Background Selection in Partially Selfing Populations

Denis Roze

Background selection in partially selfing populations

Denis Roze[*,§]

[*] CNRS, UMI 3614, Evolutionary Biology and Ecology of Algae, Roscoff, France

[§] Sorbonne Universités, UPMC Université Paris VI, Roscoff, France

Running title: Selfing and background selection

Keywords: deleterious mutation, genetic drift, effective population size, multilocus population genetics, self-fertilization

Address for correspondence:

Denis Roze

Station Biologique de Roscoff

Place Georges Teissier, CS90074

29688 Roscoff Cedex

France

Phone: (+33) 2 98 29 23 20

Fax: (+33) 2 98 29 23 24

email: roze@sb-roscoff.fr

ABSTRACT

Self-fertilizing species often present lower levels of neutral polymorphism than their outcrossing relatives. Indeed, selfing automatically increases the rate of coalescence per generation, but also enhances the effects of background selection and genetic hitchhiking by reducing the efficiency of recombination. Approximations for the effect of background selection in partially selfing populations have been derived previously assuming tight linkage between deleterious alleles and neutral loci. However, loosely linked deleterious mutations may have important effects on neutral diversity in highly selfing populations. In this paper, I use a general method based on multilocus population genetics theory to express the effect of a deleterious allele on diversity at a linked neutral locus in terms of moments of genetic associations between loci. Expressions for these genetic moments at equilibrium are then computed for arbitrary rates of selfing and recombination. An extrapolation of the results to the case where deleterious alleles segregate at multiple loci is checked using individual-based simulations. At high selfing rates, the tight linkage approximation underestimates the effect of background selection in genomes with moderate to high map length; however, another simple approximation can be obtained for this situation, and provides accurate predictions as long as the deleterious mutation rate is not too high.

INTRODUCTION

Understanding the evolutionary consequences of transitions between reproductive systems has been the focus of an important number of theoretical and empirical studies. In particular, the shift from biparental sexual reproduction to self-fertilization has occurred frequently in plants and animals (Goodwillie et al., 2005; Jarne and Auld, 2006), but the phylogenetic distribution of selfing lineages suggest that these are often relatively short-lived, and may thus correspond to an "evolutionary dead end" or "blind alley" (e.g., Stebbins, 1957; Williams, 1992; Takebayashi and Morrell, 2001; Goldberg et al., 2010; Igic and Busch, 2013). A possible reason for the lack of macroevolutionary success of selfing species may be their reduced capacity to produce novel genotypes (in particular, genotypes adapted to new environmental conditions), due to a reduced efficiency of recombination. Furthermore, self-fertilization lowers the effective size of populations and should thereby decrease the efficiency of natural selection against deleterious alleles, which may lead to mutation accumulation and population extinction (Lynch et al., 1995; Schultz and Lynch, 1997). Analyses based on molecular data show little evidence for increased ratios of non-synonymous to synonymous substitutions (d_N/d_S) in selfing lineages, that would indicate a reduced efficiency of purifying selection (Glémin and Muyle, 2014; Hartfield, 2015 and references therein): this may be due to the recent origin of those lineages, or to the low rates of outcrossing maintained by most predominantly selfing species (Wright et al., 2013). However, several recent studies showed elevated ratios of non-synonymous to synonymous polymorphism (π_N/π_S) in various selfing species (compared with their outcrossing relatives), suggesting that deleterious alleles may reach higher frequencies in selfers (e.g., Brandvain et al., 2013;

Burgarella et al., 2015, and other references listed in Table 1 of Hartfield, 2015).

The lower effective size of selfing populations has been demonstrated empirically using neutral diversity data from a variety of species (e.g., Charlesworth, 2003; Glémin et al., 2006), and is thought to result from two types of effects. The first is an automatic increase in the rate of coalescence per generation, since the ancestral lineages of the two homologous copies of a gene in an individual may coalesce in a single generation (with probability 1/2) if this individual has been produced by selfing. Due to this effect, the effective population size is given by $N_{\mathrm{e}} = N/(1+F)$ (Pollak, 1987; Nordborg, 2000), where N is the census size and where the inbreeding coefficient F equals $\alpha/(2-\alpha)$ in a population in which a proportion α of individuals are produced by selfing. Therefore, N_{e} is expected to decline linearly from N to $N/2$ as α increases from 0 to 1. However, N_{e} may be further decreased by selective sweeps or by selection against deleterious alleles (background selection, Charlesworth et al., 1993; Charlesworth, 2012), whose effects are amplified by the lower effective recombination rates of selfing populations.

Several models have computed the effect of background selection on neutral diversity in randomly mating populations, using different approaches (Hudson and Kaplan, 1995; Nordborg et al., 1996; Santiago and Caballero, 1998; Charlesworth, 2012). These showed in particular that a deleterious allele at mutation-selection balance reduces the expected diversity at a linked neutral locus by a factor $\approx 1-(ush)/(r+sh)^2$, where u is the mutation rate towards the deleterious allele, sh the heterozygous fitness effect of this allele (assumed different from zero) and r the recombination rate between the two loci. The case of partially selfing populations has been addressed by Nordborg (1997) using a structured coalescent model and a separation of timescales argument. Indeed, assuming that recombination and coalescence of lineages present

in different individuals occur at a much lower rate than the coalescence of lineages present in the same individual due to selfing, the population can be described in terms of haplotypes instead of diploid genotypes, which considerably simplifies the analysis. Under this assumption, the effect of a deleterious allele on linked neutral diversity is given by a similar expression as in the panmictic case, replacing sh by $s\left[h\left(1-F\right)+F\right]$ (measuring the strength of selection against the deleterious allele in a partially selfing population), and r by the effective recombination rate $r\left(1-F\right)$ (see also Nordborg, 2000). Extrapolating this result to the case of deleterious alleles segregating at many loci, Glémin (2007) and Glémin and Ronfort (2012) showed that the effective size of highly selfing populations may be strongly reduced by background selection effects.

Strictly, Nordborg (1997)'s result holds for tightly linked loci, since the separation of timescales argument supposes a low recombination rate. While the effective population size at a given locus should be little affected by loosely linked loci as long as the selfing rate remains moderate, this may be less so when the selfing rate is high, so that linkage disequilibria may extend over relatively large genetic distances (e.g., Nordborg et al., 2002). Using multilocus individual-based simulations of partially selfing populations, Kamran-Disfani and Agrawal (2014) observed discrepancies between the estimated N_{e} and predictions obtained by extrapolating Nordborg (1997)'s result over a whole genetic map. These may be caused by the fact that the effects of loosely linked loci are not sufficiently well predicted by the separation of timescales approximation, and become important at high selfing rates.

In this paper, I construct a model of background selection in partially selfing populations by extending the multilocus population genetics framework previously developed by Barton and Turelli (1991) and Kirkpatrick et al. (2002). As we will see,

a strength of this approach is that it allows one to decompose evolutionary processes (here the background selection effect) into different terms involving linkage disequilibria and other forms of genetic associations, for which intuitive interpretation can be given. Expressions valid for any value of the recombination rate will be derived, and shown to converge to Nordborg (1997)'s result when linkage is tight. However, this tight linkage approximation may significantly underestimate the strength of background selection when the selfing rate is high (but below 1); we will see that another approximation yielding better predictions at high selfing rates can be obtained from the general model. A good match between the analytical predictions and multilocus simulation results is observed as long as the genomic deleterious mutation rate U is not too high (0.1 per haploid genome), while discrepancies appear at higher values of U: those are likely due to genetic associations between deleterious alleles at different loci, which are neglected in the analysis.

MODEL

General method. As in previous models of background selection (e.g., Hudson and Kaplan, 1995; Nordborg et al., 1996), I will first consider the effect of a single deleterious allele maintained at mutation-selection balance at a given locus on the dynamics of genetic diversity at a linked neutral locus. This effect can be quantified by computing the expected change in neutral diversity over one generation, which is affected by various moments of genetic associations between the two loci (e.g., the variance in linkage disequilibrium, and other moments of associations between genes present either on the same haplotype or on different haplotypes of a diploid individual). Assuming recurrent

mutation at the neutral locus, the effective population size N_e at the neutral locus can be deduced from the expected neutral diversity at equilibrium (Nordborg et al., 1996). Alternatively, we may ignore mutation at the neutral locus and calculate N_e by equating the expected rate of loss in diversity per generation to $-1/(2N_e)$, since diversity is eroded at a rate $-1/(2N)$ per generation in a Wright-Fisher population. This is the approach that will be used here. Strictly, it relies on a quasi-equilibrium approximation, since moments of genetic associations (for example, the variance in linkage disequilibrium) will be expressed in terms of diversity at the neutral locus and of the frequency of the deleterious allele, implying that these moments of genetic associations equilibrate fast relative to changes in allele frequencies. However, this approximation is justified when population size is sufficiently large (so that changes in allele frequencies due to drift remain small) and yields the same expression for N_e as what would be obtained by calculating the equilibrium neutral diversity under recurrent mutation.

I consider the following life cycle: N individuals are present at the start of each generation, and produce a very large (effectively infinite) number of juveniles in proportion to their fitness. A proportion α of offspring is produced by selfing, while the remaining $1-\alpha$ is produced by random fusion of gametes. Finally, N individuals are sampled randomly among all juveniles produced, to form the next generation (drift). Fitness depends on genotype at the selected locus, where two alleles (denoted 0 and 1) are segregating: allele 1 is deleterious, reducing fitness by a factor $1-sh$ in the heterozygous state, and $1-s$ in the homozygous state. The deleterious mutation rate (from allele 0 to allele 1) is denoted u. As in previous treatments (Hudson and Kaplan, 1995; Nordborg et al., 1996) I will assume that $sN \gg 1$ and $u \ll s$, so that the frequency of the deleterious allele remains small and can be approximated

by the deterministic mutation-selection balance frequency (strictly, this also assumes $Np_{\mathrm{del}} \gg 1$, where p_{del} is the frequency of the deleterious allele). Finally, r measures the recombination rate between the two loci.

In the following I use a general method to compute the effects of selection, reproduction and drift on moments of genetic associations. This method is based on a previous formalism for the analysis of multilocus models (Barton and Turelli, 1991; Kirkpatrick et al., 2002), extended to include genetic drift. It was used previously to study selection for sex in finite diploid populations undergoing both sexual and asexual reproduction (Roze and Michod, 2010), and is described in Appendix A for the case of a partially selfing population. To simplify the notation, the examples shown in Appendix A concern the case of a biallelic neutral locus; however the method extends to multiple alleles, yielding the same expression for the decay of neutral diversity per generation. The analysis of the two-locus model will then proceed in three steps. First, I will express the expected decay of genetic diversity per generation in terms of various genetic moments involving both loci. Then, recurrence equations describing the dynamics of these two-locus moments will be derived (to the first order in $1/N$, and assuming that the deleterious allele stays at low frequency). Finally, these equations will be solved to obtain expressions for two-locus moments at (quasi-)equilibrium, in terms of diversity at the neutral locus and of the different parameters of the model. Injecting these solutions into the equation describing the decay of neutral diversity will yield an expression for the effect of the deleterious allele on N_{e} at the neutral locus.

The results of this two-locus model will then be extrapolated to a situation where deleterious alleles segregate at a large number of loci, located at various genetic distances from the neutral locus. For this, I will assume that the effects of the different

selected loci on diversity at the neutral locus are multiplicative, thereby neglecting genetic associations between selected loci. In the absence of epistasis between deleterious alleles, this approximation is expected to yield correct results under random mating in the regime considered here (where selection against deleterious alleles is stronger than drift, e.g., Hudson and Kaplan, 1995) but may be less accurate under partial selfing, as inbreeding generates different forms of associations between loci (correlations in homozygosity in particular, e.g., Roze, 2015). Nonetheless, we will see that this assumption of multiplicative effects often generates accurate predictions, as long as the genomic deleterious mutation rate is not too high.

Defining genetic associations. The parameters and variables of the model are summarized in Table 1. Throughout the following, the neutral locus is denoted A, while the selected locus is denoted B. Two alleles denoted 0 and 1 segregate at each locus (we will see below how the notation can be extended to deal with multiple neutral alleles), allele 1 at locus B being the deleterious allele. Indicator variables X_i^{M} and X_i^{P} describe the genotype of an individual at locus i: these variables equal 1 if allele 1 is present on the maternally or paternally (respectively) inherited chromosome of this individual at locus i, and 0 otherwise. The frequency of allele 1 at locus i (denoted p_i) is thus:

$$p_i = \mathrm{E}\left[\frac{X_i^{\mathrm{M}} + X_i^{\mathrm{P}}}{2}\right] \tag{1}$$

where E stands for the average over all individuals in the population. Neglecting drift, the frequency of the deleterious allele at mutation-selection balance (denoted $\tilde{p}_B$) is given by:

$$\tilde{p}_B \approx \frac{u}{s\left[h\left(1-F\right)+F\right]} \tag{2}$$

with $F = \alpha/(2-\alpha)$ (e.g., Glémin, 2003).

For each locus i, centered variables ζ_i^{M} and ζ_i^{P} are defined as:

$$\zeta_i^{\mathrm{M}} = X_i^{\mathrm{M}} - p_i, \qquad \zeta_i^{\mathrm{P}} = X_i^{\mathrm{P}} - p_i\,. \tag{3}$$

Following Kirkpatrick et al. (2002), the association between the sets $\mathbb{S}$ and $\mathbb{T}$ of loci present on the two haplotypes of the same individual is given by:

$$D_{\mathbb{S},\mathbb{T}} = \mathrm{E}\left[\zeta_{\mathbb{S},\mathbb{T}}\right] \tag{4}$$

where

$$\begin{aligned} \zeta_{\mathbb{S},\mathbb{T}} &= \frac{\zeta_{\mathbb{S}}^{\mathrm{M}}\,\zeta_{\mathbb{T}}^{\mathrm{P}} + \zeta_{\mathbb{S}}^{\mathrm{P}}\,\zeta_{\mathbb{T}}^{\mathrm{M}}}{2}, \\ \zeta_{\mathbb{S}}^{\mathrm{M}} &= \prod_{i\in\mathbb{S}} \zeta_i^{\mathrm{M}}, \quad \zeta_{\mathbb{T}}^{\mathrm{P}} = \prod_{i\in\mathbb{T}} \zeta_i^{\mathrm{P}} \end{aligned} \tag{5}$$

(note that $D_{\mathbb{S},\mathbb{T}} = D_{\mathbb{T},\mathbb{S}}$), and where sets $\mathbb{S}$ and $\mathbb{T}$ may be the empty set $\emptyset$, A, B or AB. Associations between genes present on the same haplotype of an individual ($D_{\mathbb{S},\emptyset}$) will be simply denoted $D_{\mathbb{S}}$. For example, $D_{A,A} = \mathrm{E}\left[\left(X_A^{\mathrm{M}} - p_A\right)\left(X_A^{\mathrm{P}} - p_A\right)\right]$ is a measure of the departure from Hardy-Weinberg equilibrium at locus A, while $D_{AB} = \frac{1}{2}\mathrm{E}\left[\left(X_A^{\mathrm{M}} - p_A\right)\left(X_B^{\mathrm{M}} - p_B\right) + \left(X_A^{\mathrm{P}} - p_A\right)\left(X_B^{\mathrm{P}} - p_B\right)\right]$ represents the linkage disequilibrium between loci A and B (genetic association between alleles present on the same haplotype, maternal or paternal). Similarly, $D_{A,B} = \frac{1}{2}\mathrm{E}\big[\left(X_A^{\mathrm{M}} - p_A\right)\left(X_B^{\mathrm{P}} - p_B\right) + \left(X_A^{\mathrm{P}} - p_A\right)\left(X_B^{\mathrm{M}} - p_B\right)\big]$ measures the association between alleles at loci A and B present on different haplotypes of the same individual.

Because population size is finite, allele frequencies and genetic associations are random variables. Throughout the paper, I will use the notation $\langle\mathcal{M}\rangle$ for the expected value of the genetic moment $\mathcal{M}$ (a product of allele frequencies, genetic associations, or both) at a given generation: for example, $\langle {D_{AB}}^2\rangle$ is the expected

squared linkage disequilibrium between the two loci. In the following, the moment $\langle D_{AA}\rangle = \left\langle \frac{1}{2}\mathrm{E}\left[\left(X_A^{\mathrm{M}} - p_A\right)^2 + \left(X_A^{\mathrm{P}} - p_A\right)^2\right]\right\rangle$ will be of particular importance: indeed, using the fact that $\left(X_A^{\mathrm{M}}\right)^2 = X_A^{\mathrm{M}}$ and $\left(X_A^{\mathrm{P}}\right)^2 = X_A^{\mathrm{P}}$ (since these variables equal 0 or 1), one obtains that $\langle D_{AA}\rangle = \langle p_A q_A\rangle$ (where $q_i = 1 - p_i$), thus representing the expected genetic diversity at locus A. Therefore, the effective population size at the neutral locus can be quantified by computing the rate of decay of $\langle D_{AA}\rangle$ per generation.

As mentioned above, the model can be extended to an arbitrary number n of alleles segregating at the neutral locus. In this case, we can define indicator variables $X_{A,k}^{\mathrm{M}}$ and $X_{A,k}^{\mathrm{P}}$ that equal 1 if the maternally (respectively, paternally) inherited chromosome of a given individual carries allele k at locus A (and 0 otherwise), with $k = 1, \ldots, n$. Vectors $\mathbf{X}_A^{\mathrm{M}}$ and $\mathbf{X}_A^{\mathrm{P}}$ are defined as $\mathbf{X}_A^{\mathrm{M}} = \left(X_{A,1}^{\mathrm{M}}, X_{A,2}^{\mathrm{M}}, \ldots, X_{A,n}^{\mathrm{M}}\right)$ and $\mathbf{X}_A^{\mathrm{P}} = \left(X_{A,1}^{\mathrm{P}}, X_{A,2}^{\mathrm{P}}, \ldots, X_{A,n}^{\mathrm{P}}\right)$; in each of these vectors (and for a given individual) a single element equals 1 while all other elements equal zero. Finally, the vector $\mathbf{p}_A = \mathrm{E}\left[\left(\mathbf{X}_A^{\mathrm{M}} + \mathbf{X}_A^{\mathrm{P}}\right)/2\right]$ holds the frequencies of the different alleles at locus A in the population. Defining $\boldsymbol{\zeta}_A^{\mathrm{M}} = \mathbf{X}_A^{\mathrm{M}} - \mathbf{p}_A$ and $\boldsymbol{\zeta}_A^{\mathrm{P}} = \mathbf{X}_A^{\mathrm{P}} - \mathbf{p}_A$, genetic associations may be defined in the same way as above, associations with two "A" subscripts involving a dot product between the corresponding $\boldsymbol{\zeta}_A$ vectors. In particular:

$$D_{AA} = \frac{1}{2}\mathrm{E}\left[\boldsymbol{\zeta}_A^{\mathrm{M}}.\boldsymbol{\zeta}_A^{\mathrm{M}} + \boldsymbol{\zeta}_A^{\mathrm{P}}.\boldsymbol{\zeta}_A^{\mathrm{P}}\right] = 1 - \sum_{k=1}^{n} p_{A,k}^2 \tag{6}$$

where $p_{A,k}$ is the frequency of allele k at locus A in the population. Similarly, $D_{A,A} = \mathrm{E}\left[\boldsymbol{\zeta}_A^{\mathrm{M}}.\boldsymbol{\zeta}_A^{\mathrm{P}}\right]$, while $D_{AB,A} = \frac{1}{2}\mathrm{E}\left[\left(\boldsymbol{\zeta}_A^{\mathrm{M}}.\boldsymbol{\zeta}_A^{\mathrm{P}}\right)\left(\zeta_B^{\mathrm{M}} + \zeta_B^{\mathrm{P}}\right)\right]$. Equation 6 shows that, as in the biallelic case, $\langle D_{AA}\rangle$ represents the expected genetic diversity at the neutral locus. As mentioned earlier, the method for computing multilocus moments is explained in Appendix A for the case of a biallelic neutral locus, but the results shown

below are valid for any number of alleles segregating at this locus.

Multilocus simulations. Analytical predictions will be tested using individual based, multilocus simulations. The simulation program (written in C++, and available from Dryad) is very similar to the program used in Roze (2015), representing of population of N diploid individuals whose genome consists of a linear chromosome with total map length R. Relatively large values of R (usually 10 Morgans) will be used in most simulations in order to mimic a whole genome with multiple chromosomes. Each generation, deleterious alleles occur at rate U per haploid genome; all deleterious alleles have the same selection and dominance coefficient, and have multiplicative effects on fitness (no epistasis). As we will see, variable selection coefficients have been implemented in a different version of the program. Offspring are formed by selfing with probability α, and by random fusion of gametes with probability $1-\alpha$. A neutral locus with an infinite number of possible alleles is located at the mid-point of the chromosome, mutating at a rate $\mu = 10^{-3}$ per generation. The program runs for 2×10^6 generations, genetic diversity at the neutral locus being recorded every 50 generations, and measured as $D = 1 - \sum_i {p_i}^2$ (where p_i is the frequency of neutral allele i). The effective population size at the neutral locus is then estimated from $\overline{D} = 4N_e\mu/(1+4N_e\mu)$ where $\overline{D}$ is the average neutral diversity at equilibrium, yielding:

$$N_e = \frac{\overline{D}}{4\mu\left(1-\overline{D}\right)} \tag{7}$$

(the exact expression including terms in μ^2 yields undistinguishable results for the parameter values used here). In the simulations, $\overline{D}$ is obtained by averaging after a burn-in period of 15000 generations, which was sufficient for diversity to reach equi-

librium with $\mu = 10^{-3}$. As we will see, a different version of the program including a neutral sequence with an infinite number of sites was also used, in which case diversity takes longer to equilibrate.

Data availability: DRYAD DOI: doi:10.5061/dryad.p3r01

RESULTS

All results are obtained using the method presented in Appendix A for computing recursions for multilocus moments, implemented in a *Mathematica* notebook (Supplementary File 1). All terms are derived to the first order in $1/N$ and to the first order in $\tilde{p}_B$ (the frequency of the deleterious allele at mutation-selection balance, given by equation 2).

General results. The expected change in genetic diversity at locus A per generation can be written as:

$$\langle \Delta D_{AA} \rangle = \langle \Delta_{\mathrm{s}} D_{AA} \rangle + \langle \Delta_{\mathrm{d}} D_{AA} \rangle \tag{8}$$

where $\langle \Delta_{\mathrm{s}} D_{AA} \rangle$ is the change in diversity due to selection and $\langle \Delta_{\mathrm{d}} D_{AA} \rangle$ the change in diversity due to drift, given by:

$$\langle \Delta_{\mathrm{d}} D_{AA} \rangle = -\frac{1}{2N} \left(\langle D_{AA} \rangle + \langle D_{A,A} \rangle \right) \tag{9}$$

to leading order. In the absence of selection, $\langle \Delta_{\mathrm{s}} D_{AA} \rangle$ equals zero while $\langle D_{A,A} \rangle$ at quasi-equilibrium is given by (from equation A10):

$$\langle D_{A,A} \rangle = F \langle D_{AA} \rangle \tag{10}$$

with $F = \alpha/(2-\alpha)$, and the expected change in neutral diversity per generation thus becomes:

$$\langle \Delta D_{AA} \rangle = \langle \Delta_{\mathrm{d}} D_{AA} \rangle = -\frac{1}{2N}(1+F)\langle D_{AA} \rangle . \tag{11}$$

This corresponds to the classical result that N_{e} equals $N/(1+F)$ under partial selfing (Pollak, 1987): the increased homozygosity caused by selfing amplifies the effect of drift, since the same allele is sampled twice every time a homozygote is sampled. When selection acts at locus B, one obtains the following expression for $\langle \Delta_{\mathrm{s}} D_{AA} \rangle$ to the first order in s:

$$\begin{aligned} \langle \Delta_{\mathrm{s}} D_{AA} \rangle = & -sh\left(\langle D_{AAB} \rangle + \langle D_{AA,B} \rangle\right) \\ & - s(1-2h)\left(\langle D_{AAB,B} \rangle - \langle D_{AA} D_{B,B} \rangle\right) . \end{aligned} \tag{12}$$

An expression to the second order in s is provided in Appendix B; however both expressions generally yield very similar quantitative results, although adding the terms in s^2 may slightly improve the predictions under loose linkage. The different terms that appear in equation 12 may be interpreted as follows. From the definitions given in the previous section, the term $\langle D_{AAB} \rangle + \langle D_{AA,B} \rangle$ may also be written as $\mathrm{E}\left[\frac{1}{2}\left(\zeta_{AA}^{\mathrm{M}} + \zeta_{AA}^{\mathrm{P}}\right)\left(\zeta_{B}^{\mathrm{M}} + \zeta_{B}^{\mathrm{P}}\right)\right]$ (where again E stands for the average over all individuals), thus measuring a covariance between $\frac{1}{2}\left(\zeta_{AA}^{\mathrm{M}} + \zeta_{AA}^{\mathrm{P}}\right)$ and $\zeta_{B}^{\mathrm{M}} + \zeta_{B}^{\mathrm{P}}$ (note that $\mathrm{E}\left[\zeta_{B}^{\mathrm{M}} + \zeta_{B}^{\mathrm{P}}\right] = 0$). The quantity $\frac{1}{2}\left(\zeta_{AA}^{\mathrm{M}} + \zeta_{AA}^{\mathrm{P}}\right)$ is higher in individuals carrying rarer alleles at the neutral locus: for example in the case of a biallelic neutral locus, it equals ${p_A}^2$, $\left({p_A}^2 + {q_A}^2\right)/2$ and ${q_A}^2$ in 00, 01 and 11 individuals, where p_A is the frequency of allele 1. Furthermore, the quantity $\zeta_{B}^{\mathrm{M}} + \zeta_{B}^{\mathrm{P}}$ is higher in individuals carrying more deleterious alleles at the selected locus (it is nearly 0, 1 and 2 in individuals carrying 0, 1 and 2 deleterious alleles at locus B, assuming p_B is small). Therefore, a positive value of $\langle D_{AAB} \rangle + \langle D_{AA,B} \rangle$ indicates that rarer alleles at the neutral locus tend to

be found in individuals carrying higher numbers of deleterious alleles at the selected locus, while a negative value of $\langle D_{AAB} \rangle + \langle D_{AA,B} \rangle$ indicates the opposite. Recursions for the moments $\langle D_{AAB} \rangle$ and $\langle D_{AA,B} \rangle$ over one generation are given in Appendix B, to the first order in s, $\tilde{p}_B$ and $1/N$. In the absence of selfing ($\alpha = 0$), one obtains that $\langle D_{AA,B} \rangle = 0$ at quasi-equilibrium, while $\langle D_{AAB} \rangle$ is generated by the variance in linkage disequilibrium $\langle {D_{AB}}^2 \rangle$ and by the effect of selection, and is positive when $sh > 0$. Indeed, when a given neutral allele becomes associated (by chance) to the deleterious allele at locus B, this neutral allele tends to decrease in frequency, generating a positive $\langle D_{AAB} \rangle$ (the deleterious allele at locus B tends to become associated with rarer alleles at locus A). This in turns reduces genetic diversity at the neutral locus (as shown by equation 12), since these rarer alleles will further decrease in frequency due to their association with the deleterious allele. Because partial selfing generates cross-haplotype associations (between genes present on different haplotypes of a diploid), $\langle D_{AAB} \rangle$ and $\langle D_{AA,B} \rangle$ are given by more complicated expressions when $\alpha > 0$, involving moments such as $\langle D_{AB} D_{A,B} \rangle$, $\langle D_{AB} D_{AB,B} \rangle$ or $\langle {D_{AB,B}}^2 \rangle$ (see Appendix B).

Similarly, the quantity $\langle D_{AAB,B} \rangle - \langle D_{AA} D_{B,B} \rangle$ that appears on the second line of equation 12 measures a covariance between $\frac{1}{2}\left(\zeta_{AA}^{\mathrm{M}} + \zeta_{AA}^{\mathrm{P}}\right)$ and $\zeta_{B,B}$, the quantity $\zeta_{B,B}$ being higher in homozygotes at locus B than in heterozygotes. A positive value of $\langle D_{AAB,B} \rangle - \langle D_{AA} D_{B,B} \rangle$ thus indicates that rarer alleles at locus A tend to be found more often in homozygotes at locus B than in heterozygotes. As shown by equation 12, this would reduce neutral diversity when the deleterious allele is partially recessive ($h < 0.5$), since in this case homozygotes at the selected locus have a lower fitness than heterozygotes. A recursion for $\langle D_{AAB,B} \rangle - \langle D_{AA} D_{B,B} \rangle$ is given in Appendix B (equation B14); remarkably, it shows that identity disequilibrium (covariance in homozygosity)

between the two loci generates negative $\langle D_{AAB,B}\rangle - \langle D_{AA}D_{B,B}\rangle$ (that is, rarer alleles at locus A tend to be found more often in heterozygotes at locus B) even in the absence of selection. Indeed, setting $s = 0$ in equation B14 yields at quasi-equilibrium:

$$\langle D_{AAB,B}\rangle - \langle D_{AA}D_{B,B}\rangle = -\frac{2}{N\,(2-\alpha)}\,G_{AB}\,\tilde{p}_B\,\langle D_{AA}\rangle \tag{13}$$

where G_{AB} is the identity disequilibrium between loci A and B (Weir and Cockerham, 1969). It is given by $G_{AB} = \phi_{AB} - F^2$, where

$$\phi_{AB} = \frac{\alpha}{2-\alpha}\frac{2-\alpha-2\,(2-3\alpha)\,r\,(1-r)}{2-\alpha\,[1-2r\,(1-r)]} \tag{14}$$

is the probability of joint identity-by-descent at the two loci. This result may be interpreted as follows. Due to identity disequilibrium, the frequency of heterozygotes at locus A is higher among heterozygotes at locus B than among homozygotes. Furthermore, the frequency of rarer neutral alleles is higher among heterozygotes than among homozygotes at locus A (this is easily seen in the case of a biallelic locus): therefore, the frequency of rarer alleles at locus A should be higher among heterozygotes at locus B. As shown by equation 12, this effect tends to *increase* neutral diversity, as long as the deleterious allele is partially recessive ($h < 0.5$), so that heterozygotes at locus B have a higher fitness than homozygotes. As we will see in the next subsection, the effect of identity disequilibrium dominates over all other effects when the effective recombination rate $r\,(1-F)$ is sufficiently high (more precisely, $r\,(1-F) \gg s$), in which case partially recessive deleterious alleles tend to increase neutral diversity (this effect usually stays rather small). However, weaker effective recombination increases the relative importance of the terms on the first line of equation 12, that tend to decrease neutral diversity. Furthermore, weak effective recombination changes the sign of $\langle D_{AAB,B}\rangle - \langle D_{AA}D_{B,B}\rangle$ (through the terms in s in equation B14), due to the fact

that the association between rarer alleles at locus A and the deleterious allele at locus B (represented by moments $\langle D_{AAB} \rangle$, $\langle D_{AA,B} \rangle$) generate an association between those rarer alleles and homozygosity for the deleterious allele at locus B.

As shown in Appendix B, calculating the different terms of equation 12 at quasi-equilibrium requires computing 6 two-locus moments that are generated by finite population size: $\langle {D_{AB}}^2 \rangle$, $\langle D_{AB} D_{A,B} \rangle$, $\langle {D_{A,B}}^2 \rangle$, $\langle D_{AB} D_{AB,B} \rangle$, $\langle D_{A,B} D_{AB,B} \rangle$ and $\langle {D_{AB,B}}^2 \rangle$. Recursions for these moments are also given in Appendix B. Although the solutions obtained are rather complicated, they are readily computed numerically using *Mathematica* (see Supplementary File 2); furthermore, we will see that they can be approximated by simpler expressions in several cases. Importantly, all expressions obtained are in $1/N$, and thus vanish when N tends to infinity.

Besides its effect on $\langle \Delta_{\mathrm{s}} D_{AA} \rangle$, selection at locus B also affects the average excess homozygosity at locus A $\langle D_{A,A} \rangle$, and thus the term $\langle \Delta_{\mathrm{d}} D_{AA} \rangle$ in equation 8. Because $\langle D_{A,A} \rangle$ is multiplied by $1/N$ in equation 9, it is sufficient to compute this moment in the limit as N tends to infinity, in order to obtain an expression for $\langle \Delta D_{AA} \rangle$ to the first order in $1/N$. Using the results in Roze, 2015 (also derived in Supplementary File 1), we have at quasi-equilibrium, to the first order in s and $\tilde{p}_B$:

$$\langle D_{A,A} \rangle = F \left[1 - s \left(1 - 2h \right) G_{AB} \, \tilde{p}_B \right] \langle D_{AA} \rangle \tag{15}$$

where again G_{AB} is the identity disequilibrium between loci A and B. Indeed, homozygosity at locus A is reduced when the deleterious allele at locus B is partially recessive ($h < 1/2$), due to the fact that homozygotes at locus A tend to be also homozygous at locus B, while homozygotes at locus B have a lower fitness than heterozygotes.

Putting everything together, one obtains an expression for the change in diver-

sity at locus A over one generation of the form:

$$\Delta \langle D_{AA} \rangle = -\frac{1}{2N} (1 + F) (1 + T \tilde{p}_B) \langle D_{AA} \rangle \tag{16}$$

where T is a function of s, h, r and α. To the first order in $\tilde{p}_B$, we thus have:

$$N_{\mathrm{e}} = \frac{N}{1 + F} B_{sel} \tag{17}$$

where $B_{sel} = 1 - T \tilde{p}_B$ represents the effect of background selection. Although the expression obtained for T from the equations given in Appendix B is complicated, we will now see that simple approximations can be obtained in several regimes (in particular, high effective recombination, tight linkage, and high selfing).

High effective recombination (with partial selfing). Under partial selfing and when the effective recombination rate $r (1 - F)$ is high (and assuming that the dominance coefficient h of the deleterious allele is significantly different from 0.5), equation 12 is dominated by the term on the second line, since $\langle D_{AAB,B} \rangle - \langle D_{AA} D_{B,B} \rangle$ is generated by drift even in the absence of selection and is thus proportional to $1/N$ (see equation 13), while the terms $\langle D_{AAB} \rangle$ and $\langle D_{AA,B} \rangle$ on the first line are generated by selection and drift, and are thus proportional to s/N. Neglecting the term on the first line of equation 12 and using equations 9 and 15 yields the following expression for the change in neutral diversity, to the first order in s, $\tilde{p}_B$ and $1/N$:

$$\Delta \langle D_{AA} \rangle = -\frac{1}{2N} (1 + F) \left[1 - s (1 - 2h) \frac{4 + \alpha}{2} G_{AB} \tilde{p}_B \right] \langle D_{AA} \rangle . \tag{18}$$

Using equation 2, one obtains for the effective population size at the neutral locus:

$$N_{\mathrm{e}} \approx \frac{N}{1 + F} \left[1 + \frac{u (1 - 2h)}{h (1 - F) + F} \frac{4 + \alpha}{2} G_{AB} \right] , \tag{19}$$

independent of s. Equation 19 shows that identity disequilibrium between the neutral and the selected locus (G_{AB}) increases the effective population size at the neutral locus when the deleterious allele is partially recessive. This is caused by two effects: (i) identity disequilibrium reduces the excess homozygosity at locus A caused by selfing (equation 15), since homozygotes at locus A tends to be also homozygous at locus B, while homozygotes at locus B have a lower fitness than heterozygotes when the deleterious allele is partially recessive; (ii) the higher fitness of heterozygotes at locus B increases diversity at locus A since heterozygotes at locus B tend to be also heterozygous at locus A (equation 13).

This increase in effective population size caused by identity disequilibrium usually stays modest, however (since it is only expected to occur for high effective recombination), and is thus difficult to observe in simulations. Background selection has stronger effects when the effective recombination rate becomes low, in which case the term $\langle D_{AAB} \rangle + \langle D_{AA,B} \rangle$ on the first line of equation 12 becomes of the same order of magnitude than the term on the second line (indeed, one can show that the denominator of $\langle D_{AAB} \rangle$ and $\langle D_{AA,B} \rangle$ is proportional to ϵ when both $r\,(1-F)$ and s are of order ϵ), while the sign of $\langle D_{AAB,B} \rangle - \langle D_{AA} D_{B,B} \rangle$ changes due to the effect of selection. Two approximations for this regime are given below (tight linkage, high selfing).

Random mating. In the absence of selfing ($\alpha = 0$), equation B4 – B9 yield the following expressions for $\langle {D_{AB}}^2 \rangle$ and $\langle {D_{A,B}}^2 \rangle$ at quasi-equilibrium:

$$\left\langle {D_{AB}}^2 \right\rangle = \frac{1 + r^2\,(1 - 2sh)}{2N\left[1 - (1-r)^2\,(1-2sh)\right]}\,\tilde{p}_B\,\langle D_{AA} \rangle \tag{20}$$

$$\left\langle {D_{A,B}}^2 \right\rangle = \frac{1}{2N}\,\tilde{p}_B\,\langle D_{AA} \rangle \tag{21}$$

while $\langle D_{AB}D_{A,B}\rangle$, $\langle D_{AB}D_{AB,B}\rangle$, $\langle D_{A,B}D_{AB,B}\rangle$ and $\langle {D_{AB,B}}^2\rangle$ equal zero. Furthermore, equations B10 – B14 yield:

$$\langle D_{AAB}\rangle = \frac{2sh\left[(1-r)\langle {D_{AB}}^2\rangle + r\langle {D_{A,B}}^2\rangle\right]}{1-(1-r)(1-sh)} \tag{22}$$

while $\langle D_{AA,B}\rangle$ and $\langle D_{AAB,B}\rangle - \langle D_{AA}D_{B,B}\rangle$ equal zero. Finally, the rate of decay of neutral diversity is given by:

$$\langle \Delta D_{AA}\rangle = -sh\langle D_{AAB}\rangle - (sh)^2\left(\left\langle {D_{AB}}^2\right\rangle + \left\langle {D_{A,B}}^2\right\rangle\right) - \frac{1}{2N}\langle D_{AA}\rangle . \tag{23}$$

Under tight linkage (*i.e.*, both r and s are of order ϵ, where ϵ is a small term), one obtains from equations 20 – 22:

$$\left\langle {D_{AB}}^2\right\rangle \approx \frac{\tilde{p}_B\langle D_{AA}\rangle}{4N(r+sh)} \gg \left\langle {D_{A,B}}^2\right\rangle , \tag{24}$$

$$\langle D_{AAB}\rangle \approx \frac{sh\,\tilde{p}_B\langle D_{AA}\rangle}{2N(r+sh)^2} \tag{25}$$

giving:

$$\langle \Delta D_{AA}\rangle \approx -\frac{1}{2N}\left[1+\left(\frac{sh}{r+sh}\right)^2\tilde{p}_B\right]\langle D_{AA}\rangle \tag{26}$$

in agreement with the results obtained by Hudson and Kaplan (1995) and Nordborg et al. (1996). Under loose linkage ($r \gg s$), equations 20 – 23 yield:

$$\langle \Delta D_{AA}\rangle \approx -\frac{1}{2N}\left[1+\frac{(sh)^2(1+4r^2)}{r^2}\tilde{p}_B\right]\langle D_{AA}\rangle . \tag{27}$$

Setting $r = 1/2$ in equation 27 and replacing $\tilde{p}_B$ by $u/(sh)$ yields $N_e \approx N(1-8shu)$, which is equivalent to Robertson (1961)'s heuristic result that selection at unlinked loci decreases the effective population size by four times the additive variance in fitness — indeed, the variance in fitness caused by selection against the deleterious allele is approximately $2shu$ (see also Charlesworth, 2012).

Tight linkage. When r is small, Nordborg (1997)'s separation of timescales argument can be used to express associations between genes present on different haplotypes of a diploid in terms on associations between genes present on the same haplotype (see also Nordborg, 2000; Padhukasahasram et al., 2008). Consider for example the association $D_{A,B}$, between two genes sampled at loci A and B from different haplotypes of an individual. Going backwards in time, two different events may happen to the ancestral lineages of these genes: they may find themselves on the same haplotype (which may take only a few generations if both lineages stay in the same individual due to selfing), or move to different individuals due to an outcrossing event, in which case it will take a long time before find themselves again in the same individual (assuming N is large). To leading order, the probability that these lineages join on the same haplotype before moving to different individuals is F. Considering now all possible pairs of genes at loci A and B on different haplotypes of the same individual (in all individuals of the population) and going backwards in time, we may assume that a proportion F of such pairs find themselves on the same haplotype after a small number of generations, while the remaining $1 - F$ have moved to different individual lineages (and thus become independent): therefore, $D_{A,B} \approx F\, D_{AB}$ — note that this approximation assumes tight linkage, as it neglects recombination events separating genes that have joined on the same haplotype, over the small number of generations considered. Using this approximation, we have:

$$\langle D_{AB} D_{A,B} \rangle \approx F \left\langle {D_{AB}}^2 \right\rangle, \quad \left\langle {D_{A,B}}^2 \right\rangle \approx F^2 \left\langle {D_{AB}}^2 \right\rangle. \tag{28}$$

Similarly, $D_{AB,B} \approx F\, D_{ABB} = F\,(1 - 2p_B)\, D_{AB}$ (from equation A9), which yields

(assuming that p_B is small):

$$\langle D_{AB} D_{AB,B} \rangle \approx F \left\langle {D_{AB}}^2 \right\rangle, \quad \langle D_{A,B} D_{AB,B} \rangle \approx \left\langle {D_{AB,B}}^2 \right\rangle \approx F^2 \left\langle {D_{AB}}^2 \right\rangle. \tag{29}$$

Last, $D_{AAB,B} \approx F\, D_{AABB} = F\left(D_{AA} D_{BB} + (1 - 2p_B)\, D_{AAB}\right)$, from which one obtains (using the fact that the moment $\langle p_B D_{AAB} \rangle$ is negligible, as shown in Supplementary File 1):

$$\langle D_{AAB,B} \rangle - \langle D_{AA} D_{B,B} \rangle \approx \langle D_{AA,B} \rangle \approx F \langle D_{AAB} \rangle. \tag{30}$$

Equations 28 – 30 can also be obtained from equations B4 – B14, under the assumption that r is small (see Supplementary File 1). Plugging equations 28 – 30 into equation 12 yields:

$$\langle \Delta_{\mathrm{s}} D_{AA} \rangle \approx -s \left[h\,(1 - F) + F\right] \langle D_{AAB} \rangle \tag{31}$$

Furthermore, plugging equations 29 – 30 into equations B4 and B10 and assuming that r is small yields the same expressions as equations 24 and 25 (obtained for random mating) for $\langle {D_{AB}}^2 \rangle$ and $\langle D_{AAB} \rangle$, replacing h by $h\,(1 - F) + F$, r by $r\,(1 - F)$ and N by $N/\,(1 + F)$. Therefore, the present model converges to Nordborg (1997)'s result when loci are tightly linked.

Figure 1A shows the decrease in effective population size at the neutral locus caused by the deleterious allele (B_{sel}, see equation 17), as a function of the recombination rate between the two loci (on a log scale). When r tends to zero, B_{sel} tends to $1 - \tilde{p}_B$, and selfing increases N_{e} (as it decreases the frequency of the deleterious allele). When r is sufficiently high, however (roughly, higher than 0.01 for the parameter values used in Figure 1), increased selfing causes stronger background selection. Under complete selfing ($\alpha = 1$), B_{sel} becomes independent of r and is approximately $1 - u/s$. As can be seen on Figure 1A, the tight linkage approximation accurately predicts the

solution obtained from equations B4 – B14, except when r and α are high (in which case it underestimates the strength of background selection). As we will see, this discrepancy for high values of r may lead to substantial differences when integrating over a large genetic map, as the majority of deleterious alleles are only loosely linked to the neutral locus, yet significantly affect N_e at this locus when selfing is high (this is more visible on Figure 1C, showing the strength of background selection as a function of the position of the deleterious allele along the genetic map). Note however that the tight linkage approximation yields the same prediction as the more general model when $\alpha = 1$ ($B_{sel} \approx 1 - u/s$).

High selfing. Simple approximations can also be obtained when the selfing rate is high, for any value of the recombination rate r. From equations B4 – B9 and assuming that α is close to 1, one obtains:

$$\begin{aligned} \left\langle {D_{AB}}^2 \right\rangle &\approx \langle D_{AB} D_{A,B} \rangle \approx \left\langle {D_{A,B}}^2 \right\rangle \approx \langle D_{AB} D_{AB,B} \rangle \\ &\approx \langle D_{A,B} D_{AB,B} \rangle \approx \left\langle {D_{AB,B}}^2 \right\rangle \approx \frac{1+2r}{2N\left[s+2r\left(1-\alpha+s\right)\right]} \tilde{p}_B \left\langle D_{AA} \right\rangle . \end{aligned} \tag{32}$$

Furthermore, equations B10 – B14 yield:

$$\begin{aligned} \langle D_{AAB} \rangle \approx \langle D_{AA,B} \rangle &\approx \langle D_{AAB,B} \rangle - \langle D_{AA} D_{B,B} \rangle \\ &\approx \frac{\left(1+2r\right)^2 s}{N\left[s+2r\left(1-\alpha+s\right)\right]^2} \tilde{p}_B \left\langle D_{AA} \right\rangle , \end{aligned} \tag{33}$$

which, using $\tilde{p}_B \approx u/s$, generates the following approximation for the rate of decay of neutral diversity:

$$\left\langle \Delta D_{AA} \right\rangle \approx -\frac{1}{2N}\left(1+F\right)\left[1+\left(\frac{\left(1+2r\right)s}{s+2r\left(1-\alpha+s\right)}\right)^2 \tilde{p}_B\right]\left\langle D_{AA} \right\rangle . \tag{34}$$

Note that equation 34 again yields $B_{sel} \approx 1 - u/s$ when $\alpha = 1$. As shown by Figure 1D, this approximation better matches the general model than the tight linkage ap-

proximation when the selfing rate is high (but lower than 1) and when linkage is loose. However, Figure 1B shows that it differs substantially from the general model under tight linkage, and is thus expected to perform poorly when the selfing rate is not high (since in this case background selection is mainly caused by deleterious alleles that are tightly linked to the neutral locus).

Multilocus extrapolation and simulations. Following previous work (e.g., Hudson and Kaplan, 1995; Nordborg, 1997; Glémin, 2007; Kamran-Disfani and Agrawal, 2014), results from the two-locus model may be extrapolated to the case of deleterious alleles segregating at multiple loci by assuming multiplicative effects of the different deleterious alleles on neutral diversity. When the neutral locus is located at the midpoint of a linear genome with total map length R, and when all deleterious alleles have the same selection and dominance coefficients (as in the simulation program), this yields (using equation 2):

$$N_{\mathrm{e}} \approx \frac{N}{1+F} \exp\left[-\frac{U}{s\left[h\left(1-F\right)+F\right]} \times \frac{2}{R} \int_0^{\frac{R}{2}} T(x)\, dx \right] \tag{35}$$

where U is the deleterious mutation rate per haploid genome and the $T(x)$ corresponds to the term T in equation 16, in which the recombination rate r is expressed in terms of the genetic distance x (in Morgans) between the neutral locus and the deleterious allele using Haldane's mapping function $r = \frac{1}{2}\left[1 - e^{-2x}\right]$ (Haldane, 1919). This integral can be computed numerically as shown in Supplementary File 2.

Figure 2 shows the effective size of a population of census size 20000 as a function of the selfing rate, for a haploid genomic deleterious mutation rate $U = 0.1$ and genome map length $R = 10$ Morgans. As can be seen on the figure, the analytical

model (solid curves) fits well with the simulation results for all values of α. Note that in Figures 2 – 5, solid curves have been obtained by calculating numerically the integral in equation 35, using a system of recursions expressed to the second order in s for the different moments shown in Appendix B (see Supplementary Files 1 and 2); however, the results obtained using the recursions given in Appendix B (expressed to the first order in s) are undistinguishable in most cases (results not shown). Explicit forms can be obtained for the integral in equation 35 (as a function of R and the different parameters of the model) when using either the tight linkage approximation or the high selfing approximation, and are given in Supplementary File 2. The results shown on Figure 2 confirm that the tight linkage approximation underestimate the effect of background selection when the selfing rate is high, as loosely linked deleterious alleles significantly affect neutral diversity. In this case, the high selfing approximation is more accurate, closely matching the simulation results when $\alpha > 0.85$.

Figures 3 and 4 show that the full model provides accurate predictions for N_e for different values of map length R (from 0.1 to 100) and strength of selection against deleterious alleles s (from 0.005 to 0.5). As expected, the tight linkage approximation works better at lower values of R. When the selfing rate is low, increasing s magnifies the strength of background selection, while it has the opposite effect under high selfing (due to the fact that mutation-free individuals are more abundant when selection against deleterious alleles is stronger). Finally, Figure 5 shows that increasing the deleterious mutation rate U (to 0.5 per haploid genome) increases the strength of background selection, leading to very low values of N_e at high selfing rates (N_e estimated from the simulations is close to 45 when $\alpha = 1$). When $U = 0.5$, discrepancies between the full model and the simulations appear at low values of h ($h = 0.1$ in particular) and

intermediate values of α: these are probably caused by genetic associations between selected loci (identity disequilibrium in particular), which are neglected in the analysis. Using an expression for the frequency of deleterious allele $\tilde{p}_B$ that takes into account the effect of identity disequilibria (Roze, 2015) does not significantly improve the results (not shown). Discrepancies also appear at higher values of h and high selfing rate ($\alpha \approx 0.8 - 0.9$), the analytical model underestimating the strength of background selection. Finally, the model overestimates the effect of background selection when the selfing rate is very high ($\alpha > 0.95$), for all values of h: although this is not visible on Figure 5 (as N_e is very small when α is high), it becomes apparent when N_e is plotted on a log scale, as shown by Supplementary Figure 1. For these parameter values ($\alpha = 0.98, 1$), selection against deleterious alleles is no longer efficient and these accumulate over time in the population (results not shown).

Equation 35 can be extended to the more realistic case where deleterious alleles at different loci have different fitness effects (assuming that selection remains sufficiently strong at most loci so that deleterious alleles stay near mutation-selection balance), by replacing s and h by functions of map position x (e.g., Charlesworth, 2012). If the number of deleterious loci is large and if the distribution of fitness effects of mutations does not depend on genomic location, the integral in equation 35 can be replaced by a double integral, over map position x and over the joint distribution of s and h. In order to test this, the simulation program was modified to include a log-normal distribution of selection coefficients s across loci, assuming a constant heterozygous effect of mutations (sh), generating a negative covariance between s and h (see Figure 5 in Roze, 2015). Figure 6 shows results for $\alpha = 0.9$ and for different values of the variance of deleterious effects of mutations across loci, setting the av-

erage of the log-normal distribution and the heterozygous effect of mutations so that $\overline{s} = 0.05$ and $\overline{h} = 0.25$ in all cases. As can be seen on Figure 6, increasing the variance of s slightly increases N_{e}, and this effect is captured by integrating equation 34 (high selfing approximation) over the distribution of s.

Temporal change in population size. It is worth emphasizing that the moment-based method presented in this paper yields an expression for the "inbreeding effective size" (as it is based on the rate of decay of neutral diversity, or the instantaneous rate of coalescence), while coalescent-based methods (e.g., Hudson and Kaplan, 1995; Nordborg, 1997) yield an expression for the "coalescence effective size", corresponding to half the average coalescence time between two sequences randomly sampled from the population. After a single change in population size, and if selection and recombination rate are sufficiently large relative to $1/N_{\mathrm{e}}$, the different moments computed in Appendix B should equilibrate quickly relative to the change in neutral diversity, and the inbreeding effective population size should thus rapidly converge to its equilibrium value for the new census population size. By contrast, expected coalescence times will take longer to equilibrate. Assume that the inbreeding effective size instantaneously reaches its new equilibrium value, the expected coalescence time t generations after the change is given by:

$$\mathrm{E}\left[T\right]_t = 2N_{\mathrm{e}} + \left(1 - \frac{1}{2N_{\mathrm{e}}}\right)^t \left(\mathrm{E}\left[T\right]_0 - 2N_{\mathrm{e}}\right) \tag{36}$$

where N_{e} is the new inbreeding effective size and $\mathrm{E}\left[T\right]_0$ the average coalescence time at the time of the change in population size. In order to test this prediction, the simulation program was modified to include a neutral sequence with an infinite number of sites at

the mid-point of the chromosome. Indeed, the expected coalescence time between two sequences is simply given by $D/(2\mu)$, where μ is the mutation rate within the sequence and D the average number of differences between two sequences, given by $D = 2\sum_i p_i q_i$ (where the sum is over all segregating neutral sites within the sequence). Figure 7 shows an example in which $N = 20000$ during the first 10^5 generations (starting from a monomorphic population, which is equivalent to the coalescence of all lineages at generation zero), while $N = 40000$ during the last 10^5 generations. As can be seen on the figure, the quasi-equilibrium argument leading to equation 36 correctly predicts the dynamics of $\mathrm{E}[T]$ both during the initial phase and after the change in population size. The quasi-equilibrium argument could also possibly be used in situations where population size changes continuously over time (*i.e.*, exponential population growth), although this was not explored here.

DISCUSSION

We have seen how multilocus population genetics theory can be used to express the effect of a deleterious allele at mutation-selection balance on the dynamics of diversity at a linked neutral locus in terms of moments of linkage disequilibrium and other genetic associations between these two loci. This provides an alternative to methods based on computing expected coalescence times of pairs of genes present on different types of genetic backgrounds (e.g., Nordborg, 1997; Agrawal and Hartfield, 2016), and allows one to decompose the background selection effect into different terms for which intuitive interpretation may be given. In a panmictic population, background selection is driven by the variance in linkage disequilibrium $\langle {D_{AB}}^2 \rangle$ between the two loci: neutral

alleles that become associated to the deleterious allele tend to decrease in frequency, eventually causing their loss from the population. Under partial selfing, background selection is reinforced by random associations between a given neutral allele and selected alleles present on the other haplotype of the same individual ($D_{A,B}$), and associations between neutral alleles and homozygosity at the selected locus ($D_{AB,B}$). Furthermore, we have seen that identity disequilibrium (correlation in heterozygosity between loci) has the opposite effect and enhances neutral diversity when the deleterious allele is partially recessive, but this effect usually stays modest.

Under tight linkage, the results converge to the approximation derived by Nordborg (1997) using a separation of timescales argument. In that case, the population behaves approximately as a panmictic population of size $N/(1+F)$, in which the dominance coefficient of the deleterious allele and the recombination rate are replaced by the effective parameters $h(1-F)+F$ and $r(1-F)$. Similarly, Agrawal and Hartfield (2016) showed that in a population reproducing sexually at rate σ and asexually (by mitosis) at rate $1-\sigma$, the effect of a deleterious allele located at a small recombination distance r from the neutral locus takes the same form as in a panmictic (fully sexual) population, replacing r by the effective recombination rate $r\sigma$ (as shown in Appendix C, the effect of partial asexuality can also be derived using the methods of the present article). However, Agrawal and Hartfield (2016) also showed that this tight linkage approximation underestimates the strength of background selection caused by loosely linked loci, which becomes important when sex is rare. We have seen that a similar result holds under partial selfing: at high selfing rates, neutral diversity may be significantly affected by loosely linked deleterious alleles, whose effect is underestimated by the tight linkage approximation. In that case, a more precise approximation is

provided by equation 34 above.

As mentioned previously (Nordborg, 1997), assuming multiplicative effects of deleterious alleles at different loci on neutral diversity (equation 35) should be less accurate under partial selfing than under random mating, since inbreeding generates different forms of genetic associations between those loci (e.g., Kamran-Disfani and Agrawal, 2014; Roze, 2015). Nevertheless, we have seen that equation 35 yields accurate predictions as long as the genomic deleterious mutation rate stays moderate ($U = 0.1$, Figures 2 – 4), so that the average number of deleterious alleles per genome is not too large. Another important assumption of the model is that selection against deleterious alleles is sufficiently strong relative to drift, so that these alleles are maintained near their deterministic mutation-selection equilibrium frequency. In the regime where $N_e s \approx 1$ or lower, sometimes called "interference selection" or "weak Hill-Robertson interference" regime (e.g., McVean and Charlesworth, 2000; Comeron and Kreitman, 2002; Good et al., 2014), background selection models assuming $N_e s \gg 1$ may overestimate the effect of deleterious alleles on neutral diversity by several orders of magnitude (Kaiser and Charlesworth, 2008; Good et al., 2014). This may possibly be due to negative linkage disequilibria between deleterious alleles generated by the Hill-Robertson effect (Hill and Robertson, 1966), reducing the variance in fitness in the population and thus increasing coalescence times. As illustrated by Supplementary Figure 1, this regime may be particularly important in highly selfing populations and when the deleterious mutation rate U is high, since N_e may be sufficiently reduced to affect the efficiency of selection at an important proportion of selected sites. This is confirmed by empirical observations of higher ratios of non-synonymous to synonymous polymorphism (π_N/π_S) in selfing lineages than in their outcrossing relatives (Glémin

and Muyle, 2014; Hartfield, 2015 and references therein), indicating that a significant fraction of deleterious alleles may increase in frequency due to drift (in which case the assumption that $N_e s \gg 1$ at most loci is not valid). It is thus important to keep in mind that background selection models such as the one presented here overestimate the reduction in N_e in such situations, and it would be interesting (although probably challenging) to obtain analytical predictions for neutral diversity in populations undergoing low rates of sex or high selfing rates, and in which selection against a high proportion of deleterious mutations is rendered ineffective.

More generally, the strong reduction in effective population size of highly selfing (or asexual) populations caused by background selection may give an important influence to stochastic processes that would play a more marginal role in populations with larger N_e. For example, identity disequilibrium between selected loci in a partially selfing population generates positive linkage disequilibrium between these loci, but as shown by Kamran-Disfani and Agrawal (2014), stochastic forces generating negative linkage disequilibrium become stronger than this deterministic effect when the selfing rate is high. In a similar way, deterministic forces acting on the evolution of recombination rates (e.g., Roze and Lenormand, 2005) or mutation rates (e.g., Lynch, 2010) may be overwhelmed by stochastic forces under strong inbreeding. Developing analytical models that could scale these different types of effects may thus help us to better understand how mating systems affect the evolution of genetic architecture. The methods presented in the present paper could possibly be used to explore such questions.

Acknowledgements. I thank the bioinformatics and computing service of Roscoff's

651 Biological Station for computing time, and two anonymous reviewers for helpful com-
652 ments. This work was supported by the French Agence Nationale de la Recherche
653 (project TRANS, ANR-11-BSV7-013 and project Clonix, ANR-11-BSV7-007).

654 LITERATURE CITED

655 Agrawal, A. F. and M. Hartfield. 2016. Coalescence with background and balanc-
656 ing selection in systems with bi- and uniparental reproduction: contrasting partial
657 asexuality and selfing. Genetics 202:313–326.

658 Barton, N. H. and M. Turelli. 1991. Natural and sexual selection on many loci. Genetics
659 127:229–255.

660 Brandvain, Y., T. Slotte, K. M. Hazzouri, S. I. Wright, and G. Coop. 2013. Ge-
661 nomic identification of founding haplotypes reveals the history of the selfing species
662 *Capsella rubella*. PLoS Genetics 9:e1003754.

663 Burgarella, C., P. Gayral, M. Ballenghien, A. Bernard, P. David, P. Jarne, A. Correa,
664 S. Hurtrez-Boussès, J. Escobar, N. Galtier, and S. Glémin. 2015. Molecular evolution
665 of freshwater snails with contrasting mating systems. Mol. Biol. Evol. 32:2403–2416.

666 Charlesworth, B. 2012. The effects of deleterious mutations on evolution at linked
667 sites. Genetics 190:5–22.

668 Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1993. The effect of deleterious
669 mutations on neutral molecular variation. Genetics 134:1289–1303.

670 Charlesworth, D. 2003. Effects of inbreeding on the genetic diversity of populations.
671 Phil. Trans. Roy. Soc. (Lond.) B 358:1051–1070.

672 Comeron, J. M. and M. Kreitman. 2002. Population, evolutionary and genomic con-
673 sequences of interference selection. Genetics 161:389–410.

674 Glémin, S. 2003. How are deleterious mutations purged? Drift versus nonrandom
675 mating. Evolution 57:2678–2687.

676 ———. 2007. Mating systems and the efficacy of selection at the molecular level.
677 Genetics 177:905–916.

678 Glémin, S., E. Bazin, and D. Charlesworth. 2006. Impact of mating systems on patterns
679 of sequence polymorphism on flowering plants. Proc. Roy. Soc. (Lond.) B 273:3011–
680 3019.

681 Glémin, S. and A. Muyle. 2014. Mating systems and selection efficacy: a test using
682 chloroplastic sequence data in Angiosperms. J. Evol. Biol. 27:1386–1399.

683 Glémin, S. and J. Ronfort. 2012. Adaptation and maladaptation in selfing in outcross-
684 ing species: new mutations versus standing variation. Evolution 67:225–240.

685 Goldberg, E. E., J. R. Kohn, R. Lande, K. A. Robertson, S. A. Smith, and B. Igić.
686 2010. Species selection maintains self-incompatibility. Science 330:493–495.

687 Good, B. H., A. M. Walczak, R. A. Neher, and M. M. Desai. 2014. Genetic diversity
688 in the interference selection limit. PLoS Genetics 10:e1004222.

689 Goodwillie, C., S. Kalisz, and C. G. Eckert. 2005. The evolutionary enigma of mixed
690 mating systems in plants: occurence, theoretical explanations, and empirical evi-
691 dence. Ann. Rev. Ecol. Evol. Syst. 36:47–79.

692 Haldane, J. B. S. 1919. The combination of linkage values and the calculation of
693 distances between the loci of linked factors. J. Genet. 8:299–309.

694 Hartfield, M. 2015. Evolutionary genetic consequences of facultative sex and outcross-
695 ing. J. Evol. Biol. 29:5–22.

696 Hill, W. G. and A. Robertson. 1966. The effect of linkage on limits to artificial selection.
697 Genet. Res. 8:269–294.

698 Hudson, R. R. and N. L. Kaplan. 1995. Deleterious background selection with recom-
699 bination. Genetics 141:1605–1617.

700 Igic, B. and J. W. Busch. 2013. Is self-fertilization an evolutionary dead end? New
701 Phytol. 198:386–397.

702 Jarne, P. and J. R. Auld. 2006. Animals mix it up too: the distribution of self-
703 fertilization among hermaphroditic animals. Evolution 60:1816–1824.

704 Kaiser, V. B. and B. Charlesworth. 2008. The effects of deleterious mutations on
705 evolution in non-recombining genomes. Trends Genet. 25:9–12.

706 Kamran-Disfani, A. and A. F. Agrawal. 2014. Selfing, adaptation and background
707 selection in finite populations. J. Evol. Biol. 27:1360–1371.

708 Kirkpatrick, M., T. Johnson, and N. H. Barton. 2002. General models of multilocus
709 evolution. Genetics 161:1727–1750.

710 Lynch, M. 2010. Evolution of the mutation rate. Trends Genet. 26:345–352.

711 Lynch, M., J. Conery, and R. Bürger. 1995. Mutational meltdowns in sexual popula-
712 tions. Evolution 49:1067–1080.

713 McVean, G. A. and B. Charlesworth. 2000. The effects of Hill-Robertson interference

between weakly selected mutations on patterns of molecular evolution and variation. Genetics 155:929–944.

Nordborg, M. 1997. Structured coalescent processes on different time scales. Genetics 146:1501–1514.

———. 2000. Linkage disequilibrium, gene trees and selfing: and ancestral recombination graph with partial self-fertilization. Genetics 154:923–929.

Nordborg, M., J. O. Borevitz, J. Bergelson, C. C. Berry, J. Chory, J. Hagenblad, M. Kreitman, J. N. Maloof, T. Noyes, P. J. Oefner, E. A. Stahl, and D. Weigel. 2002. The extent of linkage disequilibrium in *Arabidopsis thaliana*. Nat. Genet. 30:190–193.

Nordborg, M., B. Charlesworth, and D. Charlesworth. 1996. The effect of recombination on background selection. Genet. Res. 67:159–174.

Padhukasahasram, B., P. Marjoram, J. D. Wall, C. D. Bustamante, and M. Nordborg. 2008. Exploring population genetics models with recombination using efficient forward-time simulations. Genetics 178:2417–2427.

Pollak, E. 1987. On the theory of partially inbreeding finite populations. I. Partial selfing. Genetics 117:353–360.

Robertson, A. 1961. Inbreeding in artificial selection programmes. Genet. Res. 2:189–194.

Roze, D. 2009. Diploidy, population structure and the evolution of recombination. Am. Nat. 174:S79–S94.

735 ———. 2015. Effects of interference between selected loci on the mutation load, in-
736 breeding depression and heterosis. Genetics 201:745–757.

737 Roze, D. and T. Lenormand. 2005. Self-fertilization and the evolution of recombination.
738 Genetics 170:841–857.

739 Roze, D. and R. E. Michod. 2010. Deleterious mutations and selection for sex in finite,
740 diploid populations. Genetics 184:1095–1112.

741 Roze, D. and F. Rousset. 2008. Multilocus models in the infinite island model of
742 population structure. Theor. Popul. Biol. 73:529–542.

743 Santiago, E. and A. Caballero. 1998. Effective size and polymorphism of linked neutral
744 loci in populations under selection. Genetics 149:2105–2117.

745 Schultz, S. T. and M. Lynch. 1997. Mutation and extinction: the role of variable mu-
746 tational effects, synergistic epistasis, beneficial mutations and degree of outcrossing.
747 Evolution 51:1363–1371.

748 Stebbins, G. L. 1957. Self fertilization and population variability in higher plants. Am.
749 Nat. 91:337–354.

750 Takebayashi, N. and P. L. Morrell. 2001. Is self-fertilization an evolutionary dead
751 end? Revisiting an old hypothesis with genetic theories and a macroevolutionary
752 approach. Am. J. Bot. 88:1143–1150.

753 Weir, B. S. and C. C. Cockerham. 1969. Group inbreeding with two linked loci.
754 Genetics 63:711–742.

755 Williams, G. C. 1992. Natural Selection: Domains, Levels, and Challenges. Oxford
756 University Press, New York, NY.

757 Wright, S. I., S. Kalisz, and T. Slotte. 2013. Evolutionary consequences of self-
758 fertilization in plants. Proc. Roy. Soc. (Lond.) B 280:20130133.

APPENDIX A: DERIVING RECURSIONS FOR MULTILOCUS MOMENTS

In the following, I use the notation $D'_{\mathbb{S},\mathbb{T}}$ for genetic associations measured at the next generation, while $D^{\mathrm{juv}}_{\mathbb{S},\mathbb{T}}$ and $D^{\mathrm{par}}_{\mathbb{S},\mathbb{T}}$ represent associations measured among juveniles (after reproduction, before drift) and among parents after selection (that is, weighting each parent by its relative fitness). Therefore, selection changes associations $D_{\mathbb{S},\mathbb{T}}$ to $D^{\mathrm{par}}_{\mathbb{S},\mathbb{T}}$, recombination and fertilization (with partial selfing) change associations $D^{\mathrm{par}}_{\mathbb{S},\mathbb{T}}$ to $D^{\mathrm{juv}}_{\mathbb{S},\mathbb{T}}$, while drift changes associations $D^{\mathrm{juv}}_{\mathbb{S},\mathbb{T}}$ to $D'_{\mathbb{S},\mathbb{T}}$. Similarly, $p_i{}'$ will denote allele frequencies at the next generation, while p_i^{juv} will represent allele frequencies among juveniles (which are the same as among selected parents, as recombination and fertilization do not change allele frequencies). In the following I show how to derive equations representing these different steps. For this, I focus on examples rather than presenting general (and necessarily cumbersome) equations; however, general expressions are implemented in a *Mathematica* notebook (Supplementary File 1). Throughout the paper, I assume that selection is weak (s small), and that population size is sufficiently large so that $1/N \ll s, r$; however, no assumption is done on the relative orders of magnitude of r and s. Finally, I assume that $u \ll s$ so that the frequency of the deleterious allele at mutation-selection balance ($\tilde{p}_B$, given by equation 2) is small.

Drift. In order to illustrate the effect of drift on genetic moments, we will consider the expected diversity at the neutral locus, $\langle D_{AA} \rangle$. By definition, $\langle D'_{AA} \rangle = \left\langle \frac{1}{2} \mathrm{E}' \left[\left(X_A^{\mathrm{M}} - p_A{}' \right)^2 + \left(X_A^{\mathrm{P}} - p_A{}' \right)^2 \right] \right\rangle$, where E' is the average over all individuals of the next generation, while $p_A{}'$ is the frequency of allele 1 at locus A among these individuals. Writing $\Delta_{\mathrm{d}} p_A = p_A{}' - p_A^{\mathrm{juv}}$ the change in allele frequency due to drift, we

have:

$$\langle D'_{AA} \rangle = \left\langle \frac{1}{2} \mathrm{E}' \left[\left(X_A^{\mathrm{M}} - p_A^{\mathrm{juv}} - \Delta_{\mathrm{d}} p_A \right)^2 + \left(X_A^{\mathrm{P}} - p_A^{\mathrm{juv}} - \Delta_{\mathrm{d}} p_A \right)^2 \right] \right\rangle . \tag{A1}$$

Expanding and grouping terms in $\Delta_{\mathrm{d}} p_A$, this is:

$$\begin{aligned} \langle D'_{AA} \rangle = & \left\langle \frac{1}{2} \mathrm{E}' \left[\left(X_A^{\mathrm{M}} - p_A^{\mathrm{juv}} \right)^2 + \left(X_A^{\mathrm{P}} - p_A^{\mathrm{juv}} \right)^2 \right] \right\rangle \\ & - 2 \left\langle \Delta_{\mathrm{d}} p_A \, \frac{1}{2} \mathrm{E}' \left[\left(X_A^{\mathrm{M}} - p_A^{\mathrm{juv}} \right) + \left(X_A^{\mathrm{P}} - p_A^{\mathrm{juv}} \right) \right] \right\rangle + \left\langle (\Delta_{\mathrm{d}} p_A)^2 \right\rangle . \end{aligned} \tag{A2}$$

In the following, I use the notation $D^{\mathrm{dft}}_{\mathbb{S},\mathbb{T}}$ for moments measured among individuals of the next generation (after drift), but using the values of allele frequencies (called "reference values" in Kirkpatrick et al., 2002) before drift (among juveniles): in particular, $D^{\mathrm{dft}}_{AA} = \frac{1}{2}\mathrm{E}' \left[\left(X_A^{\mathrm{M}} - p_A^{\mathrm{juv}} \right)^2 + \left(X_A^{\mathrm{P}} - p_A^{\mathrm{juv}} \right)^2 \right]$, while we have $D^{\mathrm{dft}}_{A} = \frac{1}{2}\mathrm{E}' \left[\left(X_A^{\mathrm{M}} - p_A^{\mathrm{juv}} \right) + \left(X_A^{\mathrm{P}} - p_A^{\mathrm{juv}} \right) \right]$. Noting that $D^{\mathrm{dft}}_{A} = \Delta_{\mathrm{d}} p_A$, equation A2 can be written as:

$$\langle D'_{AA} \rangle = \left\langle D^{\mathrm{dft}}_{AA} \right\rangle - \left\langle \left(D^{\mathrm{dft}}_{A} \right)^2 \right\rangle . \tag{A3}$$

Next, we can note that products of genetic associations may also be viewed as associations between genes present in two individuals, sampled with replacement from the whole population: for example, $\left(D^{\mathrm{dft}}_{A} \right)^2$ is the association between one gene at locus A from a first individual, and one gene at locus A from a second individual sampled with replacement from the population at the next generation (and using allele frequencies among juvenile as reference values). Following previous works (Roze and Rousset, 2008; Roze, 2009; Roze and Michod, 2010), I will thus write such products as single associations, using the symbol $\widehat{/}$ to separate sets of genes present in different individuals sampled with replacement from the population. Using this notation, $\left(D^{\mathrm{dft}}_{A} \right)^2$ is written as $D^{\mathrm{dft}}_{A\widehat{/}A}$, and equation A3 becomes:

$$\langle D'_{AA} \rangle = \left\langle D^{\mathrm{dft}}_{AA} \right\rangle - \left\langle D^{\mathrm{dft}}_{A\widehat{/}A} \right\rangle . \tag{A4}$$

Equation A4 corresponds to the first step of the computation of the effect of drift on genetic moments (*i.e.*, taking into account changes in allele frequencies due to drift). The second (and last) step consists in expressing associations between genes present in different individuals sampled with replacement from the population (at the next generation) in terms of associations between genes in individuals sampled *without* replacement. For example, we have:

$$\left\langle D^{\mathrm{dft}}_{\widehat{A/A}} \right\rangle = \frac{1}{2N}\left(\left\langle D^{\mathrm{dft}}_{AA}\right\rangle + \left\langle D^{\mathrm{dft}}_{A,A}\right\rangle\right) + \left(1-\frac{1}{N}\right)\left\langle D^{\mathrm{dft}}_{A/A}\right\rangle \tag{A5}$$

where $D^{\mathrm{dft}}_{A/A}$ is the association between two genes at locus A from two individuals sampled without replacement, at the next generation (and using allele frequencies before drift as reference values). Indeed, two genes sampled with replacement from the population may be the same gene with probability $1/(2N)$, the two homologous copies of the same individual with probability $1/(2N)$, or come from two different individuals with probability $1-1/N$. Finally, because the different individuals of the next generation have been sampled independently from an infinite population of juveniles, associations between genes present in different individuals (and using as reference values allele frequencies among juveniles) are the same as associations measured among juveniles: in particular, $\left\langle D^{\mathrm{dft}}_{AA}\right\rangle = \left\langle D^{\mathrm{juv}}_{AA}\right\rangle$, $\left\langle D^{\mathrm{dft}}_{A,A}\right\rangle = \left\langle D^{\mathrm{juv}}_{A,A}\right\rangle$ and $\left\langle D^{\mathrm{dft}}_{A/A}\right\rangle = \left\langle D^{\mathrm{juv}}_{A/A}\right\rangle$. Noting that $\left\langle D^{\mathrm{juv}}_{A/A}\right\rangle = 0$ (indeed, because the population of juveniles is infinite, we have $D^{\mathrm{juv}}_{A/A} = D^{\mathrm{juv}}_{\widehat{A/A}} = \left(D^{\mathrm{juv}}_{A}\right)^2$, while $D^{\mathrm{juv}}_{A} = 0$ from the definition of genetic associations), one finally obtains:

$$\left\langle D'_{AA}\right\rangle = \left\langle D^{\mathrm{juv}}_{AA}\right\rangle - \frac{1}{2N}\left(\left\langle D^{\mathrm{juv}}_{AA}\right\rangle + \left\langle D^{\mathrm{juv}}_{A,A}\right\rangle\right), \tag{A6}$$

representing the effect of drift on the moment $\langle D_{AA}\rangle$. Because all results will be computed to the first order in $1/N$, it will be sufficient to express the moments that

are multiplied by $1/(2N)$ in equation A6 in the limit as N tends to infinity.

The same method can be used to compute the effect of drift on other genetic moments. For example, using the same reasoning as for deriving equation A4 above, one obtains for the expected squared linkage disequilibrium:

$$\left\langle {D'_{AB}}^2 \right\rangle = \left\langle D^{\mathrm{dft}}_{AB\widehat{/}AB} \right\rangle - 2\left\langle D^{\mathrm{dft}}_{AB\widehat{/}A\widehat{/}B} \right\rangle + \left\langle D^{\mathrm{dft}}_{A\widehat{/}A\widehat{/}B\widehat{/}B} \right\rangle \tag{A7}$$

while computing these moments in terms of associations among juveniles yields:

$$\begin{aligned} \left\langle {D'_{AB}}^2 \right\rangle = \left\langle D^{\mathrm{juv}\,2}_{AB} \right\rangle + \frac{1}{2N}\Bigg(& \left\langle D^{\mathrm{juv}}_{AABB} \right\rangle + \left\langle D^{\mathrm{juv}}_{AB,AB} \right\rangle \\ & - 2\left\langle D^{\mathrm{juv}}_{AB} D^{\mathrm{juv}}_{A,B} \right\rangle - 4\left\langle D^{\mathrm{juv}\,2}_{AB} \right\rangle \Bigg) + o\left(\frac{1}{N}\right). \end{aligned} \tag{A8}$$

Again, because all results are computed to the first order in $1/N$, it will be sufficient to compute the four moments within parentheses in equation A8 in the limit as population size tends to infinity (we will see below that the moments $\left\langle D^{\mathrm{juv}}_{AB} D^{\mathrm{juv}}_{A,B} \right\rangle$ and $\left\langle D^{\mathrm{juv}\,2}_{AB} \right\rangle$ become negligible in this limit, while the moment $\left\langle D^{\mathrm{juv}}_{AB,AB} \right\rangle$ is generated by selfing).

Equation A8 shows that genetic associations with repeated "B" indices (such as D_{AABB}) may appear within recursions (see also equation A12). Because locus B is biallelic, these repeated indices can be eliminated using the relation (e.g., equation 5 in Kirkpatrick et al., 2002):

$$D_{\mathbb{S}ii} = p_i q_i D_{\mathbb{S}} + (1 - 2p_i)\, D_{\mathbb{S}i} \tag{A9}$$

where $\mathbb{S}$ is any set of loci, and i is a biallelic locus. However, repeated "A" indices will not be eliminated when computing recursions, so that the equations obtained still hold for the case where more than two alleles segregate at the neutral locus.

Recombination and fertilization. Computing moments measured among juveniles in terms of moments measured among selected parents is somewhat simpler, as recombination and fertilization do not change allele frequencies: therefore, one only has to consider the different possible modes of transmission of genes between generations. In particular, $\left\langle D^{\mathrm{juv}}_{AA}\right\rangle = \langle D^{\mathrm{par}}_{AA}\rangle$, while:

$$\left\langle D^{\mathrm{juv}}_{A,A}\right\rangle = \frac{\alpha}{2}\left(\langle D^{\mathrm{par}}_{AA}\rangle + \left\langle D^{\mathrm{par}}_{A,A}\right\rangle\right). \tag{A10}$$

Indeed, the two homologous genes of a juvenile at locus A come from the same parent if this juvenile has been produced by selfing (probability α), in which case they are copies of the same parental gene with probability 1/2, while they come from the two homologous genes of the parent with probability 1/2. With probability $1-\alpha$ the juvenile has been produced by outcrossing, in which case its two homologous genes are sampled with replacement from the parental population: the association thus becomes $D^{\mathrm{par}}_{\widehat{A/A}} = (D^{\mathrm{par}}_{A})^2$, which equals zero (since $D^{\mathrm{par}}_{A} = 0$). Similarly, we have:

$$\begin{aligned}\left\langle D^{\mathrm{juv}\,2}_{AB}\right\rangle = \left\langle D^{\mathrm{juv}}_{\widehat{AB/AB}}\right\rangle = (1-r)^2\left\langle D^{\mathrm{par}}_{\widehat{AB/AB}}\right\rangle + 2r\,(1-r)\left\langle D^{\mathrm{par}}_{\widehat{AB/A,B}}\right\rangle \\ + r^2\left\langle D^{\mathrm{par}}_{\widehat{A,B/A,B}}\right\rangle,\end{aligned} \tag{A11}$$

while

$$\begin{aligned}\left\langle D^{\mathrm{juv}}_{AB,AB}\right\rangle = \frac{\alpha}{2}\Big[&(1-r)^2\left(\langle D^{\mathrm{par}}_{AABB}\rangle + \left\langle D^{\mathrm{par}}_{AB,AB}\right\rangle\right) \\ &+ 2r\,(1-r)\left(\left\langle D^{\mathrm{par}}_{AAB,B}\right\rangle + \left\langle D^{\mathrm{par}}_{ABB,A}\right\rangle\right) \\ &+ r^2\left(\langle D^{\mathrm{par}}_{AABB}\rangle + \left\langle D^{\mathrm{par}}_{AB,AB}\right\rangle\right)\Big] + (1-\alpha)\left\langle D^{\mathrm{juv}}_{\widehat{AB/AB}}\right\rangle\end{aligned} \tag{A12}$$

where $\left\langle D^{\mathrm{juv}}_{\widehat{AB/AB}}\right\rangle$ is given by equation A11.

Selection. As for the effect of drift, computing the effect of selection on genetic moments can be decomposed into two steps (see also Barton and Turelli, 1991; Kirkpatrick et al., 2002). The first step consists in taking into account the change in reference values (*i.e.*, allele frequencies that appear within associations) due to selection: denoting $D^{\text{sel}}_{\mathbb{S},\mathbb{T}}$ moments measured among selected parents, but using allele frequencies before selection (p_i) as reference values (instead of allele frequencies after selection p_i^{juv}), we have (following the same reasoning as for the derivation of equations A4 and A7):

$$\langle D^{\text{par}}_{AA} \rangle = \langle D^{\text{sel}}_{AA} \rangle - \left\langle D^{\text{sel}}_{A\widehat{/}A} \right\rangle \tag{A13}$$

while

$$\left\langle D^{\text{par}}_{AB\widehat{/}AB} \right\rangle = \left\langle D^{\text{sel}}_{AB\widehat{/}AB} \right\rangle - 2\left\langle D^{\text{sel}}_{AB\widehat{/}A\widehat{/}B} \right\rangle + \left\langle D^{\text{sel}}_{A\widehat{/}A\widehat{/}B\widehat{/}B} \right\rangle. \tag{A14}$$

Finally, computing associations measured after selection (using as reference values allele frequencies before selection) $D^{\text{sel}}_{\mathbb{S},\mathbb{T}}$, in terms of associations measured before selection $D_{\mathbb{S},\mathbb{T}}$ is done by weighting each individual by its relative fitness:

$$D^{\text{sel}}_{\mathbb{S},\mathbb{T}} = \text{E}\left[\frac{W}{\overline{W}}\,\zeta_{\mathbb{S},\mathbb{T}}\right] \tag{A15}$$

where E is the average over all individuals before selection, W is the fitness of the individual while $\overline{W} = \text{E}[W]$ is the mean fitness of the population. In order to express the right hand side of equation A15 in terms of genetic associations, it is useful to write $W/\overline{W}$ in terms of $\zeta_{\mathbb{S},\mathbb{T}}$ variables (e.g., Barton and Turelli, 1991; Kirkpatrick et al., 2002). The fitness of an individual can be written as:

$$W = 1 - sh\left(X_B^{\text{M}} + X_B^{\text{P}}\right) - s\left(1-2h\right)X_B^{\text{M}}X_B^{\text{P}} \tag{A16}$$

which, after rearranging and averaging over all individuals yields:

$$\frac{W}{\overline{W}} = \frac{1 - C - sh\left(\zeta_B^{\text{M}} + \zeta_B^{\text{P}}\right) - s\left(1-2h\right)\left[\zeta_{B,B} + p_B\left(\zeta_B^{\text{M}} + \zeta_B^{\text{P}}\right)\right]}{1 - C - s\left(1-2h\right)D_{B,B}} \tag{A17}$$

with $C = 2sh\,p_B + s\,(1-2h)\,{p_B}^2$. Equation A17 thus takes the form of a polynomial of $\zeta_{\mathbb{S},\mathbb{T}}$ variables. However, because terms in p_B and $D_{B,B}$ appear in the denominator, obtaining expressions in terms of moments of genetic associations and allele frequencies requires developing $W/\overline{W}$ as a Taylor series in s. To the first order in s, we have:

$$\frac{W}{\overline{W}} = 1 - sh\left(\zeta_B^{\mathrm{M}} + \zeta_B^{\mathrm{P}}\right) - s\,(1-2h)\left[\zeta_{B,B} - D_{B,B} + p_B\left(\zeta_B^{\mathrm{M}} + \zeta_B^{\mathrm{P}}\right)\right]. \tag{A18}$$

From this, one obtains, to the first order in s:

$$\begin{aligned}\left\langle D_{AA}^{\mathrm{sel}}\right\rangle &= \left\langle \mathrm{E}\left[\frac{W}{\overline{W}}\,\zeta_{AA}\right]\right\rangle \\ &= \langle D_{AA}\rangle - sh\left(\langle D_{AAB}\rangle + \langle D_{AA,B}\rangle\right) \\ &\quad - s\,(1-2h)\left(\langle D_{AAB,B}\rangle - \langle D_{AA}D_{B,B}\rangle + \langle p_B D_{AAB}\rangle + \langle p_B D_{AA,B}\rangle\right).\end{aligned} \tag{A19}$$

Given that moments $\langle p_B D_{AAB}\rangle$ and $\langle p_B D_{AA,B}\rangle$ equal zero at quasi-equilibrium to the first order in $1/N$ and $\tilde{p}_B$ (see Supplementary File 1), equation A19 yields equation 12 in the main text. Associations between genes present in different individuals are obtained similarly. For example, $\left\langle D^{\mathrm{sel}}_{A\widehat{\ }A}\right\rangle$ equals zero to the first order in s (because $D_A = 0$), while equation A18 yields the following expression to the second order in s:

$$\begin{aligned}\left\langle D^{\mathrm{sel}}_{A\widehat{\ }A}\right\rangle &= \left\langle \mathrm{E}\left[\frac{W}{\overline{W}}\,\zeta_A\right]\mathrm{E}\left[\frac{W}{\overline{W}}\,\zeta_A\right]\right\rangle \\ &= (sh)^2\left(\left\langle D_{AB\widehat{\ }AB}\right\rangle + 2\left\langle D_{AB\widehat{\ }A,B}\right\rangle + \left\langle D_{A,B\widehat{\ }A,B}\right\rangle\right) \\ &\quad + 2s^2h\,(1-2h)\left(\left\langle D_{AB\widehat{\ }AB,B}\right\rangle + \left\langle D_{A,B\widehat{\ }AB,B}\right\rangle\right. \\ &\qquad \left. + \left\langle p_B D_{AB\widehat{\ }AB}\right\rangle + 2\left\langle p_B D_{AB\widehat{\ }A,B}\right\rangle + \left\langle p_B D_{A,B\widehat{\ }A,B}\right\rangle\right) \\ &\quad + s^2\,(1-2h)^2\left(\left\langle D_{AB,B\widehat{\ }AB,B}\right\rangle + 2\left\langle p_B D_{AB\widehat{\ }AB,B}\right\rangle + 2\left\langle p_B D_{A,B\widehat{\ }AB,B}\right\rangle\right. \\ &\qquad \left. + \left\langle {p_B}^2 D_{AB\widehat{\ }AB}\right\rangle + 2\left\langle {p_B}^2 D_{AB\widehat{\ }A,B}\right\rangle + \left\langle {p_B}^2 D_{A,B\widehat{\ }A,B}\right\rangle\right).\end{aligned} \tag{A20}$$

Equations A19 and A20 illustrate the fact that computing the effect of selection on a given genetic moment introduces more complicated moments, with a higher number

of "B" indices. This may lead to an infinite system of recursions, as the effect of selection on these moments will introduce yet other moments with even more "B" indices. However, assuming $Ns \gg 1$ and $s \gg u$ (so that the frequency of the deleterious allele at mutation-selection balance, $\tilde{p}_B$, remains small), we may neglect moments that are proportional to ${\tilde{p}_B}^2$ — more precisely, moments that are $o(\tilde{p}_B)$ — in order to obtain expressions to the first order in $\tilde{p}_B$. As we will see, using this "rare allele approximation" yields closed systems of recursions for genetic moments.

Rare allele approximation. Deriving expressions to the first order in $\tilde{p}_B$ can be done using the following general rule. Because all genetic associations involving at least one "B" index are proportional to p_B, moments involving two elements with a "B" index (where an element is either an allele frequency p_i or an association $D_{\mathbb{S},\mathbb{T}}$ among genes present in the same individual) are of order ${\tilde{p}_B}^2$ in the limit as population size tends to infinity. For example, $\langle p_B D_{AB} \rangle$, $\langle {D_{AB}}^2 \rangle$ and $\langle D_{AB} D_{AB,B} \rangle$ are all of order ${\tilde{p}_B}^2$ as N tends to infinity, while $\langle D_{B,B} \rangle$ and $\langle D_{AB,AB} \rangle$ are of order $\tilde{p}_B$ in the same limit. Now, taking the effect of drift into account when computing recursions for genetic moments generally introduces moments with one less element carrying a "B" index, multiplied by $1/N$. This is illustrated by equation A8 showing the effect of drift on the moment $\langle {D_{AB}}^2 \rangle$ (two elements with a "B" index): drift introduces two terms involving moments carrying a single element with a "B" index ($\langle D_{AABB} \rangle$ and $\langle D_{AB,AB} \rangle$). Because these moments are multiplied by $1/N$ in equation A8, it is sufficient to express them in the limit as N tends to infinity: in this limit, these terms are thus proportional to $\tilde{p}_B$ (by contrast, the terms in $\langle {D_{AB}}^2 \rangle / N$ and $\langle D_{AB} D_{A,B} \rangle / N$ in equation A8 are proportional to ${\tilde{p}_B}^2$, and may thus be neglected). More generally,

moments carrying x elements with a "B" index are thus of order ${\tilde{p}_B}^x$ in the limit as N tends to infinity, and of order ${\tilde{p}_B}^{x-1}$ to the first order in $1/N$ (for $x > 1$). In order to compute expressions to the first order in $\tilde{p}_B$, we may thus neglect all moments that must be expressed to the first order in $1/N$ and that carry more than two elements with a "B" index (such as $\langle p_B D_{AB} D_{AB,B} \rangle$), and all moments expressed in the limit as N tends to infinity carrying more than one element with a "B" index (such as $\langle {D_{AB}}^2 \rangle$). Using this approximation, equation A20 simplifies to:

$$\begin{aligned} \left\langle D^{\text{sel}}_{\widehat{A/A}} \right\rangle = (sh)^2 \left(\left\langle {D_{AB}}^2 \right\rangle + 2 \left\langle D_{AB} D_{A,B} \right\rangle + \left\langle {D_{A,B}}^2 \right\rangle \right) \\ + 2s^2 h \left(1 - 2h\right) \left(\left\langle D_{AB} D_{AB,B} \right\rangle + \left\langle D_{A,B} D_{AB,B} \right\rangle \right) \\ + s^2 \left(1 - 2h\right)^2 \left\langle {D_{AB,B}}^2 \right\rangle , \end{aligned} \tag{A21}$$

Furthermore, one obtains for the moments $\left\langle D^{\text{par}}_{\widehat{AB/AB}} \right\rangle$, $\left\langle D^{\text{par}}_{\widehat{AB/A,B}} \right\rangle$ and $\left\langle D^{\text{par}}_{\widehat{A,B/A,B}} \right\rangle$ (that are needed to compute a recursion for $\langle {D_{AB}}^2 \rangle$, as shown by equation A11):

$$\left\langle D^{\text{par}}_{\widehat{AB/AB}} \right\rangle = \left(1 - 2sh\right) \left\langle {D_{AB}}^2 \right\rangle - 2s \left(1 - h\right) \left\langle D_{AB} D_{AB,B} \right\rangle \tag{A22}$$

$$\left\langle D^{\text{par}}_{\widehat{AB/A,B}} \right\rangle = \left(1 - 2sh\right) \left\langle D_{AB} D_{A,B} \right\rangle - s \left(1 - h\right) \left(\left\langle D_{AB} D_{AB,B} \right\rangle + \left\langle D_{A,B} D_{AB,B} \right\rangle \right) \tag{A23}$$

$$\left\langle D^{\text{par}}_{\widehat{A,B/A,B}} \right\rangle = \left(1 - 2sh\right) \left\langle {D_{A,B}}^2 \right\rangle - 2s \left(1 - h\right) \left\langle D_{A,B} D_{AB,B} \right\rangle . \tag{A24}$$

APPENDIX B: RECURSIONS FOR TWO-LOCUS MOMENTS

Using the method shown in Appendix A, one obtains for the change in neutral diversity during selection, to the second order in s:

$$\begin{aligned}\langle \Delta_{\mathrm{s}} D_{AA} \rangle = & -sh\left(\langle D_{AAB}\rangle + \langle D_{AA,B}\rangle\right) \\ & - s\left(1-2h\right)\left(\langle D_{AAB,B}\rangle - \langle D_{AA}D_{B,B}\rangle\right) \\ & - \left(sh\right)^2\left(\left\langle {D_{AB}}^2\right\rangle + 2\left\langle D_{AB}D_{A,B}\right\rangle + \left\langle {D_{A,B}}^2\right\rangle\right) \\ & - 2s^2h\left(1-2h\right)\left(\langle D_{AB}D_{AB,B}\rangle + \langle D_{A,B}D_{AB,B}\rangle\right) \\ & - s^2\left(1-2h\right)^2\left\langle {D_{AB,B}}^2\right\rangle\end{aligned} \tag{B1}$$

while the change in neutral diversity during drift is given by equation 9 in the main text. The same method can be used to compute recursions for the different moments that appear in equation B1, from which solutions at quasi-equilibrium can be obtained. For example, equations A8, A11 and A22 – A24 yield the following recursion for $\langle {D_{AB}}^2\rangle$:

$$\begin{aligned}\left\langle {D'_{AB}}^2\right\rangle = & \left(1-2sh\right)\left[\left(1-r\right)^2\left\langle {D_{AB}}^2\right\rangle + 2r\left(1-r\right)\left\langle D_{AB}D_{A,B}\right\rangle + r^2\left\langle {D_{A,B}}^2\right\rangle\right] \\ & - 2s\left(1-h\right)\left[\left(1-r\right)\left\langle D_{AB}D_{AB,B}\right\rangle + r\ \left\langle D_{A,B}D_{AB,B}\right\rangle\right] \\ & + \frac{1}{2N}\left(\langle D_{AABB}\rangle + \langle D_{AB,AB}\rangle\right).\end{aligned} \tag{B2}$$

In order to obtain an expression to the first order in $1/N$, it is sufficient to express $\langle D_{AABB}\rangle$ and $\langle D_{AB,AB}\rangle$ in equation B2 in the limit when N tends to infinity. Using equation A9, we have $D_{AABB} = p_B q_B D_{AA} + \left(1-2p_B\right)D_{AAB}$. Furthermore, D_{AAB} equals zero at quasi-equilibrium in an infinite population (indeed, the solution obtained for $\langle D_{AAB}\rangle$ from the equations below is in $1/N$): therefore, $\langle D_{AABB}\rangle \approx \tilde{p}_B\langle D_{AA}\rangle$ when population size tends to infinity. From equation A12, an recursion for $\langle D_{AB,AB}\rangle$ when

928 N tends to infinity and $s = 0$ is given by:

$$\begin{aligned}\left\langle D'_{AB,AB}\right\rangle = \frac{\alpha}{2}\Big[&\left[1-2r\left(1-r\right)\right]\left(\left\langle D_{AABB}\right\rangle + \left\langle D_{AB,AB}\right\rangle\right)\\ &+ 2r\left(1-r\right)\left(\left\langle D_{AAB,B}\right\rangle + \left\langle D_{ABB,A}\right\rangle\right)\Big].\end{aligned} \tag{B3}$$

929 Using $\left\langle D_{AABB}\right\rangle \approx \tilde{p}_B \left\langle D_{AA}\right\rangle$ and $\left\langle D_{AAB,B}\right\rangle = \left\langle D_{ABB,A}\right\rangle \approx F\,\tilde{p}_B \left\langle D_{AA}\right\rangle$ yields $\left\langle D_{AB,AB}\right\rangle =$
930 $\phi_{AB}\,\tilde{p}_B \left\langle D_{AA}\right\rangle$, where ϕ_{AB} is given by equation 14, and corresponds to the probability
931 of joint coalescence of two pairs of genes sampled at loci A and B due to selfing, in an
932 infinite population. It is possible de compute ϕ_{AB} to the first order in s, but the term
933 in s is always negligible when selection is weak, and is thus ignored here. Using the
934 expressions just derived for $\left\langle D_{AABB}\right\rangle$ and $\left\langle D_{AB,AB}\right\rangle$, equation B2 becomes:

$$\begin{aligned}\left\langle {D'_{AB}}^2\right\rangle = &\left(1-2sh\right)\left[\left(1-r\right)^2\left\langle {D_{AB}}^2\right\rangle + 2r\left(1-r\right)\left\langle D_{AB}D_{A,B}\right\rangle + r^2\left\langle {D_{A,B}}^2\right\rangle\right]\\ &- 2s\left(1-h\right)\left[\left(1-r\right)\left\langle D_{AB}D_{AB,B}\right\rangle + r\,\left\langle D_{A,B}D_{AB,B}\right\rangle\right]\\ &+ \frac{1}{2N}\left(1+\phi_{AB}\right)\tilde{p}_B\left\langle D_{AA}\right\rangle\end{aligned} \tag{B4}$$

935 Similarly, one obtains the following recursions for moments $\left\langle D_{AB}D_{A,B}\right\rangle$, $\left\langle {D_{A,B}}^2\right\rangle$,
936 $\left\langle D_{AB}D_{AB,B}\right\rangle$, $\left\langle D_{A,B}D_{AB,B}\right\rangle$ and $\left\langle {D_{AB,B}}^2\right\rangle$ (which are also generated by finite pop-
937 ulation size), to the first order in s, $\tilde{p}_B$ and $1/N$ (recursions to the second order in
938 s are derived in Supplementary File 1, but yield very similar quantitative results in
939 most cases):

$$\begin{aligned}\left\langle D'_{AB}D'_{A,B}\right\rangle = \frac{\alpha}{2}\Big[&\left(1-2sh\right)\left[\left(1-r\right)\left\langle {D_{AB}}^2\right\rangle + \left\langle D_{AB}D_{A,B}\right\rangle + r\,\left\langle {D_{A,B}}^2\right\rangle\right]\\ &- s\left(1-h\right)\Big[\left(2\left(1-r\right)+1\right)\left\langle D_{AB}D_{AB,B}\right\rangle\\ &\qquad + \left(2r+1\right)\left\langle D_{A,B}D_{AB,B}\right\rangle\Big]\Big] + \frac{1}{N}F\,\tilde{p}_B\left\langle D_{AA}\right\rangle\end{aligned} \tag{B5}$$

940

$$\begin{aligned}\left\langle {D'_{A,B}}^2\right\rangle = \left(\frac{\alpha}{2}\right)^2\Big[&\left(1-2sh\right)\left(\left\langle {D_{AB}}^2\right\rangle + 2\left\langle D_{AB}D_{A,B}\right\rangle + \left\langle {D_{A,B}}^2\right\rangle\right)\\ &- 4s\left(1-h\right)\left(\left\langle D_{AB}D_{AB,B}\right\rangle + \left\langle D_{A,B}D_{AB,B}\right\rangle\right)\Big]\\ &+ \frac{1}{2N}\left(1+\phi_{AB}\right)\tilde{p}_B\left\langle D_{AA}\right\rangle\end{aligned} \tag{B6}$$

941

$$\left\langle D'_{AB} D'_{AB,B} \right\rangle = \frac{\alpha}{2}\Big[(1-2sh)\left[(1-r)^2\left\langle {D_{AB}}^2\right\rangle + 2r\,(1-r)\left\langle D_{AB}D_{A,B}\right\rangle + r^2\left\langle {D_{A,B}}^2\right\rangle\right] + \left[1-s\,(3-h)\right]\left[(1-r)\left\langle D_{AB}D_{AB,B}\right\rangle + r\,\left\langle D_{A,B}D_{AB,B}\right\rangle\right] - s\,(1-h)\left\langle {D_{AB,B}}^2\right\rangle\Big] + \frac{1}{2N}\left(F+\phi_{AB}\right)\tilde{p}_B\left\langle D_{AA}\right\rangle \tag{B7}$$

942

$$\left\langle D'_{A,B} D'_{AB,B} \right\rangle = \left(\frac{\alpha}{2}\right)^2\Big[(1-2sh)\left[(1-r)\left\langle {D_{AB}}^2\right\rangle + \left\langle D_{AB}D_{A,B}\right\rangle + r\,\left\langle {D_{A,B}}^2\right\rangle\right] + (1-2s)\left(\left\langle D_{AB}D_{AB,B}\right\rangle + \left\langle D_{A,B}D_{AB,B}\right\rangle\right) - 2s\,(1-h)\left[(1-r)\left\langle D_{AB}D_{AB,B}\right\rangle + r\,\left\langle D_{A,B}D_{AB,B}\right\rangle\right] - 2s\,(1-h)\left\langle {D_{AB,B}}^2\right\rangle\Big] + \frac{1}{2N}\left(F+\phi_{AB}\right)\tilde{p}_B\left\langle D_{AA}\right\rangle \tag{B8}$$

943

$$\left\langle {D'_{AB,B}}^2 \right\rangle = \left(\frac{\alpha}{2}\right)^2\Big[(1-2sh)\left[(1-r)^2\left\langle {D_{AB}}^2\right\rangle + 2r\,(1-r)\left\langle D_{AB}D_{A,B}\right\rangle + r^2\left\langle {D_{A,B}}^2\right\rangle\right] + 2\,(1-2s)\left[(1-r)\left\langle D_{AB}D_{AB,B}\right\rangle + r\,\left\langle D_{A,B}D_{AB,B}\right\rangle\right] + \left[1-2s\,(2-h)\right]\left\langle {D_{AB,B}}^2\right\rangle\Big] + \frac{1}{2N}\left(F+\phi_{AB}\right)\tilde{p}_B\left\langle D_{AA}\right\rangle \tag{B9}$$

944 Equations B4 to B9 can be solved at quasi-equilibrium (setting $\langle \mathcal{M}' \rangle = \langle \mathcal{M} \rangle$ for each
945 moment) to obtain expressions for the 6 moments of the form $Z\,\tilde{p}_B \langle D_{AA} \rangle / N$, where
946 Z is a function (that differs for each moment) of s, h, r and α. Although these
947 expressions are complicated, they are easily computed numerically using *Mathematica*
948 (see Supplementary File 2).

949 Recursions for the moments $\langle D_{AAB} \rangle$ and $\langle D_{AA,B} \rangle$ that appear on the first line

of equation 12 are given by (to the first order in s, $\tilde{p}_B$ and $1/N$):

$$\begin{aligned}\langle D'_{AAB}\rangle = &(1-sh)\left[(1-r)\langle D_{AAB}\rangle + r\langle D_{AA,B}\rangle\right]\\ &+2sh\left[(1-r)\left\langle {D_{AB}}^2\right\rangle + \langle D_{AB}D_{A,B}\rangle + r\left\langle {D_{A,B}}^2\right\rangle\right]\\ &+2s(1-2h)\left[(1-r)\langle D_{AB}D_{AB,B}\rangle + r\ \langle D_{A,B}D_{AB,B}\rangle\right]\\ &-s(1-h)\left(\langle D_{AAB,B}\rangle - \langle D_{AA}D_{B,B}\rangle\right) - \frac{1}{N}\ \langle D_{AB,A}\rangle\end{aligned} \tag{B10}$$

$$\begin{aligned}\left\langle D'_{AA,B}\right\rangle = \frac{\alpha}{2}\Big[&(1-sh)\left(\langle D_{AAB}\rangle + \langle D_{AA,B}\rangle\right)\\ &+2sh\left(\left\langle {D_{AB}}^2\right\rangle + 2\langle D_{AB}D_{A,B}\rangle + \left\langle {D_{A,B}}^2\right\rangle\right)\\ &+2s(1-2h)\left(\langle D_{AB}D_{AB,B}\rangle + \langle D_{A,B}D_{AB,B}\rangle\right)\\ &-2s(1-h)\left(\langle D_{AAB,B}\rangle - \langle D_{AA}D_{B,B}\rangle\right)\Big] - \frac{1}{N}\ \langle D_{AB,A}\rangle\,.\end{aligned} \tag{B11}$$

It is sufficient to express the moment $\langle D_{AB,A}\rangle$ that appears in equations B10 and B11 in the limit as N tends to infinity. To the first order in s, we have:

$$\left\langle D'_{AB,A}\right\rangle = \frac{\alpha}{2}\left[\langle D_{AB,A}\rangle - s(1-h)\left(\langle D_{AB,AB}\rangle - \langle D_{A,A}D_{B,B}\rangle\right)\right] \tag{B12}$$

giving at quasi-equilibrium:

$$\langle D_{AB,A}\rangle = -s(1-h)\,F\,G_{AB}\,\tilde{p}_B\,\langle D_{AA}\rangle \tag{B13}$$

(see also Roze, 2015), where $G_{AB} = \phi_{AB} - F^2$ is the identity disequilibrium between loci A and B. Finally, a recursion for $\langle D_{AAB,B}\rangle - \langle D_{AA}D_{B,B}\rangle$ (second line of equation

12) to the first order in s, $\tilde{p}_B$ and $1/N$ is given by:

$$\begin{aligned}\left\langle D'_{AAB,B}\right\rangle - \left\langle D'_{AA}D'_{B,B}\right\rangle = \frac{\alpha}{2}\Big[&[1-s(2-h)]\left(\langle D_{AAB,B}\rangle - \langle D_{AA}D_{B,B}\rangle\right)\\ &+(1-sh)\left[(1-r)\langle D_{AAB}\rangle + r\langle D_{AA,B}\rangle\right]\\ &+2sh\left[(1-r)\left\langle {D_{AB}}^2\right\rangle + \langle D_{AB}D_{A,B}\rangle + r\left\langle {D_{A,B}}^2\right\rangle\right]\\ &+2s(1-2h)\left[(1-r)\langle D_{AB}D_{AB,B}\rangle + r\langle D_{A,B}D_{AB,B}\rangle\right]\\ &+2sh\left(\langle D_{AB}D_{AB,B}\rangle + \langle D_{A,B}D_{AB,B}\rangle\right)\\ &+2s(1-2h)\left\langle D_{AB,B}{}^2\right\rangle\Big] - \frac{1}{2N}G_{AB}\,\tilde{p}_B\,\langle D_{AA}\rangle\,.\end{aligned} \tag{B14}$$

Equations B10, B11 and B14 can be solved to obtain expressions for $\langle D_{AAB}\rangle$, $\langle D_{AA,B}\rangle$ and $\langle D_{AAB,B}\rangle - \langle D_{AA}D_{B,B}\rangle$ at quasi-equilibrium (using the expressions for $\langle {D_{AB}}^2\rangle$, $\langle D_{AB}D_{A,B}\rangle$, $\langle {D_{A,B}}^2\rangle$, $\langle D_{AB}D_{AB,B}\rangle$, $\langle D_{A,B}D_{AB,B}\rangle$ and $\langle D_{AB,B}{}^2\rangle$ obtained from equations B4 to B9); again, *Mathematica* commands to obtain numerical solutions can be found in Supplementary File 2. Approximations for high selfing are obtained by assuming that s and $o = 1-\alpha$ are of order ϵ and expressing equations B4 – B14 to the first order in ϵ. Similarly, tight linkage approximations are obtained by assuming that s and r are of order ϵ (see Supplementary File 1).

APPENDIX C: PARTIAL ASEXUALITY

Background selection under partial asexuality in diploids has been explored recently by Agrawal and Hartfield (2016) using coalescence models. The method presented in Appendix A can also be used for the case of a population in which a proportion σ of offspring are produced by sexual reproduction each generation (with random gamete fusion) and a proportion $1-\sigma$ by asexual reproduction (mitosis), as shown in Roze and Michod (2010). In particular, one obtains for the change in neutral diversity over one generation (to the second order in s, and assuming that h is significantly different from zero):

$$\begin{aligned}\langle \Delta D_{AA} \rangle = &-\frac{\langle D_{AA} \rangle}{2N} - sh\left(\langle D_{AAB} \rangle + \langle D_{AA,B} \rangle\right) \\ &- (sh)^2 \left(\left\langle {D_{AB}}^2 \right\rangle + 2\left\langle D_{AB} D_{A,B} \right\rangle + \left\langle {D_{A,B}}^2 \right\rangle\right).\end{aligned} \tag{C1}$$

Recursions for the different moments that appear in equation C1 are given by (to the first order in $1/N$ and $\tilde{p}_B \approx u/(sh)$):

$$\begin{aligned}\left\langle {D'_{AB}}^2 \right\rangle = (1-sh)^2 \Big[&(1-r\sigma)^2 \left\langle {D_{AB}}^2 \right\rangle + 2r\sigma(1-r\sigma)\left\langle D_{AB} D_{A,B} \right\rangle \\ &+ (r\sigma)^2 \left\langle {D_{A,B}}^2 \right\rangle\Big] + \frac{1}{2N}\tilde{p}_B \langle D_{AA} \rangle\end{aligned} \tag{C2}$$

$$\left\langle D'_{AB} D'_{A,B} \right\rangle = (1-sh)^2 (1-\sigma)\left[(1-r\sigma)\left\langle D_{AB} D_{A,B} \right\rangle + r\sigma \left\langle {D_{A,B}}^2 \right\rangle\right] \tag{C3}$$

$$\left\langle {D'_{A,B}}^2 \right\rangle = (1-sh)^2 (1-\sigma)^2 \left\langle {D_{A,B}}^2 \right\rangle + \frac{1}{2N}\tilde{p}_B \langle D_{AA} \rangle \tag{C4}$$

$$\begin{aligned}\langle D'_{AAB} \rangle = &(1-sh)\left[(1-r\sigma)\langle D_{AAB} \rangle + r\sigma \langle D_{AA,B} \rangle\right] \\ &+ 2sh\left[(1-r\sigma)\left\langle {D_{AB}}^2 \right\rangle + 2\left\langle D_{AB} D_{A,B} \right\rangle + r\sigma \left\langle {D_{A,B}}^2 \right\rangle\right]\end{aligned} \tag{C5}$$

$$\left\langle D'_{AA,B} \right\rangle = (1-\sigma)\left[(1-sh)\langle D_{AA,B} \rangle + 2sh\left(\left\langle D_{AB} D_{A,B} \right\rangle + \left\langle {D_{A,B}}^2 \right\rangle\right)\right]. \tag{C6}$$

Expressions for $\langle {D_{AB}}^2 \rangle$, $\langle D_{AB} D_{A,B} \rangle$, $\langle {D_{A,B}}^2 \rangle$, $\langle D_{AAB} \rangle$ and $\langle D_{AA,B} \rangle$ at quasi-equilibrium can be obtained from equations C2 – C6, and plugged into equation C1. When the rate of sex σ is small, this yields equations 5 and 6 in Agrawal and Hartfield, 2016.

Table 1: Parameters and variables.

N	Population size
α	Selfing rate
s	Strength of selection against the deleterious allele
h	Dominance coefficient of the deleterious allele
u	Mutation rate towards the deleterious allele
r	Recombination rate between the neutral locus and selected locus
U	Genomic deleterious mutation rate (per haploid genome)
R	Genome map length
p_A	Frequency of allele 1 at locus A (neutral locus)
p_B	Frequency of allele 1 (deleterious allele) at locus B
$\tilde{p}_B$	Deleterious allele frequency at mutation-selection balance
$D_{\mathbb{S},\mathbb{T}}$	Genetic association between the sets $\mathbb{S}$ and $\mathbb{T}$ of loci present on different haplotypes of an individual (see equation 4)
D_{AA}	Genetic diversity at the neutral locus
F	Inbreeding coefficient (probability of identity-by-descent between the maternal and paternal copies of a gene, due to selfing)
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
$\mathrm{E}[X]$	Average of the quantity X over all individuals
$\langle \mathcal{M} \rangle$	Expected value of the moment $\mathcal{M}$ over the stochastic process

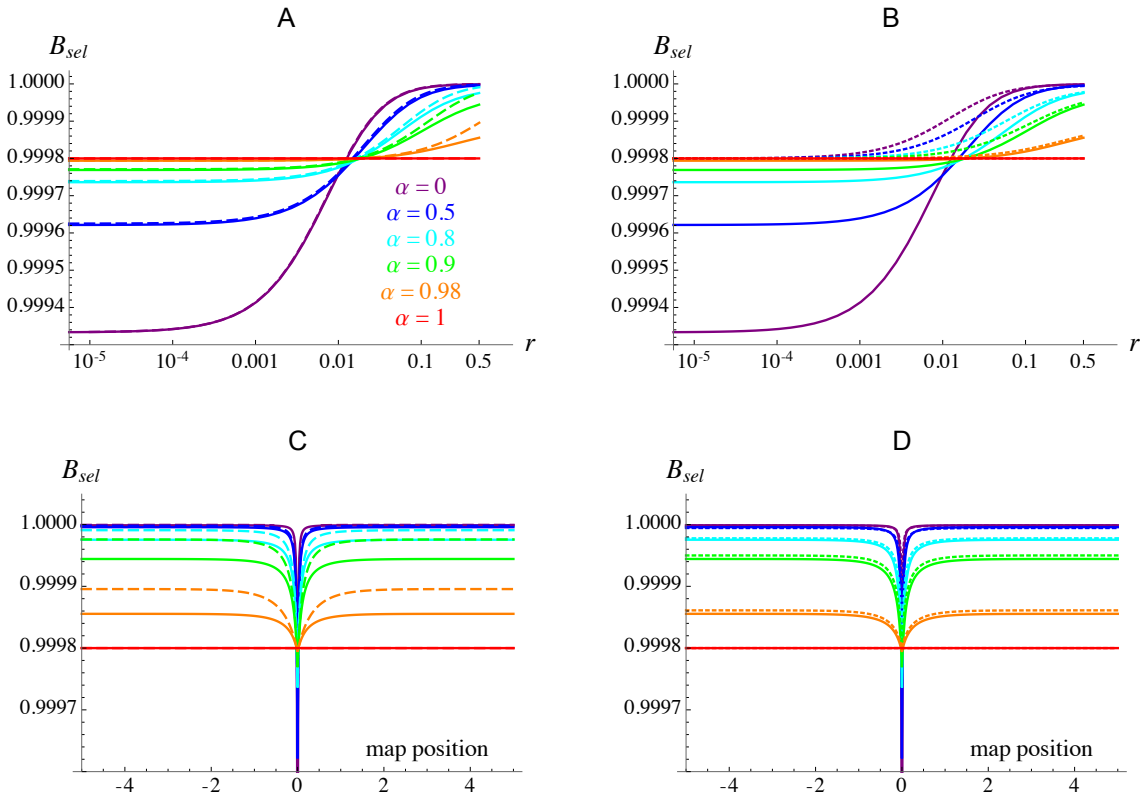
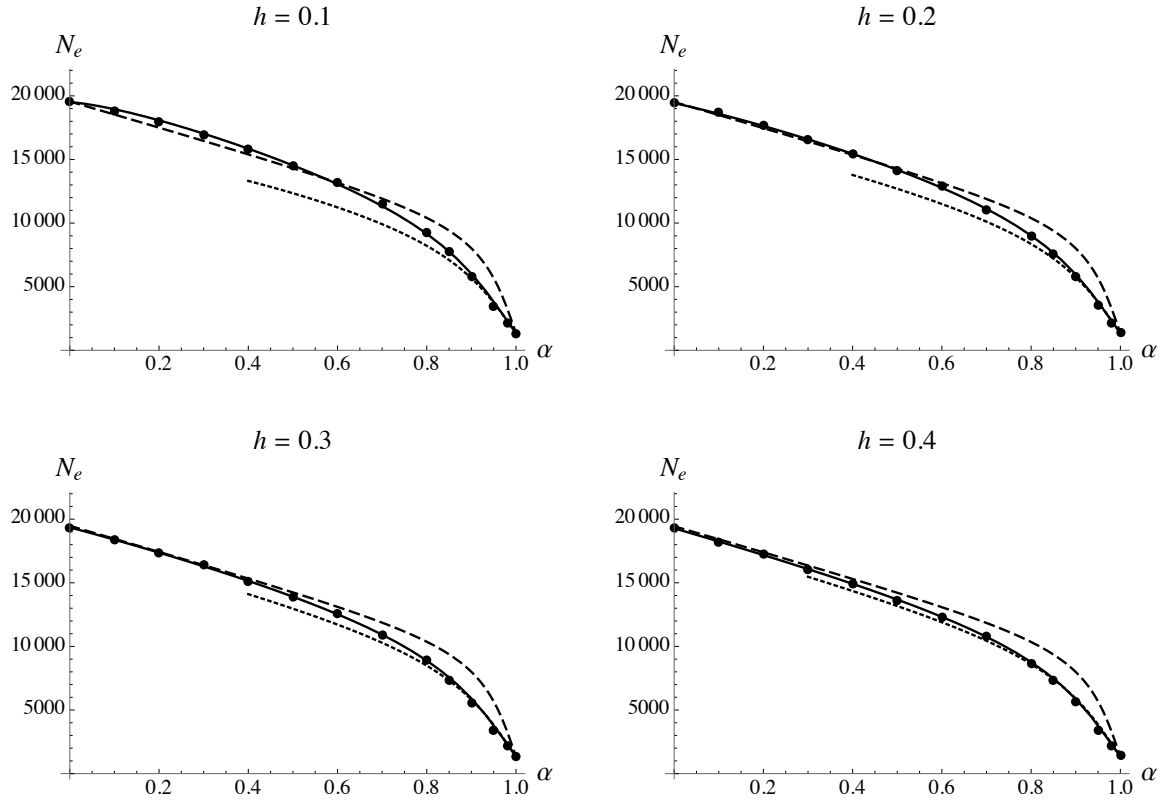
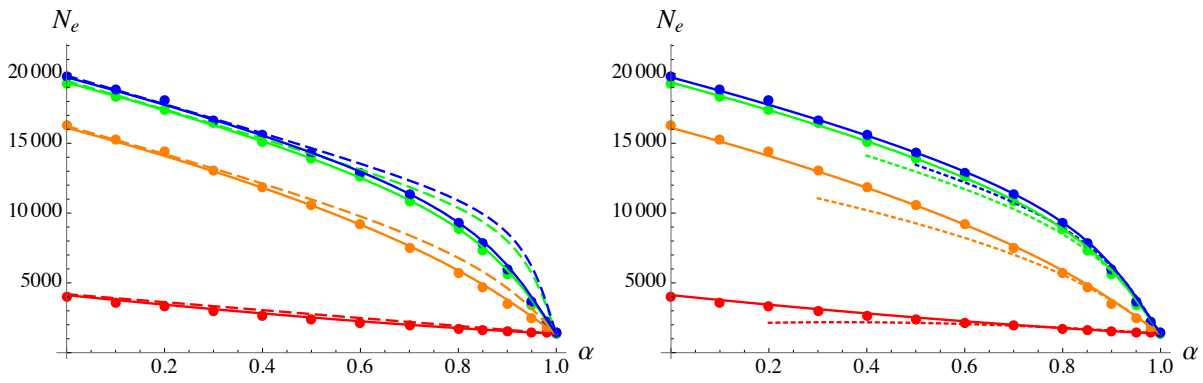


Figure 1. Background selection effect generated by a single deleterious allele (B_{sel}, see equation 17) on N_e at the neutral locus, as a function of the recombination rate r between the two loci (A, B), and of the position of the deleterious allele along a linear chromosome of total map length 10 Morgans (C, D, the neutral locus is located at position 0). Different colors correspond to different values of the selfing rate as indicated in A. Solid curves correspond to the results obtained from the equations in Appendix B, dashed-curves in A, C to the tight linkage approximation (equation 26, replacing h by $h\,(1-F)+F$ and r by $r\,(1-F)$), while dotted curves in B, D correspond to the high selfing approximation (equation 34). Parameter values: $s = 0.05$, $h = 0.3$, $u = 10^{-5}$.



998

999 **Figure 2.** Effective population size N_e as a function of the selfing rate α, for different
1000 values of the dominance coefficient of deleterious alleles h. Dots: simulation results (in
1001 this and the following figures, error bars are smaller than the size of dots); solid curves:
1002 analytical predictions from the complete model (see Supplementary File 2); dashed
1003 curves: tight linkage approximation (equation 26, replacing h by $h(1-F)+F$ and
1004 r by $r(1-F)$); dotted curves: high selfing approximation (equation 34). Parameter
1005 values: $N = 20000$, $s = 0.05$, $U = 0.1$, $R = 10$.



1006

1007 **Figure 3.** Effective population size N_e as a function of the selfing rate α, for different
1008 values of genome map length R: 0.1 (red), 1 (orange), 10 (green) and 100 (blue).
1009 Dots: simulation results; solid curves: analytical predictions from the complete model;
1010 dashed curves (left): tight linkage approximation; dotted curves (right): high selfing
1011 approximation. Parameter values: $N = 20000$, $s = 0.05$, $h = 0.3$, $U = 0.1$.

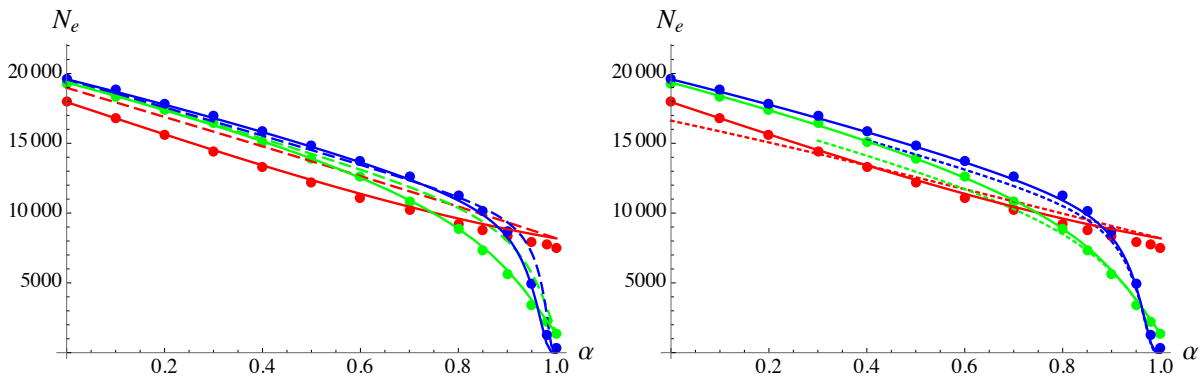


Figure 4. Effective population size N_e as a function of the selfing rate s, for different values of the strength of selection against deleterious alleles s: 0.005 (blue), 0.05 (green) and 0.5 (red). Dots: simulation results; solid curves: analytical predictions from the complete model; dashed curves (left): tight linkage approximation; dotted curves (right): high selfing approximation. Parameter values: $N = 20000$, $h = 0.3$, $U = 0.1$, $R = 10$.

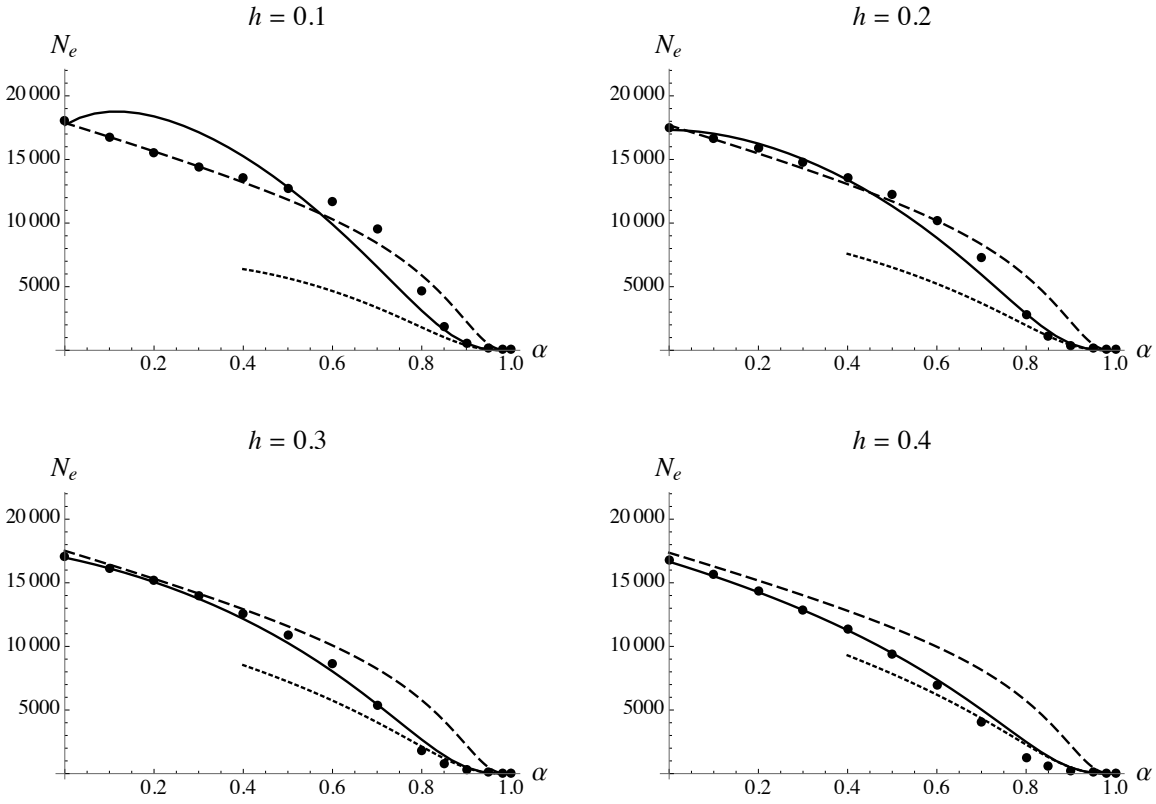


Figure 5. Effective population size N_e as a function of the selfing rate α, for different values of the dominance coefficient of deleterious alleles h: same as Figure 2 with $U = 0.5$.

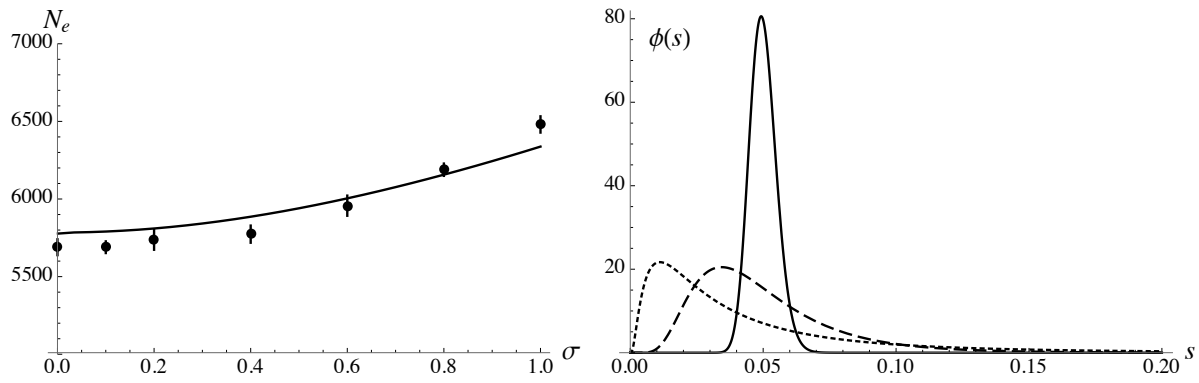


Figure 6. Left: Effective population size N_e as a function of the standard deviation σ of $\ln s$ across loci, assuming a log-normal distribution of s and setting the average of $\ln s$ so that $\overline{s} = 0.05$ for all values of σ. Curve: prediction obtained by integrating equation 34 over the distribution of s and over the genetic map; dots: simulation results. Parameter values: $N = 20000$, $\alpha = 0.9$, $U = 0.1$, $R = 10$. In the simulations the heterozygous fitness effect hs is the same for all mutations, and adjusted so that $\overline{h} = 0.25$ for all values of σ. The right figure shows the distribution of s for $\sigma = 0.1$ (solid), 0.5 (dashed) and 1 (dotted).

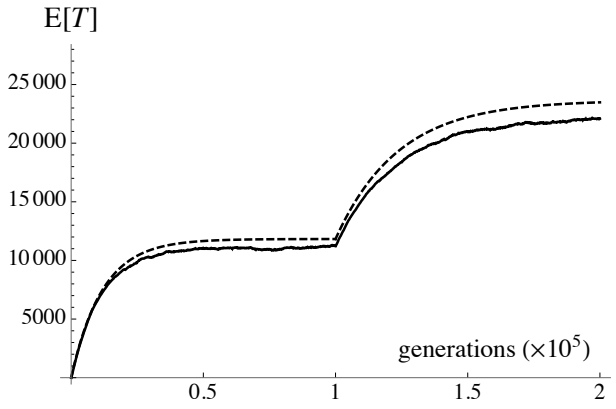


Figure 7. Average coalescence time between two randomly sampled sequences at a neutral locus located at the mid-point of the chromosome, as a function of time. The population is monomorphic at time zero (all lineages coalesce), $N = 20000$ during the first 10^5 generations while $N = 40000$ during the last 10^5 generations. Parameter values: $\alpha = 0.9$, $s = 0.05$, $h = 0.3$, $U = 0.1$, $R = 10$. Dashed curve: prediction from equation 36, where the inbreeding effective size is obtained from equation 35, and assumed to reach instantaneously its equilibrium value for a given N. Continuous curve: simulation results (average over 4000 replicate simulations), where $\mathrm{E}[T]$ is estimated from the diversity of a neutral sequence with an infinite number of sites (see text) and mutation rate $\mu = 10^{-4}$ per generation.