DOI: 10.1534/genetics.109.108258

# Deleterious Mutations and Selection for Sex in Finite Diploid Populations

**Denis Roze*,†,[1] and Richard E. Michod‡**

*Centre National de la Recherche Scientifique, Unité Mixte de Recherche 7144, Adaptation et Diversité en Milieu Marin and †Université Pierre et Marie Curie (Paris VI), 29682 Roscoff, France and ‡Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721

Manuscript received August 6, 2009
Accepted for publication January 12, 2010

ABSTRACT

In diploid populations, indirect benefits of sex may stem from segregation and recombination. Although it has been recognized that finite population size is an important component of selection for recombination, its effects on selection for segregation have been somewhat less studied. In this article, we develop analytical two- and three-locus models to study the effect of recurrent deleterious mutations on a modifier gene increasing sex, in a finite diploid population. The model also incorporates effects of mitotic recombination, causing loss of heterozygosity (LOH). Predictions are tested using multilocus simulations representing deleterious mutations occurring at a large number of loci. The model and simulations show that excess of heterozygosity generated by finite population size is an important component of selection for sex, favoring segregation when deleterious alleles are nearly additive to dominant. Furthermore, sex tends to break correlations in homozygosity among selected loci, which disfavors sex when deleterious alleles are either recessive or dominant. As a result, we find that it is difficult to maintain costly sex when deleterious alleles are recessive. LOH tends to favor sex when deleterious mutations are recessive, but the effect is relatively weak for rates of LOH corresponding to current estimates (of the order $10^{-4}$–$10^{-5}$).

ALTHOUGH sex is a costly (and often risky) enterprise, most eukaryotes engage in sex at least sometimes during their life cycle (reproduction being exclusively sexual in many species). Different theoretical models (reviewed in Agrawal 2006; Otto 2009) have explored the possible evolutionary advantages of sex due to its effects on genetic variation. Whether such benefits can maintain high rates of sex despite its important costs remains unclear, but modern computers allow more and more quantitative predictions to be obtained from simulation programs representing selection occurring at large numbers of loci (*e.g.*, Keightley and Otto 2006; Salathé *et al.* 2006). Eukaryotic sex is the combination of two complementary events—meiosis and syngamy—resulting in the alternation of a diploid and a haploid phase during the life cycle. In diploids, sex affects genetic variation through two effects: segregation and recombination. The genetic effect of recombination is to erode linkage disequilibrium (LD) among loci, and different models have explored the strength and direction of selection for recombination under various possible sources of LD: epistasis between beneficial or deleterious alleles (Charlesworth 1990, 1993; Barton 1995; Otto and Feldman 1997; Roze and Lenormand 2005); fluctuating epistasis over time, generated, for example, by host–parasite interactions (Barton 1995; Peters and Lively 1999; Otto and Nuismer 2004; Gandon and Otto 2007; Salathé *et al.* 2008), spatial heterogeneity in selection (Pylkov *et al.* 1998; Lenormand and Otto 2000); and LD generated by the interplay between selection and genetic drift, known as the Hill–Robertson effect (Felsenstein and Yokohama 1976; Otto and Barton 1997, 2001; Iles *et al.* 2003; Barton and Otto 2005; Keightley and Otto 2006; Martin *et al.* 2006; Roze and Barton 2006).

The possible benefits of segregation have been studied more recently. "Segregation" in the strict sense involves the separation of homologous chromosomes at meiosis. However, following others (*e.g.*, Otto 2003; Agrawal 2006), we use the term segregation to refer to the whole process of separation of homologous chromosomes at meiosis into separate gametes along with the eventual fusion (fertilization) of gametes taken from different individuals (outcrossing). In this article we consider only outcrossing and do not study selfing or other aspects of the mating system. In the same way as recombination "shuffles" alleles at different loci, eroding linkage disequilibria, segregation with outcrossing shuffles alleles at the same locus and brings populations closer to Hardy–Weinberg (HW) equilibrium. A first possible source of departure from HW equilibrium is dominance between alleles at a selected locus (in particular when sex is rare). Several models have explored

Supporting information is available online at http://www.genetics.org/cgi/content/full/genetics.109.108258/DC1.

[1]*Corresponding author:* Station Biologique de Roscoff, Place Georges Teissier, BP 74, 29682 Roscoff Cedex, France.
E-mail: roze@sb-roscoff.fr

the effect of dominance on selection for sex due to the segregation effect, in infinite populations. Uyenoyama and Bengtsson (1989) derived conditions for invasion of a modifier gene increasing the rate of sexual (*vs.* clonal) reproduction, when a recurrent lethal mutation occurs at a second locus. They found that under random mating, intermediate rates of sex are favored when the dominance coefficient $h$ of the lethal mutation is $< \frac{1}{3}$, while complete sex is favored when $h > \frac{1}{3}$. Otto (2003) considered a similar model, with arbitrary strength of selection at the second locus, and showed that the parameter range where sex is favored shrinks dramatically when selection becomes weak (in particular, deleterious mutations have to be nearly additive). Both Uyenoyama and Bengtsson (1989) and Otto (2003) studied the effects of partial self-fertilization, which generates departures from HW equilibrium (excess of homozygotes), and showed that selfing increases the parameter range where sex is favored. The effects of host–parasite coevolution on selection for segregation have been investigated by Agrawal and Otto (2006), who found that coevolutionary dynamics are less favorable to segregation than they are to recombination. Finally, Agrawal (2009) showed that spatial heterogeneity may favor segregation whenever locally beneficial alleles tend to be dominant, because of the excess homozygosity caused by spatial selection.

The effect of finite population size (or population structure) on selection for sex due to segregation has been little explored. Kirkpatrick and Jenkins (1989) showed that beneficial mutations may generate an advantage for sex in diploid populations, since mutations remain in the heterozygous state in asexuals (until a second mutation occurs), while in sexuals segregation quickly generates homozygotes for the beneficial allele (note that this benefit of sex does not occur in infinite populations, because a small proportion of double mutants would be produced every generation). Haag and Roze (2007) compared the equilibrium mutation load in sexual and asexual finite diploid populations and showed that the load may be far greater in asexuals in the parameter range where deleterious mutations reach fixation in the heterozygous state in asexuals, while they are still eliminated efficiently in sexuals due to segregation. However, the fate of a modifier affecting the rate of sex in such a system was not investigated. Finite population size may have stronger effects on selection for segregation than it has on selection for recombination: indeed in the case of recombination, linkage disequilibrium between loci is generated by the interplay between random drift and selection (Hill–Robertson effect) and should be of order $s_A s_B/N$ between two loci A and B, where $s_A$ and $s_B$ measure the strength of selection acting at these loci and $N$ is population size (*e.g.*, Barton and Otto 2005; Martin *et al.* 2006). By contrast, finite population size tends to generate an excess of heterozygotes (particularly in partially clonal populations) even in the absence of selection (*e.g.*, Balloux *et al.* 2003; Haag and Roze 2007), and the resulting departure from Hardy–Weinberg equilibrium should thus be of order $1/N$. However, the relative forces of selection for segregation and recombination in finite diploid populations have never been explored.

Another potentially important factor affecting selection for segregation is mitotic gene conversion and/or recombination, which occur as a consequence of DNA repair and tend to generate homozygosity (*e.g.*, Mandegar and Otto 2007). A double-strand break (DSB) is perhaps the most serious damage that can happen to a DNA molecule. Eukaryotes use two main pathways to repair such damage: nonhomologous end joining (NHEJ) involves the detection and ligation of the two broken ends, but is prone to error and chromosomal aberrations, while homologous recombination (HR) uses an homologous sequence as a template for repair, preferentially from the sister chromatid if the chromosome is replicated or from the homologous chromosome if the cell is diploid (Liang *et al.* 1998; Pâques and Haber 1999; Johnson and Jasin 2001; Helleday 2003; Aylon and Kupiec 2004). Note that HR may also occur in response to DNA lesions other than DSBs (Lettier *et al.* 2006). Indeed, one of the potential advantages of HR as a repair process is that in principle HR can cope with any kind of DNA damage. A consequence of HR is that a small fragment of the genome is transferred from the donor template to the repaired chromosome with heteroduplex DNA being generated to either side of the break, and this process may cause gene conversion when the template is the homologous chromosome. Another possible consequence of HR is crossing over (exchange of chromosome arms between homologs). Although HR occurs during both meiosis and mitosis, there are important differences between meiotic and mitotic HR: for example, the length of gene conversion tracts may differ: in yeast mitotic gene conversion tracts are often >4 kb (Judd and Petes 1988), longer than meiotic tracts (~1–2 kb), while in mammals mitotic tracts are usually short, ~200 bp−1 kb (Chen *et al.* 2007). Furthermore, a large proportion of meiotic HR events result in crossing over [up to 66% in yeast (Prado *et al.* 2003)], while this proportion is usually a lot smaller in the case of mitotic HR [<8% in several model systems (Richardson *et al.* 1998; Virgin *et al.* 2001; Ira *et al.* 2006; Chen *et al.* 2007)]. Crossing over occurring at mitosis (hereafter called "mitotic recombination," MR) may render daughter cells homozygous over large portions of the chromosome (Mandegar and Otto 2007). Such loss of heterozygosity (LOH) events represent a frequent step in oncogenesis (Gupta *et al.* 1997; Tischfield 1997; Hagstrom and Dryja 1999; Holt *et al.* 1999; Sieber *et al.* 2002), which may partly explain why mitotic crossing over is suppressed. Nevertheless, mitotic recombination is thought to occur at rate

$\sim 0.8 \times 10^{-4}$ per cell per generation in yeast (Mandegar and Otto 2007 and references therein), while LOH is often observed at a frequency of $10^{-4}$–$10^{-5}$ in normal cells *in vivo*, in mice and in humans (Tischfield 1997; Holt *et al.* 1999; Shao *et al.* 1999; Carr and Gottschling 2008). Furthermore, Chamnanpunt *et al.* (2001) measured rates of mitotic gene conversion ranging from $3 \times 10^{-2}$ to $10^{-5}$ in hybrids of the oomycete *Phytophthora sojae*. Finally, LOH has been estimated to occur at rate $\sim 10^{-4}$ per locus per generation in an experiment involving mutation-accumulation lines of asexual Daphnia (Omilian *et al.* 2006).

To our knowledge, the only model studying the effect of mitotic recombination on selection for sex was done by Mandegar and Otto (2007), who showed that the benefit of sex described in Kirkpatrick and Jenkins (1989) (creation of homozygotes for beneficial alleles, by segregation) disappears in the presence of a low rate of mitotic recombination (which allows homozygotes to be generated in asexuals). Loss of heterozygosity due to HR is also a central component of the DNA repair and mutation complementation hypothesis for the evolution of sex (Bernstein *et al.* 1985, 1988). This hypothesis distinguishes between the recombinational and the outcrossing (segregational) aspects of sex, because it assumes that these two aspects of sex have different adaptive functions. According to this hypothesis, the main function of recombination, in the sense of breakage and rejoining of DNA molecules, is to repair DSBs and other kinds of DNA damage, while the function of segregation with outcrossing is to restore heterozygosity lost by the repair process (LOH across large portions of the genome leads to the expression of recessive deleterious alleles). This hypothesis has been criticized, notably because recombination without (or with little) crossing over is possible (mitotic HR only rarely leads to crossing over) and because double-strand breaks are actively generated during the initiation of meiotic recombination, making it doubtful that the main role of meiosis is to repair DSBs (Maynard Smith 1988; Keeney *et al.* 1997; Barton and Charlesworth 1998; Keeney and Neale 2006). Nevertheless, loss of heterozygosity occurring within mitotic lineages may generate a selective force for segregation and outcrossing (due to the expression of recessive deleterious mutations). Quantifying this force (which has never been done) is one of the goals of our study.

In this article, we use a two-locus analytical model to explore the effects of finite population size and of mitotic recombination on selection for outcrossed sex and segregation, focusing on the effect of recurrent deleterious mutations. This model shows that finite population size generates a selective pressure on a modifier affecting the rate of sexual *vs.* asexual reproduction, mainly due to the fact that drift tends to generate an excess of heterozygotes in the population (in particular when the rate of sex is low). This selective force is often stronger than the deterministic force that has been described in the case of infinite, randomly mating populations (Otto 2003), in particular when deleterious alleles remain at low frequency. Furthermore, we show that loss of heterozygosity (due to mitotic gene conversion and recombination) may have important positive effects on selection for sex if it occurs sufficiently frequently (at rate $\geq \sim 10^{-3}$ per generation), while rates of LOH $\leq 10^{-4}$ have little effect. We then extend our model to introduce a second selected locus and show that selection for sex due to the interaction between selected loci may become important when the deleterious mutation rate is sufficiently high. These results are confirmed by multilocus simulations in which deleterious mutations occur at a large number of loci and where the rate of sex is free to evolve. Finally, from these simulation results it appears unlikely that deleterious mutations alone can allow the maintenance of high rates of sex when the cost of sex is twofold.

## ANALYTICAL MODEL

**General setting:** Table 1 summarizes the different parameters and variables of the model. We consider two loci M and A in a diploid population with nonoverlapping generations. At the start of a generation, the population consists of $N$ adults. These adults produce a large number of juveniles, with different fecundities depending on their genotype at locus A. We assume that two alleles $A$ and $a$ segregate at this locus and that genotypes $aa$, $Aa$, and $AA$ have (relative) fecundities 1, $1 - hs$, and $1 - s$ (where fecundity is simply the number of juveniles produced). Allele $A$ is thus deleterious, and we assume that mutations from $a$ to $A$ occur at a rate $u$ per generation and back mutations occur at rate $v$. Note that uppercase letters do not refer to dominant alleles, as we consider the case where $A$ is recessive ($h < \frac{1}{2}$) and the case where it is dominant ($h > \frac{1}{2}$). Locus M controls the proportion of juveniles produced sexually and asexually: two alleles $M$ and $m$ segregate at this locus, and individuals of genotype $mm$, $Mm$, and $MM$ produce a proportion $\sigma$, $\sigma + h_M\delta\sigma$, and $\sigma + \delta\sigma$ of offspring sexually, respectively, and the remainder asexually (*i.e.*, by mitosis). Locus M is thus a modifier locus affecting the rate of sex (Otto 2003), $\delta\sigma$ measures the modifier effect, and $h_M$ the dominance of allele $M$. Sexual reproduction involves gamete production by meiosis (we call $r$ the recombination rate between loci M and A), gamete release in the environment, and random union of gametes in the whole population. Except in the simulations, we assume no direct cost of sex, so that individuals bearing the same genotype at locus A produce the same number of juveniles, whatever their genotype at locus M (the modifier has no direct effect on fitness). Finally, $N$ individuals are sampled randomly among the large number of juveniles produced to form the next adult generation. To incorpo-

**TABLE 1**

**Parameters and variables**

| | |
|---|---|
| $N$ | Population size |
| $s$ | Strength of selection against allele $A$ |
| $h$ | Dominance of allele $A$ |
| $u$ | Mutation rate from allele $a$ to allele $A$ |
| $v$ | Back mutation rate (from $A$ to $a$) |
| $r$ | Rate of recombination between loci A and M |
| $\sigma$ | Baseline rate of sex in the population |
| $\delta\sigma$ | Effect of the modifier (allele $M$) on the rate of sex |
| $h_M$ | Dominance of allele $M$ |
| $\gamma$ | Rate of mitotic gene conversion |
| $\chi$ | Rate of mitotic crossing over (in simulations) |
| $U$ | Deleterious mutation rate per chromosome (in simulations) |
| $L$ | Average number of crossovers at meiosis (in simulations) |
| $c$ | Cost of sex (in simulations): relative fitness of asexuals in the absence of deleterious mutation |
| $\sigma_{init}$ | Initial rate of sex (in simulations) |
| $p_A$, $p_M$ | Frequencies of alleles $A$ and $M$ |
| $q_A$, $q_M$ | Frequencies of alleles $a$ and $m$ |
| $D_{A,A}$ | Departure from Hardy–Weinberg equilibrium (excess of homozygotes) at locus A |
| $D_{MA,A}$ | Association between the modifier and homozygotes at locus A |
| $D_{MA}$, $D_{M,A}$ | Association between the modifier and the deleterious allele at locus A, on the same ($D_{MA}$) or on the homologous chromosome ($D_{M,A}$) |
| $\langle X\rangle_t$ | Expected value of random variable $X$ at time $t$ |

rate effects of mitotic recombination, we assume that a proportion $\gamma$ of juveniles produced asexually have undergone (unbiased) gene conversion at locus A: therefore, a proportion $\gamma/2$ of offspring produced asexually by *Aa* heterozygotes are *AA* homozygotes, while a proportion $\gamma/2$ are *aa* homozygotes. Although gene conversion occurs at much higher rates during meiosis than during mitosis (Chen *et al.* 2007), meiotic gene conversion would not affect the dynamics (as long as it is not biased) because it would have no effect on the frequencies of homozygotes and heterozygotes among offspring. Loss of heterozygosity does not affect the modifier locus in the analytical model, but it will do so in the simulations representing mitotic recombination, as explained in the next section.

The population is described in terms of allele frequencies and genetic associations. We call $p_A$ and $p_M$ the frequencies of alleles $A$ and $M$, respectively, and $q_A$, $q_M$ the frequencies of alleles $a$ and $m$. Genetic associations measure covariances in allelic state between genes present at different loci and/or on different chromosomes of an individual (Barton and Turelli 1991; Kirkpatrick *et al.* 2002). They are represented by variables $D_{\mathbb{U},\mathbb{V}}$, where $\mathbb{U}$ and $\mathbb{V}$ are sets of loci present on the first and second haplotypes of a diploid individual. In the context of the present model, one of the most important association is $D_{A,A}$, measuring the covariance in allelic state between the two genes at locus A in an individual,

$$D_{A,A} = p_{AA} - p_A^2, \tag{1}$$

where $p_{AA}$ is the frequency of *AA* homozygotes in the population (and $p_A^2$ is the probability of sampling two $A$ alleles with replacement from the population). When $D_{A,A} = 0$, the population is at Hardy–Weinberg equilibrium, while a positive (negative) value of $D_{A,A}$ indicates an excess (deficit) of homozygotes in the population (note that $D_{A,A}$ and the more widely used inbreeding coefficient $F$ are linked by the relation $D_{A,A} = Fp_Aq_A$). Because population size is finite, allele frequencies and genetic associations are random variables. In the following, we use the notation $\langle X\rangle$ to denote the expected value of random variable $X$. A method for obtaining recursions over the different phases of the life cycle (selection, gene conversion, reproduction, and drift) for expected values of allele frequencies, genetic associations, and moments of allele frequencies and associations is presented in supporting information, Appendix SA—note that this method differs from the method used by Barton and Otto (2005) and by Martin *et al.* (2006) in that it does not assume that the population remains close to a deterministic trajectory; however, we assume that population size is large. Recursions for the different moments that affect the change in frequency of the modifier are given in Appendix SB. Throughout, we assume that selection is weak, gene conversion is rare, and population size is large: $s$, $\gamma$, and $1/N$ are all of order $\varepsilon$, where $\varepsilon$ is a small term (in the simulations, we explore cases where $s$, $\gamma \gg 1/N$). We also assume that the rates of sex and recombination ($\sigma$, $r$) are not too small. Under these conditions, it can be shown that the population quickly reaches a state of quasi-equilibrium, where expected values of genetic associations are small and change very slowly over time (Nagylaki 1993; Bürger 2000; see Appendix SA). In that case, associations can be expressed as functions of allele frequencies and of the different parameters of the model (*e.g.*, Barton and Turelli 1991; Kirkpatrick *et al.* 2002). Finally, we assume that the modifier has only a small effect on the rate of sex ($\delta\sigma$ small) and calculate terms to the first order in $\delta\sigma$ only (we thus neglect terms in $\delta\sigma^2$, $\delta\sigma^3$, ...). These different assumptions are relaxed in the simulation program presented in the next section.

**Departure from Hardy–Weinberg equilibrium:** Most of the results of the model can be understood by the effects of different forces generating departure from Hardy–Weinberg equilibrium at the selected locus (measured by $D_{A,A}$); indeed, the main effect of increasing sex in our model is to bring the population closer to HW equilibrium. From Appendix SB, the expected value of $D_{A,A}$ at quasi-equilibrium (measured after selection and gene conversion), to the first order in $\varepsilon$, is given by

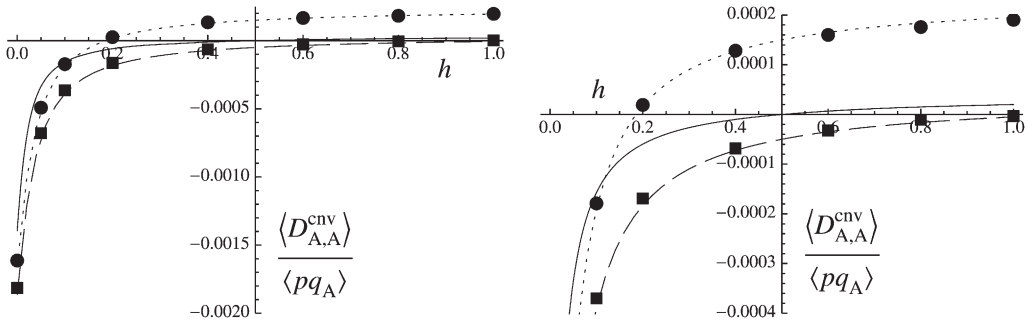



FIGURE 1.—Average departure from Hardy–Weinberg equilibrium at the selected locus, after selection and gene conversion, divided by the average genetic variance at the same locus, as a function of the dominance coefficient of deleterious mutations $h$ (the right side is a blow-up of the left). Curves correspond to analytical predictions from Equation 2 for different parameter values. Solid curve, deterministic limit ($N\to\infty$), no gene conversion ($\gamma = 0$), the value of $p_A$ at mutation–selection equilibrium is obtained by solving $-sp_A[h + (1 - 2h)p_A] + u = 0$; dashed curve, $N = 20{,}000$, $\gamma = 0$; dotted curve, $N = 20{,}000$, $\gamma = 10^{-4}$. For both the dashed and dotted curves, $\langle pq_A^2\rangle/\langle pq_A\rangle$ is obtained by numerical integration over Wright's distribution, $\phi(p_A) = Ke^{-2Ns p_A[2h+(1-2h)p_A]}p_A^{4Nu-1}q_A^{4Nv-1}$, where $K$ is a constant. Squares and circles: simulation results for $N = 20{,}000$, $\gamma = 0$ (squares) and $N = 20{,}000$, $\gamma = 10^{-4}$ (circles). Other parameter values: $s = 0.05$, $\sigma = 0.5$, $u = 10^{-5}$, $v = 10^{-6}$.

$$\langle D^{cnv}_{A,A}\rangle = \frac{1}{\sigma}\left[\left(\gamma - \frac{1}{2N}\right)\langle p_A q_A\rangle - s(1-2h)\langle p_A^2 q_A^2\rangle\right], \quad (2)$$

where "cnv" indicates that the association is measured after selection and gene conversion (see Appendix SA). Equation 2 isolates the effects of three forces generating departure from HW equilibrium: selection, gene conversion, and random drift. As was shown previously (OTTO 2003; AGRAWAL 2009), selection generates an excess of heterozygotes (negative $D_{A,A}$) whenever $h < \frac{1}{2}$ (that is, when the deleterious allele is partly recessive): indeed in that case, the fitness of heterozygotes at locus A is higher than the average fitness of homozygotes. When the deleterious allele is partly dominant ($h > \frac{1}{2}$), however, selection generates an excess of homozygotes (positive $D_{A,A}$). Because gene conversion generates homozygotes from heterozygotes, it produces a positive $D_{A,A}$ (term in $\gamma$ in Equation 2). Finally, drift tends to generate an excess of heterozygotes (negative $D_{A,A}$), through the term in $1/(2N)$ (note that this term does not involve $s$ and is thus not equivalent to the Hill–Robertson effect). This effect can be understood by noting that $D_{A,A}$ can also be written as freq($AA$)freq($aa$) − [freq($Aa$)/2][freq($Aa$)/2]. Random sampling from a population initially at HW equilibrium (without selection) may generate an excess either of homozygotes or of heterozygotes; however, when one homozygote (say $AA$) is oversampled, the effect on $D_{A,A}$ is weakened by the fact that the other homozygote ($aa$) is likely to be rarer due to the oversampling of $AA$. By contrast, when $Aa$ is oversampled, both terms of the second part of the expression of $D_{A,A}$ equally contribute toward negative $D_{A,A}$. This may be seen most easily by considering a fully asexual population in which $A$ and $a$ are neutral, and mutation rates are so small that the population is fixed most of the time for one of the three genotypes: $D_{A,A} = 0$ when $AA$ or $aa$ is fixed, while $D_{A,A} = -0.25$ when $Aa$ is fixed (and thus $D_{A,A}$ is negative, on average). Note that this asymmetry is not present with LD between two loci, because the two halves of LD = freq($AB$)freq($ab$) − freq($Ab$)freq($aB$) both involve two different haplotypes (we are grateful to Sally Otto for offering this explanation). Finally, Equation 2 also indicates that when the deleterious allele remains rare ($p_A$ small), so that the term $\langle p_A^2 q_A^2\rangle \approx \langle p_A^2\rangle$ is small relative to $\langle p_A q_A\rangle \approx \langle p_A\rangle$, the effect of selection on the departure from HW equilibrium may be negligible relative to the effects of drift and/or gene conversion. In particular, $\langle p_A^2 q_A^2\rangle/\langle p_A q_A\rangle$ should roughly scale with $u/(sh)$ when $h > 0$ and $s \gg u$, $1/N$, so that the relative importance of the three effects should depend on the relative magnitude of $1/(2N)$, $\gamma$, and $u(1 - 2h)/h$.

Figure 1 shows the equilibrium value of $\langle D^{cnv}_{A,A}\rangle/\langle p_A q_A\rangle$ predicted from Equation 2, in the case of an infinite population (no drift) and $\gamma = 0$ (solid curve), with drift ($N = 20{,}000$) but no gene conversion (dashed curve), and with both drift and gene conversion ($N = 20{,}000$ and $\gamma = 10^{-4}$, dotted curve). In the last two cases, the value of $\langle p_A^2 q_A^2\rangle/\langle p_A q_A\rangle$ is obtained by numerical integration over Wright's distribution of allele frequency at mutation–selection–drift equilibrium (note that partial asexual reproduction does not affect this distribution as long as the rate of sex is not too small, and $D_{A,A}$ is of order $\varepsilon$). Figure 1 also shows results from a simulation program, in which the effects of selection, mutation, gene conversion, reproduction, and drift on the frequencies of the three genotypes at locus A are iterated over a large number ($10^9$) of generations, $\langle D^{cnv}_{A,A}\rangle$ and $\langle p_A q_A\rangle$ being measured every 100 generations. These simulation results show that Equation 2 provides accurate predictions of the average departure from HW equilibrium.

**Selection on the modifier: weak selection:** From the methods of Appendix SA, one obtains that the expected change in frequency of the modifier at generation $t$, to leading order in $\varepsilon$, is given by

$$\langle \Delta p_M\rangle_t \approx -s(1-2h)\langle D_{MA,A}\rangle_t, \quad (3)$$

where $D_{MA,A}$ measures the association between allele $M$ and the two alleles of the same individual at locus A. In

Appendix SC, we show that at quasi-equilibrium this association is positive whenever the frequency of allele $M$ is higher in homozygotes at locus A than in heterozygotes, while $D_{MA,A}$ is negative whenever the frequency of $M$ is higher in heterozygotes than in homozygotes. Noting that $-s(1-2h)$ is proportional to the difference between the average fitness of homozygotes and the fitness of heterozygotes at locus A, Equation 3 can be understood easily: allele $M$ can be favored either when homozygotes have a higher fitness than heterozygotes [$h > \frac{1}{2}$, so that $-s(1-2h) > 0$] and allele $M$ tends to be associated with homozygotes ($D_{MA,A} > 0$) or when heterozygotes have a higher fitness than homozygotes ($h < \frac{1}{2}$) and allele $M$ tends to be associated with heterozygotes ($D_{MA,A} < 0$). Note that Equation 3 takes the same form in an infinite population (Agrawal 2009).

In the following, we focus on the case of an additive modifier ($h_M = \frac{1}{2}$); however, expressions for arbitrary $h_M$ are given in Appendix SB. Using the quasi-equilibrium approximation, we can derive an expression for $\langle D_{MA,A}\rangle_t$ as a function of the parameters and moments of allele frequencies. One obtains

$$\langle D_{MA,A}\rangle_t \approx -\frac{\delta\sigma}{2\sigma}\left[\langle p_M q_M D^{cnv}_{A,A}\rangle_t - \frac{1+2r(1-r)-r\sigma}{2N\sigma[1+r(1-\sigma)](2-r\sigma)}\langle pq_{MA}\rangle_t\right], \quad (4)$$

where $pq_{MA} = p_M q_M p_A q_A$. The first term of Equation 4 is equivalent to the expression obtained in the case of an infinite population: in that case, $D_{MA,A}$ is generated by the modifier effect ($\delta\sigma$) and by the departure from HW equilibrium ($D_{A,A}$). This simply stems from the fact that a modifier increasing sex tends to bring the population closer to HW equilibrium; therefore if $D_{A,A} > 0$ (excess of homozygotes in the population), a modifier allele increasing sex tends to create more heterozygotes and thus tends to be more frequent in heterozygotes than in homozygotes ($D_{MA,A} < 0$ if $\delta\sigma > 0$). If $D_{A,A} < 0$ (excess of heterozygotes in the population), a modifier increasing sex creates more homozygotes and is thus more frequent among homozygotes ($D_{MA,A} > 0$ if $\delta\sigma > 0$). The second term of Equation 4 is generated by drift and by the modifier effect and involves associations other than $D_{A,A}$ (in particular, associations $\langle D_{MA}{}^2\rangle$, $\langle D_{M,A}{}^2\rangle$ and $\langle D_{MA}D_{M,A}\rangle$, see Appendix SB). This extra effect comes from the fact that drift generates random associations between alleles $M$ and $A$ at some generations and between $M$ and $a$ at other generations. Because individuals carrying $M$ engage in sex more often (assuming $\delta\sigma > 0$), these random associations translate into an excess of homozygotes at locus A among individuals carrying $M$.

Replacing the first term of Equation 4 by its expression at quasi-equilibrium, one obtains

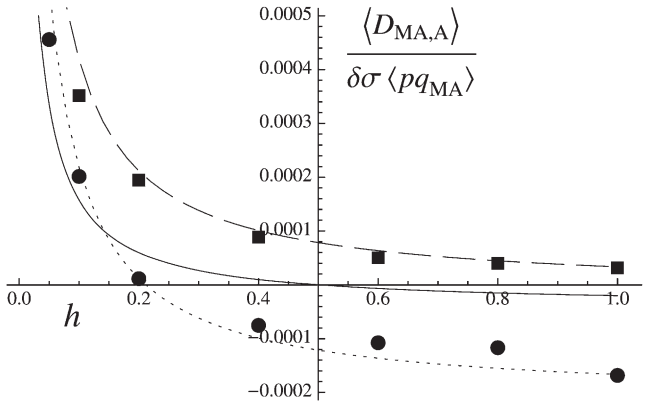



Figure 2.—Average association $D_{MA,A}$ between the modifier and homozygotes at the selected locus (additive modifier), divided by the modifier effect $\delta\sigma$ and the average product of genetic variances at both loci, as a function of the dominance coefficient of deleterious mutations $h$. Parameter values are the same as for Figure 1, with $r = 0.5$, $\delta\sigma = 0.05$.

$$\langle D_{MA,A}\rangle_t \approx \frac{\delta\sigma}{2\sigma^2}\left[\left(\frac{T_1}{2N}-\gamma\right)\langle pq_{MA}\rangle_t + s(1-2h)\langle pq_M\, pq_A^2\rangle_t\right] \quad (5)$$

with

$$T_1 = \frac{3-2r^2+r(1-\sigma)(4-r\sigma)}{[1+r(1-\sigma)](2-r\sigma)} > 0, \quad (6)$$

which simplifies to $T_1 = (6-\sigma)/(4-\sigma)$ when $r = \frac{1}{2}$ (free recombination). Note that from Equations 3 and 5, the change in frequency of the modifier is of order $\delta\sigma\varepsilon^2$. These equations predict that under weak selection (so that the quasi-equilibrium approximation holds), increased sex should never be favored in the absence of drift and gene conversion, as was found previously (*e.g.*, Agrawal 2009): when $h < \frac{1}{2}$, increasing sex tends to produce homozygotes, but homozygotes have a lower fitness than heterozygotes, while when $h > \frac{1}{2}$, increasing sex tends to produce heterozygotes, but heterozygotes have lower fitness. We see in the next section that sex can nevertheless be favored through terms in $\delta\sigma\varepsilon^3$ that may become relatively strong when dominance is weak and/or sex is rare (the parameter window favoring sex opening up as selection becomes stronger). From Equations 3 and 5, drift causes a modifier increasing sex to be associated with homozygotes at the selected locus, which favors sex whenever deleterious mutations are dominant ($h > \frac{1}{2}$), but disfavors sex when deleterious mutations are recessive ($h < \frac{1}{2}$). Finally, gene conversion favors sex when mutations are recessive (as increasing sex tends to mask mutations that have become homozygous due to gene conversion), but disfavors sex when mutations are dominant.

Figure 2 shows a test of Equation 5 against simulations. The simulation program iterates the effects of selection, mutation, gene conversion, reproduction (in-

cluding a sex modifier), and drift on the 10 two-locus diploid genotype frequencies (initial population without deleterious mutation and the modifier being in frequency 0.5) until the modifier frequency is >0.95 or <0.05, measuring $D_{MA,A}$ and $pq_{MA}$ every 10 generations. The whole process is repeated until $10^9$ points have been obtained. To generate the curves of Figure 2, we neglected any correlation in genetic variance at the two loci ($p_Aq_A$ and $p_Mq_M$) and thus replaced $\langle pq_M\ pq_A^2\rangle/\langle pq_{MA}\rangle$ by $\langle p_A^2q_A^2\rangle/\langle p_Aq_A\rangle$, which is again obtained by numerical integration over Wright's distribution.

**Selection on the modifier: weak dominance:** The term on the right-hand side of Equation 3 becomes vanishingly small as $h$ tends to $\frac{1}{2}$; in that case, other terms affecting the change in frequency of the modifier become predominant. In Appendix SB, we calculate these terms when both $s$ and $h-\frac{1}{2}$ are of order ε ("weak dominance," where "weak" is relative to the strength of selection $s$). In that case, the change in frequency of the modifier is given by

$$\langle \Delta p_M\rangle_t \approx -sh(\langle D_{MA}\rangle_t + \langle D_{M,A}\rangle_t) - s(1-2h)\langle D_{MA,A}\rangle_t, \tag{7}$$

which again takes the same form as in an infinite population (Agrawal 2009). The first term of Equation 7 represents selection on the modifier due to its association with the deleterious allele, either on the same chromosome ($D_{MA}$) or on the other chromosome of the same individual ($D_{M,A}$). As shown in Appendix SC, a positive value of $D_{MA} + D_{M,A}$ means that the frequency of $M$ is higher in individuals in which the frequency of the deleterious allele $A$ is higher (which selects against the modifier). Conversely, when $D_{MA} + D_{M,A} < 0$ the frequency of $M$ is lower in individuals carrying the deleterious allele, which favors the modifier through the first term of Equation 7. Quasi-equilibrium values of $\langle D_{MA}\rangle_t$ and $\langle D_{M,A}\rangle_t$ are given by

$$\langle D_{MA}\rangle_t = -sh\left[\frac{1+r(1-\sigma)}{r\sigma}\langle D_{MA,A}\rangle_t + \frac{\delta\sigma}{2N\sigma^3}T_3\langle pq_{MA}\rangle_t\right] \tag{8}$$

$$\langle D_{M,A}\rangle_t = -sh\left[\frac{1-\sigma}{\sigma}\langle D_{MA,A}\rangle_t + \frac{\delta\sigma}{2N\sigma^3}T_3\langle pq_{MA}\rangle_t\right] \tag{9}$$

with

$$T_3 = \frac{1}{2-\sigma}\left[1+\frac{(1-\sigma)[1+2r(1-r)-r\sigma]}{[1+r(1-\sigma)](2-r\sigma)}\right] > 0, \tag{10}$$

which simplifies to $T_3 = 3/(4-\sigma)$ when $r=\frac{1}{2}$. The first term within the brackets of Equations 8 and 9 is equivalent to the expression obtained in the case of an infinite population: in this case, $\langle D_{MA}\rangle$ and $\langle D_{M,A}\rangle$ have the opposite sign to $\langle D_{MA,A}\rangle$ at quasi-equilibrium. Indeed, when the frequency of homozygotes at locus A is higher in the subset of the population that carries allele $M$ (that is, when $D_{MA,A} > 0$), selection against the deleterious allele is more efficient in that subset (because the variance in fitness is higher). This generates a negative association between allele $M$ and the deleterious allele $A$ ($D_{MA}$, $D_{M,A} < 0$, meaning that the frequency of the deleterious allele is lower in individuals carrying the modifier allele). The magnitude of this effect increases as $r$ and σ decrease, so that the benefit of better purging remains confined to individuals carrying allele $M$. Conversely, when the frequency of heterozygotes at locus A is higher in the subset of the population that carries allele $M$ ($D_{MA,A} < 0$), purging becomes less efficient in that subset, generating positive associations $D_{MA}$ and $D_{M,A}$ between the modifier and the deleterious allele. The second term within the brackets of Equations 8 and 9 is an effect of finite population size, which tends to generate negative values of $D_{MA}$ and $D_{M,A}$ for a modifier increasing sex ($\delta\sigma > 0$). As shown in Appendix SB, this term is generated by the association $\langle D_{MA,M}\rangle$, which is itself generated by the combined action of selection and drift and is positive at quasi-equilibrium. Indeed, drift generates random associations between the deleterious allele $A$ and homozygotes at locus M at some generations ($D_{MA,M} > 0$) and between $A$ and heterozygotes at locus M at other generations ($D_{MA,M} < 0$). In the first case, selection increases the genetic variance at locus M (because it increases the frequency of heterozygotes at locus M), while in the second case it decreases the variance. Because $D_{MA,M}$ is higher in absolute value when there is more genetic variance, the average value of $D_{MA,M}$ after selection is positive. On average, the deleterious allele $A$ is thus more frequent in homozygotes at locus M than in heterozygotes; however, the fact that $MM$ individuals engage in sex more often with $Mm$ heterozygotes (in which $A$ is less frequent) than $mm$ individuals (assuming $\delta\sigma > 0$) tends to create a negative association between $M$ and $A$. Unfortunately we could not test Equations 8 and 9 by simulation, because the expected values of $D_{MA}$ and $D_{M,A}$ are very small (of order $\delta\sigma\varepsilon^2$), while the variance of these terms is much larger (of order ε, see Appendix SB). Finally, note that the first term of Equation 7 may be relatively strong even when dominance is not weak, in particular when the rate of sex is low (indeed, the first term scales with $1/\sigma^3$ while the second term scales with $1/\sigma^2$) and/or selection is sufficiently strong.

In summary, we have seen that selection on the modifier involves two effects: the effect of the modifier on the frequencies of homozygotes and heterozygotes—term $-s(1-2h)\langle D_{MA,A}\rangle$ of $\langle \Delta p_M\rangle$, hereafter called "selection through effect on genotype frequencies"—and the effect of the modifier on the frequency of the deleterious allele—term $-sh(\langle D_{MA}\rangle + \langle D_{M,A}\rangle)$ of $\langle \Delta p_M\rangle$, hereafter called "selection through effect on allele frequencies" (we prefer these terms over the commonly used "short-term" and "long-term" effects,

**TABLE 2**

**Effect of the departure from HW equilibrium generated by selection, gene conversion, and finite population size on the two components of indirect selection acting on the modifier**

| | Selection through effect on genotype frequencies | | Selection through effect on allele frequencies | |
|---|---|---|---|---|
| | $h < \frac{1}{2}$ | $h > \frac{1}{2}$ | $h < \frac{1}{2}$ | $h > \frac{1}{2}$ |
| $D_{A,A}$ generated by selection | − | − | + | − |
| Gene conversion (term in $\gamma$) | + | − | − | − |
| Finite population size (term in $1/N$) | − | + | + | + |

A " + " indicates a positive effect on the term $-s(1-2h)\langle D_{MA,A}\rangle$ of the change in frequency of the modifier (selection through effect on genotype frequencies) or on the term $-sh(\langle D_{MA}\rangle + \langle D_{M,A}\rangle)$ (selection through effect on allele frequencies), in the case of a modifier increasing sex ($\delta\sigma > 0$). A " − " indicates a negative effect.

as we think that they are more explicit). When selection is weak ($s$ very small), the second effect is expected to be negligible unless $h$ is close to $\frac{1}{2}$ or sex and recombination are low. Table 2 shows how these two terms are affected by gene conversion, by finite population size, and by the departure from HW equilibrium caused by selection, assuming that the modifier increases sex ($\delta\sigma > 0$). In the case of recessive deleterious mutations ($h < \frac{1}{2}$), finite population size and the effect of selection on $D_{A,A}$ both disfavor sex through the effect on genotype frequencies but favor sex through the effect on allele frequencies, while gene conversion has opposite effects. When mutations are dominant ($h > \frac{1}{2}$), gene conversion and the effect of selection on $D_{A,A}$ disfavor sex through both effects, while finite population size favors sex through both effects.

The evolutionarily stable rate of sex (that is, the rate of sex toward which the population should evolve, if evolution proceeds by small-step mutations) can be computed by replacing associations in Equation 7 by their expression at quasi-equilibrium and finding the values of $\sigma$ for which $\langle \Delta p_M \rangle = 0$. Solutions at which the derivative of $\langle \Delta p_M \rangle$ with respect to $\sigma$ is negative (positive) correspond to stable (unstable) equilibria. Figure 3 shows these equilibria as a function of $h$, in three different cases. In the deterministic limit ($N$ tends to infinity) and in the absence of gene conversion ($\gamma = 0$), the rate of sex converges to a stable equilibrium line when $h < \frac{1}{2}$ (solid curve), while it converges to $\sigma = 0$ when $h > \frac{1}{2}$. For $N = 20{,}000$, still in the absence of gene conversion, the situation is similar when $h < \frac{1}{2}$ (dashed curve, corresponding to a stable equilibrium line), while the rate of sex converges to 1 when $h > \frac{1}{2}$. Finally, for $N = 20{,}000$ and a rate of gene conversion $\gamma = 10^{-4}$, a stable equilibrium line appears for low values of $h$ (dotted curve) at low values of $\sigma$, with a sharp increase around $h = 0.2$, while an unstable equilibrium line appears at higher values of $h$ (~0.25–0.5, dashed-dotted curve). Starting between the dotted and dashed-dotted curves, the rate of sex should thus increase to 1, while starting on the right of the dashed-dotted curve (or below), the rate of sex should decrease toward zero. For $\gamma = 10^{-3}$ and $\gamma = 10^{-2}$, the part of the stable dotted curve at $\sigma = 1$ extends to $h = 0$, while the dashed-dotted unstable curve becomes indistinguishable from the solid curve. In that case, the population should thus evolve toward complete sex ($\sigma = 1$) from any initial value located above the solid curve and toward no sex starting on the right or below the solid curve.

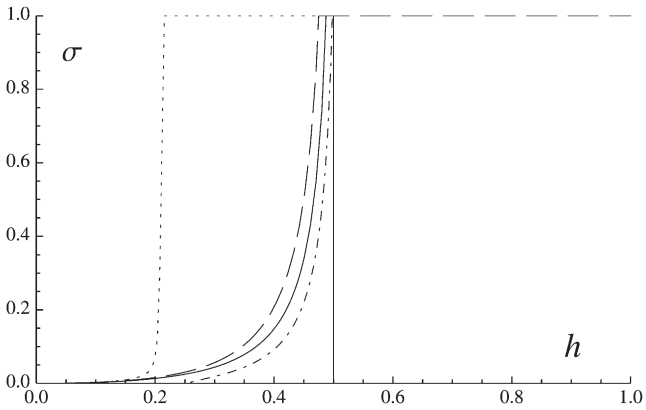



FIGURE 3.—Long-term evolution of the rate of sex (analytical prediction from the additive modifier model), as a function of the dominance of deleterious mutations $h$. Deterministic case ($N \to \infty$), no gene conversion: $\sigma$ converges to the solid curve when $h < \frac{1}{2}$ and to zero when $h > \frac{1}{2}$. $N = 20{,}000$, $\gamma = 0$: $\sigma$ converges to the dashed curve (which is at $\sigma = 1$ for $h > \frac{1}{2}$). $N = 20{,}000$, $\gamma = 10^{-4}$: the dotted curve is a stable equilibrium line, while the dashed-dotted curve is an unstable equilibrium line; $\sigma$ thus converges to a very low value when $h < 0.2$, while it converges to one starting between the dotted and the dashed-dotted curves and to zero starting on the right of the dashed-dotted curve. When $\gamma = 10^{-3}$ or $\gamma = 10^{-2}$, the part of the stable dotted curve at $\sigma = 1$ extends to $h = 0$, while the unstable dashed-dotted curve becomes indistinguishable from the solid curve for $h < \frac{1}{2}$ ($\sigma$ still converges to zero when $h > \frac{1}{2}$). Parameter values are the same as for Figures 1 and 2.

The results discussed so far assume additivity at the modifier locus ($h_M = \frac{1}{2}$). Expressions for arbitrary $h_M$ are given in Appendix SB. Analyzing these expressions shows that in many cases, the direction of selection at the modifier locus depends both on $h_M$ and on allele frequency at the modifier locus, which may lead to stable polymorphic equilibria (results not shown). Such situations (which may involve more than two modifier alleles being present simultaneously in the population) are not dealt with by our simple biallelic model (note that effects of dominance at the modifier locus have been explored by Otto 2003 in the deterministic case). However, in the next section we discuss effects of introducing modifier dominance into our multilocus simulations.

**Three-locus model:** When deleterious mutations segregate at a large number of loci, interactions between selected loci may become an important component of selection for sex (in particular, through the fact that sex allows recombination between loci). To quantify the relative effect of interactions between pairs of selected loci on selection for sex, we extended our model to include a second selected locus (denoted locus B). Two alleles $b$ and $B$ segregate at this locus; we assume that $B$ is deleterious (with the same selection and dominance coefficients as allele $A$) and that selection is multiplicative across loci (no epistasis). In such a model, different genetic associations between loci A and B are generated by drift and selection. In particular, negative linkage disequilibrium $D_{AB}$ between alleles $A$ and $B$ due to the combined action of drift and selection (the Hill–Robertson effect, *e.g.*, Hill and Robertson 1966; Barton and Otto 2005) generates a selective force for increased sex and recombination. Furthermore, genetic drift generates a correlation in homozygosity between loci A and B (measured by association $D_{AB,AB}$). Sex tends to break this correlation (decreasing the frequency of double homozygotes and double heterozygotes), which is disadvantageous whenever $h \neq \frac{1}{2}$: indeed, genotypes homozygous at one selected locus and heterozygous at the other have a lower fitness than the average of double heterozygotes and double homozygotes when $h \neq \frac{1}{2}$ (*e.g.*, Roze 2009). Other types of associations between loci A and B are also generated by selection and drift. Results from the three-locus model (for the case of an additive modifier) are presented in Appendix SD. These results show that although the selective force on sex generated by the interaction between two selected loci is smaller in magnitude than the selective force generated by each locus separately, the overall effect of interactions between loci may become important when the deleterious mutation rate $U$ is high, so that many loci segregate for deleterious alleles: indeed in that case, the number of pairs of segregating loci may by greater than the number of loci by several orders of magnitude.

To leading order in ε, the change in frequency of the modifier at quasi-equilibrium is given by

$$\langle \Delta p_M \rangle_t \approx -s(1-2h)\sum_A \langle D_{MA,A}\rangle_t + s^2(1-2h)^2 \sum_{A,B}\langle D_{MAB,AB}\rangle_t, \quad (11)$$

where the first sum is over all segregating loci, and the second sum is over all pairs of loci. The association $\langle D_{MAB,AB}\rangle$ is negative for a modifier increasing sex, reflecting the fact that sex tends to break correlations in homozygosity between loci A and B (as discussed above). Note that associations $\langle D_{MA,A}\rangle$ are also affected by interactions among pairs of loci, this effect also disfavoring sex whenever $h \neq \frac{1}{2}$ (see Appendix SD). As a result, one obtains that when $h \neq \frac{1}{2}$, sex becomes less favorable when selection becomes weaker and the mutation rate $U$ higher, so that the number of segregating loci becomes larger (see figures in Appendix SD). Note that when $U$ is sufficiently large, it is likely that higher-order associations (such as $D_{MABC,ABC}$, $D_{MABCD,ABCD}, \ldots$) would also become important; however, it seems difficult to obtain general predictions about the effects of such associations.

When dominance is weak, the change in frequency of the modifier is affected by many associations (given in Appendix SD), among which is the linkage disequilibrium $\langle D_{AB}\rangle$ generated by the Hill–Robertson effect. Again, one obtains that the overall effect of these associations increases with $U$ and may widen the parameter range in which sex is favored (in particular as the strength of selection increases and as rates of sex and recombination decrease). Interestingly, linkage disequilibrium $\langle D_{AB}\rangle$ seems to be only a minor component of selection for sex due to interactions between selected loci (see Figure D3 in Appendix SD), as many other associations produced by selection and drift generate selection for sex (see Appendix SD for further details).

## MULTILOCUS SIMULATIONS

**Description of the program:** Our simulation program (written in C++ and available upon request) uses a similar setting as in Roze (2009). The population is made of $N$ diploids, each possessing two copies of one chromosome. Values of $N$ chosen in the simulations (between 10,000 and 50,000) were sufficiently large so that they appear realistic at least for some species, while sufficiently small so that execution time of the program remains reasonable. Each chromosome is represented by a table of real values between 0 and 1, indicating the positions of deleterious mutations present on the chromosome (0 and 1 corresponding to the chromosome ends); the number of loci at which mutation can occur is thus effectively infinite. Each generation, the number of new mutations occurring on a given chromosome is sampled from a Poisson distribution with parameter $U$ (and the position of each mutation is sampled in a uniform distribution). A sex modifier locus is located at the midpoint of the chromosome (in position

0.5); alleles at this locus are represented by real values between 0 and 1. At the start of a generation, each individual becomes either fully sexual (that is, all of its offspring will be produced sexually) or fully asexual, the probability of becoming sexual being given by the average of the values of its two alleles at the modifier locus (in the following we also consider dominance effects between modifier alleles). Note that this scenario is slightly different from our analytical model, where a given individual may reproduce both sexually and asexually. The next generation is produced as follows. For each of the $N$ individuals of the next generation, an individual is sampled randomly among the present generation; if a random value between 0 and 1 is $< w/w_{max}$, where $w$ is the fecundity of the individual and $w_{max}$ the maximal fecundity in the population, the sampled individual is chosen to be the mother of the new individual (otherwise, another individual is sampled, until the test is satisfied). Fecundity is given by

$$w = \delta_c(1 - hs)^{n_{He}}(1 - s)^{n_{Ho}}, \qquad (12)$$

where $n_{He}$ and $n_{Ho}$ are the numbers of deleterious mutations present in the heterozygous and homozygous state in the individual, and where $\delta_c = 1$ in sexuals, while $\delta_c = c \geq 1$ in asexuals. The parameter $c$ allows us to introduce a direct cost of sex: for example, in an oogamous hermaphroditic species, and under random mating, individuals are expected to invest half of their resources into male gametes that do not bring any resource to the next generation. In this case (and all else being equal) an asexual female may thus produce twice as many eggs as an hermaphrodite (in which case $c = 2$). If the mother is sexual, then a second individual is sampled among all sexuals of the present generation using the procedure described above, to serve as a father. Recombinant chromosomes are then produced in both mother and father: the number of crossovers occurring at meiosis is sampled from a Poisson distribution with parameter $L$ (genome map length), the position of each crossover being random. If the mother is asexual, in the absence of gene conversion the genotype of the offspring is exactly the same as the mother's genotype. With gene conversion ($\gamma > 0$), each mutation present in the heterozygous state becomes homozygous for one of the two alleles with probability $\gamma/2$ (gene conversion thus occurs independently at the different loci). We also considered the case of mitotic crossing over (in a separate program): in that case, we assume that every time an individual is produced asexually, a mitotic crossover occurs with probability $\chi$ per chromosome (its position along the chromosome being random). If a crossover occurs, the offspring becomes homozygous for all loci located between the crossover position and the distal part of the chromosome (we assume that the centromere is located at position 0.25), including the modifier locus. Which of the two parental chromosomes is used as a template is random (note that the parameter $\chi$ is thus really the rate of mitotic crossovers that lead to loss of heterozygosity). Gene conversion is not implemented in the program that includes mitotic crossing over (note that the analytical model presented in the previous section does not differentiate between gene conversion and mitotic crossing over, as it represents only a single selected locus).

Each parameter combination is run one time. During the first 2000 generations of a run, all individuals have the same allele at the modifier locus, coding for a rate of sex $\sigma_{init}$. These 2000 generations are generally sufficient to reach mutation–selection–drift balance, except when $\sigma_{init}$ is too low and mutations accumulate (these cases will be discussed). Then, during the next $2 \times 10^6$ generations, mutations occur at rate $10^{-4}$ per generation at the modifier locus. When a mutation occurs, with probability 0.5 the rate of sex coded by the new allele is sampled in a uniform distribution between 0 and 1, while with probability 0.5 it is sampled in a uniform distribution between $\sigma_{old} - 0.1$ and $\sigma_{old} + 0.1$, where $\sigma_{old}$ is rate of sex coded by the parent allele. The average rate of sex in the population (average value of modifier alleles), average fitness, number of mutations per chromosome, and number of fixed mutations are recorded every 100 generations. Because we assume no back mutation, fixed mutations do not contribute to selection on the modifier locus and are removed from the population to increase execution speed. The evolutionarily stable rate of sex is then obtained by averaging over the last $1.9 \times 10^6$ generations (averaging over the last $10^6$ generations yields very similar results). Error bars are computed using Hastings' (1970) batching method, dividing the $1.9 \times 10^6$ generations into 10 batches and calculating the standard error over the 10 averages (these error bars are often of similar size as the symbols used in the figures). In the absence of selection ($U = 0$, $c = 1$), we checked that the average rate of sex at equilibrium is 0.5 (results not shown).

**General results:** Figure 4 shows the effects of the dominance coefficient of deleterious mutations ($h$) and the rate of mitotic gene conversion ($\gamma$) on the average rate of sex at equilibrium ($\sigma$) starting from $\sigma_{init} = 1$, for $U = 0.05$ (left) and $U = 0.5$ (right). In the absence of gene conversion (solid squares in Figure 4), the rate of sex evolves toward zero when deleterious alleles are completely (or nearly completely) recessive ($h \sim 0$–$0.1$); deleterious mutations then accumulate in the heterozygous state, and the simulation is stopped before $2 \times 10^6$ generations, as the program becomes very slow. Higher values of $h$ lead to greater rates of sex, $\sigma$ reaching a plateau when $0.3 \leq h \leq 1$. For $U = 0.05$ (Figure 4, left), the average rate of sex does not depart from the neutral expectation (dashed-dotted line) when $h \geq 0.2$; in this case, the selective pressure for sex is very weak unless $\sigma$ is very small, and dynamics at the modifier locus are

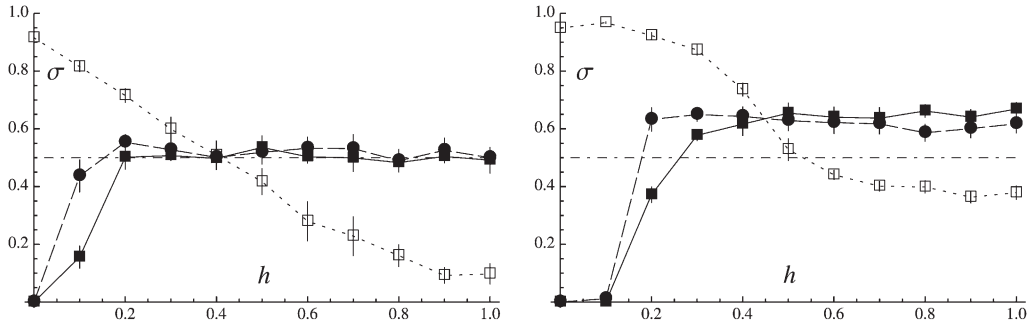



FIGURE 4.—Average rate of sex observed in multilocus simulations, as a function of the dominance coefficient of deleterious mutations ($h$) and for different rates of mitotic gene conversion ($\gamma$): solid squares, solid lines, $\gamma = 0$; solid circles, dashed lines, $\gamma = 10^{-4}$; open squares, dotted lines, $\gamma = 10^{-3}$ (note that lines simply connect simulation results and do not correspond to analytical predictions). Left, $U = 0.05$; right, $U = 0.5$. Other parameter values: $N = 20{,}000$, $s = 0.05$, $L = 10$ (this high value is chosen to mimic multiple chromosomes), $\sigma_{init} = 1$, $c = 1$ (no direct cost of sex). The dashed-dotted line represents the average rate of sex in the absence of selection (direct or indirect) at the modifier locus ($\sigma = 0.5$).

mainly driven by random drift. Because Figure 4 does not inform us about selection for sex when $h \geq 0.2$, we used a modified version of the program to obtain an estimate of this selective force when only two alleles segregate at the modifier locus and obtained results that are compatible with predictions from our two-locus model (see Appendix SE). When $U = 0.5$ (Figure 4, right), the rate of sex evolves toward relatively high values when $h \geq 0.3$, while the two-locus model predicts that high rates of sex should not be favored unless $h \geq \sim 0.45$ (Figure 3, dashed line). This discrepancy does not come from the fact that Figure 3 was assuming free recombination: indeed, modifying the program to have free recombination among loci has only very little effect on the results (not shown). Rather, selection for sex when $0.3 \leq h \leq 0.45$ is probably a consequence of interactions between selected loci: indeed, our three-locus model (Appendix SD) indicates that the effect of such interactions is relatively important when $U = 0.5$ and that it tends to widen the range of values of $h$ for which high rates of sex are favored. Furthermore, we modified our simulation program to eliminate benefits of recombination (by keeping the $\sigma L$ product constant among individuals) and obtained that sex was not favored unless $h > 0.4$ in this modified program (see Figure F2 in Appendix SF). Note that our three-locus model still does not predict that $\sigma$ should be $>0.5$ when $0.3 \leq h \leq 0.4$ (Figure D4 in Appendix SD, with $U = 0.5$, $s = 0.05$); this may be due to the fact that we computed some three-locus associations under the assumption that dominance is weak (see Appendix SD) and the resulting expressions may thus not be precise when $h < 0.4$; alternatively, it may be an effect of interactions among more than two selected loci.

It can be noted that the average rate of sex reaches a plateau at $\sigma \sim 0.6$ when $h \geq 0.5$ (for $U = 0.5$ and $\gamma = 0$), while both the two- and three-locus models predict an evolutionarily stable rate of sex at $\sigma = 1$ in this case. This difference is likely to be due to the fact that the strength of selection for sex decreases very fast as $\sigma$ increases (selection through effect on genotype frequencies decreases as $1/\sigma^2$, while selection through effect on allele frequencies decreases as $1/\sigma^3$). For example, the three-locus model predicts a selection gradient for sex [measured by $s_M = \langle \Delta p_M \rangle / \langle (\delta\sigma/2) p q_M \rangle$] $\sim 5 \times 10^{-5}$ when $h = 0.5$ and $\sigma = 0.5$ (under free recombination and other parameters as in Figure 4, right) and $\sim 3 \times 10^{-6}$ when $\sigma = 1$. Selection between alleles coding for different, high rates of sex is thus extremely weak, and the change in frequencies of these alleles will be dominated by random drift (this is confirmed by the fact that the rate of sex fluctuates widely over the course of a simulation, as shown in Appendix SF). Drift at the modifier locus thus prevents the average rate of sex from reaching 1 (reducing the mutation rate at the modifier locus from $10^{-4}$ to $10^{-5}$ has little effect on the results—not shown).

The qualitative effects of mitotic gene conversion match the predictions of the analytical model: gene conversion tends to increase sex when deleterious mutations are recessive and to decrease sex when mutations are dominant. Figure 4 shows that a rate of gene conversion of $10^{-4}$ has little effect on the results (solid circles), while a rate of $10^{-3}$ has much more of an effect (open squares). Although our analytical model (assuming $\delta\sigma$ small) predicts an unstable equilibrium when $\gamma > 0$ and $h < \frac{1}{2}$ (dashed-dotted curve in Figure 3), it is not clear how this should affect the dynamics in the simulations, since mutation at the modifier locus may generate new alleles coding for any rate of sex. For all points shown in Figure 4 (for which the initial rate of sex was set to $\sigma_{init} = 1$), we performed additional simulations with an initial rate of sex $\sigma_{init} = 0.01$. This led to very similar quantitative results when $U = 0.05$ (corresponding to Figure 4, left) except for $h = 0$ and $\gamma = 10^{-3}$, in which case $\sigma$ goes to zero (not shown). In this last case mutation–selection equilibrium is not reached during the preliminary generations of the simulation, as mutations accumulate in the heterozygous state within nearly clonal lineages, generating a very strong segregation load that prevents sex from increasing. Similar results are obtained when $U = 0.5$ and $\sigma_{init} = 0.01$, except that

mutation accumulation during the preliminary generations occurs for all values of $h$ between 0 and 0.3, preventing sex from increasing. When $h \geq 0.4$, the rate of sex reaches similar average values as when $\sigma_{init} = 1$ (not shown). When $\sigma_{init}$ is set to 0.05, results for $h = 0.3$ become similar to results obtained when $\sigma_{init} = 1$ (while mutations accumulate during preliminary generations when $h \leq 0.2$).

Results presented in Figure 4 (and in the next figures) assume additivity at the modifier locus. Incorporating dominance effects between modifier alleles in a very general way is difficult in the framework of our simulation model, because the dominance relationships between all segregating alleles need to be specified. To have an idea of how dominance at the modifier locus may affect our results, we modified our simulation program so that each modifier allele $i$ codes for a rate of sex $\sigma_i$ and has a "dominance coefficient" $\theta_i$, which is sampled from a uniform distribution between 0 and 1 for each new allele. The probability that an individual reproduces sexually is then given by $(\theta_1\sigma_1 + \theta_2\sigma_2)/(\theta_1 + \theta_2)$, where $\sigma_1$, $\theta_1$ and $\sigma_2$, $\theta_2$ refer to the two modifier alleles of the individual. We ran this modified program for the parameter values corresponding to the solid squares in Figure 4 ($\gamma = 0$, $U = 0.05$, and $U = 0.5$); however, the results obtained from the two programs were almost undistinguishable (results not shown). Still, it is possible that dominance at the modifier locus would have more effect if it was modeled differently (in particular, our modified program still implies some form of average additivity).

Figure 5 shows the effect of the rate of mitotic crossing over $\chi$ on the average rate of sex, when $U = 0.5$. One can see that $\chi = 10^{-3}$ has less effect than $\gamma = 10^{-3}$ (in particular in the case of recessive mutations). Indeed, mitotic crossing over leads to LOH at loci located between the crossover and the distal part of the chromosome only: therefore, the average rate of LOH per locus and per asexual generation is lower for $\chi = 10^{-3}$ than for $\gamma = 10^{-3}$.

**Effects of *N*, *L*, *s*:** Figure 6 shows the effect of varying population size $N$, genome map length $L$, and the strength of selection against deleterious mutations $s$, when $\gamma = \chi = 0$. Changing $N$ from $10^4$ to $5 \times 10^4$ has only little effect on the average rate of sex when $h \geq 0.3$ (Figure 6, top left). This may be due to the fact that increasing $N$ has two opposite effects: it decreases the strength of indirect selection acting on the modifier (that favors high rates of sex when $h \geq 0.3$), but it also decreases the effect of drift at the modifier locus (that tends to bring the rate of sex closer to 0.5), and the two effects may cancel each other. When $h < 0.3$, decreasing $N$ increases the range of values of $h$ for which the population evolves toward asexuality. This effect may be due to the fact that when a modifier allele coding for a low rate of sex reaches high frequency (either by selection or by drift), deleterious mutations

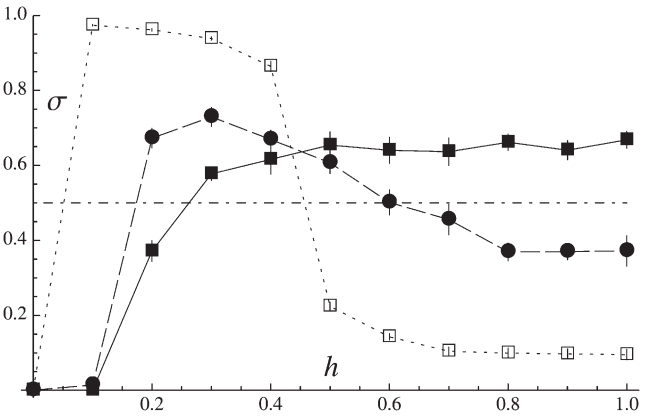



FIGURE 5.—The same as Figure 4 for $U = 0.5$ and for different rates of mitotic crossing over: solid squares, solid lines, $\chi = 0$; solid circles, dashed lines, $\chi = 10^{-3}$; open squares, dotted lines, $\chi = 10^{-2}$.

can accumulate in the heterozygous state (in particular when $h$ is low), which prevents the rate of sex from increasing back (as it generates a strong segregation load). This effect should be stronger for smaller population sizes (because mutation accumulation occurs faster). Effects of map length $L$ are shown in Figure 6, top right. For intermediate values of $h$, decreasing map length $L$ tends to increase the rate of sex. This is predicted by the two-locus model, as selection on the modifier through its effect on allele frequency (term in $D_{MA} + D_{M,A}$) becomes stronger when linkage is tighter, while recombination has less effect on selection through effect on genotype frequency (term in $D_{MA,A}$); note that the three-locus model also predicts more selection for sex with tighter linkage (Appendix SD). When $h$ is low, however, low recombination increases the speed of mutation accumulation, which tends to trap the population in the asexual state (as discussed above): for $h = 0.2$, the population evolves toward asexuality when $L = 0.1$ and $L = 1$, while $L = 10$ allows sex to be maintained (results under free recombination are very similar to $L = 10$, not shown). Note that for $L = 0.1$ and $h \geq 0.2$, deleterious mutations fix in the population, at relatively high rates for high values of $h$ (see additional results in Appendix SF). Finally, Figure 6 shows that increasing $s$ leads to higher rates of sex, which is expected since the relative effect of selection through effect on allele frequencies increases with $s$ (Equations 7–9). Furthermore, low values of $s$ facilitate the accumulation of mutations in the heterozygous state when $h$ is low, which again may trap the population in the asexual state. Finally, Figure 6 shows that the population evolves toward low sex when $s$ is low and $h$ is high (dotted line in Figure 6, bottom). This effect is not expected from the two-locus model, but can be explained in terms of interactions between pairs of selected loci (three-locus model): indeed, we have seen that sex tends to break correlations in homo-

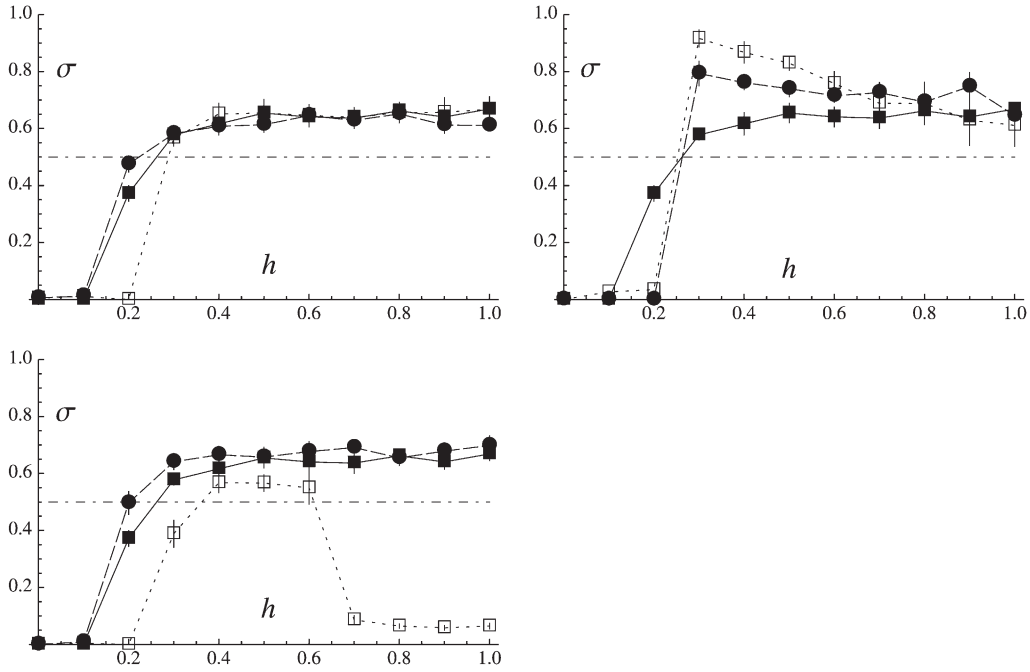



FIGURE 6.—Average rate of sex observed in multilocus simulations, as a function of the dominance coefficient of deleterious mutations ($h$), in the absence of gene conversion or mitotic recombination ($\gamma = \chi = 0$). Top left, different values of population size $N$: dotted line, $N = 10{,}000$; solid line, $N = 20{,}000$; dashed line, $N = 50{,}000$. Top right, different values of genome map length $L$: solid line, $L = 10$; dashed line, $L = 1$; dotted line, $L = 0.1$. Bottom: different values of the strength of selection against deleterious alleles $s$: dotted line, $s = 0.01$; solid line, $s = 0.05$; dashed line, $s = 0.1$. Other parameter values are $U = 0.5$, $c = 1$, $N = 20{,}000$ (top right and bottom), $L = 10$ (top left and bottom), and $s = 0.05$ (top). Note that when $s = 0.01$ (bottom, dotted line), deleterious mutations fix in the population at a high rate for $h = 0.7$–1 (see Appendix SF).

zygosity among selected loci (which disfavors sex whenever $h \neq \frac{1}{2}$) and that this effect should predominate over other effects of interactions between loci when $s$ is sufficiently small (see Appendix SD).

**Epistasis:** Interactions between selected loci may also have important effects on selection for sex in the presence of epistasis. Appendix SG shows simulation results incorporating epistasis, assuming epistasis is the same among all pairs of loci and neglecting higher-order epistasis (*i.e.*, epistatic effects between more than two loci). These results indicate that negative additive-by-additive epistasis favors sex (even when large in absolute value), while negative additive-by-dominance and dominance-by-dominance epistasis tends to disfavor sex. These results are discussed more at length in Appendix SG.

**Costly sex:** The previous results assumed that sex had no intrinsic cost. Figure 7 presents the average rate of sex as a function of the cost of sex $c$ (recall that $c = 1$ means no cost, while a twofold cost corresponds to $c = 2$), for a genomic mutation rate $U = 1$. In Figure 7, left, mutant alleles at the modifier locus may code for any rate of sex between 0 and 1 (as in previous results). In the case of additive deleterious alleles ($h = 0.5$), $\sigma$ decreases sharply as $c$ becomes >1, but then decreases slowly as $c$ increases to 2. When the average rate of sex is <0.2, however (that is, when $c \geq \sim 1.2$), deleterious mutations accumulate in the population: the numbers of fixed mutations after $2 \times 10^6$ generations are ~130, 2700, 9000, and 57,000 for $c = 1.2$, 1.3, 1.5, and 2, respectively (Appendix SF provides more detailed results), and mutational meltdown may thus lead to the

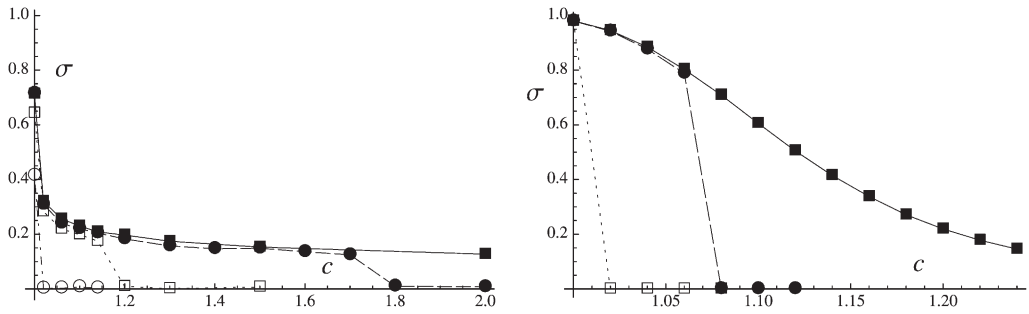



FIGURE 7.—Average rate of sex observed in multilocus simulations, as a function of the cost of sex ($c$) and for different values of the dominance coefficient of deleterious mutations ($h$): solid squares, solid lines, $h = 0.5$; solid circles, dashed lines, $h = 0.4$; open squares, dotted lines, $h = 0.3$; open circles, dashed-dotted line, $h = 0.2$ (left panel only). Left, mutant alleles at the modifier locus can take any value between 0 and 1; right, competition between fully sexuals ($\sigma = 1$) and fully asexuals ($\sigma = 0$). Other parameter values: $U = 1$, $\gamma = \chi = 0$, and other parameters as in Figures 4 and 5. Error bars are smaller than the size of symbols.

extinction of the population (changing population size from $N = 20{,}000$ to $N = 50{,}000$ leads to similar results, as shown in Appendix SF). When deleterious alleles are recessive ($h < 0.5$), $\sigma$ is similar to the value obtained for $h = 0.5$ when $c$ is not too high, but for high costs of sex the population becomes fully asexual. As discussed previously, this is probably due to the fact that when a mutation coding for no or very little sex happens to reach a high frequency (for example, because it occurred on a good genetic background), mutations start accumulating in the heterozygous state, leading to an irreversible process (sex cannot increase again because the segregation load is too strong). Note that it is possible that asexual mutants would eventually invade at lower values of $c$, if the simulation could run for a larger number of generations. Figure 7, right, shows the average frequency of fully sexual individuals ($\sigma = 1$), when fully asexuals ($\sigma = 0$) occur at rate $10^{-4}$ per generation (back mutations also occurring at rate $10^{-4}$). Patterns are similar, except that when $h < 0.5$ asexual mutants invade at lower values of $c$ (which may simply be due to the fact that asexual mutants appear more rapidly in the population than in the previous case). Additional results for $U = 0.5$, $\gamma = 0$, and $\gamma = 10^{-3}$ are presented in Appendix SF, showing that mitotic gene conversion has little effect on these results. Introducing negative additive-by-additive epistasis increases the mean rate of sex for low to intermediate values of $c$ (see Appendix SF); it also prevents the accumulation of deleterious alleles, allowing sex to be maintained at intermediate levels when mutations are recessive and the cost of sex is high (Figure F3, bottom right, in Appendix SF).

## DISCUSSION

**Two-locus model:** Deleterious mutations have been repeatedly presented as a potentially important factor favoring sex and recombination, either due to deterministic interactions among mutations (synergistic epistasis, *e.g.*, Kondrashov 1988) or due to stochastic effects (*e.g.*, Keightley and Otto 2006); however, many of these earlier models focused on the case of haploid organisms and did not investigate what should be the evolutionarily stable rate of sex in the population (in particular when sex is costly). In this article, we have explored the effects of deleterious mutations on the direction and strength of selection acting on a modifier gene affecting the rate of sexual (*vs.* asexual) reproduction, in a finite diploid population. Some of these effects stem from the fact that sex tends to bring the population closer to HW equilibrium at each selected locus (segregation). Otto (2003) has shown that in an infinite, randomly mating population, dominance generates departures from HW equilibrium that may favor sex when deleterious mutations are weakly recessive. Finite population size also generates departures from HW equilibrium (excess of heterozygotes), and we have shown that this effect favors sex when deleterious mutations are dominant or weakly recessive. It is important to stress that this stochastic effect is not the intralocus equivalent of the Hill–Robertson effect that generates negative linkage disequilibrium between loci, as the Hill–Robertson effect requires the interaction between selection and drift and tends to be smaller in magnitude than departures from HW equilibrium generated by drift alone. Because the effect of finite population size is proportional to the genetic variance at each selected locus, it may dominate over the deterministic departure from HW equilibrium generated by selection, which is proportional to the square of genetic variance. In particular, our expression for $\langle D_{A,A} \rangle$ at quasi-equilibrium (Equation 2) shows that the stochastic effect of finite population size should dominate over the deterministic effect of selection whenever $1/(2N) \gg s(1-2h)\langle p_A^2 q_A^2 \rangle / \langle p_A q_A \rangle$. In the deterministic limit and when $h > 0$, $\langle p_A^2 q_A^2 \rangle / \langle p_A q_A \rangle \approx u/(hs)$, and the last condition thus becomes independent of $s$. With finite $N$, numerical integration over Wright's distribution shows that $s(1-2h)\langle p_A^2 q_A^2 \rangle / \langle p_A q_A \rangle$ is also unaffected by $s$ over a wide range of values of $s$ (results not shown): therefore, increasing $s$ will not increase the relative effect of the deterministic term over the stochastic term. We measured deleterious allele frequencies at segregating loci in our multilocus simulations for $N = 20{,}000$, $U = 0.5$, $s = 0.05$, and $h = 0.2$, 0.3, and 0.4 and found that $\langle p_A \rangle \approx 3.6 \times 10^{-4}$, $2.7 \times 10^{-4}$, and $2.2 \times 10^{-4}$, while $\langle p_A^2 q_A^2 \rangle / \langle p_A q_A \rangle \approx 1.4 \times 10^{-3}$, $9.3 \times 10^{-4}$, and $7.2 \times 10^{-4}$ for $h = 0.2$, 0.3, and 0.4 (respectively). From this (and from Equation 2), one finds that the deterministic and stochastic terms are of similar orders of magnitude (for these parameter values), the deterministic term being slightly stronger than the stochastic term for $h = 0.2$, while it is weaker for $h = 0.3$ and 0.4. Note, however, that other associations than $D_{A,A}$ generate selection for sex in finite populations (such as the variance in $D_{MA}$, see Equation B2 in Appendix SB).

**Heterozygote excess:** This work raises the question of the occurrence of heterozygote excess within populations. Equation 2 indicates that $D_{A,A}$ should be only very slightly negative in fully sexual, randomly mating populations, while it may by more strongly negative when sex becomes rare. Negative $F_{IS}$ values (indicating local excess of heterozygotes) have indeed been measured in some partly or fully clonal populations (*e.g.*, Prugnolle *et al.* 2005; Guillemin *et al.* 2008). Many other studies have reported positive $F_{IS}$ values, often interpreted as "inbreeding". However, positive $F_{IS}$ may be generated by different processes, with different consequences of the evolution of sex modifiers. Self-fertilization (or mating within families before dispersal) will generate excess of homozygotes, favoring sex when deleterious mutations are recessive (Otto 2003; Agrawal 2009) because sex tends to mask these mutations. Positive $F_{IS}$ may also result from population substructure into smaller panmictic units (the Wahlund effect). In this last case,

heterozygote excess may occur at the level of these units (in particular in facultative sexuals), generating a selective force on sex similar to the one described in this article (assuming that competition occurs at least partly at the scale of the panmictic units). The effects of population structure on selection for sex in diploids will be explored further in a future article.

**Interactions between selected loci:** Indirect selection on a sex modifier locus is also affected by associations involving the modifier and several selected loci. In particular, finite population size tends to generate correlations in homozygosity that are broken down by sex: in the case of two selected loci, sex tends to decrease the frequency of double homozygotes (hom–hom) and double heterozygotes (het–het) at these loci and to increase the frequency of homozygotes at one locus and heterozygotes at the other (hom–het). In the absence of epistasis, the average fitness of hom–het genotypes is lower than the average fitness of hom–hom and het–het genotypes whenever $h \neq 0.5$ (Figure 4 in Roze 2009), selecting against sex through the term $s^2(1-2h)^2\langle D_{\mathrm{MAB,AB}}\rangle_t$ in Equation 11. In particular in the case of partially recessive deleterious alleles, finite population size tends to generate associative overdominance between chromosomes carrying deleterious mutations at different loci. In the same way as overdominance at a single locus selects against sex (due to segregation load), associative overdominance also disfavors sex because sex generates lower fitness hom–het mixtures from parents heterozygous at many loci (we thank Sally Otto for pointing out this analogy). The effect of such associations involving two or more selected loci may become important when deleterious alleles segregate at many loci (high $U$). It becomes also particularly important when sex occurs very rarely and deleterious mutations accumulate in the population. As was found before (Charlesworth *et al.* 1993a,b; Charlesworth and Charlesworth 1997), we observed in our simulations that when $h$ is sufficiently low, deleterious mutations accumulate in the population without reaching fixation: different types of strongly mutated chromosomes are then maintained in the population by associative overdominance, which strongly selects against sex. This generates an irreversible process, for when an allele coding for no sex (or very little sex) reaches high frequency, the population starts to accumulate recessive deleterious alleles in the heterozygous state, preventing sex from increasing back (this accumulation occurs faster when $s$ and $h$ and $N$ and $L$ are low and when $U$ is high). This is particularly problematic when sex has a strong direct cost, in which case a mutation coding for low sex may quickly reach high frequency if it occurs on a genetic background carrying only few mutations. Indeed, our multilocus simulations show that when $h < 0.4$, the rate of sex always goes to zero when the cost of sex is sufficiently high. Estimates of average dominance coefficients of deleterious alleles vary widely, ranging from 0.1 to 0.4 in mutation-accumulation studies (*e.g.*, Houle *et al.* 1997; García-Dorado and Caballero 2000; Vassilieva *et al.* 2000; García-Dorado *et al.* 2004; Halligan and Keightley 2009), while fitness assays of spontaneous mutations in yeast indicate an average $h$ of ~0.2 (Szafraniec *et al.* 2003). Taken together, these results suggest that $h$ is likely to be <0.4 on average.

Other types of interactions among selected loci favor sex, however, and may be stronger than the effect of breaking correlations in homozygosity, in particular when $h > \sim 0.2$–$0.3$, selection is not too weak, and/or rates of sex and recombination are small. Among these interactions are negative linkage disequilibria generated by the Hill–Robertson effect (Hill and Robertson 1966; Barton and Otto 2005; Keightley and Otto 2006), favoring increased recombination. However, our quasi-equilibrium results indicate that selection for sex is also driven by many other associations between pairs of selected loci (also generated by selection and drift), the linkage disequilibrium $\langle D_{\mathrm{AB}}\rangle$ being only a minor component of selection for sex (see Figures D2 and D3 in Appendix SD). Similar results were obtained in a previous model on the evolution of recombination modifiers in diploid, spatially structured populations (Roze 2009), the major difference being that associations $\langle D_{\mathrm{MA,A}}\rangle$, $\langle D_{\mathrm{MB,B}}\rangle$ are much stronger and play a more important role in the evolution of sex modifiers than in the evolution of recombination modifiers.

Previous multilocus simulation models on the evolution of sex in finite, haploid populations found that selection for sex (measured by the probability of fixation of a modifier increasing sex or recombination relative to the fixation probability of a neutral mutation) increases with population size (Keightley and Otto 2006; Gordo and Campos 2008). Here, we found that increasing population size may increase the average rate of sex when deleterious mutations are partially recessive (because increasing population size slows the buildup of associative overdominance that disfavors sex), but not when mutations are additive (compare, for example, Figure 7, left, and Figure F3, bottom left, in Appendix SF). Hence, increasing population size may have different effects on relative fixation probabilities of sex modifiers and on the evolutionarily stable rate of sex (note that the diffusion approximation for the relative fixation probability is $2s_{\mathrm{M}}N_{\mathrm{e}}$ for $s_{\mathrm{M}} \gg 1/N_{\mathrm{e}}$, which may increase as $N$ increases, even if $s_{\mathrm{M}}$ decreases).

**Effects of loss of heterozygosity:** We have also explored the effects of loss of heterozygosity due to mitotic gene conversion and/or mitotic recombination and found that it tends to favor sex when deleterious mutations are recessive (because sex tends to mask mutations that have been made homozygous by mitotic gene conversion/recombination). However, our results indicate that without a cost of sex ($c = 1$), this effect is important only for rates of mitotic gene conversion of

the order $\geq 10^{-3}$ and rates of mitotic crossing over of the order $\geq 10^{-2}$ (Figures 4 and 5). These values are higher than most available estimates on rates of loss of heterozygosity (Tischfield 1997; Shao *et al.* 1999; Holt *et al.* 1999; Omilian *et al.* 2006; Mandegar and Otto 2007; Carr and Gottschling 2008), suggesting that the consequence of such events on the evolution of sex in the presence of deleterious mutations may be limited. Furthermore, we found that mitotic gene conversion does not affect much our results when sex is costly (see Appendix SF). Nevertheless, it would be interesting to obtain more data about rates of mitotic recombination, as only few estimates are available (in particular one could imagine that mitotic recombination could be more frequent in more stressful environments, if DNA damage occurs more frequently). Finally, effects of LOH on selection for sex are very similar to the effects of self-fertilization (described in Otto 2003; Agrawal 2009) and should become negligible in partly selfing populations (as rates of LOH are typically much lower than selfing rates).

**Conclusion:** Taken together, our results cast doubt on the hypothesis that deleterious mutations could allow the maintenance of high rates of sex in the face of strong costs (in diploids). Although the idea that the twofold cost of sex can be paid if $U \geq 1$ is often encountered in the literature, it is not always remembered that this result assumes synergistic epistasis among mutations (Kondrashov 1988), while the data available do not show any clear trend toward synergism among mutations (*e.g.*, Elena and Lenski 1997; Rice 2002). Our simulation results for $U = 1$ and $c = 2$ indicate that a low rate of sex can be maintained when deleterious alleles are additive, while sex is not maintained when deleterious alleles are partially recessive (Figure 7, Appendix SD). Note, however, that deleterious mutations may select for higher rates of sex when they also affect the mating success of males, in species in which sexual selection is sufficiently strong (Hadany and Beker 2007). Furthermore, deleterious mutations can act in combination with other factors such as host–parasite interactions (Peters and Lively 1999; Otto and Nuismer 2004; Gandon and Otto 2007; Salathé *et al.* 2008), local adaptation (Pylkov *et al.* 1998; Lenormand and Otto 2000; Agrawal 2009), or adaptive evolution (Otto and Barton 2001). Testing to which extent a combination of these factors can favor high rates of costly sex would thus be interesting.

Finally, data about the occurrence and fitness effects of mutations leading to asexuality are badly needed: indeed, it may well be that in many cases, asexuals do not benefit from the theoretical twofold advantage over sexuals, due to pleiotropic effects of the mutation leading to loss of sex. In mammals, for example, asexual females may not be able to produce viable offspring due to the genomic imprinting system, and similar constraints may operate in other groups. In the species in which purely asexual forms coexist with sexuals, determining the relative fitnesses of sexuals and asexuals (and disentangling direct effects from effects of the genetic background) represents a promising way of evaluating current theories on the evolution of sex.

We thank Aneil Agrawal, Sally Otto, François Rousset, and two anonymous reviewers for very helpful comments and the genomics and computing service of Roscoff's Biological Station for computing time. This work has been supported by an Accueil de Compétences en Bretagne fellowship from Région Bretagne (project CycleVie) and by a research mobility grant from the Centre National de la Recherche Scientifique to D.R.

## LITERATURE CITED

Agrawal, A. F., 2006 Evolution of sex: Why do organisms shuffle their genotypes? Curr. Biol. **16:** R696–R704.

Agrawal, A. F., 2009 Spatial heterogeneity and the evolution of sex in diploids. Am. Nat. **174:** S54–S70.

Agrawal, A. F., and S. P. Otto, 2006 Host-parasite coevolution and selection on sex through the effects of segregation. Am. Nat. **168:** 617–629.

Aylon, Y., and M. Kupiec, 2004 DSB repair: the yeast paradigm. DNA Repair **3:** 797–815.

Balloux, F., L. Lehmann and T. de Meuûs, 2003 The population genetics of clonal and partially clonal diploids. Genetics **164:** 1635–1644.

Barton, N. H., 1995 A general model for the evolution of recombination. Genet. Res. **65:** 123–144.

Barton, N. H., and B. Charlesworth, 1998 Why sex and recombination? Science **281:** 1986–1990.

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

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

Bernstein, H., H. C. Byerly, F. A. Hopf and R. E. Michod, 1985 Genetic damage, mutation, and the evolution of sex. Science **229:** 1277–1281.

Bernstein, H., F. A. Hopf and R. E. Michod, 1988 Is meiotic recombination an adaptation for repairing DNA, producing genetic variation, or both? pp. 139–160 in *The Evolution of Sex: An Examination of Current Ideas*, edited by R. E. Michod and B. R. Levin. Sinauer Associates, Sunderland, MA.

Bürger, R., 2000 *The Mathematical Theory of Selection, Recombination, and Mutation.* Wiley, Chichester, UK.

Carr, L. L., and D. E. Gottschling, 2008 Does age influence loss of heterozygosity? Exp. Gerontol. **43:** 123–129.

Chamnanpunt, J., W. Shan and B. M. Tyler, 2001 High frequency mitotic gene conversion in genetic hybrids of the oomycete *Phytophthora sojae.* Proc. Natl. Acad. Sci. USA **98:** 14530–14535.

Charlesworth, B., 1990 Mutation-selection balance and the evolutionary advantage of sex and recombination. Genet. Res. **55:** 199–221.

Charlesworth, B., 1993 Directional selection and the evolution of sex and recombination. Genet. Res. **61:** 205–224.

Charlesworth, B., and D. Charlesworth, 1997 Rapid fixation of deleterious alleles can be caused by Muller's ratchet. Genet. Res. **70:** 63–73.

Charlesworth, D., M. T. Morgan and B. Charlesworth, 1993a Mutation accumulation in finite outbreeding and inbreeding populations. Genet. Res. **61:** 39–56.

Charlesworth, D., M. T. Morgan and B. Charlesworth, 1993b Mutation accumulation in finite populations. J. Hered. **84:** 321–325.

Chen, J.-M., D. N. Cooper, N. Chuzhanova, C. Férec and G. P. Patrinos, 2007 Gene conversion: mechanisms, evolution and human disease. Nat. Rev. Genet. **8:** 762–775.

Elena, S. F., and R. E. Lenski, 1997 Test of synergistic interactions among deleterious mutations in bacteria. Nature **390:** 395–397.

Felsenstein, J., and S. Yokohama, 1976 The evolutionary advantage of recombination. II. Individual selection for recombination. Genetics **83:** 845–859.

Gandon, S., and S. P. Otto, 2007 The evolution of sex and recombination in response to abiotic or coevolutionary fluctuations in epistasis. Genetics **175:** 1835–1863.

García-Dorado, A., and A. Caballero, 2000 On the average coefficient of dominance of deleterious spontaneous mutations. Genetics **155:** 1991–2001.

García-Dorado, A., C. López-Fanjul and A. Caballero, 2004 Rates and effects of deleterious mutations and their evolutionary consequences, pp. 20–32 in *Evolution of Molecules and Ecosystems*, edited by A. Moya and E. Font. Oxford University Press, Oxford.

Gordo, I., and P. R. A. Campos, 2008 Sex and deleterious mutations. Genetics **179:** 621–626.

Guillemin, M.-L., S. Faugeron, C. Destombe, F. Viard, J. A. Correa *et al.*, 2008 Genetic variation in wild and cultivated populations of the haploid-diploid red alga *Gracilaria chilensis*: how farming practices favor asexual reproduction and heterozygosity. Evolution **62:** 1500–1519.

Gupta, P. K., A. Sahota, S. A. Boyadjiev, S. Bye, S. Shao *et al.*, 1997 High frequency *in vivo* loss of heterozygosity is primarily a consequence of mitotic recombination. Cancer Res. **57:** 1188–1193.

Haag, C. R., and D. Roze, 2007 Genetic load in sexual and asexual diploids: segregation, dominance and genetic drift. Genetics **176:** 1663–1678.

Hadany, L., and T. Beker, 2007 Sexual selection and the evolution of obligatory sex. BMC Evol. Biol. **7:** 245.

Hagstrom, S. A., and T. P. Dryja, 1999 Mitotic recombination map of 13cen-13q14 derived from an investigation of loss of heterozygosity in retinoblastomas. Proc. Natl. Acad. Sci. USA **96:** 2952–2957.

Halligan, D. L., and P. D. Keightley, 2009 Spontaneous mutation accumulation studies in evolutionary genetics. Ann. Rev. Ecol. Evol. Syst. **40:** 151–172.

Hastings, W. K., 1970 Monte Carlo sampling methods using Markov chains and their applications. Biometrika **57:** 97–109.

Helleday, T., 2003 Pathways for mitotic homologous recombination in mammalian cells. Mutat. Res. **532:** 103–115.

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

Holt, D., M. Dreimanis, M. Pfeiffer, F. Firgaira, A. Morley *et al.*, 1999 Interindividual variation in mitotic recombination. Am. J. Hum. Genet. **65:** 1423–1427.

Houle, D., K. A. Hughes, S. Assimacopoulos and B. Charlesworth, 1997 The effect of spontaneous mutation on quantitative traits. II. Dominance of mutations with effects on life-history traits. Genet. Res. **70:** 27–34.

Iles, M. M., K. Walters and C. Cannings, 2003 Recombination can evolve in large finite populations given selection on sufficient loci. Genetics **165:** 2249–2258.

Ira, G., D. Satory and J. E. Haber, 2006 Conservative inheritance of newly synthesized DNA in double-strand-break-induced gene conversion. Mol. Cell. Biol. **26:** 9424–9429.

Johnson, R. D., and M. Jasin, 2001 Double-strand-break-induced homologous recombination in mammalian cells. Biochem. Soc. Trans. **29:** 196–201.

Judd, S. R., and T. D. Petes, 1988 Physical lengths of meiotic and mitotic gene conversion tracts in *Saccharomyces cerevisiae*. Genetics **118:** 401–410.

Keeney, S., and M. J. Neale, 2006 Initiation of meiotic recombination by formation of DNA double-strand breaks: mechanisms and regulation. Biochem. Soc. Trans. **34:** 523–525.

Keeney, S., C. N. Giroux and N. Kleckner, 1997 Meiosis-specific double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell **88:** 375–384.

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

Kirkpatrick, M., and C. D. Jenkins, 1989 Genetic segregation and the maintenance of sexual reproduction. Nature **339:** 300–301.

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

Kondrashov, A. S., 1988 Deleterious mutations and the evolution of sexual reproduction. Nature **336:** 435–440.

Lenormand, T., and S. P. Otto, 2000 The evolution of recombination in a heterogeneous environment. Genetics **156:** 423–438.

Lettier, G., Q. Feng, A. Antùnez de Mayolo, N. Erdeniz, R. J. D. Reid *et al.*, 2006 The role of DNA double-strand breaks in spontaneous homologous recombination in *S. cerevisiae*. PLoS Genet. **2:** 1773–1786.

Liang, F., M. Han, P. J. Romanienko and M. Jasin, 1998 Homology-directed repair is a major double-strand break repair pathway in mammalian cells. Proc. Natl. Acad. Sci. USA **95:** 5172–5177.

Mandegar, M. A., and S. P. Otto, 2007 Mitotic recombination counteracts the benefits of genetic segregation. Proc. R. Soc. Lond. B Biol. Sci. **274:** 1301–1307.

Martin, G., S. P. Otto and T. Lenormand, 2006 Selection for recombination in structured populations. Genetics **172:** 593–609.

Maynard Smith, J., 1988 The evolution of recombination, pp. 106–125 in *The Evolution of Sex: An Examination of Current Ideas*, edited by R. E. Michod and B. R. Levin. Sinauer Associates, Sunderland, MA.

Nagylaki, T., 1993 The evolution of multilocus systems under weak selection. Genetics **134:** 627–647.

Omilian, A. R., M. E. A. Cristescu, J. L. Dudycha and M. Lynch, 2006 Ameiotic recombination in asexual lineages of *Daphnia*. Proc. Natl. Acad. Sci. USA **103:** 18638–18643.

Otto, S. P., 2003 The advantages of segregation and the evolution of sex. Genetics **164:** 1099–1118.

Otto, S. P., 2009 The evolutionary enigma of sex. Am. Nat. **174:** S1–S14.

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

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

Otto, S. P., and M. W. Feldman, 1997 Deleterious mutations, variable epistatic interactions, and the evolution of recombination. Theor. Popul. Biol. **51:** 134–147.

Otto, S. P., and S. L. Nuismer, 2004 Species interactions and the evolution of sex. Science **304:** 1018–1020.

Pâques, F., and J. E. Haber, 1999 Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. Microbiol. Mol. Biol. Rev. **63:** 349–404.

Peters, A. D., and C. M. Lively, 1999 The red queen and fluctuating epistasis: a population genetic analysis of antagonistic coevolution. Am. Nat. **154:** 393–405.

Prado, F., F. Cortés-Ledesma, P. Huertas and A. Aguilera, 2003 Mitotic recombination in *Saccharomyces cerevisiae*. Curr. Genet. **42:** 185–198.

Prugnolle, F., D. Roze, A. Théron and T. de Meeûs, 2005 *F*-statistics under alternation of sexual and asexual reproduction: a model and data from schistosomes (Platyhelminth parasites). Mol. Ecol. **14:** 1355–1365.

Pylkov, K. V., L. A. Zhivotovsky and M. W. Feldman, 1998 Migration versus mutation in the evolution of recombination under multilocus selection. Genet. Res. **71:** 247–256.

Rice, W. R., 2002 Experimental tests of the adaptive significance of sexual recombination. Nat. Rev. Genet. **3:** 241–251.

Richardson, C., M. E. Moynahan and M. Jasin, 1998 Double-strand break repair by interchromosomal recombination: suppression of chromosomal translocations. Genes Dev. **12:** 3831–3842.

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

Roze, D., and N. H. Barton, 2006 The Hill-Robertson effect and the evolution of recombination. Genetics **173:** 1793–1811.

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

Salathé, M., R. Salathé, P. Schmid-Hempel and S. Bonhoeffer, 2006 Mutation accumulation in space and the maintenance of sexual reproduction. Ecol. Lett. **9:** 941–946.

Salathé, M., R. D. Kouyos, R. R. Regoes and S. Bonhoeffer, 2008 Rapid parasite adaptation drives selection for high recombination rates. Evolution **62:** 295–300.

Shao, C., L. Deng, O. Henegaru, L. Liang, N. Raikwar *et al.*, 1999 Mitotic recombination produces the majority of recessive

fibroblast variants in heterozygous mice. Proc. Natl. Acad. Sci. USA **96:** 9230–9235.

Sieber, O. M., K. Heinimann, P. Gorman, H. Lamlum, M. Crabtree *et al.*, 2002 Analysis of chromosomal instability in human colorectal adenomas with two mutational hits at APC. Proc. Natl. Acad. Sci. USA **99:** 16910–16915.

Szafraniec, K., D. M. Wloch, P. Sliwa, R. H. Borts and R. Korona, 2003 Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous loci of the yeast *Saccharomyces cerevisiae*. Genet. Res. **82:** 19–31.

Tischfield, J. A., 1997 Loss of heterozygosity or: how I learned to stop worrying and love mitotic recombination. Am. J. Hum. Genet. **61:** 995–999.

Uyenoyama, M. K., and B. O. Bengtsson, 1989 On the origin of meiotic reproduction: a genetic modifier model. Genetics **123:** 873–885.

Vassilieva, L. L., A. M. Hook and M. Lynch, 2000 The fitness effect of spontaneous mutations in *Caenorhabditis elegans*. Evolution **54:** 1234–1246.

Virgin, J. B., J. P. Bailey, F. Hasta, J. Neville, A. Cole *et al.*, 2001 Crossing over is rarely associated with intragenic mitotic recombination in Schizosaccharomyces pombe. Genetics **157:** 63–77.

Communicating editor: M. K. Uyenoyama

GENETICS

Supporting Information
http://www.genetics.org/cgi/content/full/genetics.109.108258/DC1

# Deleterious Mutations and Selection for Sex in Finite Diploid Populations

Denis Roze and Richard E. Michod


DOI: 10.1534/genetics.109.108258

# File S1: Online Appendices

## ONLINE APPENDIX A: GETTING THE RECURSIONS

Recursions for genetic associations in a finite, diploid population are obtained as follows (all recursions have been implemented in *Mathematica* notebooks that are available on request). Throughout this appendix, set letters (such as $\mathbb{U}$, $\mathbb{V}$, $\mathbb{S}$, $\mathbb{T}$...) stand for sets of loci. In the context of our two-locus model these sets may be the empty set $\emptyset$, M, A, or MA; however the notation is more general and applies to an arbitrary number of loci.

**Variables.** The population is described in terms of allele frequencies and genetic associations (Barton and Turelli, 1991; Kirkpatrick *et al.*, 2002). The frequency of the capital-letter allele ($M$ or $A$) at locus $X$ (where $X$ may be M or A) in the two haplotypes of the diploid individual $j$ are denoted $p_{X(j1)}$ and $p_{X(j2)}$ (these variables equal 0 or 1). The average frequency of the capital-letter allele at locus $X$ in the whole population is denoted $p_X$. Centered variables $\zeta$ are defined as:

$$\zeta_{X(j1)} = p_{X(j1)} - \wp_X, \qquad \zeta_{X(j2)} = p_{X(j2)} - \wp_X \tag{A1}$$

where $\wp_X$ is called a "reference value" (Kirkpatrick *et al.*, 2002). Unless otherwise specified, this reference value will be $p_X$ (the frequency of the capital-letter allele at locus $X$ in the population). The association between the sets $\mathbb{S}$ and $\mathbb{T}$ of loci present in the two haplotypes of the same individual is defined as:

$$D_{\mathbb{S},\mathbb{T}} = \mathrm{E}\left[\zeta_{\mathbb{S},\mathbb{T}(j)}\right] \tag{A2}$$

where E stands for the average over the whole population (average over all $j$), and where

$$\zeta_{\mathbb{S},\mathbb{T}(j)} = \frac{\zeta_{\mathbb{S}(j1)}\,\zeta_{\mathbb{T}(j2)} + \zeta_{\mathbb{S}(j2)}\,\zeta_{\mathbb{T}(j1)}}{2}, \qquad \zeta_{\mathbb{S}(j1)} = \prod_{X\in\mathbb{S}} \zeta_{X(j1)}, \qquad \zeta_{\mathbb{T}(j2)} = \prod_{X\in\mathbb{T}} \zeta_{X(j2)}\,. \tag{A3}$$

(note that $D_{\mathbb{S},\mathbb{T}} = D_{\mathbb{T},\mathbb{S}}$). Associations between genes present on the same haplotype of an individual ($D_{\mathbb{S},\emptyset}$) will be simply denoted $D_{\mathbb{S}}$.

In a finite population, allele frequencies and genetic associations are random variables. The expected value of the random variable $x$ at generation $t$ will be denoted $\langle x\rangle_t$; for example, $\langle D_{\mathrm{MA,A}}\rangle_t$ is the expectation of $D_{\mathrm{MA,A}}$ at generation $t$. We will also need to express moments of associations, such as $\langle {D_{\mathrm{MA}}}^2\rangle_t$ or $\langle D_{\mathrm{MA}} D_{\mathrm{M,A}}\rangle_t$. In the following, it will be useful to express these moments under the form of associations between genes present in two (or more) individuals sampled with replacement from the population. For example, $D_{\mathbb{S},\mathbb{T}} D_{\mathbb{U},\mathbb{V}}$ can be considered as the association between genes in the sets $\mathbb{S}$ and $\mathbb{T}$ from one individual, and genes in the sets $\mathbb{U}$ and $\mathbb{V}$ from another individual, sampled with replacement from the population. This association will be denoted $D_{\mathbb{S},\mathbb{T}\widehat{/}\mathbb{U},\mathbb{V}}$, where the $\widehat{/}$ symbol separates sets of genes from the two individuals. We thus have:

$$D_{\mathbb{S},\mathbb{T}\widehat{/}\mathbb{U},\mathbb{V}} = \mathrm{E}\left[\zeta_{\mathbb{S},\mathbb{T}(j)}\,\zeta_{\mathbb{U},\mathbb{V}(k)}\right] \tag{A4}$$

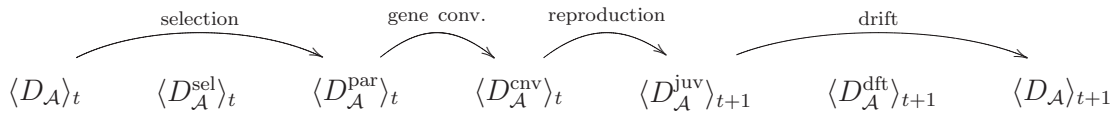
where E stands for the average over all possible pairs of individuals $j$ and $k$ (including $j = k$). Associations between genes from more than two individuals sampled with replacement (such as $D_{\mathbb{S},\mathbb{T}\widehat{/}\mathbb{U},\mathbb{V}\widehat{/}\mathbb{X},\mathbb{Y}}$) can be defined similarly. Finally, we will also need to define associations between genes from two or more individuals sampled *without* replacement. For this, we will use the slash symbol to separate sets of genes from

different individuals. For example, $D_{\mathbb{S},\mathbb{T}/\mathbb{U},\mathbb{V}}$ denotes the association between the sets $\mathbb{S}$ and $\mathbb{T}$ of genes from one individual, and the sets $\mathbb{U}$ and $\mathbb{V}$ from a different individual (averaged over all possible pairs of individuals). Note that we have:

$$D_{\mathbb{S},\widehat{\mathbb{T}/\mathbb{U},\mathbb{V}}} = \frac{D_{\mathbb{S}\mathbb{U},\mathbb{T}\mathbb{V}} + D_{\mathbb{S}\mathbb{V},\mathbb{T}\mathbb{U}}}{2N} + \left(1 - \frac{1}{N}\right) D_{\mathbb{S},\mathbb{T}/\mathbb{U},\mathbb{V}}\,. \tag{A5}$$

as sampling with replacement involves either the same individual sampled twice (first term, see equation A3) or different individuals (second term). Associations between genes present in more than two individuals sampled with replacement can be similarly expressed in terms of associations involving individuals sampled without replacement.

In the following, we derive recursions for expectations of genetic associations over the different phases of the life-cycle (selection, gene conversion, reproduction and sampling of the next generation). Note that only the last phase (random sampling) is stochastic, as we assume that selection occurs through differences in fecundity (number of gametes and/or asexual eggs produced) and that fecundity is effectively infinite. We will use different notations to denote genetic associations measured at different phases of the life cycle, as summarized by the following figure:




As we will see, $\langle D^{\mathrm{sel}}_{\mathcal{A}}\rangle_t$ and $\langle D^{\mathrm{dft}}_{\mathcal{A}}\rangle_{t+1}$ correspond to intermediate steps in calculating the effects of selection and drift (respectively) on expected values of associations. $\langle D_{\mathcal{A}}\rangle_{t+1}$ measures the expectation of the association among genes in the set $\mathcal{A}$ (which may include genes present in several individuals) among adults of generation $t+1$ (just after sampling). $\langle D^{\mathrm{dft}}_{\mathcal{A}}\rangle_{t+1}$ measures the same association, but using as reference

values (the $\wp_X$ in equation A1) allele frequencies before sampling (that is, among juveniles). $\langle D_{\mathcal{A}}^{\text{juv}}\rangle_{t+1}$ denotes expected associations measured at the juvenile stage (after reproduction, before sampling), while $\langle D_{\mathcal{A}}^{\text{cnv}}\rangle_t$ denotes associations measured in their parents (generation $t$) after gene conversion, that is, in the parents diploid cells that give rise to asexual eggs (by mitosis) or to gametes (by meiosis). $\langle D_{\mathcal{A}}^{\text{par}}\rangle_t$ then denotes associations among parents before gene conversion, but taking into account the effect of selection by weighting zeta variables by the relative fecundity of the individual; for example:

$$\left\langle D_{\mathbb{S},\mathbb{T}}^{\text{par}}\right\rangle_t = \left\langle \text{E}\left[\frac{f_j}{f}\,\zeta_{\mathbb{S},\mathbb{T}(j)}\right]\right\rangle_t \tag{A6}$$

where $f_j$ is the fecundity of parent $j$, $f$ the average fecundity in the population, and E stands for the average over all parents (adults of generation $t$). Similarly,

$$\left\langle D_{\mathbb{S},\mathbb{T}\widehat{/}\mathbb{U},\mathbb{V}}^{\text{par}}\right\rangle_t = \left\langle \text{E}\left[\left(\frac{f_j}{f}\,\zeta_{\mathbb{S},\mathbb{T}(j)}\right)\left(\frac{f_k}{f}\,\zeta_{\mathbb{U},\mathbb{V}(k)}\right)\right]\right\rangle_t \tag{A7}$$

where E is the average over all $j$ and $k$ (including $j = k$). Associations $\langle D_{\mathcal{A}}^{\text{par}}\rangle_t$ will be called "associations after selection". The same associations, but using as reference values allele frequencies before selection will be denoted $\langle D_{\mathcal{A}}^{\text{sel}}\rangle_t$.

**Drift.** The last phase of the life cycle corresponds to the random sampling of $N$ individuals among the infinite number of juveniles produced. The association between the sets of genes $\mathbb{S}$ and $\mathbb{T}$ on the two haplotypes of the same individual, after sampling, is defined as (from equation A3):

$$\begin{aligned}\langle D_{\mathbb{S},\mathbb{T}}\rangle_{t+1} = \Big\langle \frac{1}{2}\,\text{E}\Big[&\prod_{X\in\mathbb{S}}\left(p_{X(j1)} - p_X\right)\prod_{Y\in\mathbb{T}}\left(p_{Y(j2)} - p_Y\right)\\ &+\prod_{X\in\mathbb{S}}\left(p_{X(j2)} - p_X\right)\prod_{Y\in\mathbb{T}}\left(p_{Y(j1)} - p_Y\right)\Big]\Big\rangle_{t+1}\end{aligned} \tag{A8}$$

where E is the average over all individuals $j$ after sampling, and the reference values $p_X$, $p_Y$ correspond to allele frequencies after sampling. We first express this association in terms of associations measured after sampling, but using as reference values allele frequencies among juveniles (before sampling), denoted $p_X^{\text{juv}}$. We have $p_X = p_X^{\text{juv}} + \Delta_{\text{d}} p_X$, where$\Delta_{\text{d}} p_X$ is the change in allele frequency $p_X$ due to drift (*i.e.*, to sampling). We can rewrite equation A8 as:

$$\begin{aligned}\langle D_{\mathbb{S},\mathbb{T}}\rangle_{t+1} = &\Bigg\langle \frac{1}{2}\,\mathrm{E}\Bigg[\prod_{X\in\mathbb{S}}\Big(p_{X(j1)} - p_X^{\text{juv}} - \Delta_{\text{d}} p_X\Big)\prod_{Y\in\mathbb{T}}\Big(p_{Y(j2)} - p_Y^{\text{juv}} - \Delta_{\text{d}} p_Y\Big) \\ &+ \prod_{X\in\mathbb{S}}\Big(p_{X(j2)} - p_X^{\text{juv}} - \Delta_{\text{d}} p_X\Big)\prod_{Y\in\mathbb{T}}\Big(p_{Y(j1)} - p_Y^{\text{juv}} - \Delta_{\text{d}} p_Y\Big)\Bigg]\Bigg\rangle_{t+1}.\end{aligned} \tag{A9}$$

Denoting $D^{\text{dft}}$ associations measured after sampling, but using as reference values allele frequencies before sampling (the $p_X^{\text{juv}}$), equation A9 can be written as:

$$\begin{aligned}\langle D_{\mathbb{S},\mathbb{T}}\rangle_{t+1} = \langle D^{\text{dft}}_{\mathbb{S},\mathbb{T}}\rangle_{t+1} - \sum_{X\in\mathbb{S}}\langle \Delta_{\text{d}} p_X\, D^{\text{dft}}_{(\mathbb{S}\backslash X),\mathbb{T}}\rangle_{t+1} - \sum_{Y\in\mathbb{T}}\langle \Delta_{\text{d}} p_Y\, D^{\text{dft}}_{\mathbb{S},(\mathbb{T}\backslash Y)}\rangle_{t+1} \\ + \sum_{X\in\mathbb{S}}\sum_{Y\in\mathbb{T}}\langle \Delta_{\text{d}} p_X \Delta_{\text{d}} p_Y\, D^{\text{dft}}_{(\mathbb{S}\backslash X),(\mathbb{T}\backslash Y)}\rangle_{t+1} + \ldots\end{aligned} \tag{A10}$$

where $\mathbb{S}\backslash X$ means "the set $\mathbb{S}$, from which $X$ is removed", and other terms include sums over all possible subsets of $\mathbb{S}$ and $\mathbb{T}$. A more compact expression is provided by equation 15 in Kirkpatrick *et al.*, 2002:

$$\langle D_{\mathcal{A}}\rangle_{t+1} = \sum_{\mathcal{B}\subset\mathcal{A}}\left\langle D^{\text{dft}}_{\mathcal{A}\backslash\mathcal{B}}\prod_{X\in\mathcal{B}}(-\Delta_{\text{d}} p_X)\right\rangle_{t+1} \tag{A11}$$

where $\mathcal{A}$ is a set of loci that may be present in different individuals (for example we can have $\mathcal{A} = \mathbb{S}, \mathbb{T}\widehat{/}\mathbb{U}, \mathbb{V}$), and the sum is over all possible subsets $\mathcal{B}$ of $\mathcal{A}$. Finally, we

can note that$\Delta_{\mathrm{d}} p_X = D_X^{\mathrm{dft}}$, and thus:

$$\begin{aligned}\langle D_{\mathcal{A}}\rangle_{t+1} &= \sum_{\mathcal{B}\subset\mathcal{A}}(-1)^{|\mathcal{B}|}\left\langle D^{\mathrm{dft}}_{\mathcal{A}\setminus\mathcal{B}}\prod_{X\in\mathcal{B}} D_X^{\mathrm{dft}}\right\rangle_{t+1}\\ &= \sum_{\mathcal{B}\subset\mathcal{A}}(-1)^{|\mathcal{B}|}\left\langle D^{\mathrm{dft}}_{(\mathcal{A}\setminus\mathcal{B})\underbrace{\widehat{/}X\widehat{/}Y\widehat{/}Z\ldots}_{X,\ Y,\ Z\ldots\in\mathcal{B}}}\right\rangle_{t+1}\end{aligned} \tag{A12}$$

where $|\mathcal{B}|$ is the number of genes in the set $\mathcal{B}$ (equation A12 can be implemented in *Mathematica*). The change from the first to the second line of equation A12 is simply a change of notation, as $\langle D_X D_Y D_Z\rangle$ can also be noted $\left\langle D_{X\widehat{/}Y\widehat{/}Z}\right\rangle$ (see A4). For example, equation A12 yields:

$$\langle D_{\mathrm{A,A}}\rangle_{t+1} = \left\langle D^{\mathrm{dft}}_{\mathrm{A,A}}\right\rangle_{t+1} - \left\langle D^{\mathrm{dft}}_{\mathrm{A}\widehat{/}\mathrm{A}}\right\rangle_{t+1} \tag{A13}$$

$$\begin{aligned}\langle D_{\mathrm{MA,A}}\rangle_{t+1} = \left\langle D^{\mathrm{dft}}_{\mathrm{MA,A}}\right\rangle_{t+1} &- \left\langle D^{\mathrm{dft}}_{\mathrm{MA}\widehat{/}\mathrm{A}}\right\rangle_{t+1} - \left\langle D^{\mathrm{dft}}_{\mathrm{M,A}\widehat{/}\mathrm{A}}\right\rangle_{t+1}\\ &- \left\langle D^{\mathrm{dft}}_{\mathrm{A,A}\widehat{/}\mathrm{M}}\right\rangle_{t+1} + 2\left\langle D^{\mathrm{dft}}_{\mathrm{M}\widehat{/}\mathrm{A}\widehat{/}\mathrm{A}}\right\rangle_{t+1}\end{aligned} \tag{A14}$$

$$\left\langle D_{\mathrm{MA}\widehat{/}\mathrm{MA}}\right\rangle_{t+1} = \left\langle D^{\mathrm{dft}}_{\mathrm{MA}\widehat{/}\mathrm{MA}}\right\rangle_{t+1} - 2\left\langle D^{\mathrm{dft}}_{\mathrm{MA}\widehat{/}\mathrm{M}\widehat{/}\mathrm{A}}\right\rangle_{t+1} + \left\langle D^{\mathrm{dft}}_{\mathrm{M}\widehat{/}\mathrm{M}\widehat{/}\mathrm{A}\widehat{/}\mathrm{A}}\right\rangle_{t+1}. \tag{A15}$$

The next step is to express associations between genes from individuals sampled with replacement (after drift) in terms of associations involving individuals sampled *without* replacement. For example, we have:

$$\left\langle D^{\mathrm{dft}}_{\mathrm{A}\widehat{/}\mathrm{A}}\right\rangle_{t+1} = \frac{\left\langle D^{\mathrm{dft}}_{\mathrm{AA}}\right\rangle_{t+1} + \left\langle D^{\mathrm{dft}}_{\mathrm{A,A}}\right\rangle_{t+1}}{2N} + \left(1-\frac{1}{N}\right)\left\langle D^{\mathrm{dft}}_{\mathrm{A/A}}\right\rangle_{t+1}. \tag{A16}$$

A more general expression allowing us to express associations involving an arbitrary number of individuals sampled with replacement in terms of associations between individuals sampled without replacement has been implemented in *Mathematica*.

Finally, we can note that because adults are sampled independently from an infinite number of juveniles, the expectation of an association between genes present in the same or in different adults, after sampling (and using as reference values allele frequencies before sampling) equals the same association measured among juveniles (*i.e.*, before sampling). Therefore

$$\left\langle D^{\mathrm{dft}}_{\mathcal{A}} \right\rangle_{t+1} = \left\langle D^{\mathrm{juv}}_{\mathcal{A}} \right\rangle_{t+1} \tag{A17}$$

when genes in the set $\mathcal{A}$ are present in a single individual, or in several individuals sampled *without* replacement. We thus have a method to express expected values of associations after drift in terms of associations measured at the juvenile stage, after reproduction. For example, combining equations A13 and A16 yields:

$$\langle D_{\mathrm{A,A}} \rangle_{t+1} = \left\langle D^{\mathrm{juv}}_{\mathrm{A,A}} \right\rangle_{t+1} - \frac{\left\langle D^{\mathrm{juv}}_{\mathrm{AA}} \right\rangle_{t+1} + \left\langle D^{\mathrm{juv}}_{\mathrm{A,A}} \right\rangle_{t+1}}{2N} - \left(1 - \frac{1}{N}\right) \left\langle D^{\mathrm{juv}}_{\mathrm{A/A}} \right\rangle_{t+1} \tag{A18}$$

where associations on the right-hand-side are measured at the juvenile stage. Finally, we can note that the last term of equation A18 equals zero: indeed, because the number of juveniles is assumed to be infinite, we have $\left\langle D^{\mathrm{juv}}_{\mathrm{A/A}} \right\rangle_{t+1} = \left\langle D^{\mathrm{juv}}_{\mathrm{A}\widehat{/}\mathrm{A}} \right\rangle_{t+1} = \left\langle D^{\mathrm{juv}}_{\mathrm{A}} D^{\mathrm{juv}}_{\mathrm{A}} \right\rangle_{t+1} = 0$ (since $D^{\mathrm{juv}}_{\mathrm{A}} = 0$).

Recursions for products of allele frequencies and associations can be obtained similarly. For example, a recursion for $\langle p_{\mathrm{M}} D_{\mathrm{A,A}} \rangle$ over sampling is obtained as follows (dropping indices $t+1$):

$$\begin{aligned}
\langle p_{\mathrm{M}} D_{\mathrm{A,A}} \rangle &= \left\langle \left( p^{\mathrm{juv}}_{\mathrm{M}} + \Delta_{\mathrm{d}} p_{\mathrm{M}} \right) \mathrm{E} \left[ \left( p_{\mathrm{A}(j1)} - p^{\mathrm{juv}}_{\mathrm{A}} - \Delta_{\mathrm{d}} p_{\mathrm{A}} \right) \left( p_{\mathrm{A}(j2)} - p^{\mathrm{juv}}_{\mathrm{A}} - \Delta_{\mathrm{d}} p_{\mathrm{A}} \right) \right] \right\rangle \\
&= \left\langle p^{\mathrm{juv}}_{\mathrm{M}} D^{\mathrm{dft}}_{\mathrm{A,A}} \right\rangle + \left\langle D^{\mathrm{dft}}_{\mathrm{A,A}\widehat{/}\mathrm{M}} \right\rangle - \left\langle p^{\mathrm{juv}}_{\mathrm{M}} D^{\mathrm{dft}}_{\mathrm{A}\widehat{/}\mathrm{A}} \right\rangle - \left\langle D^{\mathrm{dft}}_{\mathrm{A}\widehat{/}\mathrm{A}\widehat{/}\mathrm{M}} \right\rangle \\
&= \left\langle p^{\mathrm{juv}}_{\mathrm{M}} D^{\mathrm{juv}}_{\mathrm{A,A}} \right\rangle + \frac{1}{2N} \left( 2 \left\langle D^{\mathrm{juv}}_{\mathrm{MA,A}} \right\rangle - \left\langle p^{\mathrm{juv}}_{\mathrm{M}} D^{\mathrm{juv}}_{\mathrm{AA}} \right\rangle - \left\langle p^{\mathrm{juv}}_{\mathrm{M}} D^{\mathrm{juv}}_{\mathrm{A,A}} \right\rangle \right) \\
&\qquad - \frac{1}{4N^2} \left( 2 \left\langle D^{\mathrm{juv}}_{\mathrm{MA,A}} \right\rangle - \left\langle D^{\mathrm{juv}}_{\mathrm{MAA}} \right\rangle - \left\langle D^{\mathrm{juv}}_{\mathrm{M,AA}} \right\rangle \right) .
\end{aligned} \tag{A19}$$

A general expression for deriving recursions for any product of allele frequencies and associations has been implemented in *Mathematica*.

**Reproduction (no sex modifier).** We then need to express associations measured after reproduction in terms of associations measured in the parental generation. We first consider the case where the modifier has no effect ($\delta\sigma = 0$). The association $D^{\text{juv}}_{\mathbb{S},\mathbb{T}}$ (measured after reproduction) can be expressed in terms of associations in the parental generation by:

$$\left\langle D^{\text{juv}}_{\mathbb{S},\mathbb{T}} \right\rangle_{t+1} = (1-\sigma)\left\langle D^{\text{cnv}}_{\mathbb{S},\mathbb{T}} \right\rangle_t + \sigma \sum_{\mathbb{U}\mathbb{V}=\mathbb{S}} \sum_{\mathbb{X}\mathbb{Y}=\mathbb{T}} t_{\mathbb{U}|\mathbb{V}}\, t_{\mathbb{X}|\mathbb{Y}} \left\langle D^{\text{cnv}}_{\mathbb{U},\mathbb{V}\widehat{/}\mathbb{X},\mathbb{Y}} \right\rangle_t \tag{A20}$$

where associations on the right-hand-side are measured in the parental population, after selection and gene conversion. This equation reads as follows. $D^{\text{juv}}_{\mathbb{S},\mathbb{T}}$ is the association between the sets $\mathbb{S}$ and $\mathbb{T}$ of loci from the two haplotypes of a randomly sampled juvenile. With probability $1-\sigma$, this juvenile has been produced asexually, in which case the association equals the same association measured in the parental generation (first term of equation A20). With probability $\sigma$, the juvenile has been produced sexually; in that case, genes present on the same haplotype are rearranged by recombination at meiosis. The effect of recombination is described by coefficients $t_{\mathbb{U}|\mathbb{V}}$ (see Barton and Turelli, 1991; Kirkpatrick *et al.*, 2002), where $(\mathbb{U},\mathbb{V})$ is a partition of the set $\mathbb{S}$, and $t_{\mathbb{U}|\mathbb{V}}$ is the probability that loci in the set $\mathbb{U}$ come from one of the haplotypes of the parent, and loci in $\mathbb{T}$ from the other; the sum is over all possible subsets of $\mathbb{S}$, including $(\mathbb{S},\emptyset)$. For example, for $\mathbb{S} = \text{MA}$, we have $t_{\text{MA}|\emptyset} = 1-r$, and $t_{\text{M}|\text{A}} = r$. Finally, because we assume random mating, genes present on different haplotypes of a juvenile come from two parents, sampled with replacement from the

parental generation (when selection occurs, the probability that a parent is sampled is proportional to its fecundity). Equation A20 can be generalized to express associations between genes present in different juveniles, for example:

$$\begin{aligned}\left\langle D^{\text{juv}}_{\mathbb{S},\mathbb{T}/\mathbb{U},\mathbb{V}}\right\rangle_{t+1} &= (1-\sigma)^2\left\langle D^{\text{cnv}}_{\mathbb{S},\mathbb{T}\widehat{/}\mathbb{U},\mathbb{V}}\right\rangle_t \\ &+ \sigma(1-\sigma)\sum_{\mathbb{AB}=\mathbb{S}}\sum_{\mathbb{CD}=\mathbb{T}} t_{\mathbb{A}|\mathbb{B}}\, t_{\mathbb{C}|\mathbb{D}}\left\langle D^{\text{cnv}}_{\mathbb{A},\mathbb{B}\widehat{/}\mathbb{C},\mathbb{D}\widehat{/}\mathbb{U},\mathbb{V}}\right\rangle_t \\ &+ \sigma(1-\sigma)\sum_{\mathbb{AB}=\mathbb{U}}\sum_{\mathbb{CD}=\mathbb{V}} t_{\mathbb{A}|\mathbb{B}}\, t_{\mathbb{C}|\mathbb{D}}\left\langle D^{\text{cnv}}_{\mathbb{S},\mathbb{T}\widehat{/}\mathbb{A},\mathbb{B}\widehat{/}\mathbb{C},\mathbb{D}}\right\rangle_t \\ &+ \sigma^2\sum_{\mathbb{AB}=\mathbb{S}}\sum_{\mathbb{CD}=\mathbb{T}}\sum_{\mathbb{EF}=\mathbb{U}}\sum_{\mathbb{GH}=\mathbb{V}} t_{\mathbb{A}|\mathbb{B}}\, t_{\mathbb{C}|\mathbb{D}}\, t_{\mathbb{E}|\mathbb{F}}\, t_{\mathbb{G}|\mathbb{H}}\left\langle D^{\text{cnv}}_{\mathbb{A},\mathbb{B}\widehat{/}\mathbb{C},\mathbb{D}\widehat{/}\mathbb{E},\mathbb{F}\widehat{/}\mathbb{G},\mathbb{H}}\right\rangle_t .\end{aligned} \tag{A21}$$

A general expression (valid for an arbitrary number of loci and individuals) has been implemented in *Mathematica*. For example in the absence of selection and gene conversion ($s=\gamma=0$) so that $\langle D^{\text{cnv}}_{\mathcal{A}}\rangle_t = \langle D_{\mathcal{A}}\rangle_t$, we have:

$$\left\langle D^{\text{juv}}_{\text{A,A}}\right\rangle_{t+1} = (1-\sigma)\langle D_{\text{A,A}}\rangle_t + \sigma\left\langle D_{\text{A}\widehat{/}\text{A}}\right\rangle_t \tag{A22}$$

Because $D_{\text{A}\widehat{/}\text{A}} = (D_{\text{A}})^2 = 0$, this simplifies to:

$$\left\langle D^{\text{juv}}_{\text{A,A}}\right\rangle_{t+1} = (1-\sigma)\langle D_{\text{A,A}}\rangle_t \tag{A23}$$

Similarly, one obtains (still for $s=\gamma=0$):

$$\begin{aligned}\left\langle D^{\text{juv}}_{\text{MA/MA}}\right\rangle_{t+1} &= (1-\sigma)^2\left\langle D_{\text{MA}\widehat{/}\text{MA}}\right\rangle_t \\ &+ 2\sigma(1-\sigma)\left[(1-r)\left\langle D_{\text{MA}\widehat{/}\text{MA}}\right\rangle_t + r\left\langle D_{\text{MA}\widehat{/}\text{M,A}}\right\rangle_t\right] \\ &+ \sigma^2\Big[(1-r)^2\left\langle D_{\text{MA}\widehat{/}\text{MA}}\right\rangle_t + 2r(1-r)\left\langle D_{\text{MA}\widehat{/}\text{M,A}}\right\rangle_t \\ &\qquad + r^2\left\langle D_{\text{M,A}\widehat{/}\text{M,A}}\right\rangle_t\Big] .\end{aligned} \tag{A24}$$

**Quasi-equilibrium.** The form of the recursions obtained for genetic associations indicates that when population size $N$ is large, and when rates of sex and recombination

are sufficiently large ($\sigma \gg 1/N$, $r \gg 1/N$), the population will quickly reach a quasi-equilibrium state in which associations are small, and change very slowly over time. Consider for example the association $D_{\mathrm{A,A}}$ (measuring departure from Hardy-Weinberg equilibrium at locus A) in the absence of selection and gene conversion ($s = \gamma = 0$). From equations A18 and A23, we have:

$$\langle D_{\mathrm{A,A}} \rangle_{t+1} = (1-\sigma) \langle D_{\mathrm{A,A}} \rangle_t - \frac{1}{2N} \left[ \langle D_{\mathrm{AA}} \rangle_t + (1-\sigma) \langle D_{\mathrm{A,A}} \rangle_t \right] \tag{A25}$$

Note that $\langle D_{\mathrm{AA}} \rangle_t$ is the association between a gene at locus A and itself, given by $\left( \mathrm{E}\left[ {\zeta_{\mathrm{A}(j1)}}^2 \right] + \mathrm{E}\left[ {\zeta_{\mathrm{A}(j2)}}^2 \right] \right)/2$. Because we consider biallelic loci, repeated indices in associations can be eliminated using equation 5 in Kirkpatrick *et al.*, 2002:

$$D_{\mathbb{S}XX,\mathbb{T}} = p_X q_X D_{\mathbb{S},\mathbb{T}} + (1-2p_X) D_{\mathbb{S}X,\mathbb{T}} \tag{A26}$$

where $q_X = 1 - p_X$. In particular, we have $D_{\mathrm{AA}} = p_{\mathrm{A}} q_{\mathrm{A}}$. Using a similar reasoning as in Nagylaki, 1993, one can show that when $\sigma$ is sufficiently large, $\langle D_{\mathrm{A,A}} \rangle_t$ quickly becomes of order $1/N$, while the rate of change of $\langle D_{\mathrm{A,A}} \rangle$ per generation becomes of order $1/N^2$. In particular, equation A25 indicates that $\langle D_{\mathrm{A,A}} \rangle_t$ should be of order $1/N$ when $t \geq t_1 \sim \ln(1/N)/\ln(1-\sigma)$. Once $\langle D_{\mathrm{A,A}} \rangle_t$ is of order $1/N$, we have (still from equation A25):

$$\Delta_{t+1} \langle D_{\mathrm{A,A}} \rangle = (1-\sigma) \Delta_t \langle D_{\mathrm{A,A}} \rangle - \frac{\Delta_t \langle p_{\mathrm{A}} q_{\mathrm{A}} \rangle}{2N} + o(1/N) \tag{A27}$$

where$\Delta_{\ t} Z = Z_{t+1} - Z_t$. Note that the rate of change of allele frequencies is of order $1/N$, and thus the term$\Delta_{\ t} \langle p_{\mathrm{A}} q_{\mathrm{A}} \rangle / (2N)$ is of order $1/N^2$. Equation A27 thus indicates that when $t > 2t_1$, $\Delta_t \langle D_{\mathrm{A,A}} \rangle$ should be of order $1/N^2$. Once this quasi-equilibrium has been reached, we have from equation A25:

$$\langle D_{\mathrm{A,A}} \rangle_t + \Delta_t \langle D_{\mathrm{A,A}} \rangle = (1-\sigma) \langle D_{\mathrm{A,A}} \rangle_t - \frac{\langle p_{\mathrm{A}} q_{\mathrm{A}} \rangle_t}{2N} + o(1/N) \tag{A28}$$

and we may neglect$\Delta_t \langle D_{\mathrm{A,A}} \rangle$ (which is of order $1/N^2$) and solve equation A28 to express $\langle D_{\mathrm{A,A}} \rangle_t$ as a function of $\langle p_{\mathrm{A}} q_{\mathrm{A}} \rangle_t$:

$$\langle D_{\mathrm{A,A}} \rangle_t = -\frac{\langle p_{\mathrm{A}} q_{\mathrm{A}} \rangle_t}{2N\sigma} + o(1/N) . \tag{A29}$$

The quasi-equilibrium approximation also holds when other forces (selection, gene conversion, modifier effect) affect changes in allele frequencies, as long as these forces are weak relative to rates of sex and recombination (Barton and Turelli, 1991; Nagylaki, 1993; Bürger, 2000; Kirkpatrick *et al.*, 2002).

The other neutral associations that play a role in the model are $\langle D_{\mathrm{MA,MA}} \rangle$, $\left\langle D_{\mathrm{MA}\widehat{/}\mathrm{MA}} \right\rangle$, $\left\langle D_{\mathrm{MA}\widehat{/}\mathrm{M,A}} \right\rangle$ and $\left\langle D_{\mathrm{M,A}\widehat{/}\mathrm{M,A}} \right\rangle$. These associations can also be expressed at quasi-equilibrium, to the first order in $1/N$, using the above methods. For example, the recursion for $\left\langle D_{\mathrm{M,A}\widehat{/}\mathrm{M,A}} \right\rangle$ is given by (after elimination of terms that quickly become of order $1/N^2$):

$$\left\langle D_{\mathrm{M,A}\widehat{/}\mathrm{M,A}} \right\rangle_{t+1} = (1-\sigma)^2 \left\langle D_{\mathrm{M,A}\widehat{/}\mathrm{M,A}} \right\rangle_t + \frac{\langle D_{\mathrm{MM,AA}} \rangle_t}{2N} + o(1/N) . \tag{A30}$$

Using equation A26, we have:

$$\langle D_{\mathrm{MM,AA}} \rangle = \langle pq_{\mathrm{MA}} \rangle + \langle D_{\mathrm{M,A}} \rangle - 2\langle p_{\mathrm{M}} D_{\mathrm{M,A}} \rangle - 2\langle p_{\mathrm{A}} D_{\mathrm{M,A}} \rangle + 4\langle p_{\mathrm{M}} p_{\mathrm{A}} D_{\mathrm{M,A}} \rangle \tag{A31}$$

with $pq_{\mathrm{MA}} = p_{\mathrm{M}} q_{\mathrm{M}} p_{\mathrm{A}} q_{\mathrm{A}}$. One finds easily that $\langle D_{\mathrm{M,A}} \rangle = 0$ at quasi-equilibrium, to the first order in $1/N$. Indeed, the recursion for $D_{\mathrm{M,A}}$ is given by:

$$\langle D_{\mathrm{M,A}} \rangle_{t+1} = (1-\sigma) \langle D_{\mathrm{M,A}} \rangle_t + o(1/N) . \tag{A32}$$

Recursions for $\langle p_{\mathrm{M}} D_{\mathrm{M,A}} \rangle$, $\langle p_{\mathrm{A}} D_{\mathrm{M,A}} \rangle$ and $\langle p_{\mathrm{M}} p_{\mathrm{A}} D_{\mathrm{M,A}} \rangle$ can be obtained using the methods developed above. One finds that, to the first order in $1/N$, these moments equal zero at quasi-equilibrium. Therefore, we have:

$$\langle D_{\mathrm{MM,AA}} \rangle_t = \langle pq_{\mathrm{MA}} \rangle_t + o(1/N) . \tag{A33}$$

A similar argument gives $\langle D_{\mathrm{MMAA}} \rangle_t = \langle pq_{\mathrm{MA}} \rangle_t + o(1/N)$. Equation A30 thus gives at quasi-equilibrium:

$$\left\langle D_{\mathrm{M,A}\widehat{/}\mathrm{M,A}} \right\rangle_t = \frac{\langle pq_{\mathrm{MA}} \rangle_t}{2N\left[1-(1-\sigma)^2\right]} + o(1/N) \ . \tag{A34}$$

**Reproduction, with sex modifier.** To incorporate an effect of locus M on the rate of sex (proportion of sexually produced offspring) so that genotypes $mm$, $Mm$ and $MM$ at locus M have rates of sex $\sigma$, $\sigma + h_{\mathrm{M}}\,\delta\sigma$ and $\sigma + \delta\sigma$, we can write the rate of sex of parent $j$ under the form (e.g. Barton, 1995):

$$\sigma_j = \overline{\sigma} + d\sigma_{\mathrm{M}}\left(\zeta_{\mathrm{M}(j1)} + \zeta_{\mathrm{M}(j2)}\right) + d\sigma_{\mathrm{M,M}}\left(\zeta_{\mathrm{M,M}(j)} - D_{\mathrm{M,M}}\right) \tag{A35}$$

with:

$$\begin{aligned} \overline{\sigma} &= \sigma + \delta\sigma\left[2h_{\mathrm{M}}\,p_{\mathrm{M}} + (1-2h_{\mathrm{M}})\left({p_{\mathrm{M}}}^2 + D_{\mathrm{M,M}}\right)\right] \\ d\sigma_{\mathrm{M}} &= \delta\sigma\left[h_{\mathrm{M}} + (1-2h_{\mathrm{M}})\,p_{\mathrm{M}}\right] \\ d\sigma_{\mathrm{M,M}} &= \delta\sigma\,(1-2h_{\mathrm{M}}) \end{aligned} \tag{A36}$$

(note that $\overline{\sigma}$ is the average rate of sex in the population). The probability that a random juvenile has been produced asexually is $1-\overline{\sigma}$, while the probability that it has been produced sexually is $\overline{\sigma}$. If it has been produced asexually, the probability that it has been produced by parent $j$ is $(1-\sigma_j)\,/\,(1-\overline{\sigma})$. If it has been produced sexually, the probability that one of its haplotypes comes from individual $j$, and the other from individual $k$ is $(\sigma_j/\overline{\sigma})\,(\sigma_k/\overline{\sigma})$. The same reasoning as for equation A20 above gives us:

$$\begin{aligned} \left\langle D^{\mathrm{juv}}_{\mathbb{S},\mathbb{T}} \right\rangle_{t+1} = \Bigg\langle \mathrm{E}\Big[&(1-\overline{\sigma})\,\frac{1-\sigma_j}{1-\overline{\sigma}}\,\zeta_{\mathbb{S},\mathbb{T}(j)} \\ &+ \overline{\sigma}\left(\frac{\sigma_j}{\overline{\sigma}}\right)\left(\frac{\sigma_k}{\overline{\sigma}}\right)\sum_{\mathbb{UV}=\mathbb{S}}\sum_{\mathbb{XY}=\mathbb{T}} t_{\mathbb{U}|\mathbb{V}}\,t_{\mathbb{X}|\mathbb{Y}}\,\zeta_{\mathbb{U},\mathbb{V}(j)}\zeta_{\mathbb{X},\mathbb{Y}(k)}\Big]\Bigg\rangle_t \end{aligned} \tag{A37}$$

where again E stands for the average over all parental individuals $j$ and $k$ (including $j = k$). Some rearranging gives:

$$\left\langle D^{\text{juv}}_{\mathbb{S},\mathbb{T}} \right\rangle_{t+1} = \left\langle D^{\text{par}}_{\mathbb{S},\mathbb{T}} \right\rangle_t - \left\langle \text{E}\left[\sigma_j\, \zeta_{\mathbb{S},\mathbb{T}(j)}\right] \right\rangle_t + \left\langle \frac{1}{\overline{\sigma}} \sum_{\mathbb{U}\mathbb{V}=\mathbb{S}} \sum_{\mathbb{X}\mathbb{Y}=\mathbb{T}} t_{\mathbb{U}|\mathbb{V}}\, t_{\mathbb{X}|\mathbb{Y}}\, \text{E}\left[\left(\sigma_j\, \zeta_{\mathbb{U},\mathbb{V}(j)}\right)\left(\sigma_k\, \zeta_{\mathbb{X},\mathbb{Y}(k)}\right)\right] \right\rangle_t . \tag{A38}$$

Equation A35 can then be used to replace $\sigma_j$ and $\sigma_k$ by functions of $\delta\sigma$, $h_{\text{M}}$, $p_{\text{M}}$ and $\zeta$-variables. Further simplification is obtained by assuming that the sex modifier has a weak effect ($\delta\sigma$ small), and deriving recursions to the first order in $\delta\sigma$ only. This allows us to express equation A38 in terms of expectations of genetic associations among parents $\langle D^{\text{cnv}}_{\mathcal{A}} \rangle_t$ (after selection and gene conversion), and products of allele frequencies and genetic associations.

For example, one obtains for $D^{\text{juv}}_{\text{MA,A}}$ (in the absence of selection and gene conversion, and keeping only terms that are of first order in $\delta\sigma$ and in $1/N$ at quasi-equilibrium):

$$\left\langle D^{\text{juv}}_{\text{MA,A}} \right\rangle_{t+1} = (1-\sigma)\left\langle D_{\text{MA,A}} \right\rangle_t - d\sigma_{\text{M,M}} \left\langle (1-2p_{\text{M}})\, D_{\text{MA,MA}} \right\rangle_t + \left\langle d\sigma_{\text{M}} \Big[ (1-r)\, D_{\text{MA}\widehat{/}\text{MA}} + D_{\text{MA}\widehat{/}\text{M,A}} + r\, D_{\text{M,A}\widehat{/}\text{M,A}} - p_{\text{M}} q_{\text{M}} D_{\text{A,A}} - D_{\text{MA,MA}} \Big] \right\rangle_t . \tag{A39}$$

Note that when $h_{\text{M}} \neq 1/2$ (dominance at the modifier locus), $d\sigma_{\text{M}}$ depends on $p_{\text{M}}$ (see equation A36).

**Gene conversion.** Gene conversion occurring at locus A only affects associations involving this locus. In particular, we have:

$$\left\langle D^{\text{cnv}}_{\mathbb{U}\text{A},\mathbb{V}} \right\rangle_t = \left(1 - \gamma + \frac{\gamma}{2}\right) \left\langle D^{\text{par}}_{\mathbb{U}\text{A},\mathbb{V}} \right\rangle_t + \frac{\gamma}{2} \left\langle D^{\text{par}}_{\mathbb{U},\mathbb{V}\text{A}} \right\rangle_t \tag{A40}$$

$$\left\langle D^{\mathrm{cnv}}_{\mathbb{U}\mathrm{A},\mathbb{V}\mathrm{A}}\right\rangle_t = (1-\gamma)\left\langle D^{\mathrm{par}}_{\mathbb{U}\mathrm{A},\mathbb{V}\mathrm{A}}\right\rangle_t + \frac{\gamma}{2}\left(\left\langle D^{\mathrm{par}}_{\mathbb{U}\mathrm{A}\mathrm{A},\mathbb{V}}\right\rangle_t + \left\langle D^{\mathrm{par}}_{\mathbb{U},\mathbb{V}\mathrm{A}\mathrm{A}}\right\rangle_t\right) \tag{A41}$$

where $\gamma$ is the rate of gene conversion, and where the sets $\mathbb{U}$ and $\mathbb{V}$ do not contain A. In our model, gene conversion only appears in the derivation of $\langle D_{\mathrm{A,A}}\rangle$ at quasi-equilibrium. Indeed, we have:

$$\left\langle D^{\mathrm{cnv}}_{\mathrm{A,A}}\right\rangle_t = (1-\gamma)\left\langle D^{\mathrm{par}}_{\mathrm{A,A}}\right\rangle_t + \gamma\left\langle D^{\mathrm{par}}_{\mathrm{AA}}\right\rangle_t . \tag{A42}$$

**Selection.** When selection occurs (parents have different fecundities), we need to express $\langle D^{\mathrm{par}}_{\mathcal{A}}\rangle_t$ (associations among parents after selection) in terms of $\langle D_{\mathcal{A}}\rangle_t$ (associations before selection). As in Kirkpatrick *et al.* (2002) and as in the case of drift above, we proceed in two steps: first express $\langle D^{\mathrm{par}}_{\mathcal{A}}\rangle_t$ in terms of associations after selection, but using as reference values allele frequencies before selection (these associations are denoted $\left\langle D^{\mathrm{sel}}_{\mathcal{A}}\right\rangle_t$), and then express $\left\langle D^{\mathrm{sel}}_{\mathcal{A}}\right\rangle_t$ in terms of associations before selection $\langle D_{\mathcal{A}}\rangle_t$.

Changing reference values is done as above in the case of drift (equations A8 to A11). We have:

$$\langle D^{\mathrm{par}}_{\mathcal{A}}\rangle_t = \sum_{\mathcal{B}\subset\mathcal{A}}\left\langle D^{\mathrm{sel}}_{\mathcal{A}\setminus\mathcal{B}}\prod_{X\in\mathcal{B}}(-\Delta_{\mathrm{s}}p_X)\right\rangle_t \tag{A43}$$

where the sum is over all possible subsets $\mathcal{B}$ of $\mathcal{A}$, and$\Delta_{\ \mathrm{s}}p_X$ is the change in allele frequency $p_X$ due to selection. Because$\Delta_{\ \mathrm{s}}p_X = D^{\mathrm{sel}}_X$, equation A43 can be written as:

$$\langle D^{\mathrm{par}}_{\mathcal{A}}\rangle_t = \sum_{\mathcal{B}\subset\mathcal{A}}(-1)^{|\mathcal{B}|}\left\langle D^{\mathrm{sel}}_{(\mathcal{A}\setminus\mathcal{B})\underbrace{\widehat{\ }X\widehat{\ }Y\widehat{\ }Z\ldots}_{X,\ Y,\ Z\ldots\in\mathcal{B}}}\right\rangle_t . \tag{A44}$$

For example, we have:

$$\langle D^{\mathrm{par}}_{\mathrm{MA}}\rangle_t = \left\langle D^{\mathrm{sel}}_{\mathrm{MA}}\right\rangle_t - \left\langle D^{\mathrm{sel}}_{\mathrm{M}\widehat{\ }\mathrm{A}}\right\rangle_t \tag{A45}$$

$$\begin{aligned}\left\langle D^{\text{par}}_{\text{MA,M}}\right\rangle_t = \left\langle D^{\text{sel}}_{\text{MA,M}}\right\rangle_t - \left\langle D^{\text{sel}}_{\text{MA}\widehat{/}\text{M}}\right\rangle_t - \left\langle D^{\text{sel}}_{\text{M,A}\widehat{/}\text{M}}\right\rangle_t \\ - \left\langle D^{\text{sel}}_{\text{M,M}\widehat{/}\text{A}}\right\rangle_t + 2\left\langle D^{\text{sel}}_{\text{M}\widehat{/}\text{M}\widehat{/}\text{A}}\right\rangle_t .\end{aligned} \tag{A46}$$

$$\begin{aligned}\left\langle p^{\text{juv}}_{\text{M}} D^{\text{par}}_{\text{MA}}\right\rangle_t = \left\langle p_{\text{M}} D^{\text{sel}}_{\text{MA}}\right\rangle_t + \left\langle D^{\text{sel}}_{\text{MA}\widehat{/}\text{M}}\right\rangle_t - \left\langle p_{\text{M}} D^{\text{sel}}_{\text{M}\widehat{/}\text{A}}\right\rangle_t \\ - \left\langle D^{\text{sel}}_{\text{M}\widehat{/}\text{M}\widehat{/}\text{A}}\right\rangle_t .\end{aligned} \tag{A47}$$

Last, we have to express associations after selection, using as reference values allele frequencies before selection, in terms of associations before selection. We have:

$$\left\langle D^{\text{sel}}_{\mathbb{S},\mathbb{T}}\right\rangle_t = \left\langle \text{E}\left[\frac{f_j}{f}\,\zeta_{\mathbb{S},\mathbb{T}(j)}\right]\right\rangle_t \tag{A48}$$

$$\left\langle D^{\text{sel}}_{\mathbb{S},\mathbb{T}\widehat{/}\mathbb{U},\mathbb{V}}\right\rangle_t = \left\langle \text{E}\left[\left(\frac{f_j}{f}\,\zeta_{\mathbb{S},\mathbb{T}(j)}\right)\left(\frac{f_k}{f}\,\zeta_{\mathbb{U},\mathbb{V}(k)}\right)\right]\right\rangle_t \tag{A49}$$

where again $f_j$ is the fecundity of parent $j$ and $f$ the average fecundity, E stands for the average over all individuals $j$ and $k$ (including $j = k$), and reference values in the $\zeta$'s are allele frequencies before selection.

In the two-locus model, we can write the fecundity of individual $j$ under the form:

$$f_j = 1 + T_{\text{A}} + a_{\text{A}}\left(\zeta_{\text{A},\emptyset(j)} + \zeta_{\emptyset,\text{A}(j)}\right) + a_{\text{A,A}}\,\zeta_{\text{A,A}(j)} \tag{A50}$$

with

$$\begin{gathered}T_{\text{A}} = -s\,p_{\text{A}}\left[2h + (1-2h)\,p_{\text{A}}\right], \quad a_{\text{A}} = -s\left[h + (1-2h)\,p_{\text{A}}\right], \\ a_{\text{A,A}} = -s\,(1-2h)\,.\end{gathered} \tag{A51}$$

The average fecundity in the population is thus given by:

$$f = 1 + T_{\text{A}} + a_{\text{A,A}}\,D_{\text{A,A}}\,. \tag{A52}$$

The fraction $f_j/f$ can then be expressed to various orders in $s$, and the result plugged into equations A48 and A49. For example, we have:

$$\frac{f_j}{f} = 1 + a_{\mathrm{A}}\left(\zeta_{\mathrm{A},\emptyset(j)} + \zeta_{\emptyset,\mathrm{A}(j)}\right) + a_{\mathrm{A,A}}\left(\zeta_{\mathrm{A,A}(j)} - D_{\mathrm{A,A}}\right) + o\left(s\right). \tag{A53}$$

This has been implemented in a *Mathematica* notebook, in order to obtain recursions to arbitrary orders in selection coefficients.

We proceed similarly in the three-locus model, writing fecundity as:

$$\begin{aligned} f_j = &\left[1 + T_{\mathrm{A}} + a_{\mathrm{A}}\left(\zeta_{\mathrm{A},\emptyset(j)} + \zeta_{\emptyset,\mathrm{A}(j)}\right) + a_{\mathrm{A,A}}\,\zeta_{\mathrm{A,A}(j)}\right] \\ &\times \left[1 + T_{\mathrm{B}} + a_{\mathrm{B}}\left(\zeta_{\mathrm{B},\emptyset(j)} + \zeta_{\emptyset,\mathrm{B}(j)}\right) + a_{\mathrm{B,B}}\,\zeta_{\mathrm{B,B}(j)}\right] \end{aligned} \tag{A54}$$

where $T_{\mathrm{B}}$, $a_{\mathrm{B}}$ and $a_{\mathrm{B,B}}$ are given by the same expressions as $T_{\mathrm{A}}$, $a_{\mathrm{A}}$ and $a_{\mathrm{A,A}}$ above, replacing $p_{\mathrm{A}}$ by $p_{\mathrm{B}}$.

## ONLINE APPENDIX B: THE RECURSIONS

We derive here expressions for the change in frequency of a sex modifier in the two-locus model, assuming that the rate of gene conversion and selection are weak ($\gamma$ and $s$ of order $\epsilon$, where epsilon is a small term), population size is large ($1/N$ is of order $\epsilon$), and rates of sex and recombination are large relative to $\epsilon$, so that the population quickly reaches a state of quasi-equilibrium where genetic associations are small and change very slowly over time (see Appendix A for further details). We also assume that the modifier has a weak effect ($\delta\sigma$), and calculate all terms to the first order in $\delta\sigma$ only. We first consider the case where deleterious mutations have arbitrary dominance (no assumption on $h$), and then explore the case of weak dominance (*i.e.*, $h - 1/2$ of order $\epsilon$).

**Weak selection, $h$ arbitrary.** We assume here that $s$, $\gamma$ and $1/N$ are of order $\epsilon$, and do not make any assumption on $h$. From the methods presented in Appendix A, the expected change in frequency of the modifier at generation $t$ is given by:

$$\langle \Delta p_{\mathrm{M}} \rangle_t = -s\,(1-2h)\,\langle D_{\mathrm{MA,A}} \rangle_t + o\left(\epsilon^2\right) . \tag{B1}$$

At quasi-equilibrum, a positive value of $\langle D_{\mathrm{MA,A}} \rangle$ means that allele $M$ is more frequent in homozygotes at the A locus than in heterozygotes, while a negative value means that $M$ is more frequent in heterozygotes (see Appendix C). Three forces affect this association: random drift due to finite population size (which generates a positive $D_{\mathrm{MA,A}}$), gene conversion (which generates a negative $D_{\mathrm{MA,A}}$) and selection at the A locus, whose effect on $D_{\mathrm{MA,A}}$ has the sign of $s\,(1-2h)$. Using the methods of Appendix

A, the recursion for $\langle D_{\mathrm{MA,A}}\rangle_t$ once quasi-equilibrium has been reached is given by:

$$\begin{aligned}\langle D_{\mathrm{MA,A}}\rangle_{t+1} &= (1-\sigma)\langle D_{\mathrm{MA,A}}\rangle_t \\ &+ \delta\sigma h_{\mathrm{M}}\Big[(1-r)\left\langle {D_{\mathrm{MA}}}^2\right\rangle_t + \langle D_{\mathrm{MA}}D_{\mathrm{M,A}}\rangle_t + r\left\langle {D_{\mathrm{M,A}}}^2\right\rangle_t \\ &\qquad - \langle D_{\mathrm{MA,MA}}\rangle_t - \left\langle pq_{\mathrm{M}}D^{\mathrm{cnv}}_{\mathrm{A,A}}\right\rangle_t\Big] \\ &+ \delta\sigma(1-2h_{\mathrm{M}})\Big[(1-r)\left\langle p_{\mathrm{M}}{D_{\mathrm{MA}}}^2\right\rangle_t + \langle p_{\mathrm{M}}D_{\mathrm{MA}}D_{\mathrm{M,A}}\rangle_t + r\left\langle p_{\mathrm{M}}{D_{\mathrm{M,A}}}^2\right\rangle_t \\ &\qquad - \langle D_{\mathrm{MA,MA}}\rangle_t + \langle p_{\mathrm{M}}D_{\mathrm{MA,MA}}\rangle_t - \left\langle {p_{\mathrm{M}}}^2q_{\mathrm{M}}D^{\mathrm{cnv}}_{\mathrm{A,A}}\right\rangle_t\Big]\end{aligned} \tag{B2}$$

where $D^{\mathrm{cnv}}_{\mathrm{A,A}}$ corresponds to $D_{\mathrm{A,A}}$ measured after selection and gene conversion, before reproduction (see Appendix A). An expression for $\langle D_{\mathrm{MA,A}}\rangle_t$ at quasi-equilibrium can be obtained by setting $\langle D_{\mathrm{MA,A}}\rangle_t = \langle D_{\mathrm{MA,A}}\rangle_{t+1}$ in equation B2, and solving for $\langle D_{\mathrm{MA,A}}\rangle_t$. In the case of an additive modifier ($h_{\mathrm{M}} = 1/2$), the last term of equation B2 equals zero, and $\langle D_{\mathrm{MA,A}}\rangle_t$ is generated by associations $\langle {D_{\mathrm{MA}}}^2\rangle_t$, $\langle D_{\mathrm{MA}}D_{\mathrm{M,A}}\rangle_t$, $\langle {D_{\mathrm{M,A}}}^2\rangle_t$, $\langle D_{\mathrm{MA,MA}}\rangle_t$ and $\left\langle pq_{\mathrm{M}}D^{\mathrm{cnv}}_{\mathrm{A,A}}\right\rangle_t$ (second term of equation B2). The quasi-equilibrium values of these associations are obtained by solving the following recursions (expressed to the first order in $\epsilon$):

$$\begin{aligned}\left\langle pq_{\mathrm{M}}D^{\mathrm{cnv}}_{\mathrm{A,A}}\right\rangle_{t+1} &= (1-\sigma)\left\langle pq_{\mathrm{M}}D^{\mathrm{cnv}}_{\mathrm{A,A}}\right\rangle_t - s(1-2h)\left\langle pq_{\mathrm{M}}(pq_{\mathrm{A}})^2\right\rangle_t + \gamma\langle pq_{\mathrm{MA}}\rangle_t \\ &\quad - \frac{\langle pq_{\mathrm{MA}}\rangle_t}{2N}\end{aligned} \tag{B3}$$

$$\begin{aligned}\langle D_{\mathrm{MA,MA}}\rangle_{t+1} &= (1-\sigma)\langle D_{\mathrm{MA,MA}}\rangle_t + \sigma\Big[(1-r)^2\left\langle {D_{\mathrm{MA}}}^2\right\rangle_t \\ &\quad + 2r(1-r)\langle D_{\mathrm{MA}}D_{\mathrm{M,A}}\rangle_t + r^2\left\langle {D_{\mathrm{M,A}}}^2\right\rangle_t\Big]\end{aligned} \tag{B4}$$

$$\begin{aligned}\left\langle {D_{\mathrm{MA}}}^2\right\rangle_{t+1} &= \frac{\langle pq_{\mathrm{MA}}\rangle_t}{2N} + (1-r\sigma)^2\left\langle {D_{\mathrm{MA}}}^2\right\rangle_t \\ &\quad + 2r\sigma(1-r\sigma)\langle D_{\mathrm{MA}}D_{\mathrm{M,A}}\rangle_t + r^2\sigma^2\left\langle {D_{\mathrm{M,A}}}^2\right\rangle_t\end{aligned} \tag{B5}$$

$$\langle D_{\mathrm{MA}}D_{\mathrm{M,A}}\rangle_{t+1} = (1-\sigma)\left[(1-r\sigma)\langle D_{\mathrm{MA}}D_{\mathrm{M,A}}\rangle_t + r\sigma\left\langle {D_{\mathrm{M,A}}}^2\right\rangle_t\right] \tag{B6}$$

$$\left\langle {D_{\mathrm{M,A}}}^2 \right\rangle_{t+1} = (1-\sigma)^2 \left\langle {D_{\mathrm{M,A}}}^2 \right\rangle_t + \frac{\langle pq_{\mathrm{MA}} \rangle_t}{2N} \tag{B7}$$

giving at quasi-equilibrium:

$$\left\langle pq_{\mathrm{M}} D^{\mathrm{cnv}}_{\mathrm{A,A}} \right\rangle_t = \frac{1}{\sigma} \left[ \left( \gamma - \frac{1}{2N} \right) \langle pq_{\mathrm{MA}} \rangle_t - s\,(1-2h) \left\langle pq_{\mathrm{M}} \left(pq_{\mathrm{A}}\right)^2 \right\rangle_t \right] \tag{B8}$$

$$\langle D_{\mathrm{MA,MA}} \rangle_t = \frac{[1-2r\,(1-r)]\,\langle pq_{\mathrm{MA}} \rangle_t}{2Nr\sigma\,(2-r\sigma)} \tag{B9}$$

$$\left\langle {D_{\mathrm{MA}}}^2 \right\rangle_t = \frac{\langle pq_{\mathrm{MA}} \rangle_t}{2Nr\sigma\,(2-r\sigma)} \left[ 1 + \frac{r^2\sigma}{2-\sigma} \frac{1+(1-\sigma)\,(1-r\sigma)}{1-(1-\sigma)\,(1-r\sigma)} \right] \tag{B10}$$

$$\langle D_{\mathrm{MA}} D_{\mathrm{M,A}} \rangle_t = \frac{(1-\sigma)\,r\,\langle pq_{\mathrm{MA}} \rangle_t}{2N\sigma\,(2-\sigma)\,[1+r\,(1-\sigma)]} \tag{B11}$$

$$\left\langle {D_{\mathrm{M,A}}}^2 \right\rangle_t = \frac{\langle pq_{\mathrm{MA}} \rangle_t}{2N\sigma\,(2-\sigma)}\,. \tag{B12}$$

In the case of a non-additive modifier ($h_{\mathrm{M}} \neq 1/2$), $\langle D_{\mathrm{MA,A}} \rangle_t$ is also affected by the moments $\left\langle {p_{\mathrm{M}}}^2 q_{\mathrm{M}} D^{\mathrm{cnv}}_{\mathrm{A,A}} \right\rangle_t$, $\langle p_{\mathrm{M}} D_{\mathrm{MA,MA}} \rangle_t$, $\left\langle p_{\mathrm{M}} {D_{\mathrm{MA}}}^2 \right\rangle_t$, $\langle p_{\mathrm{M}} D_{\mathrm{MA}} D_{\mathrm{M,A}} \rangle_t$ and $\left\langle p_{\mathrm{M}} {D_{\mathrm{M,A}}}^2 \right\rangle_t$ (third term of equation B2). However, applying the methods of Appendix A shows that expressions for these associations at quasi-equilibrium are given by equations B8 to B12, replacing $\langle pq_{\mathrm{MA}} \rangle_t$ by $\langle p_{\mathrm{M}} pq_{\mathrm{MA}} \rangle_t$. Putting everything together, one obtains:

$$\begin{aligned} \langle D_{\mathrm{MA,A}} \rangle_t = \frac{\delta\sigma}{\sigma^2} \Bigg[ & \left( \frac{T_1}{2N} - \gamma \right) \left( h_{\mathrm{M}} \langle pq_{\mathrm{MA}} \rangle_t + (1-2h_{\mathrm{M}}) \langle p_{\mathrm{M}} pq_{\mathrm{MA}} \rangle_t \right) \\ & - \frac{T_2}{2N} (1-2h_{\mathrm{M}}) \left( \langle pq_{\mathrm{MA}} \rangle_t - 2 \langle p_{\mathrm{M}} pq_{\mathrm{MA}} \rangle_t \right) \\ & + s\,(1-2h) \left( h_{\mathrm{M}} \left\langle pq_{\mathrm{M}} \left(pq_{\mathrm{A}}\right)^2 \right\rangle_t + (1-2h_{\mathrm{M}}) \left\langle p_{\mathrm{M}} pq_{\mathrm{M}} \left(pq_{\mathrm{A}}\right)^2 \right\rangle_t \right) \Bigg] \end{aligned} \tag{B13}$$

with:

$$T_1 = \frac{3 - 2r^2 + r\,(1-\sigma)\,(4-r\sigma)}{[1+r\,(1-\sigma)]\,(2-r\,\sigma)} \tag{B14}$$

$$T_2 = \frac{1 - 2r\,(1-r)}{r\,(2 - r\,\sigma)}\,. \tag{B15}$$

When $r = 1/2$ (free recombination), we have $T_1 = (6-\sigma)\,/\,(4-\sigma)$ and $T_2 = 2/\,(4-\sigma)$. The term on the last line of equation B13 is equivalent to the result obtained in the case of an infinite population, without gene conversion (e.g., Agrawal, 2009). The effect of gene conversion appears in the first line of equation B13 (term in $\gamma$), while the effect of drift is represented by the terms in $T_1$ and $T_2$.

**Weak dominance.** When dominance is weak ($h - 1/2$ of order $\epsilon$), the change in frequency of the modifier is given by:

$$\langle \Delta_t p_{\mathrm{M}} \rangle = -s\,h\left(\langle D_{\mathrm{MA}} \rangle_t + \langle D_{\mathrm{M,A}} \rangle_t\right) - s\,(1-2h)\,\langle D_{\mathrm{MA,A}} \rangle_t + o\left(\epsilon^3\right)\,. \tag{B16}$$

The expected value of the association $D_{\mathrm{MA,A}}$ at quasi-equilibrium is still given by equation B13, except that the term in $s\,(1-2h)$ (third line of equation B13) is now of order $\epsilon^2$ and may be neglected. Associations $D_{\mathrm{MA}}$ and $D_{\mathrm{M,A}}$ are of order $\epsilon^2$ at quasi-equilibrium, and are generated by the combined action of random drift, selection and the modifier effect. Recursions for these associations at quasi-equilibrium are given by the following expressions:

$$\langle D_{\mathrm{MA}} \rangle_{t+1} = (1 - r\,\sigma)\,\langle D_{\mathrm{MA}} \rangle_t + r\,\sigma\,\langle D_{\mathrm{M,A}} \rangle_t - s\,h\,\langle D_{\mathrm{MA,A}} \rangle_t \tag{B17}$$

$$\begin{aligned}\langle D_{\mathrm{M,A}} \rangle_{t+1} = {} & (1-\sigma)\left(\langle D_{\mathrm{M,A}} \rangle_t - s\,h\,\langle D_{\mathrm{MA,A}} \rangle_t\right) - \delta\sigma h_{\,\mathrm{M}}\left\langle D^{\mathrm{par}}_{\mathrm{MA,M}} \right\rangle_t \\ & - \delta\sigma\,(1-2h_{\mathrm{M}})\left[\left\langle pq_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{MA}} \right\rangle_t + \left\langle pq_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{M,A}} \right\rangle_t + \left\langle D^{\mathrm{par}}_{\mathrm{MA,M}} \right\rangle_t - \left\langle p_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{MA,M}} \right\rangle_t\right].\end{aligned} \tag{B18}$$

where "par" denotes associations measured after selection (see Appendix A). Note that the last term of equation B18 equals zero in the case of an additive modifier

($h_{\mathrm{M}} = 1/2$). These expressions are functions of the association $\langle D_{\mathrm{MA,A}} \rangle$ generated by drift and/or gene conversion and by the modifier effect (given by equation B13) and the associations $\left\langle D^{\mathrm{par}}_{\mathrm{MA,M}} \right\rangle$, $\left\langle p_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{MA,M}} \right\rangle$, $\langle p q_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{MA}} \rangle$ and $\left\langle p q_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{M,A}} \right\rangle$ generated by the combined effects of selection and drift. In the following, we derive quasi-equilibrium expressions for these last associations. A recursion for $\left\langle D^{\mathrm{par}}_{\mathrm{MA,M}} \right\rangle$ is given by:

$$\begin{aligned} \left\langle D^{\mathrm{par}}_{\mathrm{MA,M}} \right\rangle_{t+1} = (1-\sigma) \left\langle D^{\mathrm{par}}_{\mathrm{MA,M}} \right\rangle_t - s\,h \Big( & \langle D_{\mathrm{MA,MA}} \rangle_t - \left\langle {D_{\mathrm{MA}}}^2 \right\rangle_t \\ & - 2 \langle D_{\mathrm{MA}} D_{\mathrm{M,A}} \rangle_t - \left\langle {D_{\mathrm{M,A}}}^2 \right\rangle_t \Big) \end{aligned} \tag{B19}$$

giving at quasi-equilibrium:

$$\left\langle D^{\mathrm{par}}_{\mathrm{MA,M}} \right\rangle_t = \frac{s\,h \left[ 3 - \sigma + r \left[ 4 - \sigma \left( 6 - \sigma \right) - r \left( 2 - \sigma \left( 1 + \sigma \right) \right) \right] \right] \langle p q_{\mathrm{MA}} \rangle_t}{N \sigma^2 \left( 2 - \sigma \right) \left( 2 - r\,\sigma \right) \left[ 1 + r \left( 1 - \sigma \right) \right]} \,. \tag{B20}$$

Under free recombination ($r = 1/2$), this simplifies to:

$$\left\langle D^{\mathrm{par}}_{\mathrm{MA,M}} \right\rangle_t = \frac{3\,s\,h \,\langle p q_{\mathrm{MA}} \rangle_t}{N \sigma^2 \left( 4 - \sigma \right)} \,. \tag{B21}$$

$\left\langle p_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{MA,M}} \right\rangle_t$ at quasi-equilibrium is given by the same expression, replacing $\langle p q_{\mathrm{MA}} \rangle_t$ by $\langle p_{\mathrm{M}} p q_{\mathrm{MA}} \rangle_t$. Finally,the moments $\langle p q_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{MA}} \rangle$ and $\left\langle p q_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{M,A}} \right\rangle$ are obtained by solving:

$$\begin{aligned} \langle p_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{MA}} \rangle_{t+1} = & (1 - r\,\sigma) \langle p_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{MA}} \rangle_t + r\,\sigma \left\langle p_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{M,A}} \right\rangle_t \\ & - s\,h \left( \left\langle {D_{\mathrm{MA}}}^2 \right\rangle_t + \langle D_{\mathrm{MA}} D_{\mathrm{M,A}} \rangle_t \right) \end{aligned} \tag{B22}$$

$$\begin{aligned} \left\langle {p_{\mathrm{M}}}^2 D^{\mathrm{par}}_{\mathrm{MA}} \right\rangle_{t+1} = & (1 - r\,\sigma) \left\langle {p_{\mathrm{M}}}^2 D^{\mathrm{par}}_{\mathrm{MA}} \right\rangle_t + r\,\sigma \left\langle {p_{\mathrm{M}}}^2 D^{\mathrm{par}}_{\mathrm{M,A}} \right\rangle_t \\ & - 2\,s\,h \left( \left\langle p_{\mathrm{M}} {D_{\mathrm{MA}}}^2 \right\rangle_t + \langle p_{\mathrm{M}} D_{\mathrm{MA}} D_{\mathrm{M,A}} \rangle_t \right) \end{aligned} \tag{B23}$$

$$\left\langle p_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{M,A}} \right\rangle_{t+1} = (1 - \sigma) \left\langle p_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{M,A}} \right\rangle_t - s\,h \left( \langle D_{\mathrm{MA}} D_{\mathrm{M,A}} \rangle_t + \left\langle {D_{\mathrm{M,A}}}^2 \right\rangle_t \right) \tag{B24}$$

$$\left\langle {p_{\mathrm{M}}}^2 D^{\mathrm{par}}_{\mathrm{M,A}} \right\rangle_{t+1} = (1 - \sigma) \left\langle {p_{\mathrm{M}}}^2 D^{\mathrm{par}}_{\mathrm{M,A}} \right\rangle_t - 2\,s\,h \left( \langle p_{\mathrm{M}} D_{\mathrm{MA}} D_{\mathrm{M,A}} \rangle_t + \left\langle p_{\mathrm{M}} {D_{\mathrm{M,A}}}^2 \right\rangle_t \right) \tag{B25}$$

Expressions for associations $\langle D_{\mathrm{MA}} \rangle$ and $\langle D_{\mathrm{M,A}} \rangle$ are obtained by solving equations B17 and B18 at equilibrium, and plugging quasi-equilibrium values of other moments that appear in equations B17 and B18. In the case of an additive modifier ($h_{\mathrm{M}} = 1/2$), no gene conversion ($\gamma = 0$) and free recombination ($r = 1/2$), one obtains:

$$\langle D_{\mathrm{MA}} \rangle_t = -\left(\frac{\delta\sigma}{2}\right) \frac{s\,h\,(24 - 9\sigma + \sigma^2)\,\langle pq_{\mathrm{MA}} \rangle_t}{2N\sigma^3\,(4 - \sigma)} \tag{B26}$$

$$\langle D_{\mathrm{M,A}} \rangle_t = -\left(\frac{\delta\sigma}{2}\right) \frac{s\,h\,(3 - \sigma)\,\langle pq_{\mathrm{MA}} \rangle_t}{2N\sigma^3}\,. \tag{B27}$$

## ONLINE APPENDIX C: INTERPRETATION OF ASSOCIATIONS

We give here some intuitive interpretation of the different components of selection on the modifier. Denote $x_i$ ($i$ going from 0 to 9) the frequencies of the different diploid genotypes in the population, with $x_0 = \mathrm{fr}\,(ma/ma)$, $x_1 = \mathrm{fr}\,(Ma/ma)$, $x_2 = \mathrm{fr}\,(Ma/Ma)$, $x_3 = \mathrm{fr}\,(mA/ma)$, $x_4 = \mathrm{fr}\,(MA/ma)$, $x_5 = \mathrm{fr}\,(Ma/mA)$, $x_6 = \mathrm{fr}\,(MA/Ma)$, $x_7 = \mathrm{fr}\,(mA/mA)$, $x_8 = \mathrm{fr}\,(MA/mA)$, $x_9 = \mathrm{fr}\,(MA/MA)$. The relative fecundity $f_i$ of type $i$ is determined by its genotype at locus A, and is given by $f_0 = f_1 = f_2 = 1$, $f_3 = f_4 = f_5 = f_6 = 1 - h\,s$, $f_7 = f_8 = f_9 = 1 - s$. Calling $p_{M,i}$ the frequency of allele $M$ in genotype $i$ (which is either 0, 1/2 or 1), the expected change in frequency of the modifier at generation $t$ is given by:

$$\langle \Delta p_{\mathrm{M}} \rangle_t = \left\langle \frac{\sum_i f_i\, x_i\, p_{M,i}}{\sum_i f_i\, x_i} - p_{\mathrm{M}} \right\rangle_t . \tag{C1}$$

To the first order in $s$, this can be written as a sum of two terms:

$$\langle \Delta p_{\mathrm{M}} \rangle_t = -\frac{s}{2} \left\langle 2\, p_A\, p_a \left( p_{M|A} - p_{M|a} \right) \right\rangle_t - \frac{s}{2} (1 - 2h) \left\langle p_{\mathrm{Ho}}\, p_{\mathrm{He}} \left( p_{M|\mathrm{Ho}} - p_{M|\mathrm{He}} \right) \right\rangle_t . \tag{C2}$$

The first term of equation C2 represents selection acting on $M$ due to association between $M$ and the deleterious allele $A$: $p_A$ and $p_a$ are the frequencies of alleles $A$ and $a$ in the population, while $p_{M|A} = \left[ (x_4 + x_5)/2 + x_6 + x_8 + 2x_9 \right] / (2\, p_A)$ represents the probability that when sampling one gene at each locus from the same individual (either on the same or on different haplotypes) and given that allele $A$ is sampled at the selected locus, one samples allele $M$ at the modifier locus. Similarly, $p_{M|a} = \left[ (x_4 + x_5)/2 + x_1 + x_6 + 2x_2 \right] / (2\, p_A)$ represents the probability that given allele $a$ is sampled at the selected locus, $M$ is sampled at the modifier locus. When $p_{M|A} - p_{M|a} > 0$, allele $M$ is thus more often associated with the deleterious allele

$A$ (on the same or on the other haplotype) than by chance, which selects against $M$ through the first term of equation C2. The second term of equation C2 represents selection acting on $M$ due to the association between $M$ and homozygotes or heterozygotes at locus A: $p_{\mathrm{He}} = x_3 + x_4 + x_5 + x_6$ and $p_{\mathrm{Ho}} = 1 - p_{\mathrm{He}}$ are the frequencies of heterozygotes and homozygotes at locus A, while $p_{M|\mathrm{He}} = \left[\left(x_4 + x_5\right)/2 + x_6\right]/p_{\mathrm{He}}$ and $p_{M|\mathrm{Ho}} = \left[\left(x_1 + x_8\right)/2 + x_2 + x_9\right]/p_{\mathrm{Ho}}$ measure the frequency of $M$ among heterozygotes and among homozygotes at locus A. When $p_{M|\mathrm{Ho}} - p_{M|\mathrm{He}} > 0$, allele $M$ is more frequent in homozygotes than in heterozygotes at locus A, which favors $M$ when homozygotes have a higher fecundity (note that $-\left(s/2\right)\left(1-2h\right)$ is the average fecundity of homozygotes, minus the fecundity of heterozygotes).

Assuming that $s$, $\gamma$ and $1/N$ are all of order $\epsilon$, equation C2 also gives the expected change in frequency of $M$ at quasi-equilibrium to the second order in $\epsilon$. Indeed, expanding equation C1 to the second order in $s$ yields a term in $s^2$ that equals zero in the absence of any association in the population. Because associations are of order $\epsilon$ at quasi-equilibrium, this term is of order $\epsilon^3$, and thus negligible. In order to compute the two terms of equation C2 at quasi-equilibrium, it is easier to rewrite them in terms of genetic associations. One obtains:

$$p_A\, p_a \left(p_{M|A} - p_{M|a}\right) = D_{\mathrm{MA}} + D_{\mathrm{M,A}} \tag{C3}$$

$$\frac{1}{2} p_{\mathrm{Ho}}\, p_{\mathrm{He}} \left(p_{M|\mathrm{Ho}} - p_{M|\mathrm{He}}\right) = D_{\mathrm{MA,A}} - \frac{1}{2}\left(1 - 2p_{\mathrm{A}}\right)\left(D_{\mathrm{MA}} + D_{\mathrm{M,A}}\right) \tag{C4}$$

(this can be checked by expressing $D_{\mathrm{MA}}$, $D_{\mathrm{M,A}}$ and $D_{\mathrm{MA,A}}$ in terms of genotype frequencies $x_i$). The methods developed in Appendix A can then be used to derive recursions for the expected values of these associations, and obtain expressions at

quasi-equilibrium. One finds that $\langle D_{\mathrm{MA,A}}\rangle$ is of order $\delta\sigma\,\epsilon$ at quasi-equilibrium, while $\langle D_{\mathrm{MA}}\rangle$, $\langle D_{\mathrm{M,A}}\rangle$, $\langle p_{\mathrm{A}} D_{\mathrm{MA}}\rangle$ and $\langle p_{\mathrm{A}} D_{\mathrm{M,A}}\rangle$ are of order $\delta\sigma\,\epsilon^2$. Therefore, the first term of the change in frequency of $M$ (equation C2) is negligible compared to the second term, and this second term equals $-s\,(1-2h)\,\langle D_{\mathrm{MA,A}}\rangle_t$ to the second order in $\epsilon$. This also shows that to the first order in $\epsilon$, $\langle D_{\mathrm{MA,A}}\rangle$ equals $\left\langle p_{\mathrm{Ho}}\,p_{\mathrm{He}}\left(p_{M|\mathrm{Ho}}-p_{M|\mathrm{He}}\right)/2\right\rangle$ at quasi-equilibrium (equation C4). A positive value of $\langle D_{\mathrm{MA,A}}\rangle$ therefore means that the frequency of $M$ tends to be higher in homozygotes than in heterozygotes at locus A, while a negative value of $\langle D_{\mathrm{MA,A}}\rangle$ means that the frequency of $M$ tends to be higher in heterozygotes.

When dominance is weak ($h-1/2$ is of order $\epsilon$), the two components of selection on the modifier become of the same order of magnitude. In that case, the expected change in frequency of the modifier becomes:

$$\begin{aligned}\langle \Delta p_{\mathrm{M}}\rangle_t = -\frac{s}{2}\left\langle (1+s\,p_A)\,p_A\,p_a\left(p_{M|A}-p_{M|a}\right)\right\rangle_t \\ -\frac{s}{2}\,(1-2h)\left\langle p_{\mathrm{Ho}}\,p_{\mathrm{He}}\left(p_{M|\mathrm{Ho}}-p_{M|\mathrm{He}}\right)\right\rangle_t + o\left(\epsilon^3\right).\end{aligned} \tag{C5}$$

Again, these two components can be computed at quasi-equilibrium by expressing them in terms of genetic associations. As before, $\langle D_{\mathrm{MA,A}}\rangle$ is of order $\delta\sigma\,\epsilon$ at quasi-equilibrium, while $\langle D_{\mathrm{MA}}\rangle$, $\langle D_{\mathrm{M,A}}\rangle$, $\langle p_{\mathrm{A}} D_{\mathrm{MA}}\rangle$ and $\langle p_{\mathrm{A}} D_{\mathrm{M,A}}\rangle$ are of order $\delta\sigma\,\epsilon^2$. The factor $1+s\,p_A$ in the first term of equation C5 can thus be neglected (as it generates terms of order $\delta\sigma\,\epsilon^4$), and the first term becomes equivalent to $-\,(s/2)\left(\langle D_{\mathrm{MA}}\rangle_t+\langle D_{\mathrm{M,A}}\rangle_t\right)$, while the second term is again equivalent to $-s\,(1-2h)\,\langle D_{\mathrm{MA,A}}\rangle_t$.

## ONLINE APPENDIX D: THREE-LOCUS MODEL

In order to study the effect of interactions among pairs of selected loci on selection for sex, we extended our two-locus model to include a second selected locus (called locus B). We assume that selection is multiplicative across loci A and B (no epistasis), with the same selection and dominance coefficients at both loci. Expressions for the change in frequency of the modifier as a function of genetic associations, and expressions for associations at quasi-equilibrium can be computed using the methods of Appendix A. To keep the analysis tractable, we focused on the case of an additive modifier ($h_{\mathrm{M}} = 1/2$), and do not consider effects of gene conversion ($\gamma = 0$). We will also express all terms to the first order in $1/N$ (assuming $N$ is sufficiently large so that terms in $1/N^2$ can be neglected), and to the first order in the modifier effect $\delta\sigma$. In a first part, we assume that selection is weak (of order $\epsilon$) and express the change in frequency of the modifier to leading order in $\epsilon$, while in a second part we will consider the case where both selection and dominance are weak: $s = O(\epsilon)$, $h - 1/2 = O(\epsilon)$.

**Weak selection.** To leading order in $\epsilon$, the change in frequency of the modifier is given by:

$$\langle \Delta p_{\mathrm{M}} \rangle_t \approx -s(1-2h)\left[\langle D_{\mathrm{MA,A}} \rangle_t + \langle D_{\mathrm{MB,B}} \rangle_t\right] + s^2 (1-2h)^2 \langle D_{\mathrm{MAB,AB}} \rangle_t \qquad \text{(D1)}$$

Association $\langle D_{\mathrm{MAB,AB}} \rangle$ is generated by drift and by the modifier effect. This association is negative for a modifier increasing sex, reflecting the fact that sex tends to break correlations in homozygosity between loci A and B (which disfavors sex whenever $h \neq 1/2$). The solution obtained at quasi-equilibrium takes the form:

$$\langle D_{\mathrm{MAB,AB}} \rangle_t = -\frac{\delta\sigma}{2N\sigma^2} T_4 \langle pq_{\mathrm{MAB}} \rangle_t \qquad \text{(D2)}$$

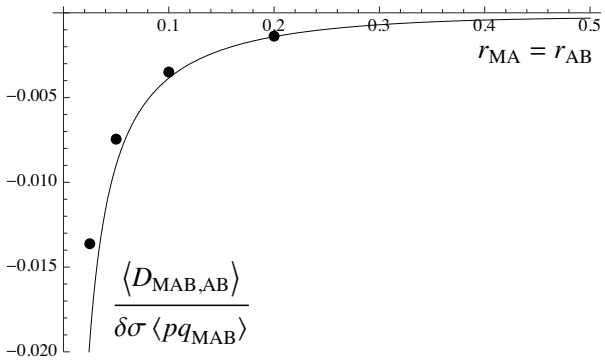
where $T_4$ is a positive (and complicated) function of recombination rates $r_{\mathrm{MA}}$ and $r_{\mathrm{AB}}$ (we assume that loci are in the order M - A - B along the chromosome) and of the baseline rate of sex $\sigma$. When $r_{\mathrm{MA}} = 1/2$, $T_4$ simplifies to

$$T_4 = \frac{[1 - 2r_{\mathrm{AB}}(1 - r_{\mathrm{AB}})]\,[2 - \sigma(1 + 2r_{\mathrm{AB}})]}{(1 + r_{\mathrm{AB}})\,(2 - \sigma\, r_{\mathrm{AB}})\,[4 - \sigma(1 + r_{\mathrm{AB}})]} \tag{D3}$$

while when $r_{\mathrm{MA}}$ and $r_{\mathrm{AB}}$ both equal 1/2 it becomes:

$$T_4 = \frac{8(1 - \sigma)}{3(4 - \sigma)(8 - 3\sigma)}\,. \tag{D4}$$

A simulation test of our expression for $\langle D_{\mathrm{MAB,AB}} \rangle$ against simulations (in which loci A and B are neutral) is shown on the following figure, for $N = 3000$, $\sigma = 0.2$, $\delta\sigma = 0.03$ (curve: analytical prediction; dots: simulation results, averages over $10^7$ points).




Note that in an infinite population, $D_{\mathrm{MAB,AB}}$ is generated by selection and by the modifier effect, and takes the form at quasi-equilibrium:

$$D_{\mathrm{MAB,AB}} = -\frac{\delta\sigma}{2\sigma^2}\, s^2 (1 - 2h)^2\, pq_{\mathrm{M}}\, (pq_{\mathrm{AB}})^2\,. \tag{D5}$$

However, in a finite population this term should be negligible compared to the term shown in equation D2 when deleterious mutations remain in low frequency, unless $N$ is really extremely large.

One can also note that the second term of equation D1 (involving $\langle D_{\mathrm{MAB,AB}}\rangle$) is of order $\delta\sigma\,\epsilon^2/N$, while the first term is of order $\delta\sigma\,(\epsilon^2+\epsilon/N)$. However, we will assume that when summed over a large number of segregating loci, the effect of interactions between pairs of selected loci may become of the same order of magnitude as the effect of individual loci (because the number of pairs of loci will be much greater than the number of individual loci).

Associations $\langle D_{\mathrm{MA,A}}\rangle$ and $\langle D_{\mathrm{MB,B}}\rangle$ are also affected by three-locus associations. A recursion for $\langle D_{\mathrm{MA,A}}\rangle$ is given by:

$$\begin{aligned}\langle D_{\mathrm{MA,A}}\rangle_{t+1} = (1-\sigma)\left[\langle D_{\mathrm{MA,A}}\rangle_t - s\,(1-2h)\,\langle D_{\mathrm{MAB,AB}}\rangle_t\right] \\ + \frac{\delta\sigma}{2}\Big[(1-r)\left\langle {D_{\mathrm{MA}}}^2\right\rangle_t + \langle D_{\mathrm{MA}}D_{\mathrm{M,A}}\rangle_t + r\left\langle {D_{\mathrm{M,A}}}^2\right\rangle_t \\ - \left\langle D^{\mathrm{par}}_{\mathrm{MA,MA}}\right\rangle_t - \left\langle pq_{\mathrm{M}}D^{\mathrm{par}}_{\mathrm{A,A}}\right\rangle_t\Big]\end{aligned} \tag{D6}$$

where associations $\left\langle D^{\mathrm{par}}_{\mathrm{MA,MA}}\right\rangle$ and $\left\langle pq_{\mathrm{M}}D^{\mathrm{par}}_{\mathrm{A,A}}\right\rangle$ are measured after selection. To leading order in $s$ and $1/N$, expressions for $\langle {D_{\mathrm{MA}}}^2\rangle$, $\langle D_{\mathrm{MA}}D_{\mathrm{M,A}}\rangle$ and $\langle {D_{\mathrm{M,A}}}^2\rangle$ at quasi-equilibrium are not affected by three-locus interactions, and are still given by equations B10 to B12. Expressions for $\left\langle D^{\mathrm{par}}_{\mathrm{MA,MA}}\right\rangle$ and $\left\langle pq_{\mathrm{M}}D^{\mathrm{par}}_{\mathrm{A,A}}\right\rangle$ are obtained by solving the recursions:

$$\begin{aligned}\left\langle pq_{\mathrm{M}}D^{\mathrm{par}}_{\mathrm{A,A}}\right\rangle_{t+1} = (1-\sigma)\left\langle pq_{\mathrm{M}}D^{\mathrm{cnv}}_{\mathrm{A,A}}\right\rangle_t - \frac{\langle pq_{\mathrm{MA}}\rangle_t}{2N} \\ - s\,(1-2h)\left\langle pq_{\mathrm{M}}\,(pq_{\mathrm{A}})^2\right\rangle_t - s\,(1-2h)\,\langle pq_{\mathrm{M}}D_{\mathrm{AB,AB}}\rangle_t\end{aligned} \tag{D7}$$

$$\begin{aligned}\left\langle D^{\mathrm{par}}_{\mathrm{MA,MA}}\right\rangle_{t+1} = (1-\sigma)\,\langle D_{\mathrm{MA,MA}}\rangle_t - s\,(1-2h)\,\langle D_{\mathrm{MAB,MAB}}\rangle_t \\ + \sigma\left[(1-r)^2\left\langle {D_{\mathrm{MA}}}^2\right\rangle_t + 2r\,(1-r)\,\langle D_{\mathrm{MA}}D_{\mathrm{M,A}}\rangle_t + r^2\left\langle {D_{\mathrm{M,A}}}^2\right\rangle_t\right]\end{aligned} \tag{D8}$$

where $\langle pq_{\mathrm{M}}D_{\mathrm{AB,AB}}\rangle$ and $\langle D_{\mathrm{MAB,MAB}}\rangle$ at quasi-equilibrium are given by:

$$\langle pq_{\mathrm{M}}D_{\mathrm{AB,AB}}\rangle_t = \frac{[1-2r_{\mathrm{AB}}\,(1-r_{\mathrm{AB}})]\,\langle pq_{\mathrm{MAB}}\rangle_t}{2Nr_{\mathrm{AB}}\,\sigma\,(2-r_{\mathrm{AB}}\,\sigma)} \tag{D9}$$

$$\langle D_{\mathrm{MAB,MAB}}\rangle_t = \frac{[1-2r_{\mathrm{MA}}(1-r_{\mathrm{MA}})][1-2r_{\mathrm{AB}}(1-r_{\mathrm{AB}})]\langle pq_{\mathrm{MAB}}\rangle_t}{2N\sigma(r_{\mathrm{MA}}+r_{\mathrm{AB}}-r_{\mathrm{MA}}r_{\mathrm{AB}})[2-\sigma(r_{\mathrm{MA}}+r_{\mathrm{AB}}-r_{\mathrm{MA}}r_{\mathrm{AB}})]} \quad \text{(D10)}$$

From equations D6 to D10, one obtains at quasi-equilibrium:

$$\begin{aligned}\langle D_{\mathrm{MA,A}}\rangle_t \approx \frac{\delta\sigma}{2\sigma^2}\left[\left(\frac{T_1}{2N}-\gamma\right)\langle pq_{\mathrm{MA}}\rangle_t + s(1-2h)\left\langle pq_{\mathrm{M}}\,{pq_{\mathrm{A}}}^2\right\rangle_t\right] \\ + \frac{\delta\sigma}{2N\sigma^3}s(1-2h)\sum_{\mathrm{B}}T_5\langle pq_{\mathrm{MAB}}\rangle\end{aligned} \quad \text{(D11)}$$

where $T_1$ is given by equation 6 in the main text, and $T_5$ is a positive (and complicated) function of $r_{\mathrm{MA}}$, $r_{\mathrm{AB}}$ and $\sigma$. This term in $T_5$ (together with the second term of equation D1) generates selection against a modifier increasing sex whenever $h \neq 1/2$. Again, although each term in $\langle pq_{\mathrm{MAB}}\rangle$ that appears in $\langle\Delta p_{\mathrm{M}}\rangle$ is smaller in magnitude than terms in $\langle pq_{\mathrm{MA}}\rangle$, $\langle pq_{\mathrm{MB}}\rangle$, their overall effect may be important when many loci segregate for deleterious alleles.

**Weak dominance.** In the case where dominance is weak ($h-1/2$ is of order $\epsilon$), the change in frequency of the modifier is affected by many associations. One obtains (dropping "t" indices):

$$\langle\Delta p_{\mathrm{M}}\rangle = \Xi_1+\Xi_2+\Xi_3+\Xi_4 \quad \text{(D12)}$$

with:

$$\Xi_1 = -sh[\langle D_{\mathrm{MA}}\rangle+\langle D_{\mathrm{M,A}}\rangle] - sh[\langle D_{\mathrm{MB}}\rangle+\langle D_{\mathrm{M,B}}\rangle] \quad \text{(D13)}$$

$$\begin{aligned}\Xi_2 = &-s(1-2h)[\langle D_{\mathrm{MA,A}}\rangle+\langle D_{\mathrm{MB,B}}\rangle] \\ &+(sh)^2[\langle D_{\mathrm{MAB}}\rangle+\langle D_{\mathrm{MA,B}}\rangle+\langle D_{\mathrm{MB,A}}\rangle+\langle D_{\mathrm{M,AB}}\rangle] \\ &-s\left(1-2h+2sh^2\right)[\langle p_{\mathrm{A}}D_{\mathrm{MA}}\rangle+\langle p_{\mathrm{A}}D_{\mathrm{M,A}}\rangle+\langle p_{\mathrm{B}}D_{\mathrm{MB}}\rangle+\langle p_{\mathrm{B}}D_{\mathrm{M,B}}\rangle]\end{aligned} \quad \text{(D14)}$$

$$\begin{aligned}\Xi_3 = {} & s^2h\,(1-2h)\left[\langle D_{\mathrm{MAB,A}}\rangle + \langle D_{\mathrm{MA,AB}}\rangle + \langle D_{\mathrm{MAB,B}}\rangle + \langle D_{\mathrm{MB,AB}}\rangle\right] \\ & + 2\,(sh)^3\left[\langle D_{\mathrm{MA}}D_{\mathrm{AB}}\rangle + \langle D_{\mathrm{M,A}}D_{\mathrm{AB}}\rangle + \langle D_{\mathrm{MA}}D_{\mathrm{A,B}}\rangle + \langle D_{\mathrm{M,A}}D_{\mathrm{A,B}}\rangle\right. \\ & \qquad \left. + \langle D_{\mathrm{MB}}D_{\mathrm{AB}}\rangle + \langle D_{\mathrm{M,B}}D_{\mathrm{AB}}\rangle + \langle D_{\mathrm{MB}}D_{\mathrm{A,B}}\rangle + \langle D_{\mathrm{M,B}}D_{\mathrm{A,B}}\rangle\right] \\ & + s^2h\left(1-2h+2sh^2\right)\left[\langle p_{\mathrm{A}}D_{\mathrm{MAB}}\rangle + \langle p_{\mathrm{A}}D_{\mathrm{MA,B}}\rangle + \langle p_{\mathrm{A}}D_{\mathrm{MB,A}}\rangle + \langle p_{\mathrm{A}}D_{\mathrm{M,AB}}\rangle\right. \\ & \qquad \left. + \langle p_{\mathrm{B}}D_{\mathrm{MAB}}\rangle + \langle p_{\mathrm{B}}D_{\mathrm{MA,B}}\rangle + \langle p_{\mathrm{B}}D_{\mathrm{MB,A}}\rangle + \langle p_{\mathrm{B}}D_{\mathrm{M,AB}}\rangle\right]\end{aligned} \tag{D15}$$

$$\begin{aligned}\Xi_4 = {} & s^2\,(1-2h)^2\,\langle D_{\mathrm{MAB,AB}}\rangle \\ & - 2\,(sh)^4\left[\langle D_{\mathrm{MAB}}D_{\mathrm{AB}}\rangle + \langle D_{\mathrm{MA,B}}D_{\mathrm{AB}}\rangle + \langle D_{\mathrm{MB,A}}D_{\mathrm{AB}}\rangle + \langle D_{\mathrm{M,AB}}D_{\mathrm{AB}}\rangle\right. \\ & \qquad \left. + \langle D_{\mathrm{MAB}}D_{\mathrm{A,B}}\rangle + \langle D_{\mathrm{MA,B}}D_{\mathrm{A,B}}\rangle + \langle D_{\mathrm{MB,A}}D_{\mathrm{A,B}}\rangle + \langle D_{\mathrm{M,AB}}D_{\mathrm{A,B}}\rangle\right].\end{aligned} \tag{D16}$$

Associations that appear in$\Xi_4$ are of order $\delta\sigma/N$, and result from the effect of the modifier on neutral associations such as $\langle D_{\mathrm{AB,AB}}\rangle$, $\langle D_{\mathrm{MAB,MAB}}\rangle$, $\langle {D_{\mathrm{MAB}}}^2\rangle$... These associations affect the change in frequency of the modifier directly through$\Xi_4$, and indirectly by affecting associations that appear in$\Xi_1$, $\Xi_2$ and$\Xi_3$. Associations that appear in$\Xi_3$ are of order $\delta\sigma\,\epsilon/N$. These are generated by the associations that appear in$\Xi_4$ and by selection, and also by the effect of the modifier on associations generated by selection and drift, such as $\langle D_{\mathrm{AB,A}}\rangle$, $\langle p_{\mathrm{A}}D_{\mathrm{AB}}\rangle$, $\langle D_{\mathrm{MAB,MA}}\rangle$... (these associations are of order $\epsilon/N$). Associations in$\Xi_3$ also affect the associations that appear in$\Xi_1$ and $\Xi_2$. As we have seen in the two-locus model (Appendix B), associations $\langle D_{\mathrm{MA,A}}\rangle$ and $\langle D_{\mathrm{MB,B}}\rangle$ that appear in$\Xi_2$ are generated by two-locus interactions (between M and A, and between M and B), resulting from the effect of the modifier on associations $\langle D_{\mathrm{A,A}}\rangle$, $\langle D_{\mathrm{MA,MA}}\rangle$, $\langle {D_{\mathrm{MA}}}^2\rangle$... These associations are also affected by three-locus interactions through terms in $\delta\sigma\,\epsilon^2/N$, due to the fact that associations $\langle D_{\mathrm{A,A}}\rangle$, $\langle D_{\mathrm{MA,MA}}\rangle$,

$\langle {D_{\mathrm{MA}}}^2 \rangle$... are affected by selection at locus B through terms in $\epsilon^2/N$, and also to the fact that selection converts some of the associations that appear in$\Xi_3$ and$\Xi_4$ into $\langle D_{\mathrm{MA,A}} \rangle$ and $\langle D_{\mathrm{MB,B}} \rangle$. The other associations that appear in$\Xi_2$ are also generated by associations that appear in$\Xi_3$ and$\Xi_4$ and by selection (in particular, the moments $\langle p_{\mathrm{A}} D_{\mathrm{MA}} \rangle$ and $\langle p_{\mathrm{A}} D_{\mathrm{M,A}} \rangle$ were negligible in the two-locus model, but are now generated by three-locus associations that appear in$\Xi_3$ and$\Xi_4$). Additionally, associations $\langle D_{\mathrm{MAB}} \rangle$, $\langle D_{\mathrm{MA,B}} \rangle$, $\langle D_{\mathrm{MB,A}} \rangle$ and $\langle D_{\mathrm{M,AB}} \rangle$ also result from the effect of the modifier on associations generated by selection and drift, such as $\langle D_{\mathrm{AB}} \rangle$, $\langle D_{\mathrm{A,B}} \rangle$, $\langle D_{\mathrm{MAB,M}} \rangle$, $\langle D_{\mathrm{MA}} D_{\mathrm{MB}} \rangle$... (these associations are of order $\epsilon^2/N$). Finally, we have seen in Appendix B that associations that appear in$\Xi_1$ are generated by two-locus interactions, through the combined action of associations $\langle D_{\mathrm{MA,A}} \rangle$ and $\langle D_{\mathrm{MB,B}} \rangle$ and selection, and through the effect of the modifier on associations $\langle D_{\mathrm{MA,M}} \rangle$ and $\langle D_{\mathrm{MB,M}} \rangle$ generated by selection and drift. Associations in$\Xi_1$ are also affected by three-locus interactions (through terms in $\delta\sigma\,\epsilon^3/N$), due to the fact that selection at locus B affects $\langle D_{\mathrm{MA,M}} \rangle$ through terms in $\epsilon^3/N$ (and similarly for $\langle D_{\mathrm{MB,M}} \rangle$), to the fact that $\langle D_{\mathrm{MA,A}} \rangle$ and $\langle D_{\mathrm{MB,B}} \rangle$ are affected by three-locus interactions (as we have just discussed), and also to the fact that other associations that appear in$\Xi_2$, $\Xi_3$ and$\Xi_4$ influence associations in$\Xi_1$. As a consequence, one obtains that three-locus interactions affect the change in frequency of the modifier through terms in $\delta\sigma\,\epsilon^4 \langle pq_{\mathrm{MAB}} \rangle /N$.

For space reasons we do not provide expressions for the differents associations that appear above, but these can be found in a *Mathematica* notebook available on request, for both cases where loci are in order M − A − B and in order A − M − B. In Figures D1 to D3, we analyze different components of selection for sex due to three-locus interactions. Figure D1 shows that positive selection for sex occurs

mainly through associations $\langle D_{\mathrm{MA}} \rangle$, $\langle D_{\mathrm{M,A}} \rangle$, $\langle D_{\mathrm{MB}} \rangle$ and $\langle D_{\mathrm{M,B}} \rangle$ ($\Xi_1$ term), the$\Xi_2$ term selecting against sex unless $h$ is close to 1/2, and$\Xi_3$ and$\Xi_4$ having little effect. Note that components of selection for sex due to two-locus interactions (terms in $\langle pq_{\mathrm{MA}} \rangle$, $\langle pq_{\mathrm{MB}} \rangle$ that have been derived in the two-locus model) have been removed here, in order to focus on the effect of three-locus interactions (terms in $\langle pq_{\mathrm{MAB}} \rangle$). However, we have just seen that associations in$\Xi_1$ are generated by all associations that appear in$\Xi_2$, $\Xi_3$ and$\Xi_4$. Figure D2 shows another decomposition of the effect of three-locus interactions on the change in frequency of the modifier: the solid curves show selection on the modifier due to the breaking of neutral associations such as $\langle D_{\mathrm{AB,AB}} \rangle$, $\langle D_{\mathrm{MAB,MAB}} \rangle$, $\langle {D_{\mathrm{MAB}}}^2 \rangle$ ($\Xi_4$ term, plus the effect of associations that appear in$\Xi_4$ on$\Xi_1$, $\Xi_2$, $\Xi_3$), the dotted curve shows selection due to the breaking of associations of order $\epsilon/N$, such as $\langle D_{\mathrm{AB,A}} \rangle$, $\langle p_{\mathrm{A}} D_{\mathrm{AB}} \rangle$, $\langle D_{\mathrm{MAB,MA}} \rangle$ (through some of the terms of$\Xi_1$, $\Xi_2$, $\Xi_3$), while the dashed curve shows selection due to the breaking of associations of order $\epsilon^2/N$, such as $\langle D_{\mathrm{AB}} \rangle$, $\langle D_{\mathrm{A,B}} \rangle$, $\langle D_{\mathrm{MAB,M}} \rangle$, $\langle D_{\mathrm{MA}} D_{\mathrm{MB}} \rangle$ (through some of the terms of$\Xi_1$ and$\Xi_2$). Finally, the dashed-single-dotted curves show the effect of three-locus interactions on selection on the modifier through its effect on associations $\langle D_{\mathrm{A,A}} \rangle$, $\langle D_{\mathrm{B,B}} \rangle$, $\langle {D_{\mathrm{MA}}}^2 \rangle$, $\langle {D_{\mathrm{MB}}}^2 \rangle$... (that generate associations $\langle D_{\mathrm{MA,A}} \rangle$, $\langle D_{\mathrm{MB,B}} \rangle$), while the dashed-double-dotted curves show the effect of three-locus interactions on selection on the modifier through its effect on associations $\langle D_{\mathrm{MA,M}} \rangle$, $\langle D_{\mathrm{MB,M}} \rangle$ (that affect associations in$\Xi_1$). Figure D2 shows that when dominance is weak ($h$ close to 1/2), components of selection for sex due to the breaking of associations of order $1/N$ (solid), $\epsilon/N$ (dotted) and $\epsilon^2/N$ (dashed) are of similar strength. Figure D3 shows the strength of selection for sex generated by the linkage disequilibrium $\langle D_{\mathrm{AB}} \rangle$, relative to selection generated by other associations involving loci A and B. As can be seen on figure D3, the effect of

$\langle D_{\mathrm{AB}} \rangle$ is relatively minor, in particular when $h$ is close to 1/2.

Finally, we use our model to obtain extrapolations about selection on a sex modifier when deleterious mutations occur at a large number of loci within the genome. From our quasi-equilibrium expressions, and assuming free recombination among all loci, the selection gradient acting on the modifier takes the form:

$$s_{\mathrm{M}} \equiv \frac{\langle \Delta p_{\mathrm{M}} \rangle}{\frac{\delta\sigma}{2} \langle pq_{\mathrm{M}} \rangle} = \frac{\Omega_1}{N} \sum_{\mathrm{A}} \langle pq_{\mathrm{A}} \rangle + \Omega_2 \sum_{\mathrm{A}} \langle pq_{\mathrm{A}}{}^2 \rangle + \frac{\Omega_3}{N} \sum_{\mathrm{A,B}} \langle pq_{\mathrm{AB}} \rangle \qquad \text{(D17)}$$

where the$\Omega$ s are functions of $s$, $h$ and $\sigma$, and where the first two terms correspond to the stochastic and deterministic terms of the two-locus model, while the third term is the effect of three-locus interactions ($\Omega_1$ is of order $\epsilon^2$, while$\Omega_2$ and$\Omega_3$ are of order $\epsilon^4$); sums are over all segregating loci. We performed numerical integrations over Wright's distribution to evaluate $\sum_{\mathrm{A}} \langle pq_{\mathrm{A}} \rangle$ and $\sum_{\mathrm{A}} \langle pq_{\mathrm{A}}{}^2 \rangle$ (which converge to finite limits when the number of loci gets very large and the mutation rate per locus very small), and replaced $\sum_{\mathrm{A,B}} \langle pq_{\mathrm{AB}} \rangle$ by $(1/2) \left[\sum_{\mathrm{A}} \langle pq_{\mathrm{A}} \rangle\right]^2$ (simulations indicate that $U/(sh)$ and $(1/2)\left[U/(sh)\right]^2$ are often good approximations of $\sum_{\mathrm{A}} \langle pq_{\mathrm{A}} \rangle$ and $\sum_{\mathrm{A,B}} \langle pq_{\mathrm{AB}} \rangle$ as long as $s \gg u, 1/N$ and $h$ is not too small). With linkage, integrating the selection gradient over a given map length is difficult because some quasi-equilibrium expressions diverge when $r_{\mathrm{MA}}$, $r_{\mathrm{AB}}$ tend to zero. In order to obtain an idea of the effect of increasing linkage, we considered an artificial situation where the recombination rate is 0.1 between the modifier and each selected locus (setting $r_{\mathrm{MA}} = r_{\mathrm{MB}} = 0.1$ in the model where loci are in order $\mathrm{A} - \mathrm{M} - \mathrm{B}$). Although this situation is quite artificial, comparing the results with the case of free recombination gives an idea of how increasing linkage affects the relative strength of the different terms in equation D17.

Results are shown on figures D4, D5 and D6, for different values of $U$ and $s$.

These figures show that the effect of three-locus interactions (third term of equation D17) become relatively strong as $U$ increases (compare top and bottom panels), as $\sigma$ decreases (compare D4 and D5), and as linkage becomes tighter (compare D4 and D6). As $s$ decreases, the parameter range in which sex is favored tends to shrink, in particular when $U$ is high (compare left and right panels).

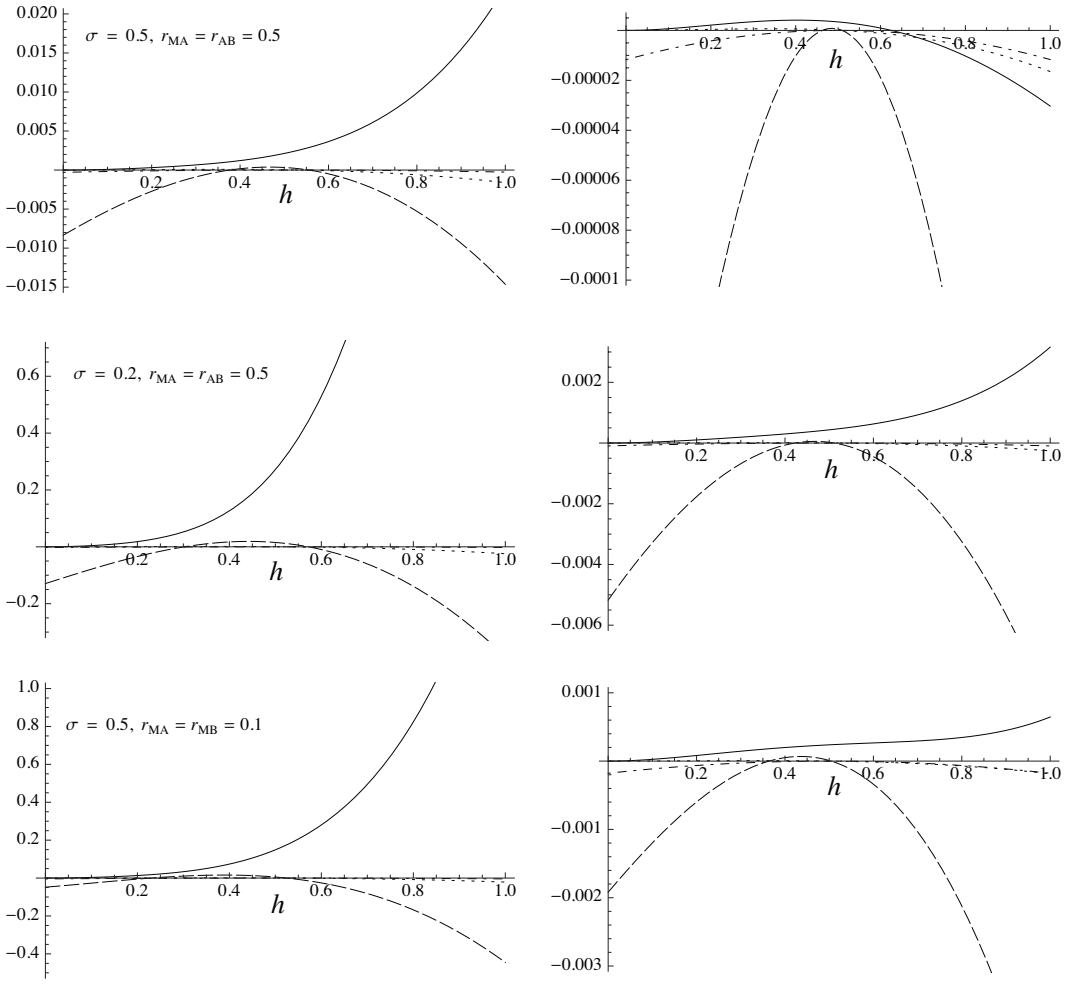



**Figure D1.** Effect of three-locus interactions on selection for sex. The different curves show different components of the effect of three-locus interactions on the change in frequency of the modifier (terms in $\langle pq_{\mathrm{MAB}} \rangle$ of $\langle \Delta p_{\mathrm{M}} \rangle$): solid curves:$\Xi_1$ term, dashed curves:$\Xi_2$, dotted curves:$\Xi_3$, dotted-dashed curves:$\Xi_4$. Note that terms generated by two-locus interactions (terms in $\langle pq_{\mathrm{MA}} \rangle$ and $\langle pq_{\mathrm{MB}} \rangle$ that that affect$\Xi_1$ and$\Xi_2$) have been removed. The different terms have been divided by $\delta\sigma \langle pq_{\mathrm{MAB}} \rangle / N$. Top row: $\sigma = 0.5$, $r_{\mathrm{MA}} = r_{\mathrm{AB}} = 0.5$, middle row: $\sigma = 0.2$, $r_{\mathrm{MA}} = r_{\mathrm{AB}} = 0.5$, bottom row: $\sigma = 0.5$, $r_{\mathrm{MA}} = r_{\mathrm{MB}} = 0.1$, loci in order A − M − B. Left: $s = 0.05$, right: $s = 0.01$.

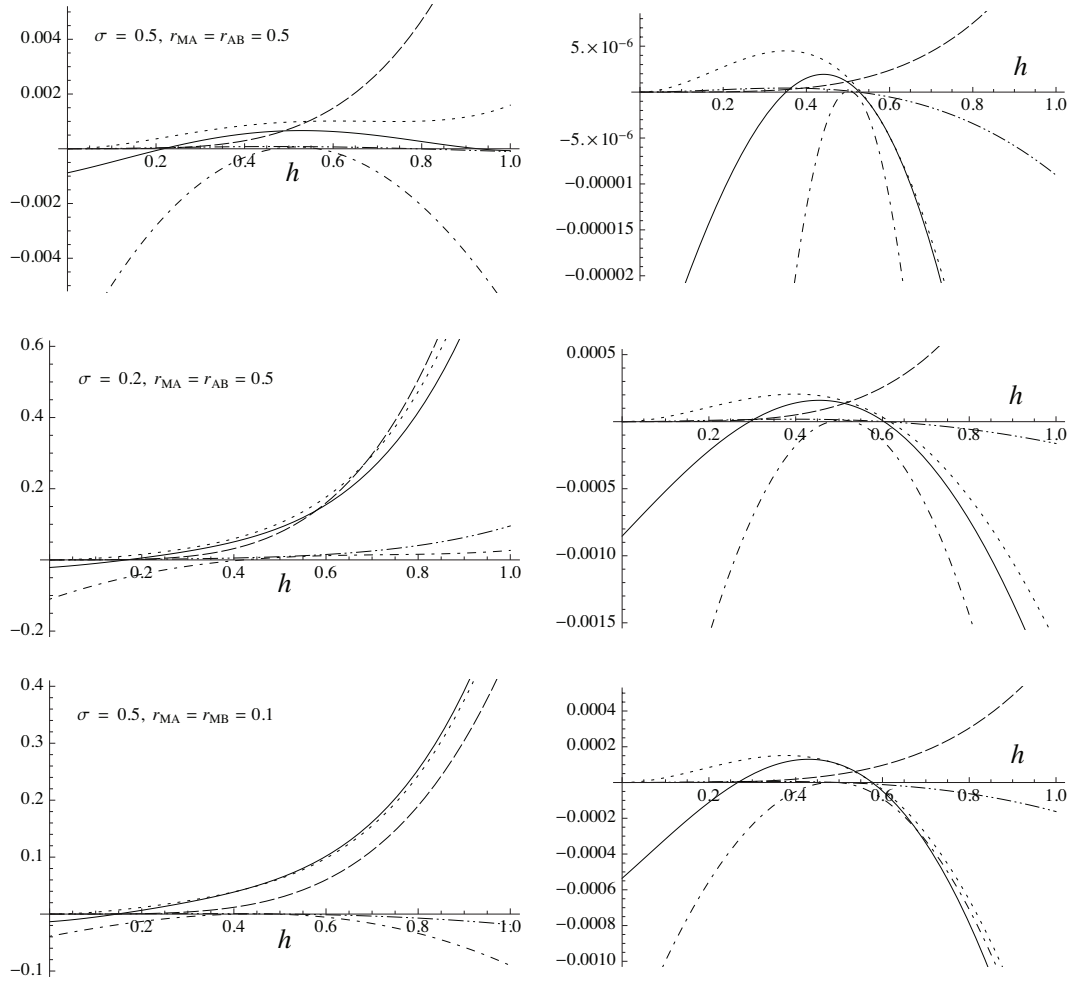



**Figure D2.** Selection generated by breaking neutral associations involving loci A and B (solid curves), associations of order $\epsilon/N$ (dotted curves) and associations of order $\epsilon^2/N$ (dashed curves). Dashed-single-dotted curves: selection generated by the effect of three locus interactions on associations $\langle D_{\mathrm{A,A}}\rangle$, $\langle D_{\mathrm{B,B}}\rangle$, $\langle {D_{\mathrm{MA}}}^2\rangle$... that affect $\langle D_{\mathrm{MA,A}}\rangle$, $\langle D_{\mathrm{MB,B}}\rangle$. Dashed-double-dotted curves: selection generated by the effect of three locus interactions on associations $\langle D_{\mathrm{MA,M}}\rangle$, $\langle D_{\mathrm{MB,M}}\rangle$ that affect $\langle D_{\mathrm{MA}}\rangle$, $\langle D_{\mathrm{MB}}\rangle$, $\langle D_{\mathrm{M,A}}\rangle$, $\langle D_{\mathrm{M,B}}\rangle$. The different components have been scaled by $\delta\sigma\left\langle pq_{\mathrm{MAB}}\right\rangle/N$. Same parameter values as in Figure D1.

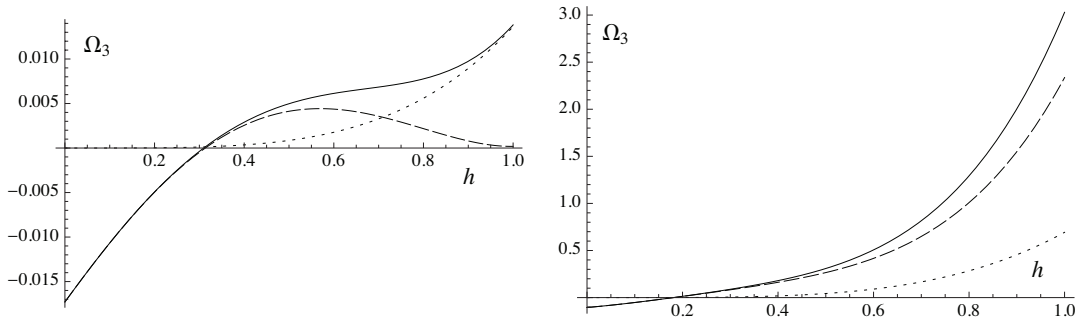



**Figure D3.** Relative effect of the linkage disequilibrium $\langle D_{\mathrm{AB}} \rangle$ on selection for sex due to three-locus interactions. In each panel the solid curve represents the term$\Omega_3$ that appears in equation D17, the dotted curve is the part of$\Omega_3$ that is generated by $\langle D_{\mathrm{AB}} \rangle$, and the dashed curve the part of$\Omega_3$ generated by associations other than $\langle D_{\mathrm{AB}} \rangle$. Left: free recombination ($r_{\mathrm{MA}} = r_{\mathrm{AB}} = 0.5$), right: loci in order A − M − B, $r_{\mathrm{MA}} = r_{\mathrm{MB}} = 0.1$. Other parameters: $s = 0.05$, $\sigma = 0.5$ ($\sigma = 0.1$ leads to similar results – not shown).

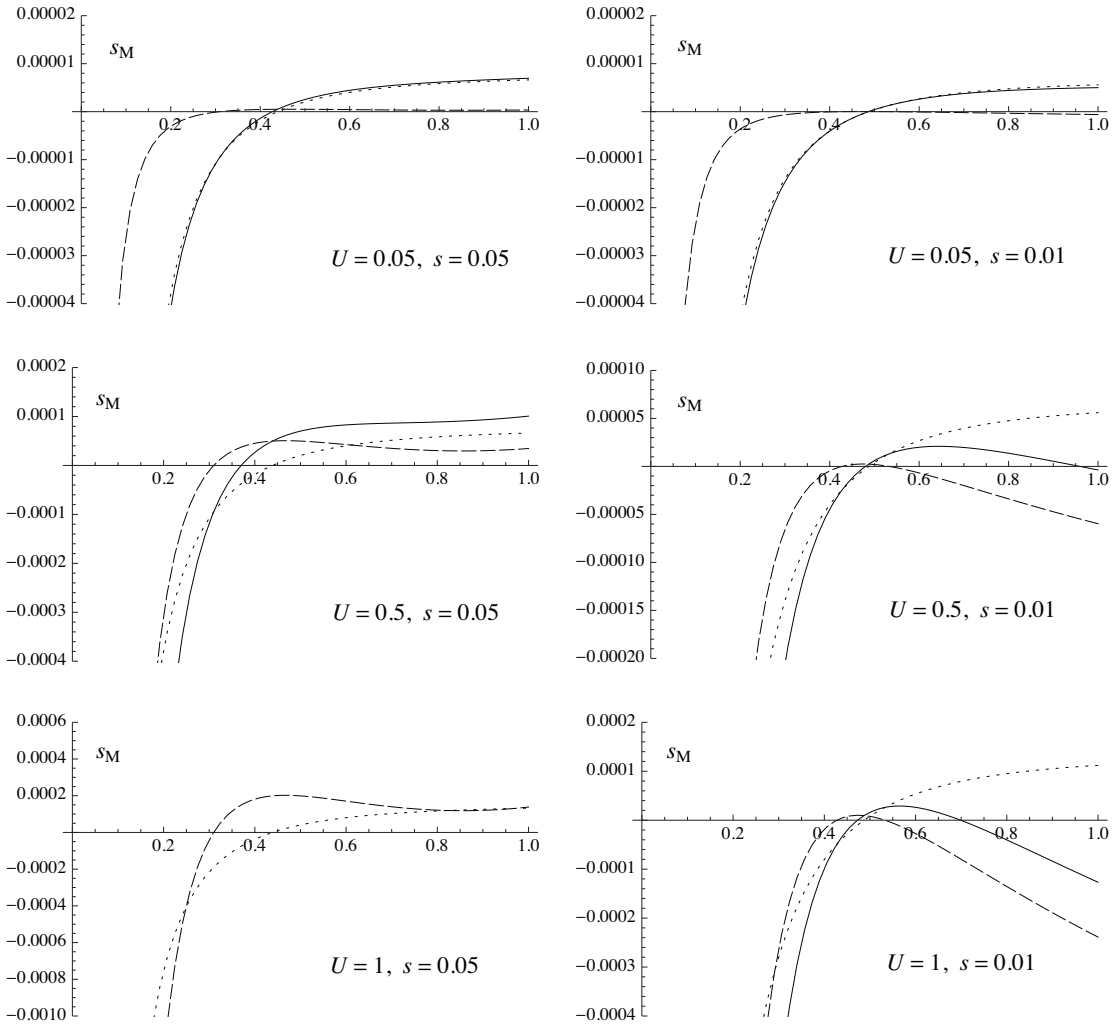



**Figure D4.** Components of selection for the modifier, as a function of $h$. In each panel, the solid curve shows the selection gradient $s_{\mathrm{M}}$ defined by equation D17, the dotted curve shows the effect of two-locus interactions (first two terms of equation D17), and the dashed curve the effect of three-locus interactions (third term of equation D17). Parameter values are $r_{\mathrm{MA}} = r_{\mathrm{AB}} = 0.5$, $\sigma = 0.5$.

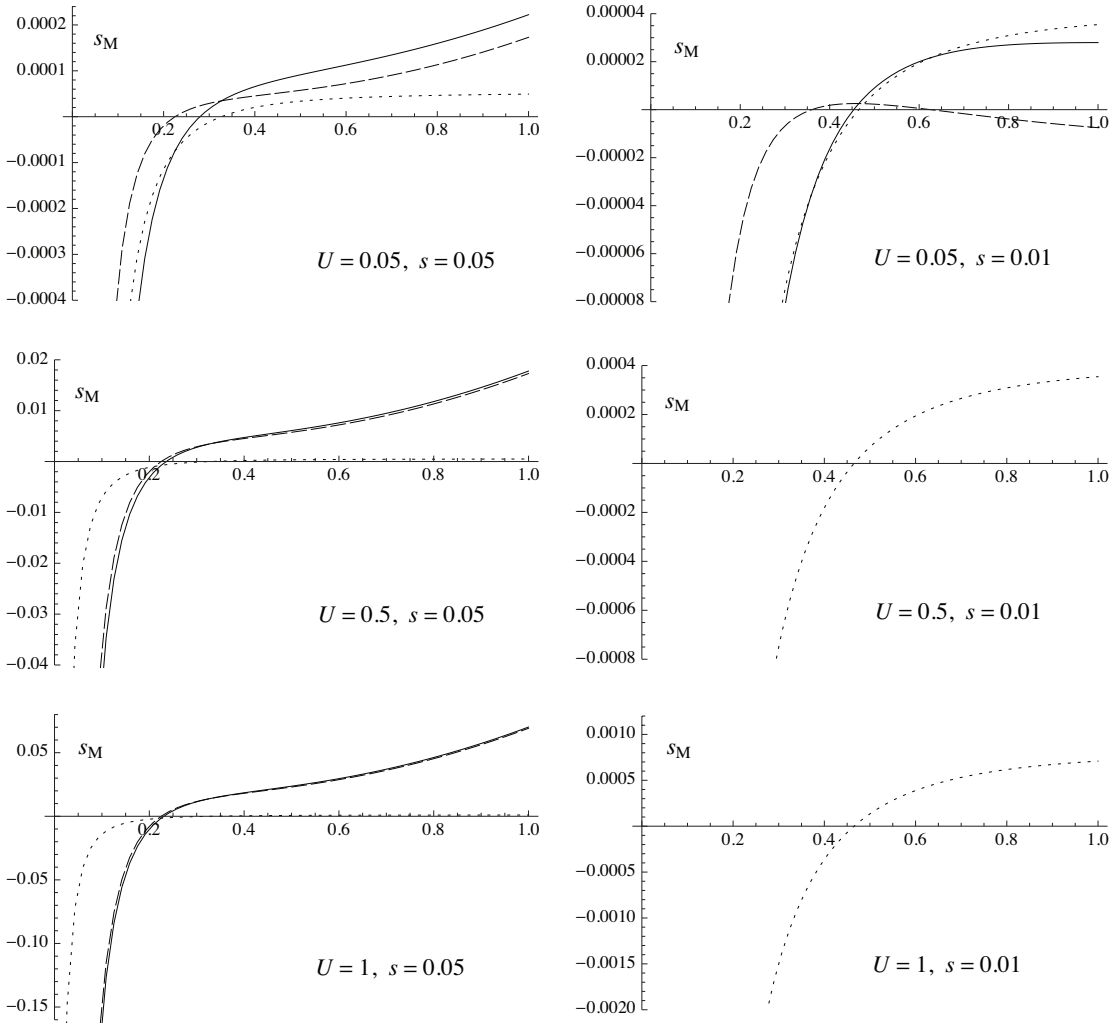



**Figure D5.** Same as figure D4, with $\sigma = 0.2$.

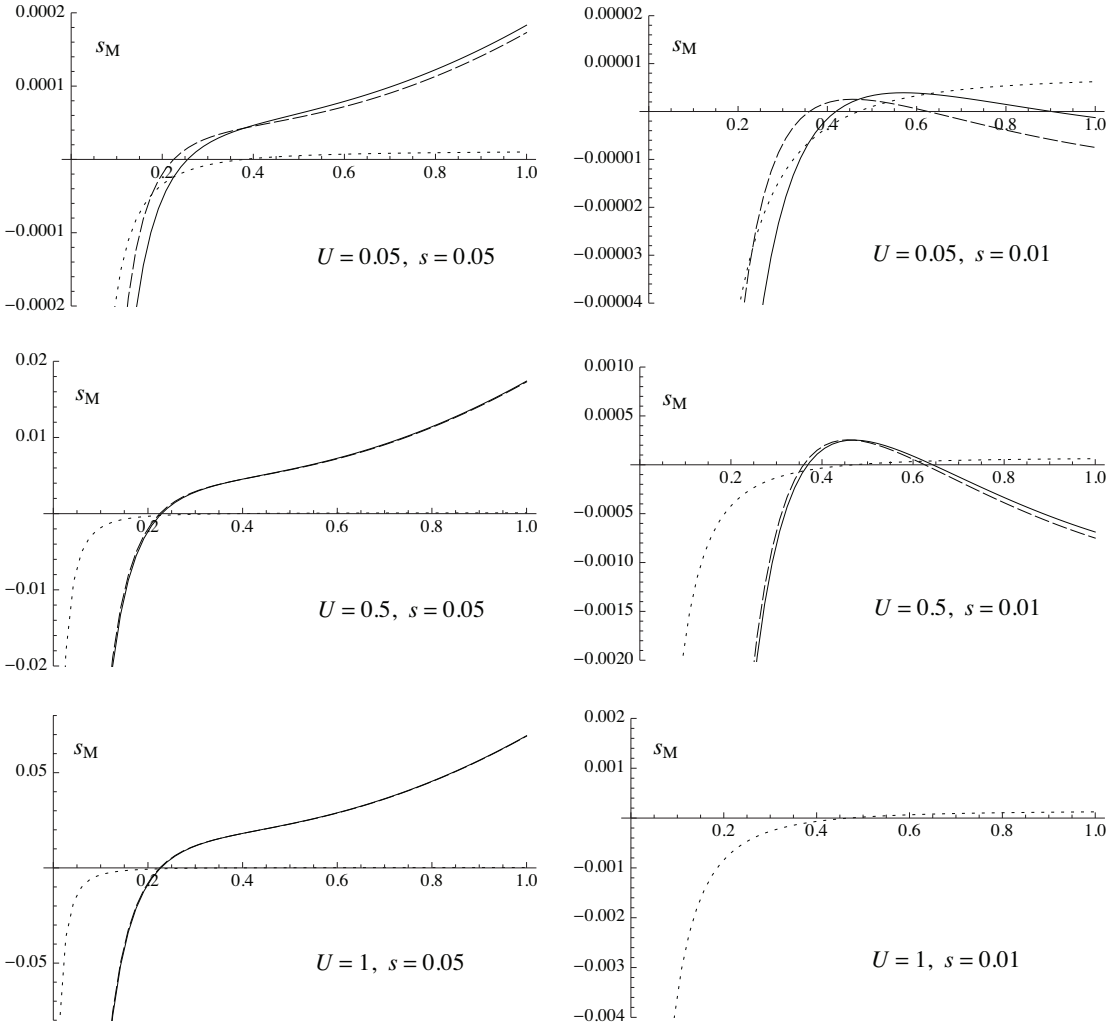



**Figure D6.** Same as figure D4, assuming that the recombination rate between the modifier and each selected locus is 0.1 (using the model where loci are in order A−M−B and setting $r_{\mathrm{MA}} = r_{\mathrm{MB}} = 0.1$); $\sigma = 0.5$.

## ONLINE APPENDIX E: SIMULATING TWO MODIFIER ALLELES

As discussed in the main text, figure 4 does not inform us about the strength and direction of selection for sex when $U = 0.05$ and $h \geq 0.2$ (as the rate of sex does does depart from the mutation - drift equilibrium at $\sigma = 0.5$). We performed additional simulations where only two alleles ($m$ and $M$) segregate at the modifier locus, with a symmetric mutation rate of $10^{-4}$ between the two alleles. The strength of selection for allele $M$ can then be deduced from its average frequency at equilibrium. In this modified program recombination is free among all loci, in order to facilitate comparisons with the analytical model. We ran 90 replications of this program for different values of $h$, for $\sigma = 0.2$, $\delta\sigma = 0.1$, $h_{\mathrm{M}} = 0.5$, $U = 0.05$, $\gamma = 0$ and other parameters as in figure 4. Each replicate ran for $2 \times 10^6$ generations (after 2000 initial generations where allele $M$ is absent), and the frequency of $M$ was measured every 10 generations. We also ran 4 replications for different values of $h$, $\sigma = 0.01$ and $\delta\sigma = 0.09$. Results are shown in figure E1: vertical bars correspond the the 95% confidence interval of the strength of selection for $M$ estimated from the simulations for the different values of $h$ (standard error is calculated over the different replicate simulations), while curves correspond to analytical predictions, given by $\langle \Delta p_{\mathrm{M}} \rangle \,/\, \left\langle \frac{1}{2}\, p q_{\mathrm{M}} \right\rangle$.

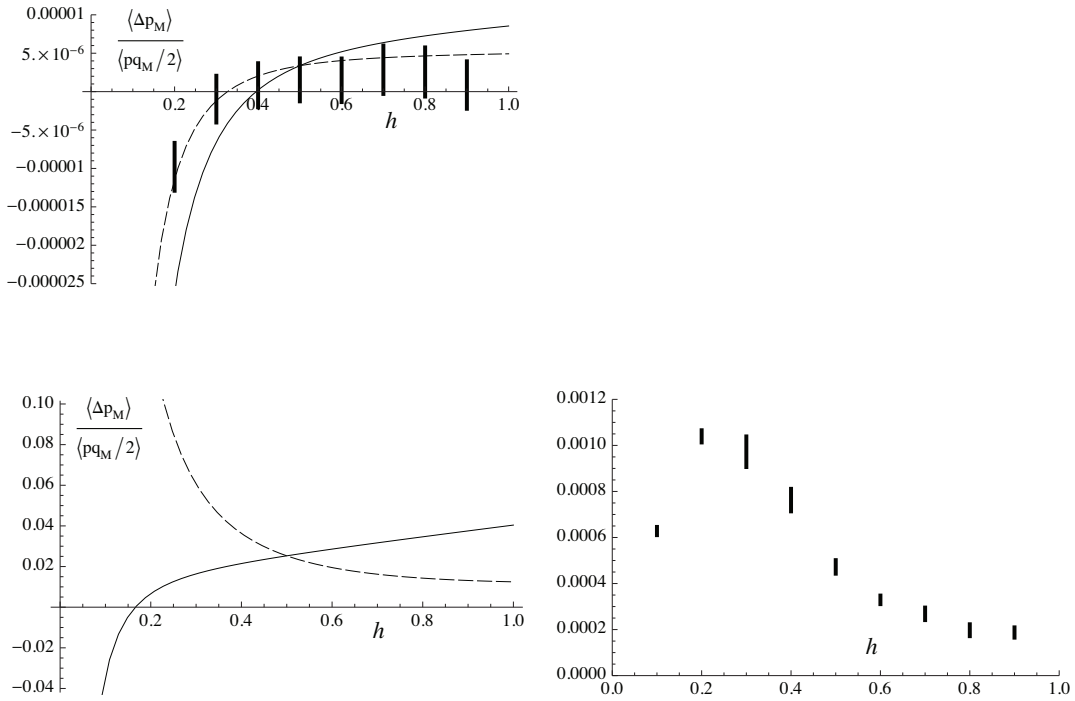



**Figure E1.** Strength of selection for allele $M$ as a function of $h$. Vertical bars correspond to 95% confidence intervals from the simulations, where the strength of selection for $M$ is estimated from the average value of $p_{\mathrm{M}}$. Solid curves correspond to the predictions from the two-locus model under weak dominance (from equations 5, 7, 8 and 9 in the main text); dashed curves correspond to expressions derived to the third order in $\epsilon$, for arbitrary $h$ (see below). Top: $\sigma = 0.2$, $\delta\sigma = 0.1$; bottom: $\sigma = 0.01$, $\delta\sigma = 0.09$. Other parameters: $N = 20000$, $s = 0.05$, $U = 0.05$, $h_{\mathrm{M}} = 0.5$, $\gamma = 0$, free recombination. Note that the third (right) panel shows the strength of selection separately for simulation results because the values were much smaller than the analytic predictions.

Figure E1 shows that for $\sigma = 0.2$ and $\delta\sigma = 0.1$ (top), allele $M$ is disfavored when $h = 0.2$, while allele $M$ seems favored when $h \geq 0.5$ (although selection is very weak and error bars reach zero). These simulation results are compatible with the two-locus model. Note that the prediction derived from equations 5, 7, 8 and 9 in the main text (solid curves in figure E1) assumes weak dominance ($h$ close to 0.5); a better approximation is obtained by expressing associations $\langle D_{\mathrm{MA}} \rangle$, $\langle D_{\mathrm{M,A}} \rangle$ and $\langle D_{\mathrm{MA,A}} \rangle$ to the second order in $\epsilon$ for arbitrary $h$ (dashed curves in figure E1). Recursions for these associations to the second order in $\epsilon$ are given below. When $\sigma = 0.01$, predictions from the analytical model are false by two orders of magnitude; this is expected, as the quasi-equilibrium approximation does not hold when the rate of sex is small. Nevertheless, it can be noted that the dashed curve predicts that selection for sex increases as $h$ decreases, which is observed in the simulations. Finally, simulations show that a rate of sex of 0.1 is favored over 0.01 for all values of $h$ between 0.1 and 0.9.

In the case of an additive modifier ($h_{\mathrm{M}} = 0.5$), recursions for $\langle D_{\mathrm{MA}} \rangle$, $\langle D_{\mathrm{M,A}} \rangle$ and $\langle D_{\mathrm{MA,A}} \rangle$ to the second order in $\epsilon$ for arbitrary $h$ are given by the following equations. More precisely, we assume that $N$ is large and keep only terms in $s/N$ and in $s^2$ (we thus neglect terms in $1/N^2$). Furthermore, we assume that the frequency of deleterious allele is small, so that terms in $\langle p {q_{\mathrm{A}}}^2 \rangle / N$ can be neglected compared to terms in $\langle p q_{\mathrm{A}} \rangle / N$. One obtains:

$$\langle D_{\mathrm{MA}} \rangle_{t+1} = (1 - r\,\sigma)\,\langle D_{\mathrm{MA}} \rangle_t + r\,\sigma\,\langle D_{\mathrm{M,A}} \rangle_t - s\,(1-h)\,\langle D_{\mathrm{MA,A}} \rangle_t \qquad \text{(E1)}$$

$$\begin{aligned}\langle D_{\mathrm{M,A}}\rangle_{t+1} &= (1-\sigma)\left[\langle D_{\mathrm{M,A}}\rangle_t - s\,(1-h)\,\langle D_{\mathrm{MA,A}}\rangle_t\right] \\ &\quad - \frac{\delta\sigma}{2}\Big[\langle D_{\mathrm{MA,M}}\rangle_t - s\,(1-h)\,\langle D_{\mathrm{MA,MA}}\rangle_t \\ &\qquad + s\,h\left(\left\langle {D_{\mathrm{MA}}}^2\right\rangle_t + 2\,\langle D_{\mathrm{MA}} D_{\mathrm{M,A}}\rangle_t + \left\langle {D_{\mathrm{M,A}}}^2\right\rangle_t\right)\Big]\end{aligned} \tag{E2}$$

$$\begin{aligned}\langle D_{\mathrm{MA,M}}\rangle_{t+1} &= (1-\sigma)\Big[\langle D_{\mathrm{MA,M}}\rangle_t - s\,(1-h)\,\langle D_{\mathrm{MA,MA}}\rangle_t \\ &\qquad + s\,h\left(\left\langle {D_{\mathrm{MA}}}^2\right\rangle_t + 2\,\langle D_{\mathrm{MA}} D_{\mathrm{M,A}}\rangle_t + \left\langle {D_{\mathrm{M,A}}}^2\right\rangle_t\right)\Big]\end{aligned} \tag{E3}$$

$$\begin{aligned}\langle D_{\mathrm{MA,A}}\rangle_{t+1} &= (1-\sigma)\,(1-s)\,\langle D_{\mathrm{MA,A}}\rangle_t \\ &\quad + \frac{\delta\sigma}{2}\Big[(1-2sh)\left[(1-r)\left\langle {D_{\mathrm{MA}}}^2\right\rangle_t + \langle D_{\mathrm{MA}} D_{\mathrm{M,A}}\rangle_t + r\,\left\langle {D_{\mathrm{M,A}}}^2\right\rangle_t\right] \\ &\qquad - (1-s)\,\langle D_{\mathrm{MA,MA}}\rangle_t - \left\langle pq_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{A,A}}\right\rangle_t\Big]\end{aligned} \tag{E4}$$

where $\left\langle pq_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{A,A}}\right\rangle_t$ is measured after selection. In the last expression, associations $\left\langle {D_{\mathrm{MA}}}^2\right\rangle_t$, $\langle D_{\mathrm{MA}} D_{\mathrm{M,A}}\rangle_t$, $\left\langle {D_{\mathrm{M,A}}}^2\right\rangle_t$ and $\langle D_{\mathrm{MA,MA}}\rangle_t$ must be expressed at quasi-equilibrium to the first order in $s$. One finds that recursions for these associations are given by equations B4 to B7, multiplying $\langle D_{\mathrm{MA,MA}}\rangle_t$ by $1-s$ and multiplying $\left\langle {D_{\mathrm{MA}}}^2\right\rangle_t$, $\langle D_{\mathrm{MA}} D_{\mathrm{M,A}}\rangle_t$, and $\left\langle {D_{\mathrm{M,A}}}^2\right\rangle_t$ by $1-2sh$. Finally, the recursion for $\left\langle pq_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{A,A}}\right\rangle_t$ is given by:

$$\begin{aligned}\left\langle pq_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{A,A}}\right\rangle_{t+1} &= (1-\sigma)\left\langle pq_{\mathrm{M}} D^{\mathrm{par}}_{\mathrm{A,A}}\right\rangle_t \\ &\quad - \left[1 - s\,\frac{1-\sigma}{\sigma}\right]\left[\frac{\langle pq_{\mathrm{MA}}\rangle}{2N} + s\,(1-2h)\left\langle pq_{\mathrm{M}}\, {pq_{\mathrm{A}}}^2\right\rangle\right].\end{aligned} \tag{E5}$$

ONLINE APPENDIX F: ADDITIONAL SIMULATION RESULTS

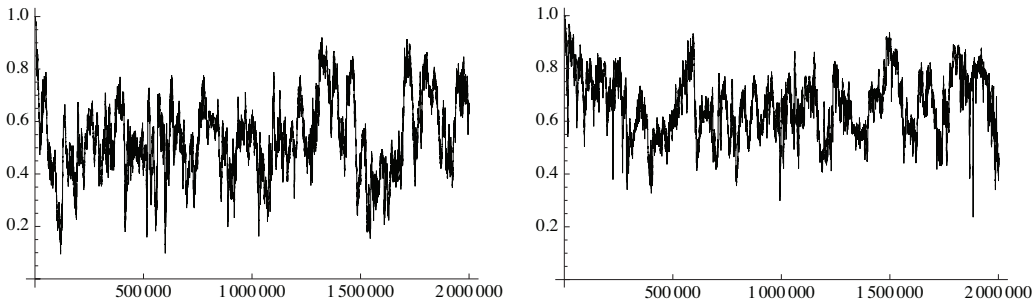



**Figure F1: trajectories.** Average rate of sex $\sigma$ in the population over time (in generations) during the course of a simulation, for $N = 20000$, $s = 0.05$, $h = 0.5$, $L = 10$, $\sigma_{\text{init}} = 1$, $c = 1$, $\gamma = \chi = 0$, $U = 0.05$ (left) and $U = 0.5$ (right). Other parameter values lead to similar dynamics (with wide fluctuations of the rate of sex) except when $\sigma$ goes to zero.

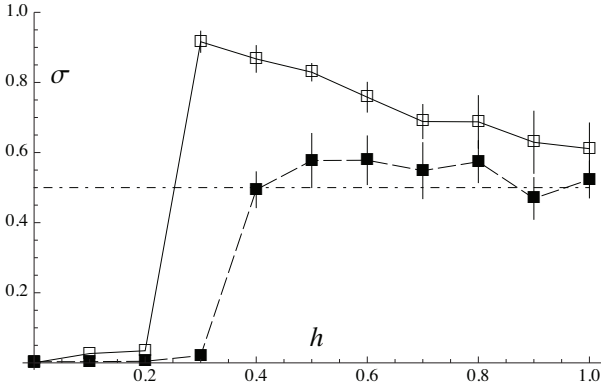



**Figure F2: eliminating benefits of recombination.** Average rate of sex observed in simulations, as a function of the dominance coefficient $h$ of deleterious mutations. Empty squares, solid line: $L = 0.1$. Filled squares, dashed lines: the $\sigma L$ product is kept constant across individuals (by adjusting map length $L$ as a function of the value of modifier alleles carried by the individual) in order to eliminate selection for recombination; $\sigma L = 0.05$ (the program is constrained so that the rate of sex cannot be less than 0.001). Other parameter values: $N = 20000$, $U = 0.5$, $s = 0.05$, $\gamma = 0$, $c = 1$, $\sigma_{\text{init}} = 1$. Note that this method for eliminating benefits of recombination will not work when $L$ is already large, so that recombination between many pairs of loci is close to 0.5 (because incresing $L$ even further will not have much effect). Indeed, results for $L = 10$ (as in figure 4) and $\sigma L$ fixed to 5 are undistinguishable (not shown).

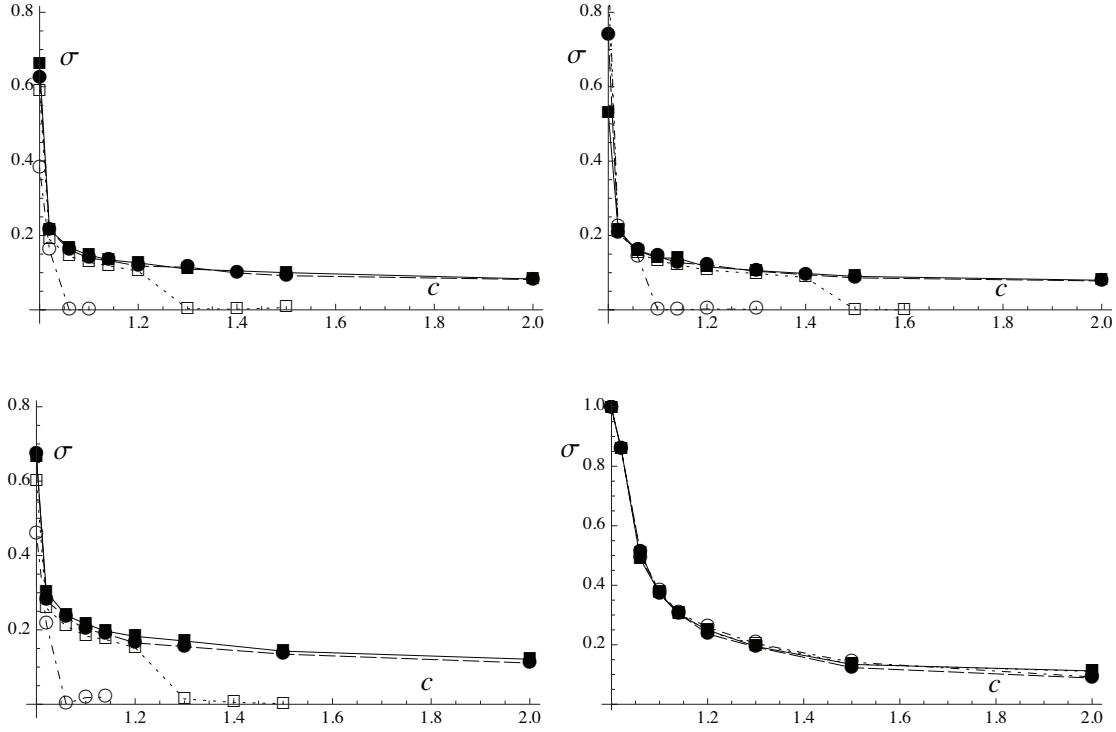



**Figure F3: effect of the cost of sex (additional results).** This figure is the equivalent of figure 7 left (average rate of sex at equilibrium, as a function of the cost of sex $c$) for $U = 0.5$, $N = 20000$, $\gamma = 0$ (top left), $U = 0.5$, $N = 20000$, $\gamma = 10^{-3}$ (top right), and $U = 1$, $N = 50000$, $\gamma = 0$ (bottom left), $U = 0.5$, $N = 20000$, $\gamma = 0$, $e_{\mathrm{a\times a}} = -0.02$ (bottom right; epistasis is zero in the three other plots). Symbols refer the to the same values of $h$ as in figure 7. Other parameters are the same as in figure 7. Error bars are smaller than the size of symbols.

**Detailed simulation results.** The following tables show the average rate of sex ($\sigma$), average fitness ($\overline{w}$), average number of mutations per haplotype ($n_{\text{mut}}$) and number of fixed mutations at the end of the simulation ($n_{\text{fix}}$) for the different parameter values considered in the figures (see figure legends for parameter values). The "–" sign means that the rate of sex goes to zero, in which case deleterious mutations accumulate in the heterozygous state and the program has to be stopped. When fixation of deleterious mutations occurs ($n_{\text{fix}} > 0$), fixed mutations are removed from the population every 20 generations. Measures of average fitness and number of mutations per haplotype do not take into account mutations that have been removed from the population. Note that when mutations do not reach fixation, mean fitness is often close to the deterministic expectation $e^{-2U}$, while the number of mutations per chromosome is also often close to the deterministic expectation $U/(sh)$.

**Simulation results corresponding to figure 4**:

$U = 0.05$, $\gamma = 0$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | – | 0.16 | 0.50 | 0.51 | 0.50 | 0.54 | 0.50 | 0.50 | 0.48 | 0.50 | 0.49 |
| $\overline{w}$ | – | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 |
| $n_{\text{mut}}$ | – | 9.87 | 4.94 | 3.28 | 2.45 | 1.95 | 1.62 | 1.38 | 1.20 | 1.06 | 0.95 |
| $n_{\text{fix}}$ | – | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$U = 0.05$, $\gamma = 10^{-4}$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | – | 0.44 | 0.55 | 0.53 | 0.50 | 0.52 | 0.53 | 0.53 | 0.49 | 0.53 | 0.50 |
| $\overline{w}$ | – | 0.91 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 |
| $n_{\text{mut}}$ | – | 9.79 | 4.94 | 3.28 | 2.45 | 1.95 | 1.62 | 1.38 | 1.20 | 1.06 | 0.95 |
| $n_{\text{fix}}$ | – | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$U = 0.05$, $\gamma = 10^{-3}$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.92 | 0.82 | 0.72 | 0.60 | 0.51 | 0.42 | 0.28 | 0.23 | 0.16 | 0.092 | 0.098 |
| $\overline{w}$ | 0.95 | 0.91 | 0.90 | 0.90 | 0.90 | 0.91 | 0.91 | 0.91 | 0.91 | 0.91 | 0.91 |
| $n_{\text{mut}}$ | 80.39 | 9.76 | 4.93 | 3.28 | 2.44 | 1.95 | 1.61 | 1.38 | 1.20 | 1.06 | 0.95 |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$U = 0.5$, $\gamma = 0$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | – | – | 0.37 | 0.58 | 0.62 | 0.65 | 0.64 | 0.64 | 0.66 | 0.64 | 0.67 |
| $\overline{w}$ | – | – | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 |
| $n_{\text{mut}}$ | – | – | 49.4 | 32.8 | 24.5 | 19.5 | 16.2 | 13.8 | 12.0 | 10.6 | 9.51 |
| $n_{\text{fix}}$ | – | – | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$U = 0.5$, $\gamma = 10^{-4}$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | – | – | 0.63 | 0.65 | 0.64 | 0.63 | 0.62 | 0.62 | 0.59 | 0.60 | 0.62 |
| $\overline{w}$ | – | – | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 |
| $n_{\text{mut}}$ | – | – | 74.6 | 32.8 | 24.5 | 19.5 | 16.2 | 13.8 | 12.0 | 10.6 | 9.50 |
| $n_{\text{fix}}$ | – | – | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$U = 0.5$, $\gamma = 10^{-3}$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.95 | 0.97 | 0.92 | 0.87 | 0.74 | 0.53 | 0.44 | 0.40 | 0.40 | 0.36 | 0.38 |
| $\overline{w}$ | 0.59 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 |
| $n_{\text{mut}}$ | 816.2 | 97.8 | 49.4 | 32.8 | 24.5 | 19.5 | 16.2 | 13.8 | 12.0 | 10.6 | 9.53 |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**Simulation results corresponding to figure 5**:

$\chi = 10^{-3}$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | – | – | 0.67 | 0.73 | 0.67 | 0.61 | 0.50 | 0.46 | 0.37 | 0.37 | 0.37 |
| $\overline{w}$ | – | – | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 |
| $n_{\text{mut}}$ | – | – | 49.4 | 32.8 | 24.5 | 19.5 | 16.2 | 13.8 | 12.0 | 10.7 | 9.56 |
| $n_{\text{fix}}$ | – | – | 0 | 0 | 0 | 0 | 0 | 0 | 10 | 1 | 77 |

$\chi = 10^{-2}$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | – | 0.97 | 0.96 | 0.94 | 0.86 | 0.22 | 0.14 | 0.10 | 0.10 | 0.097 | 0.095 |
| $\overline{w}$ | – | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.36 | 0.33 | 0.32 | 0.32 | 0.32 |
| $n_{\text{mut}}$ | – | 97.9 | 49.4 | 32.8 | 24.5 | 19.4 | 16.9 | 16.3 | 15.0 | 13.9 | 12.8 |
| $n_{\text{fix}}$ | – | 0 | 0 | 0 | 0 | 0 | 878 | 17431 | 27474 | 38062 | 41602 |

**Simulation results corresponding to figure 6**:

$N = 10000$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | – | – | – | 0.57 | 0.65 | 0.65 | 0.65 | 0.64 | 0.66 | 0.66 | 0.67 |
| $\overline{w}$ | – | – | – | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 |
| $n_{\text{mut}}$ | – | – | – | 32.8 | 24.5 | 19.5 | 16.2 | 13.8 | 12.0 | 10.6 | 9.52 |
| $n_{\text{fix}}$ | – | – | – | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$N = 50000$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | – | – | 0.47 | 0.58 | 0.61 | 0.61 | 0.65 | 0.63 | 0.65 | 0.61 | 0.61 |
| $\overline{w}$ | – | – | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 |
| $n_{\text{mut}}$ | – | – | 49.5 | 32.8 | 24.5 | 19.5 | 16.2 | 13.8 | 12.0 | 10.6 | 9.50 |
| $n_{\text{fix}}$ | – | – | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$L = 1$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | – | – | – | 0.79 | 0.76 | 0.74 | 0.72 | 0.73 | 0.70 | 0.75 | 0.65 |
| $\overline{w}$ | – | – | – | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 |
| $n_{\text{mut}}$ | – | – | – | 32.9 | 24.6 | 19.5 | 16.2 | 13.8 | 12.0 | 10.6 | 9.52 |
| $n_{\text{fix}}$ | – | – | – | 0 | 0 | 0 | 0 | 0 | 0 | 29 | 1 |

$L = 0.1$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | – | – | – | 0.92 | 0.87 | 0.83 | 0.76 | 0.69 | 0.69 | 0.63 | 0.61 |
| $\overline{w}$ | – | – | – | 0.29 | 0.30 | 0.31 | 0.31 | 0.32 | 0.33 | 0.34 | 0.34 |
| $n_{\text{mut}}$ | – | – | – | 39.6 | 29.1 | 23.1 | 19.3 | 16.3 | 14.0 | 12.2 | 10.8 |
| $n_{\text{fix}}$ | – | – | – | 109 | 1268 | 4051 | 10251 | 16554 | 17928 | 17284 | 19332 |

$s = 0.01$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | – | – | – | 0.39 | 0.57 | 0.57 | 0.55 | 0.086 | 0.064 | 0.058 | 0.065 |
| $\overline{w}$ | – | – | – | 0.36 | 0.37 | 0.37 | 0.37 | 0.29 | 0.31 | 0.34 | 0.35 |
| $n_{\text{mut}}$ | – | – | – | 170.1 | 124.8 | 99.8 | 83.2 | 101.1 | 87.5 | 74.8 | 68.3 |
| $n_{\text{fix}}$ | – | – | – | 47 | 0 | 0 | 66 | 198149 | 304529 | 346657 | 352502 |

$s = 0.1$:

| $h$ | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | – | – | 0.50 | 0.64 | 0.67 | 0.66 | 0.68 | 0.69 | 0.65 | 0.68 | 0.70 |
| $\overline{w}$ | – | – | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 |
| $n_{\text{mut}}$ | – | – | 24.5 | 16.2 | 12.0 | 9.50 | 7.84 | 6.65 | 5.75 | 5.06 | 4.50 |
| $n_{\text{fix}}$ | – | – | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**Simulation results corresponding to figure 7 left**:

$h = 0.5$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.5 | 2 |
|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.71 | 0.32 | 0.26 | 0.23 | 0.21 | 0.20 | 0.17 | 0.15 | 0.13 |
| $\overline{w}$ | 0.13 | 0.14 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.12 | 0.10 |
| $n_{\text{mut}}$ | 39.1 | 39.5 | 40.2 | 40.9 | 41.7 | 42.5 | 44.0 | 48.7 | 57.6 |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 1 | 63 | 136 | 2689 | 9076 | 57156 |

$h = 0.4$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.4 | 1.5 | 1.7 | 1.8 | 2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.71 | 0.31 | 0.24 | 0.22 | 0.20 | 0.18 | 0.16 | 0.15 | 0.15 | 0.12 | – | – |
| $\overline{w}$ | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.12 | 0.12 | 0.10 | – | – |
| $n_{\text{mut}}$ | 49.1 | 49.6 | 50.5 | 51.1 | 51.8 | 53.2 | 56.1 | 58.5 | 60.0 | 68.0 | – | – |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | 18 | 256 | 1127 | 2866 | 14951 | – | – |

$h = 0.3$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.5 |
|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.64 | 0.28 | 0.22 | 0.20 | 0.18 | – | – | – |
| $\overline{w}$ | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | – | – | – |
| $n_{\text{mut}}$ | 65.7 | 66.4 | 67.8 | 68.9 | 70.6 | – | – | – |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | – | – | – |

**Simulation results corresponding to figure 7 right**:

$h = 0.5$:

| $c$ | 1 | 1.02 | 1.04 | 1.06 | 1.08 | 1.1 | 1.12 | 1.14 | 1.16 | 1.18 | 1.2 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.98 | 0.95 | 0.88 | 0.80 | 0.71 | 0.61 | 0.51 | 0.42 | 0.34 | 0.27 | 0.22 |
| $\overline{w}$ | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 |
| $n_{\text{mut}}$ | 39.0 | 39.0 | 39.1 | 39.2 | 39.4 | 39.7 | 40.0 | 40.5 | 41 | 41.5 | 42.1 |
| $n_{\text{fix}}$ | 4 | 0 | 11 | 0 | 13 | 22 | 83 | 87 | 144 | 411 | 667 |

$h = 0.4$:

| $c$ | 1 | 1.02 | 1.04 | 1.06 | 1.08 | 1.1 | 1.12 | 1.14 | 1.16 | 1.18 | 1.2 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.98 | 0.94 | 0.88 | 0.79 | – | – | – | – | – | – | – |
| $\overline{w}$ | 0.14 | 0.14 | 0.14 | 0.14 | – | – | – | – | – | – | – |
| $n_{\text{mut}}$ | 49.0 | 49.0 | 49.1 | 49.3 | – | – | – | – | – | – | – |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | – | – | – | – | – | – | – |

**Simulation results corresponding to figure F3 top left**:

$h = 0.5$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.5 | 2 |
|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.66 | 0.22 | 0.17 | 0.15 | 0.13 | 0.13 | 0.11 | 0.10 | 0.083 |
| $\overline{w}$ | 0.37 | 0.37 | 0.38 | 0.38 | 0.39 | 0.40 | 0.42 | 0.46 | 0.51 |
| $n_{\text{mut}}$ | 19.5 | 19.7 | 20.0 | 20.3 | 20.5 | 20.8 | 21.5 | 22.6 | 25.8 |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 8 | 45 | 115 | 667 | 8041 |

$h = 0.4$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.4 | 1.5 | 2 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.62 | 0.21 | 0.16 | 0.14 | 0.13 | 0.12 | 0.11 | 0.10 | 0.092 | 0.081 |
| $\overline{w}$ | 0.37 | 0.37 | 0.38 | 0.38 | 0.39 | 0.40 | 0.43 | 0.44 | 0.45 | 0.50 |
| $n_{\text{mut}}$ | 24.5 | 24.7 | 25.1 | 25.5 | 25.6 | 26.2 | 26.6 | 27.7 | 28.6 | 33.0 |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | 6 | 26 | 164 | 325 | 5451 |

$h = 0.3$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.4 |
|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.59 | 0.19 | 0.14 | 0.13 | 0.12 | 0.10 | – | – |
| $\overline{w}$ | 0.37 | 0.37 | 0.37 | 0.38 | 0.39 | 0.39 | – | – |
| $n_{\text{mut}}$ | 32.8 | 33.2 | 33.7 | 34.2 | 34.7 | 35.1 | – | – |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | 89 | – | – |

**Simulation results corresponding to figure F3 top right**:

$h = 0.5$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.5 | 2 |
|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.53 | 0.21 | 0.16 | 0.14 | 0.14 | 0.12 | 0.11 | 0.090 | 0.080 |
| $\overline{w}$ | 0.37 | 0.37 | 0.38 | 0.38 | 0.40 | 0.40 | 0.42 | 0.45 | 0.53 |
| $n_{\text{mut}}$ | 19.5 | 19.6 | 20.0 | 20.3 | 20.4 | 20.9 | 20.6 | 22.9 | 25.4 |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 8 | 32 | 264 | 927 | 5531 |

$h = 0.4$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.4 | 1.5 | 2 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.74 | 0.21 | 0.16 | 0.14 | 0.13 | 0.12 | 0.10 | 0.094 | 0.086 | 0.078 |
| $\overline{w}$ | 0.37 | 0.37 | 0.38 | 0.39 | 0.39 | 0.41 | 0.42 | 0.44 | 0.45 | 0.52 |
| $n_{\text{mut}}$ | 24.5 | 24.6 | 24.9 | 25.2 | 25.6 | 25.9 | 26.8 | 27.6 | 28.5 | 31.6 |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | 0 | 12 | 88 | 109 | 2127 |

$h = 0.3$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.4 | 1.5 | 1.6 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.87 | 0.21 | 0.15 | 0.13 | 0.12 | 0.11 | 0.098 | 0.087 | – | – |
| $\overline{w}$ | 0.37 | 0.37 | 0.38 | 0.39 | 0.39 | 0.40 | 0.42 | 0.43 | – | – |
| $n_{\text{mut}}$ | 32.8 | 32.8 | 33.2 | 33.6 | 34.0 | 34.7 | 35.8 | 37.2 | – | – |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 50 | – | – |

**Simulation results corresponding to figure F3 bottom left**:

$h = 0.5$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.5 | 2 |
|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.66 | 0.30 | 0.24 | 0.22 | 0.20 | 0.18 | 0.17 | 0.14 | 0.12 |
| $\overline{w}$ | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.13 | 0.12 |
| $n_{\text{mut}}$ | 39.0 | 39.3 | 39.9 | 40.5 | 41.1 | 41.9 | 42.9 | 46.5 | 54.3 |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 1 | 45 | 202 | 3338 | 32413 |

$h = 0.4$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.5 | 2 |
|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.67 | 0.28 | 0.23 | 0.20 | 0.19 | 0.17 | 0.15 | 0.13 | 0.11 |
| $\overline{w}$ | 0.14 | 0.13 | 0.14 | 0.13 | 0.14 | 0.13 | 0.14 | 0.13 | 0.11 |
| $n_{\text{mut}}$ | 49.0 | 49.4 | 50.0 | 50.9 | 51.3 | 52.9 | 54.1 | 58.3 | 69.8 |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | 32 | 59 | 1268 | 15115 |

$h = 0.3$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.4 |
|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.60 | 0.26 | 0.21 | 0.18 | 0.17 | 0.15 | – | – |
| $\overline{w}$ | 0.14 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | – | – |
| $n_{\text{mut}}$ | 65.7 | 66.2 | 67.1 | 68.1 | 68.9 | 71.0 | – | – |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | 2 | – | – |

**Simulation results corresponding to figure F3 bottom right**:

$h = 0.5$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.5 | 2 |
|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.99 | 0.86 | 0.49 | 0.37 | 0.30 | 0.25 | 0.19 | 0.13 | 0.11 |
| $\overline{w}$ | 0.52 | 0.51 | 0.51 | 0.51 | 0.52 | 0.53 | 0.56 | 0.63 | 0.80 |
| $n_{\text{mut}}$ | 3.2 | 3.3 | 3.4 | 3.5 | 3.5 | 3.6 | 3.7 | 3.8 | 3.9 |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$h = 0.4$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.5 | 2 |
|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.99 | 0.86 | 0.51 | 0.37 | 0.30 | 0.24 | 0.19 | 0.12 | 0.088 |
| $\overline{w}$ | 0.53 | 0.52 | 0.51 | 0.52 | 0.53 | 0.54 | 0.57 | 0.63 | 0.80 |
| $n_{\text{mut}}$ | 3.3 | 3.3 | 3.5 | 3.6 | 3.6 | 3.7 | 3.8 | 3.9 | 4.0 |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$h = 0.3$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.5 | 2 |
|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.99 | 0.86 | 0.51 | 0.38 | 0.31 | 0.24 | 0.20 | 0.13 | 0.11 |
| $\overline{w}$ | 0.53 | 0.53 | 0.52 | 0.53 | 0.53 | 0.55 | 0.58 | 0.64 | 0.82 |
| $n_{\text{mut}}$ | 3.4 | 3.4 | 3.6 | 3.6 | 3.7 | 3.8 | 3.8 | 3.9 | 4.0 |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$h = 0.2$:

| $c$ | 1 | 1.02 | 1.06 | 1.1 | 1.14 | 1.2 | 1.3 | 1.5 | 2 |
|---|---|---|---|---|---|---|---|---|---|
| $\sigma$ | 0.99 | 0.86 | 0.49 | 0.38 | 0.31 | 0.26 | 0.20 | 0.14 | 0.092 |
| $\overline{w}$ | 0.54 | 0.54 | 0.53 | 0.53 | 0.54 | 0.56 | 0.59 | 0.65 | 0.82 |
| $n_{\text{mut}}$ | 3.4 | 3.5 | 3.6 | 3.7 | 3.8 | 3.8 | 3.9 | 4.0 | 4.2 |
| $n_{\text{fix}}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

## ONLINE APPENDIX G: EFFECTS OF EPISTASIS

Epistasis is incoporated in the simulation program as in Roze, 2009. In particular, three forms of epistasis are distinguished: additive-by-additive epistasis ($e_{\mathrm{a\times a}}$) measures the effect of the interaction between two deleterious alleles at two loci, while additive-by-dominance ($e_{\mathrm{a\times d}}$) and dominance-by-dominance ($e_{\mathrm{d\times d}}$) epistasis measure the effect of the interaction between three and four deleterious alleles (respectively) at two loci. From previous work, $e_{\mathrm{a\times a}}$ and $e_{\mathrm{d\times d}}$ should be particularly important for selection on sex modifiers. Additive-by-additive epistasis ($e_{\mathrm{a\times a}}$) should favor sex when it is weakly negative (e.g., Barton, 1995), due to the fact that it generates negative linkage disequilibrium between selected loci. Dominance-by-dominance epistasis ($e_{\mathrm{d\times d}}$, measuring the effect of the interaction between 4 deleterious alleles at two loci) may also have important effects in inbred or finite populations (Roze and Lenormand, 2005; Roze, 2009; Agrawal, 2009). In particular, we have seen that in finite populations, sex tends to increase the frequency of genotypes homozygous at one selected locus and heterozygous at another, this effect being represented by the term $\langle D_{\mathrm{MAB,AB}}\rangle$ in equation 11, which is negative (see also Appendix D). When $e_{\mathrm{d\times d}}$ is sufficiently negative, these genotypes have higher fitness (on average), generating selection for increased sex (see also figure 4 in Roze, 2009). However, our simulation results indicate that negative $e_{\mathrm{d\times d}}$ has the opposite effect, selecting against sex (see figure G1). Figure G1 also shows that negative additive-by-dominance epistasis also disfavors sex.

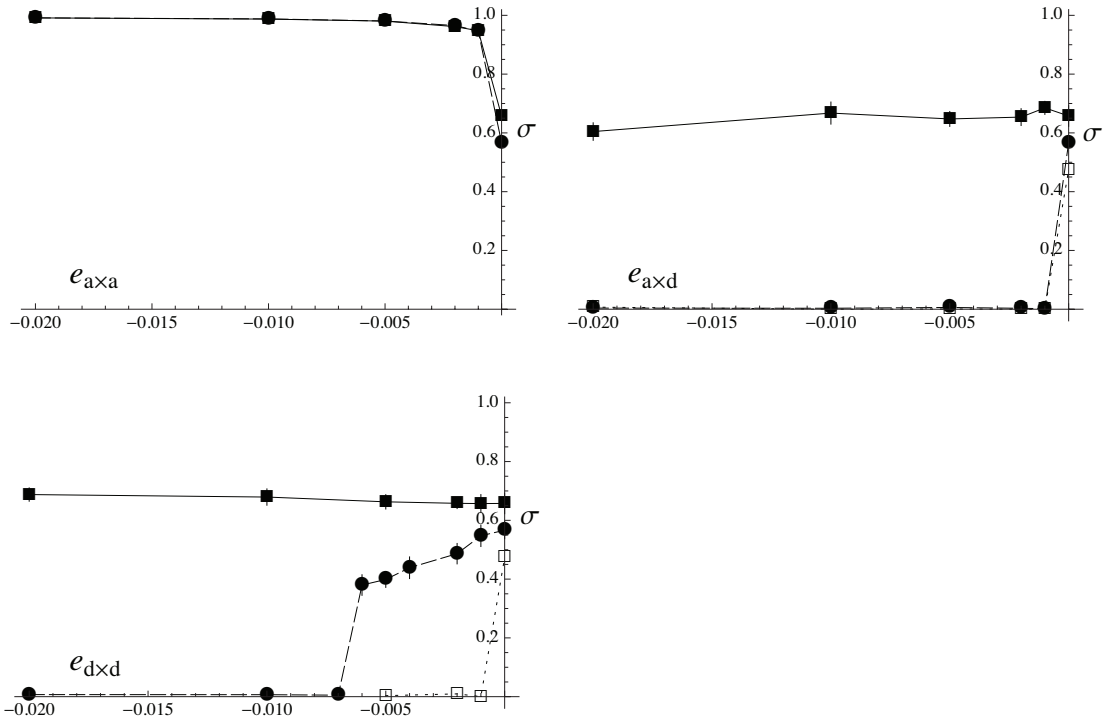



**Figure G1: effects of epistasis on the rate of sex: simulation results.** Average rate of sex $\sigma$ in the population as a function of additive-by-additive epistasis (top left), additive-by-dominance epistasis (top right) and dominance-by-dominance epistasis (bottom), for $s = 0.1$ (filled squares, solid lines), $s = 0.01$ (filled circles, dashed lines) and $s = 0.001$ (empty squares, dotted lines, top right and bottom only). Other parameter values are $N = 20000$, $U = 0.5$, $L = 10$, $h = 0.5$, $\sigma_{\text{init}} = 1$, $\gamma = 0$. In each panel, other components of epistasis are set to zero.

The fact that strongly negative $e_{\mathrm{a}\times\mathrm{a}}$ favors sex in the present model can be explained by the fact that the effect of selection at each locus becomes strong relative to epistasis when mutations segregate at many loci (see Roze, 2009). Similarly, selection against sex under negative $e_{\mathrm{d}\times\mathrm{d}}$ probably comes from the fact that negative $e_{\mathrm{d}\times\mathrm{d}}$ increases selection against homozygotes at each locus (due to interactions with other homozygous loci). Indeed, one obtains that $e_{\mathrm{d}\times\mathrm{d}}$ affects the change in frequency of the sex modifier through a term

$$e_{\mathrm{d}\times\mathrm{d}}\left(\left\langle {p_{\mathrm{B}}}^2 D_{\mathrm{MA,A}}\right\rangle + \left\langle {p_{\mathrm{A}}}^2 D_{\mathrm{MB,B}}\right\rangle + \left\langle D_{\mathrm{MAB,AB}}\right\rangle\right) \tag{G1}$$

(to first order in $e_{\mathrm{d}\times\mathrm{d}}$). The first two terms are approximately $\left\langle {p_{\mathrm{B}}}^2\right\rangle \left\langle D_{\mathrm{MA,A}}\right\rangle$ and $\left\langle {p_{\mathrm{A}}}^2\right\rangle \left\langle D_{\mathrm{MB,B}}\right\rangle$, selecting against sex under negative $e_{\mathrm{d}\times\mathrm{d}}$ due to the fact that a modifier increasing sex tends to be associated with homozygotes: $\left\langle D_{\mathrm{MA,A}}\right\rangle$, $\left\langle D_{\mathrm{MB,B}}\right\rangle$ are positive (see Appendix B). The last term, however, favors sex under negative $e_{\mathrm{d}\times\mathrm{d}}$ since $\left\langle D_{\mathrm{MAB,AB}}\right\rangle < 0$. Note that negative $e_{\mathrm{d}\times\mathrm{d}}$ has similar antagonistic effects on selection for recombination in structured, diploid populations (Roze, 2009); however, while negative $e_{\mathrm{d}\times\mathrm{d}}$ can favor recombination in a given range of parameters (figure 9 in Roze, 2009), we could not find any combination of parameters for which negative $e_{\mathrm{d}\times\mathrm{d}}$ favors higher rates of sex. Negative $e_{\mathrm{d}\times\mathrm{d}}$ may thus be more favorable to recombination than it is to sex (which is probably due to the fact that $D_{\mathrm{MA,A}}$ associations are smaller in magnitude in the case of a recombination modifier, as they are generated by the effect of selection on three-locus associations). Finally, $e_{\mathrm{a}\times\mathrm{d}}$ affects the change in frequency of the sex modifier through a term

$$e_{\mathrm{a}\times\mathrm{d}}\left(\left\langle p_{\mathrm{B}} D_{\mathrm{MA,A}}\right\rangle + \left\langle p_{\mathrm{A}} D_{\mathrm{MB,B}}\right\rangle\right) \tag{G2}$$

(to first order in $e_{\mathrm{a}\times\mathrm{d}}$), which also generates selection against sex under negative $e_{\mathrm{a}\times\mathrm{d}}$.

These results indicate that positive $e_{\mathrm{a}\times\mathrm{d}}$ and $e_{\mathrm{d}\times\mathrm{d}}$ should favor sex; however it is difficult to test this prediction using our simulation model, because positive $e_{\mathrm{a}\times\mathrm{d}}$ or $e_{\mathrm{d}\times\mathrm{d}}$ lead to the rapid fixation of mutations (as combinations of mutations become advantageous), unless $e_{\mathrm{a}\times\mathrm{d}}$, $e_{\mathrm{d}\times\mathrm{d}}$ are very weak relative to $s$ (but in this last case $e_{\mathrm{a}\times\mathrm{d}}$ and $e_{\mathrm{d}\times\mathrm{d}}$ have little effect).