,b}$). This yields:

$$\begin{aligned}
 \tilde{\alpha}_{ab} &= \phi_{ab} (2\alpha_{ab} + 2\alpha_{a,b} + 2\alpha_{ab,a} + 2\alpha_{ab,b} + \alpha_{ab,ab}) \\
 &\quad + (F - \phi_{ab}) (2\alpha_{ab} + 2\alpha_{a,b} + \alpha_{ab,a} + \alpha_{ab,b}) + (1 - 2F + \phi_{ab}) \alpha_{ab} \quad (\text{B1}) \\
 &= (1 + \phi_{ab}) \alpha_a + 2F \alpha_{a,b} + (F + \phi_{ab}) (\alpha_{ab,a} + \alpha_{ab,b}) + \phi_{ab} \alpha_{ab,ab}.
 \end{aligned}$$

APPENDIX C: GENERAL QLE APPROXIMATIONS

The following expressions combine approximations obtained under different regimes (high effective recombination, weak recombination, high selfing, epistasis of order ϵ^2 or ϵ , see Supplementary Material). The different associations affecting the change in frequency of the modifier (equation 6) are given by (with $D_{mab,ab}^0$ given by equation 10):

$$D_{mab,ab} \approx D_{mab,ab}^0 + F D_{mab} \quad (C1)$$

$$D_{mab,a} \approx D_{ma,ab} \approx F [(\alpha_b + \alpha_{b,b}) D_{mab,ab}^0 + D_{mab}] \quad (C2)$$

$$D_{m,ab} \approx D_{ma,b} \approx D_{mb,a} \approx F D_{mab} \quad (C3)$$

$$D_{ma,a} \approx F [\alpha_{b,b} D_{mab,ab}^0 + D_{ma}] \quad (C4)$$

$$D_{m,a} \approx F D_{ma} \quad (C5)$$

$$D_{mab} \approx \frac{(1 + 2F r_{ma})(1 + 2F r_{mb})}{Y_{mab}} \left[(1 + 2F r_{ab}) T_{ab} D_{mab,ab}^0 - \delta r_{ab} H_m (D_{ab} - D_{a,b}) pq_m - \delta r_{ab} (1 - H_m) (D_{mab,m} - D_{ma,mb}) \right] \quad (C6)$$

$$D_{ma} \approx \frac{(1 + 2F r_{ma}) \left[[2F (\alpha_a + \alpha_{a,a}) \alpha_{b,b} + \alpha_{ab,b} + \alpha_{ab,ab}] D_{mab,ab}^0 + (\tilde{\alpha}_b + \tilde{\alpha}_{ab}) D_{mab} \right]}{r_{ma} (1 - F) - (1 + 2F r_{ma}) \tilde{\alpha}_a} \quad (C7)$$

($D_{mab,b}$, $D_{mb,ab}$, $D_{mb,b}$, $D_{m,b}$ and D_{mb} being given by symmetric expressions), with:

$$T_{ab} = \alpha_{ab} + \alpha_{ab,a} + \alpha_{ab,b} + \alpha_{ab,ab} + 2F (\alpha_a + \alpha_{a,a}) (\alpha_b + \alpha_{b,b}) \quad (C8)$$

$$Y_{mab} = (1 - F) [r_{mab} + 2 (r_{ma} r_{mb} + r_{ma} r_{ab} + r_{mb} r_{ab}) + 6 r_{ma} r_{mb} r_{ab}] - (1 + 2F r_{ma}) (1 + 2F r_{mb}) (1 + 2F r_{ab}) (\tilde{\alpha}_a + \tilde{\alpha}_b + \tilde{\alpha}_{ab}) \quad (C9)$$

$$D_{ab} - D_{a,b} \approx \frac{(1 - F) [\tilde{\alpha}_{ab} - \tilde{\alpha}_a \tilde{\alpha}_b + 2F G_{ab} (\alpha_a + \alpha_{a,a}) (\alpha_b + \alpha_{b,b})]}{r_{ab} (1 - F) - (1 + 2F r_{ab}) (\tilde{\alpha}_a + \tilde{\alpha}_b + \tilde{\alpha}_{ab})} pq_{ab} \quad (C10)$$

$$D_{mab,m} - D_{ma,mb} \approx \frac{F (r_{ma} + r_{mb} - r_{ab} - 2r_{ma} r_{mb})}{1 + F (r_{ma} + r_{mb} + r_{ab} - 2r_{ma} r_{mb})} (D_{ab} - D_{a,b}) pq_m. \quad (C11)$$

The change in frequency of the modifier is given by $\Delta p_m = \sum_{U,V} \alpha_{U,V} D_{mU,V}$ where the double sum is over all elements of the set $\{\emptyset, a, b, ab\}$, and thus decomposes into a term generated by $D_{mab,ab}^0$ (that becomes predominant under high effective recombination) and a term generated by $D_{ab} - D_{a,b}$ (that becomes predominant under weak recombination or strong selfing).