Stabilizing selection, mutational bias, and the evolution of sex*

Eloïse Vanhoenacker, Linnéa Sandell, Denis Roze

Stabilizing selection, mutational bias and the evolution of sex

Eloïse Vanhoenacker[†,‡], Linnéa Sandell[*] and Denis Roze[†,‡]

† CNRS, UMI 3614 Evolutionary Biology and Ecology of Algae, 29688 Roscoff, France

‡ Sorbonne Université, 29688 Roscoff, France

* Department of Zoology, University of British Columbia, Vancouver BC V6T1Z4, Canada

Running title: Mutational bias and sex

Keywords: adaptive landscape, epistasis, evolutionary quantitative genetics, multilocus population genetics, reproductive systems

Address for correspondence:

Denis Roze

Station Biologique de Roscoff

Place Georges Teissier, CS90074

29688 Roscoff Cedex

France

Phone: (+33) 2 56 45 21 39

Fax: (+33) 2 98 29 23 24

email: roze@sb-roscoff.fr

Acknowledgements: We thank Brian Charlesworth for suggesting considering the effect of mutational bias, two anonymous reviewers for helpful comments, and the bioinformatics and computing service of Roscoff's Biological Station (Abims platform) for computing time. This work was supported by the French Agence Nationale de la Recherche (project SexChange, ANR-14-CE02-0001).

Author contributions: EV, LS and DR analyzed the model, EV and DR wrote the article.

Data archiving: the DOI for our data is 10.5061/dryad.50q65s0.

ABSTRACT

Stabilizing selection around a fixed phenotypic optimum is expected to disfavor sexual reproduction, since asexually reproducing organisms can maintain a higher fitness at equilibrium, while sex disrupts combinations of compensatory mutations. This conclusion rests on the assumption that mutational effects on phenotypic traits are unbiased, that is, mutation does not tend to push phenotypes in any particular direction. In this paper, we consider a model of stabilizing selection acting on an arbitrary number of polygenic traits coded by bialellic loci, and show that mutational bias may greatly reduce the mean fitness of asexual populations compared with sexual ones in regimes where mutations have weak to moderate fitness effects. Indeed, mutation and drift tend to push the population mean phenotype away from the optimum, this effect being enhanced by the low effective population size of asexual populations. In a second part, we present results from individual-based simulations showing that positive rates of sex are favored when mutational bias is present, while the population evolves towards complete asexuality in the absence of bias. We also present analytical (QLE) approximations for the selective forces acting on sex in terms of the effect of sex on the mean and variance in fitness among offspring.

INTRODUCTION

Various possible evolutionary benefits of sexual reproduction have been proposed in order to explain the widespread occurrence of this reproductive mode among eukaryotes (e.g., Agrawal, 2006; Otto, 2009; Hartfield and Keightley, 2012). These broadly fall into two categories: direct selective advantages of meiotic recombination, in particular in terms of DNA repair (e.g., Bernstein et al., 1985, 1988), or indirect benefits stemming from the disruption of linkage disequilibria and other forms of genetic associations through recombination and segregation. Breaking genetic associations affects the mean fitness of offspring when the fitness effect of alleles depends on the genetic background (dominance, epistasis); it may also affect the variance in fitness among offspring, and thus the response to selection. In the absence of dominance or epistasis and under random mating, stochastic events occurring in finite populations tend to generate negative genetic associations — negative linkage disequilibrium between selected loci (Hill and Robertson, 1966; Felsenstein, 1974) and excess heterozygosity in diploids (e.g., Balloux et al., 2003). Breaking these negative associations increases the variance in fitness among offspring and the efficiency of natural selection, favouring higher rates of sex or recombination (Otto and Barton, 1997, 2001; Barton and Otto, 2005; Roze and Barton, 2006; Martin et al., 2006; Roze and Michod, 2010). Multilocus simulation programs showed that selection for recombination generated by such stochastic effects may be strong when sex is rare, but decreases rapidly as the baseline rate of sex in the population increases (Keightley and Otto, 2006; Hartfield et al., 2010; Roze, 2014).

Genetic associations may also be produced by deterministic forces: in particular,

dominance and epistatic interactions between alleles affecting fitness are known to be widespread (e.g., de Visser and Elena, 2007; Martin et al., 2007; Halligan and Keightley, 2009; Manna et al., 2012), and represent another source of linkage disequilibria or deviations from Hardy-Weinberg equilibrium. In randomly mating populations living in a constant environment, breaking associations generated by dominance or epistasis decreases the mean fitness of offspring (segregation or recombination load), generating a short-term cost for sex and recombination (Barton, 1995; Charlesworth and Barton, 1996; Otto, 2003) — this short-term cost may turn into a short-term benefit when mating is non-random or when the environment changes in space or time (Lenormand and Otto, 2000; Otto, 2003; Roze and Lenormand, 2005; Gandon and Otto, 2007; Agrawal, 2009). In a longer term, sex is generally beneficial when interactions cause a negative curvature of the fitness function (e.g., negative epistasis, partially recessive deleterious alleles), generating negative genetic associations that limit the efficiency of selection (Barton, 1995; Otto, 2003). Multilocus simulations including fixed epistasis between loci have suggested that epistatic interactions may only play a secondary role in the evolution of recombination, however, stochastic (Hill-Robertson) effects being often stronger (Otto and Barton, 2001; Keightley and Otto, 2006). Nevertheless, epistatic interactions are known to vary across pairs of loci (e.g., Phillips et al., 2000; de Visser and Elena, 2007; Martin et al., 2007), and this variation (which should generally disfavor recombination, Otto and Feldman, 1997) has not been considered in recent multilocus simulation studies on the evolution of sex and recombination.

Models of stabilizing selection acting on quantitative phenotypic traits represent a simple way of introducing distributions on epistatic interactions (on fitness), including possible compensatory effects between mutations (indeed, a mutation displacing

a phenotypic trait away from the optimum can be compensated by another mutation having the opposite effect on the trait). Interestingly, the predicted distribution of fitness effects and epistatic interactions among mutations obtained from classical models such as Fisher's geometric model of adaptation with a Gaussian shaped fitness function have been shown to accurately describe empirical distributions of epistasis in bacteria and viruses (Martin et al., 2007), justifying the use of such models to explore the effects of the variance in epistasis. Selection for recombination under stabilizing, directional or fluctuating selection acting on one or several polygenic traits has been explored by previous simulation models (Maynard Smith, 1980, 1988; Kondrashov and Yampolsky, 1996). They showed that while recombination is disfavored under stabilizing selection around a fixed optimum, environmental change may favor recombination. A mathematical analysis based on the infinitesimal model was proposed by Charlesworth (1993) (see also Appendix 2 in Barton, 1995), showing that recombination increases the phenotypic variance by breaking negative genetic associations generated by epistatic interactions among loci, thereby increasing the speed of adaptation.

As in most evolutionary quantitative genetics models, the studies just mentioned assume unbiased mutational effects on phenotypic traits: mutations are always as likely to increase as to decrease the value of a given trait. Several authors explored the effect of mutational bias on quantitative traits (e.g., Waxman and Peck, 2003; Zhang and Hill, 2008; Charlesworth, 2013a,b), and showed that such a bias may significantly reduce the mean fitness of populations in regimes where drift has substantial effects at loci coding for the traits, by displacing mean phenotypes away from their optimal values (thereby introducing a component of directional selection). Although the effect of mutational bias has only been explored in sexual populations, it should in

principle be stronger in asexual populations, due to their reduced effective population size caused by interference effects between loci. This may generate selection for sex and recombination in the absence of environmental change.

In this paper, we explore the effect of mutational bias in a simple, isotropic model of stabilizing selection acting on an arbitrary number of phenotypic traits, in a haploid, facultatively sexual population. We first assume a fixed rate of sex in the population, and show that mutational bias may strongly reduce the mean fitness of populations in which sex is rare or absent, provided that mutations affecting phenotypic traits have weak to moderate fitness effects. We then introduce genetic variation for the rate of sex, and show that the equilibrium rate of sex is an increasing function of the degree of mutational bias. Finally, we use the methods of Barton (1995) and Charlesworth and Barton (1996) to express different components of selection for sex in terms of the effect of sex on the mean fitness and additive variance in fitness among offspring, and show that these expressions provide correct predictions when selection is sufficiently weak.

METHODS

Life cycle. The different parameters and variables of the model are summarized in Table 1. We consider a population of N haploid organisms with discrete generations. Each individual may generate a fraction of its offspring asexually (by mitosis), the remaining fraction being produced sexually. In the last case, gametes are produced by mitosis and fuse at random in the population to form zygotes, which immediately undergo meiosis to produce haploid juveniles. We will first consider that all individuals

invest equally into sexual reproduction, the parameter σ representing the rate of sex in the population (proportion of sexually produced offspring): $\sigma = 0$ corresponds to obligate asexual reproduction, and $\sigma = 1$ to obligate sex. In a second step (described below), we will introduce genetic variation for the rate of sex. We assume that individuals are hermaphroditic (generating both male and female gametes) and produce very large (effectively infinite) numbers of juveniles, among which N are sampled randomly to form the next adult generation (note that hermaphroditic haploid individuals occur in some species of mosses, ferns and algae).

Throughout the paper, fitness W denotes the overall fecundity of an individual and depends on the values of n quantitative phenotypic traits under stabilizing selection, represented by the vector $\mathbf{z} = (z_1, z_2, \dots z_n)$. In the following, we use greek letters α, β, γ... to denote phenotypic traits, while latin letters i, j, k... will denote loci. We assume that each phenotypic trait can be decomposed into a genetic and an environmental component:

$$z_\alpha = g_\alpha + e_\alpha \tag{1}$$

where g_α is the individual's genetic contribution to trait α ("breeding value"), and where the environmental effect e_α is independent of the genotype of the individual and is sampled from a Gaussian distribution with mean 0 and variance V_e (the same for all traits). Average phenotypes and breeding values in the population are denoted $\overline{z_\alpha}$ and $\overline{g_\alpha}$ (with $\overline{z_\alpha} \approx \overline{g_\alpha}$ when the population is sufficiently large). As we assume no genotype $\times$ environment interaction, the variance of trait α is given by:

$$V_\alpha = V_{\mathrm{g},\alpha} + V_\mathrm{e} \tag{2}$$

where $V_{\mathrm{g},\alpha}$ is the genetic variance for trait α (variance of g_α). The genetic covariance

between traits α and β (covariance between g_α and g_β) will be denoted $C_{\mathrm{g},\alpha\beta}$. Finally, $\langle X \rangle$ will denote the expected value of the quantity X at mutation-selection-drift equilibrium: for example, $\langle V_{\mathrm{g},\alpha} \rangle$ is the average genetic variance for trait α at equilibrium.

As we will see, some of our analytical results on the selective forces acting on the rate of sex do not depend on the specific shape of the fitness function. However, our simulation programs and some of our approximations assume an isotropic, Gaussian-shaped fitness function around the phenotypic optimum, located at $\mathbf{z} = (0, 0, \ldots 0)$:

$$W = \exp\left[-\frac{\sum_{\alpha=1}^{n} z_\alpha^2}{2\omega^2}\right], \tag{3}$$

where ω^2 represents the strength of selection. The mean fitness associated with a given genotype (obtained by averaging over the distribution of environmental effects e_α) is given by:

$$W_{\mathrm{g}} = W_{\mathrm{g,max}} \exp\left[-\frac{\sum_{\alpha=1}^{n} g_\alpha^2}{2V_{\mathrm{s}}}\right] \tag{4}$$

where $V_{\mathrm{s}} = \omega^2 + V_{\mathrm{e}}$, and where $W_{\mathrm{g,max}} = \left(\omega^2/V_{\mathrm{s}}\right)^{n/2}$ is the mean fitness of an optimal genotype (e.g., Lande, 1976a).

Genetic architecture of traits and mutational bias. We assume that selected traits are coded by ℓ loci with additive effects, so that

$$g_\alpha = \sum_{j=1}^{\ell} g_{\alpha j} \tag{5}$$

where $g_{\alpha j}$ is the contribution of the allele at locus j on trait α. Loci are assumed biallelic (although some of our results on the selective forces acting on sex are valid under more general architectures), the alleles at each locus being denoted 0 and 1. Assuming that an individual carrying allele 0 at all loci is at the phenotypic optimum,

the contribution of locus j on trait α can be written as:

$$g_{\alpha j} = r_{\alpha j}\, X_j, \tag{6}$$

where X_j is an indicator variable equal to 1 if the individual carries allele 1 at locus j (while $X_j = 0$ otherwise), and $r_{\alpha j}$ is the effect of allele 1 at locus j on trait α (note that $r_{\alpha j}$ may be negative). The frequency of allele 1 at locus j is denoted p_j, while $q_j = 1 - p_j$. At each locus, we assume that mutation occurs at the same rate u in both directions (from 0 to 1 and from 1 to 0), while $U = u\ell$ denotes the mutation rate on the whole set of loci affecting selected traits. As in previous works (Chevin et al., 2010; Lourenço et al., 2011; Roze and Blanckaert, 2014), we introduce a parameter m measuring the degree of pleiotropy of mutations: each locus only affects a subset m (sampled randomly and independently for each locus) of the n traits under selection. We assume that the distribution of $r_{\alpha j}$ over all loci affecting trait α has average b and standard deviation a — the same for all traits — without any covariance between mutational effects on the different traits. From equation 4, the average deleterious effect of mutations on $\log W_{\mathrm{g}}$ (in an optimal genotype) is given by:

$$\overline{s_d} = \frac{1}{\ell} \sum_{j=1}^{\ell} \sum_{\alpha=1}^{n} \frac{r_{\alpha j}^2}{2V_{\mathrm{s}}} = \frac{m\,(a^2 + b^2)}{2V_{\mathrm{s}}}. \tag{7}$$

The parameter b represents the degree of mutational bias, since mutation tends to displace mean phenotypes away from the optimum when $b \neq 0$. In the following, mutational bias will be measured using a scaled parameter θ, defined as $\theta = b^2/\left(a^2 + b^2\right)$ and varying between 0 and 1. For a given value of $\overline{s_d}$ (mean fitness effect of mutations), θ will thus allow us to explore a continuum between two extreme situations corresponding to two classical models: $\theta = 0$ corresponds to Fisher's geometrical model without mutational bias, with a variance a^2 of mutational effects and possible compensatory

effects among different mutations, while $\theta = 1$ corresponds to a situation where all mutations have the same fitness effect ($a^2 = 0$) and selection thus becomes directional (alleles 1 are disfavored), without any possible compensatory effect among mutations. Note that a^2 and b^2 are simply expressed in terms of $\overline{s_d}$ and θ, as $a^2 = 2V_\mathrm{s}\,(1-\theta)\,\overline{s_d}/m$ and $b^2 = 2V_\mathrm{s}\,\theta\,\overline{s_d}/m$. Furthermore, equation 4 indicates that the parameters a, b, ω^2 and V_e should only affect changes in genotype frequencies through the scaled parameters $\tilde{a} = a/\sqrt{2V_\mathrm{s}}$ and $\tilde{b} = b/\sqrt{2V_\mathrm{s}}$, since genotypic fitnesses become independent of V_s when expressed in terms of the scaled phenotypic traits $\tilde{g_\alpha} = g_\alpha/\sqrt{2V_\mathrm{s}}$. For a given choice of $\overline{s_d}$ and θ, the results should thus not depend on ω^2 and V_e.

Using the parameters $\overline{s_d}$ and θ (instead of $\tilde{a}$ and $\tilde{b}$) will allow us to change the degree of mutational bias θ (between 0 and 1) while keeping the average fitness effect of mutations $\overline{s_d}$ constant. This is equivalent to the approach used by Zhang and Hill (2008), in which the variance of mutational effects decreases as the degree of mutational bias increases in order to maintain a constant mutational variance V_M, defined as the per generation increase in phenotypic variance due to mutation (in our model, $V_\mathrm{M} = \frac{m}{n}\,U\,(a^2 + b^2) = 2V_\mathrm{s}\,\overline{s_d}\,U/n$). Finally, we can note that while the average coefficient of epistasis (on fitness) between mutations is zero in the absence of bias (e.g., Martin et al., 2007), it becomes negative when $\theta > 0$. Indeed, defining epistasis e as a deviation from additivity of mutational effects on $\log W_\mathrm{g}$, we have (assuming that the number of loci ℓ is large):

$$\overline{e} = -\frac{2}{\ell\,(\ell-1)} \sum_{j \neq k} \sum_{\alpha=1}^{n} \frac{r_{\alpha j}\, r_{\alpha k}}{2V_\mathrm{s}} = -2\,\rho\,\theta\,\overline{s_d} \tag{8}$$

with $\rho = m/n$. In the extreme case when $\theta = 1$ and $\rho = 1$ (all mutations have exactly the same phenotypic effect), epistasis becomes constant for all pairs of mutations and

equals $-2\overline{s_d}$.

Change in phenotypic basis. Due to the symmetry of our model, average trait values and genetic variances at equilibrium should be the same for all traits, while mutational bias ($\theta > 0$) will tend to displace the mean phenotype of the population in the direction of the $(1, 1, \ldots 1)$ vector. Although the effects of mutations on the different traits are not correlated, mutation generates genetic covariances $C_{\mathrm{g},\alpha\beta}$ between traits in the population (of the same magnitude for all pairs of traits), since individuals carrying more 1 alleles in their genomes tend to lie further in the direction of the $(1, 1, \ldots 1)$ vector. For analytical derivations, it is useful to define a new phenotypic basis in which the average mutational bias lies along the first axis and in which the genetic variance-covariance matrix is diagonal, thus eliminating covariances between traits (see Figure 1). This can be done by defining new breeding values ${g_\alpha}'$ as:

$$ {g_1}' = \frac{1}{\sqrt{n}} \sum_{\beta=1}^{n} g_\beta \tag{9} $$

$$ {g_\alpha}' = \frac{1}{\sqrt{(\alpha-1)\,\alpha}} \left[(\alpha-1)\, g_\alpha - \sum_{\beta=1}^{\alpha-1} g_\beta \right], \quad \alpha > 1 \tag{10} $$

(e.g., p. 380 in Anton, 2005). The fitnesses of genotypes in the new basis are still given by equation 4, replacing g_α by ${g_\alpha}'$. The average effect of mutations on $\tilde{g}_1{}' = {g_1}'/\sqrt{2V_\mathrm{s}}$ is given by:

$$ \tilde{b}_1{}' = \frac{1}{\ell} \sum_{j=1}^{\ell} \tilde{r}_{1j}{}' = \frac{1}{\ell} \sum_{j=1}^{\ell} \frac{1}{\sqrt{n}} \sum_{\beta=1}^{n} \frac{r_{\beta j}}{\sqrt{2V_\mathrm{s}}}, \tag{11} $$

yielding:

$$ \tilde{b}_1{}' = \sqrt{\rho\, \theta\, \overline{s_d}} \tag{12} $$

where again $\rho = m/n$ (note that equation 8 may thus be written as $\overline{e} = -2\tilde{b}_1{}'^2$). Due to the mutational bias, $\overline{z_1}'$ will tend to be positive, while the genetic variance along

the first axis ($V_{g,1}'$) will be larger than along the other axes (see Figure 1).

Genetic control of the rate of sex. In order to explore the selective forces acting on reproductive mode, we will assume that a given individual may invest proportions s and $1-s$ of its resources in sexual and asexual reproduction (respectively), and that genetic variation for s exists in the population. As in previous papers (Roze and Michod, 2010; Roze and Otto, 2012; Roze, 2014), we introduce a direct cost of sex c by assuming that the probabilities that an individual is the maternal parent of a juvenile through asexual and sexual reproduction are proportional to $1-s$ and s/c, respectively ($c=1$ in the absence of cost, while $c=2$ corresponds to a twofold cost of sex). This cost may be caused by anisogamy (cost of males): for example $c=2$ when half of the resources invested in sex are used to produce male gametes, assuming that the same amount of resources is needed to produce a female gamete and an asexual spore. Alternatively, the cost may result from the failure of gametes to find a partner (assuming that a proportion $1-1/c$ of gametes are lost), or to extra energetic costs associated with gamete production compared with asexual spore production. The rate of sex σ of an individual is defined as the proportion of sexually produced individuals among its maternally produced offspring, given by:

$$\sigma = \frac{s}{c\,(1-s)+s} \tag{13}$$

($\sigma = s$ in the absence of cost). We assume that, like the other traits, investment in sex can be decomposed into an additive genetic and an environmental component:

$$s = \overline{s} + g_s + e_s \tag{14}$$

where $\overline{s}$ is the average investment in sex in the population, $g_s = \sum_i g_{si}$ (g_{si} being the

effect of the allele present at locus i on investment in sex) while e_s is sampled from a centered Gaussian distribution with variance $V_{\mathrm{e},s}$. These equations assume that the distribution of values of s in the population is not too close to 0 or 1 (otherwise the assumption of additivity may not hold, as s cannot be lower than 0 or higher than 1). As above, the variance in s in the population is given by $V_{\mathrm{g},s} + V_{\mathrm{e},s}$ (where $V_{\mathrm{g},s}$ is the variance in g_s). Throughout the paper, we will assume that loci affecting investment in sex do not affect the traits under stabilizing selection.

Assuming that the variance in s in the population is sufficiently small, the rate of sex σ may also be decomposed into an additive genetic and an environmental component:

$$\sigma = \overline{\sigma} + g_\sigma + e_\sigma \tag{15}$$

where $\overline{\sigma}$ is the mean rate of sex. From equation 13 and 14 (and assuming that g_s and e_s are small, of order ϵ), we have:

$$\overline{\sigma} \approx \frac{\overline{s}}{c\,(1-\overline{s}) + \overline{s}}, \quad V_{\mathrm{g},\sigma} \approx \frac{c^2}{\left[c\,(1-\overline{s}) + \overline{s}\right]^4} V_{\mathrm{g},s} \tag{16}$$

(to leading order in ϵ) where $V_{\mathrm{g},\sigma}$ is the genetic variance for the rate of sex (variance of g_σ).

Simulation programs. Our individual-based simulation programs (written in C++) are available from Dryad, and described in Supplementary File S1. The genome of each individual consists in a single linear chromosome with map length R (average number of cross-overs at meiosis). The ℓ biallelic loci affecting the n traits under stabilizing selection are equally spaced along the chromosome, each of these loci affecting a subset of m randomly sampled traits as described above. Investment in sexual re-

production s is coded by ℓ_s multiallelic loci (with an infinite number of possible alleles per locus), which are also equally spaced along the chromosome (see Figure 2); assuming multiallelic loci ensures that all possible rates of sex between 0 and 1 may be achieved even when the number of loci affecting investment in sex is low. Mutational effects at these loci are sampled from a centered Gaussian distribution with variance a_s^2 (the mutational effect being added to the value coded by the allele before mutation). Investment in sex s is obtained by summing allelic effects at all these ℓ_s loci, and adding an environmental component drawn from a centered Gaussian distribution with variance $V_{\mathrm{e},s}$ (if the value obtained is lower than 0 or higher than 1, it is then set to 0 or 1). In a different version of the program the ℓ_s multiallelic loci do not affect investment in sex (which is fixed), but correspond to neutral loci which are used to estimate the effective population size N_{e}. For this, diversity at each of these neutral loci is computed as $D = 1 - \sum_i p_i^2$ (where p_i is the frequency of allele i), and the effective population size is estimated by $N_{\mathrm{e}} \approx \overline{D} / \left[2\mu \left(1 - \overline{D}\right)\right]$, where $\overline{D}$ is the average diversity over neutral loci and generations, and μ the mutation rate at each neutral locus (generally fixed to 10^{-3}). Simulations with a fixed rate of sex generally lasted 10^5 generations, while simulations in which investment in sex was free to evolve lasted 2×10^6 generations (however the rate of sex generally reached an equilibrium within the first 5×10^5 generations).

EFFECT OF MUTATIONAL BIAS ON MEAN FITNESS

In this section we assume that the rate of sex σ is fixed, and explore the effect of σ and of mutational bias on mean fitness. The mutation load L measures the reduction in mean fitness of the population due to the presence of deleterious alleles, and is defined as:

$$L = 1 - \frac{\overline{W}}{W_{\mathrm{g,max}}}, \tag{17}$$

where $\overline{W}$ is mean fitness and $W_{\mathrm{g,max}}$ the fitness of an optimal genotype. Throughout this section, we assume an isotropic, Gaussian-shaped fitness function (equation 3). Assuming that the variance in log-fitness in the population is small and that population size is large, we have (see Supplementary File S2):

$$\langle L \rangle \approx 1 - \exp\left[-\frac{1}{2V_{\mathrm{s}}} \sum_{\alpha=1}^{n} \left(\langle V_{\mathrm{g},\alpha} \rangle + \langle \overline{g_\alpha}^2 \rangle \right) \right]. \tag{18}$$

In the absence of mutational bias, the effect of deviations of mean phenotypes from their optimal values (the term in $\langle \overline{g_\alpha}^2 \rangle$ in equation 18) is proportional to $1/N_{\mathrm{e}}$, and should thus remain small when N_{e} is sufficiently large (Lande, 1976b; Charlesworth, 2013b). However, in the presence of mutational bias, drift may cause substantial deviations of mean phenotypes away from the optimum (Zhang and Hill, 2008). Simple approximations for the load can be obtained in the regime where selection is negligible relative to drift at all loci. Assuming that the variance of $\overline{g_\alpha}$ due to drift is small, we have $\langle \overline{g_\alpha}^2 \rangle \approx \langle \overline{g_\alpha} \rangle^2$, while $\langle \overline{g_\alpha} \rangle = \sum_{j=i}^{\ell} r_{\alpha j} \langle p_j \rangle$ in our biallelic model. Using equation 12, and the fact that $\langle p_j \rangle = 1/2$ under symmetric mutation when the effect of selection at locus j is neglected, one obtains:

$$\frac{1}{2V_{\mathrm{s}}} \sum_{\alpha=1}^{n} \langle \overline{g_\alpha}^2 \rangle \approx \frac{1}{4} \left(\ell \, \tilde{b}_1{}' \right)^2 \tag{19}$$

where $\tilde{b}_1{}' = \sqrt{\rho\,\theta\,\overline{s_d}}$ is the (scaled) magnitude of mutational bias (along the $z_1{}'$ axis). Furthermore, linkage disequilibria between loci should be close to zero on average when selection is sufficiently weak, in which case the genetic variance for trait α is given by:

$$\langle V_{\mathrm{g},\alpha}\rangle \approx \sum_{j=1}^{\ell} r_{\alpha j}^2 \langle p_j q_j\rangle \tag{20}$$

(e.g., Lynch and Walsh, 1998). Given that $\langle p_j q_j\rangle \approx Nu/\left(1+4Nu\right)$ at mutation-drift balance, one obtains from equation 20:

$$\frac{1}{2V_{\mathrm{s}}}\sum_{\alpha=1}^{n}\langle V_{\mathrm{g},\alpha}\rangle \approx \overline{s_d}\,\frac{NU}{1+4Nu}, \tag{21}$$

finally giving:

$$\langle L\rangle \approx 1-\exp\left[-\overline{s_d}\,\frac{NU}{1+4Nu}-\frac{1}{4}\left(\ell\,\tilde{b}_1{}'\right)^2\right] \tag{22}$$

Equation 22 is equivalent to equation 8 in Roze and Blanckaert (2014) in the absence of mutational bias ($\tilde{b}_1{}' = 0$). It is expected to hold only when selection (measured by $\overline{s_d}$) is so weak that its effect on the distribution of trait values in the population is negligible. As $\overline{s_d}$ increases, $\langle\overline{g_\alpha}\rangle$ and $\langle V_{\mathrm{g},\alpha}\rangle$ depart more and more from the expressions given above; however, simulations indicate that equation 21 stays valid over a wider range of values of $\overline{s_d}$ than equation 19, in agreement with previous observations that selection may have significant effects on mean trait values even when $\langle p_i q_i\rangle$ at each locus is mainly controlled by mutation and drift (Robertson, 1960; Campbell, 1984; Barton, 1989; Charlesworth, 2013a). Based on this, it is possible to derive a better approximation for low $\overline{s_d}$ by taking the effect of selection on $\langle\overline{g_\alpha}\rangle$ into account, while still neglecting the effect of selection on genetic variance (and neglecting linkage disequilibria). This yields (see Supplementary File S2 for derivation):

$$\langle L\rangle \approx 1-\exp\left[-\overline{s_d}\,\frac{NU}{1+4Nu}-\frac{\left(\ell\,\tilde{b}_1{}'\right)^2}{4\left[1+\frac{\overline{s_d}}{n}\left[1+\theta\left(m-1\right)\right]\frac{N\ell}{1+4Nu}\right]^2}\right]. \tag{23}$$

Approximations for the regime where genetic variances are significantly affected by mutation, selection and drift are more difficult to obtain. Under very strong selection against mutant alleles (so that the contribution to future generations of individuals deviating from the optimum can be neglected), the mutation load becomes (for both sexual and asexual populations):

$$L \approx 1 - e^{-U} \tag{24}$$

(e.g., Kimura and Maruyama, 1966). Under sexual reproduction, equation 24 also holds under weaker selection in the absence of mutational bias ($\theta = 0$), as long as drift and linkage disequilibria may be neglected (e.g., Bürger, 1998, Supplementary File S2). In the case of an asexual population, an expression for the load at mutation-selection balance (still in the absence of mutational bias, and neglecting drift) can be obtained assuming a Gaussian distribution of trait values in the population:

$$L \approx 1 - \exp\left[-\sqrt{\frac{n}{2} U \overline{s_d}}\right] \tag{25}$$

(Lande, 1980a; Roze and Blanckaert, 2014). Generalizing these expressions to introduce mutational bias is not straightforward in the context of our biallelic model, as the degree of mutational bias changes depending on the position of mean phenotypes; however, previous studies have shown that the effect of mutational bias is generally small in regimes where drift is negligible (Waxman and Peck, 2003; Zhang and Hill, 2008). In Supplementary File S2, we show that a deterministic approximation for the load in a sexual population under the maximum level of bias ($\theta = 1$) is given by:

$$L \approx 1 - \exp\left[-\frac{4\rho U - \overline{s_d} + \sqrt{\overline{s_d}\,(8\rho U + \overline{s_d})}}{8\rho}\right] \tag{26}$$

(see Supplementary File S2 for the same expression in terms of $\tilde{b}_1{}'$, $\overline{s_d}$ and U).

Figure 3 shows the equilibrium mutation load as a function of $\overline{s_d}$, for different rates of sex and levels of mutational bias. In the absence of mutational bias ($\theta = 0$), the load is generally higher in sexual ($\sigma = 1$) than in asexual ($\sigma = 0$) populations, due to the fact that recombination breaks combinations of alleles with compensatory effects (recombination load). This pattern reverses for high values of $\overline{s_d}$ ($\overline{s_d} = 0.1$ in Figure 3), as the frequency of deleterious alleles is increased by Hill-Robertson effects in asexual populations. While the effect of mutational bias (with $\theta = 0.1$) on the mean fitness of sexual populations stays modest, it greatly increases the load of asexual populations for small values of $\overline{s_d}$ (between 10^{-5} and 10^{-3}) — see Supplementary Figure S1 for results under stronger bias ($\theta = 0.5$). Supplementary Figure S2 shows that this increase in L is caused by deviations of mean phenotypes from the optimum, due to the combined effects of mutational bias and drift. Indeed, Figure 3 shows that the effective population size of asexual populations (estimated from the average diversity at neutral loci, see Methods) is greatly reduced by background selection effects.

As shown by Figure 4, equation 22 correctly predicts the increase in load caused by mutational bias at very low values of $\overline{s_d}$, but rapidly overestimates L as $\overline{s_d}$ increases, as it neglects the effects of selection (see Supplementary Figure S3 for the relative effects of genetic variance and of deviations of mean phenotypes from the optimum). In the case of sexual populations, equation 23 provides better predictions (dotted curves in Figure 4) but still fails when $\overline{s_d}$ is not very small, as it neglects the effect of selection on genetic variances. In agreement with previous results (Waxman and Peck, 2003; Zhang and Hill, 2008), we find that in sexual populations, the effect of mutational bias stays rather small in the deterministic regime ($N\overline{s_d} \gg 1$). Very strong

levels of bias ($\theta = 0.5,\ 1$) decrease the load in this regime, this effect being correctly predicted by our deterministic approximation for $\theta = 1$ (equation 26): this is due to the fact that mutational bias generates negative epistasis (on average) between deleterious alleles (equation 8), reducing the mutation load of sexual populations (e.g., Kimura and Maruyama, 1966; Kondrashov and Crow, 1988). Figure 5 shows that the effect of mutational bias increases as the number of loci ℓ increases (allowing stronger deviations from the fitness optimum) and as population size decreases (see Supplementary Figures S4 – S6 for results under stronger bias and for the relative effects of genetic variance and of deviations of mean phenotypes from the optimum on the load). The effects of the degree of pleiotropy of mutations m and of the total number of selected traits n are shown on Figure 6. The mutation load increases with the degree of pleiotropy (Figure 6, top panels): indeed, the magnitude of mutational bias $\tilde{b}_1{}'$ increases with $\rho\,\theta$ (with $\rho = m/n$, equation 12). Increasing m while keeping $\rho\,\theta$ constant has only little effect on the load (Figure 6, middle panels), confirming that m mostly affects the load through its effect on $\tilde{b}_1{}'$. Finally, Figure 6 shows that increasing n while keeping m/n (and thus $\tilde{b}_1{}'$) constant has little effect on the load in sexual populations, while it increases the load of asexual populations due to stronger deviations of mean phenotypes from the optimum (see Supplementary Figures S7). Indeed, increasing the dimensionality n of the fitness landscape reduces the chances that a deleterious allele can be compensated by mutations at other loci, and thus enhances the effect of mutational bias in asexuals.

Overall, these results show that the combined action of mutational bias and genetic drift may greatly reduce the mean fitness of asexual populations when the average fitness effect of mutations is small to moderate, this increase in load being

maximized for intermediate strengths of selection against deleterious alleles $\overline{s_d}$, higher values of pleiotropy m/n, number of selected traits n and number of loci ℓ, and for lower values of population size N. In the next section, we will see how this translates into selection on modifier genes affecting the rate of sex of individuals.

EVOLUTION OF SEX

Analytical approximations. Expressions for the effect of selection on the rate of sex are derived in Supplementary File S3, assuming weak selection, a Gaussian distribution of traits affecting fecundity (the z_α's) and a low variance for the rate of sex in the population. Under these assumptions, the change in the mean rate of sex over one generation ($\Delta\overline{\sigma}$) can be decomposed into two terms, representing the effect of the cost of sex (direct selection), and indirect selection caused by the effect of sex on genetic associations between loci affecting fecundity:

$$\Delta\overline{\sigma} = \Delta_{\text{cost}}\,\overline{g_\sigma} + \Delta_{\text{ind}}\,\overline{g_\sigma}\,. \tag{27}$$

As shown in Supplementary File S3, $\Delta_{\text{cost}}\,\overline{g_\sigma} \approx \beta_{\text{cost}}\,V_{\text{g},\sigma}$, where

$$\beta_{\text{cost}} \approx -\frac{c-1}{1+(c-1)\,\overline{\sigma}} \tag{28}$$

represents the direct selection gradient (selecting against sex when $c > 1$). Indirect selection in turn decomposes into two terms, sometimes called the "short-term" and "long-term" effect of breaking genetic associations (e.g., Agrawal, 2006):

$$\Delta_{\text{ind}}\,\overline{g_\sigma} = \Delta_{\text{short}}\,\overline{g_\sigma} + \Delta_{\text{long}}\,\overline{g_\sigma}\,. \tag{29}$$

The short-term effect is due to the fact that, in the presence of epistatic interactions, breaking genetic associations between loci affects the mean fitness of offspring. Under

our isotropic fitness function (equation 3), and assuming that phenotypes are measured in a basis that eliminates covariances between traits (the basis defined by equations 9 and 10), $\Delta_{\text{short}}\,\overline{g_\sigma}$ is given by:

$$\Delta_{\text{short}}\,\overline{g_\sigma} \approx \sum_{\alpha=1}^{n} \frac{\partial \ln \overline{W}}{\partial V_{\text{g},\alpha}}\, M_{\text{g},\sigma\alpha\alpha} \tag{30}$$

where $M_{\text{g},\sigma\alpha\alpha}$ is the third moment $\text{E}\left[(g_\sigma - \overline{g_\sigma})\,(g_\alpha - \overline{g_\alpha})^2\right]$ (where E stands for the average over all individuals). A more general expression for arbitrary fitness function is given in the Appendix (see Supplementary File S3 for derivation). Under stabilizing selection, $\partial \ln \overline{W}/\partial V_{\text{g},\alpha}$ is negative (mean fitness decreases as the genetic variance for selected traits increases). Furthermore, selection tends to generate associations (linkage disequilibria) between alleles at different loci with compensatory effects on selected traits, thereby reducing $V_{\text{g},\alpha}$. By breaking these associations, sex increases the genetic variance among offspring: therefore, the genetic variance tends to be higher among individuals that engage more in sex (*i.e.*, with higher values of g_σ) than among individuals that engage less in sex, translating into a positive value of $M_{\text{g},\sigma\alpha\alpha}$. The term representing the short-term effect ($\Delta_{\text{short}}\,\overline{g_\sigma}$) is thus negative, corresponding to the short-term cost of breaking genetic associations that have been generated by selection — one can show that this term is equivalent to the term in δV_g in Charlesworth's (1993) recombination modifier model, see also Appendix 2 of Barton (1995).

The long-term effect stems from the fact that increasing the genetic variance among offspring allows a better response to directional selection, and can be written as:

$$\Delta_{\text{long}}\,\overline{g_\sigma} \approx \sum_{\alpha=1}^{n} \frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}}\, C_{\text{g},\sigma\alpha} \tag{31}$$

where $C_{\text{g},\sigma\alpha} = \text{E}\left[(g_\sigma - \overline{g_\sigma})\,(g_\alpha - \overline{g_\alpha})\right]$ is the genetic covariance between the rate of sex

σ and trait α. Equation 31 corresponds to the classical expression describing the effect of selection on correlated characters (Lande, 1979): if selection favors higher values of trait α ($\partial \ln \overline{W}/\partial \overline{z_\alpha} > 0$), a positive genetic covariance between traits α and σ will lead to the evolution of higher values of σ. In our model, directional selection is caused by mutational bias displacing mean phenotypes from the optimum, and thus occurs along the first phenotypic axis of the basis defined by equations 9 and 10 ($\partial \ln \overline{W}/\partial \overline{z_\alpha} = 0$ along all other axes). Because sex increases the response to directional selection by increasing the genetic variance among offspring, trait values tend to be closer to the optimum in individuals that engage more in sex: $C_{\mathrm{g},\sigma\alpha}$ has the same sign as $\partial \ln \overline{W}/\partial \overline{z_\alpha}$, and $\Delta_{\mathrm{long}}\, \overline{g_\sigma}$ is thus positive — this term is equivalent to the term in $\delta \overline{z}$ in Charlesworth (1993).

Charlesworth (1993) and Barton (1995) showed how the short-term and long-term effect can be expressed in terms of mean trait values and genetic variances for selected traits in a recombination modifier model, neglecting the effects of genetic drift on genetic associations and using a quasi linkage equilibrium (QLE) approximation. An equivalent derivation for the case of the present model is given in Supplementary File S3, the main results being summarized in the Appendix. For this, we assume an infinite population size, large number of loci affecting fecundity, weak selection and low variance for the rate of sex in the population; we also assume that the rate of sex is not too low (for the QLE approximation to hold). One obtains that $\Delta_{\mathrm{short}}\, \overline{g_\sigma} \approx \beta_{\mathrm{short}}\, V_{\mathrm{g},\sigma}$, $\Delta_{\mathrm{long}}\, \overline{g_\sigma} \approx \beta_{\mathrm{long}}\, V_{\mathrm{g},\sigma}$, where the short and long-term selection gradients are given by:

$$\beta_{\mathrm{short}} \approx -\frac{1}{2{V_{\mathrm{s}}}^2\, r_{\mathrm{h},1}\, \overline{\sigma}^2} \left(\sum_{\alpha=1}^{n} {V_{\mathrm{g},\alpha}}^2 - \frac{\overline{z_1}^2\, {V_{\mathrm{g},1}}^2}{V_{\mathrm{s}}} \right), \tag{32}$$

$$\beta_{\mathrm{long}} \approx \left(\frac{1}{r_{\mathrm{h},2}\, \overline{\sigma}} - \frac{1}{r_{\mathrm{h},1}} \right) \frac{1}{\overline{\sigma}^2} \frac{\overline{z_1}^2\, {V_{\mathrm{g},1}}^2}{{V_{\mathrm{s}}}^3}. \tag{33}$$

Equations 32 and 33 assume that traits are measured in the phenotypic basis given by equations 9 and 10, so that only the first phenotypic trait (with average $\overline{z_1}$ and genetic variance ${V_{\mathrm{g},1}}^2$) is under directional selection. The terms $r_{\mathrm{h},1}$ and $r_{\mathrm{h},2}$ that appear in the denominators of β_{short} and β_{long} correspond to harmonic mean recombination rates among loci. Defining r_{ijk} as the probability that at least one recombination event occurs at meiosis between a locus i affecting investment in sex and loci j and k affecting selected traits, $r_{\mathrm{h},1}$ is the harmonic average of r_{ijk} over all possible triplets of loci i, j and k, while $r_{\mathrm{h},2}$ is the harmonic average of $r_{ij}\, r_{ijk}$, where r_{ij} is the recombination rate between loci i and j. The maximum possible values of $r_{\mathrm{h},1}$ and $r_{\mathrm{h},2}$ (obtained for the case of freely recombining loci) are thus 3/4 and 3/8, respectively.

Equations 32 and 33 indicate that both the short-term and long-term selection gradients increase as the mean rate of sex in the population $\overline{\sigma}$ decreases, β_{long} increasing more rapidly (due to the term in $1/\overline{\sigma}^3$). However, both expressions diverge as $\overline{\sigma}$ tends to zero, due to the QLE approximation. Equation 33 also shows that the long-term effect vanishes in the absence of mutational bias ($\overline{z_1} = 0$). The genetic architecture of investment in sex affects β_{long} and β_{long} through $r_{\mathrm{h},1}$ and $r_{\mathrm{h},2}$. Provided that the number of loci affecting fecundity is large and that their distribution over the genome is relatively uniform, the harmonic averages of r_{ijk} and $r_{ij}\, r_{ijk}$ over all j and k should be similar for all loci i affecting investment in sex, and the indirect selection gradient should thus be little affected by the number of loci coding for the rate of sex.

As we have seen in the previous section, it is difficult to obtain general analytical expressions for mean trait values ($\overline{z_1}$) and genetic variances ($V_{\mathrm{g},\alpha}$) at mutation-selection-drift equilibrium under mutational bias, for arbitrary values of $\overline{s_d}$ and $\overline{\sigma}$, and we were thus not able to express the mean rate of sex in the population at equilibrium

in terms of the different parameters of the model. One can note, however, that the approximations above for the short and long term selection gradients can be expressed in terms of the effect of sex on the average and variance in fitness among offspring, that could (at least in principle) be measured from an experimental population (Barton, 1995; Charlesworth and Barton, 1996). Indeed, denoting $\overline{W}_{\rm sex}$ and $\overline{W}_{\rm asex}$ the mean fitness of sexually and asexually produced offspring (respectively), and $\mathrm{Var}_{\mathrm{A},\,\mathrm{sex}}(\ln W)$, $\mathrm{Var}_{\mathrm{A},\,\mathrm{asex}}(\ln W)$ the additive variance in log fitness among sexually and asexually produced offspring, we have (see Supplementary File S3):

$$\beta_{\rm short} \approx \frac{\Delta_1}{r_{\rm h,1}\,\overline{\sigma}}, \quad \beta_{\rm long} \approx \left(\frac{1}{r_{\rm h,2}\,\overline{\sigma}} - \frac{1}{r_{\rm h,1}}\right)\frac{\Delta_2}{\overline{\sigma}} \tag{34}$$

with:

$$\Delta_1 = \ln \overline{W}_{\rm sex} - \ln \overline{W}_{\rm asex}, \tag{35}$$

$$\Delta_2 = \mathrm{Var}_{\mathrm{A},\,\mathrm{sex}}(\ln W) - \mathrm{Var}_{\mathrm{A},\,\mathrm{asex}}(\ln W)\,. \tag{36}$$

Equations 34 – 36 are valid in principle for any shape of the fitness function (not necessarily Gaussian), as long as selection is sufficiently weak and the number of selected loci is sufficiently large. However, as the previous results, they assume that genetic associations remain small (QLE approximation), causing them to diverge as the mean rate of sex in the population tends to zero.

Simulation results. Figure 7 shows that, in agreement with the discussion above, the number of loci affecting investment in sex has very little effect on the mean rate of sex in the population ($\overline{\sigma}$) at equilibrium (the numbers 9, 99 and 999 were chosen so that the number of loci affecting fecundity between two loci affecting sex is 1000, 100 and 10, respectively — see Figure 2). As shown by Figures 7 – 9, the population

evolves towards asexuality in the absence of mutational bias ($\theta = 0$), while increasing the magnitude of mutational bias $\tilde{b}_1{}'$ (by increasing either θ or m/n, see equation 12) increases the equilibrium rate of sex. Higher rates of sex evolve under higher values of the mutation rate U, larger numbers of selected loci ℓ and lower values of population size N, due to stronger effects of mutational bias (Figures 8 – 9). Similarly, increasing the dimensionality of the fitness landscape n while keeping m/n constant (so that $\tilde{b}_1{}'$ stays constant) enhances the effect of mutational bias in asexuals (Figures 6, S7), favoring higher rates of sex (Figure 8). The mean fitness effect of deleterious alleles has a non-monotonic effect on selection for sex, the equilibrium rate of sex being maximized for intermediate values of $\overline{s_d}$ (Figure 9). The genome map length R also has a non-monotonic effect on the equilibrium rate of sex (Figure 9): up to a certain point, increasing linkage favors sex since the long-term benefit of sex increases faster than the short-term cost as linkage becomes tighter (as can be seen from equations 32 and 33, and the fact that $r_{\mathrm{h},2}$ decreases faster than $r_{\mathrm{h},1}$ as recombination rates decrease). However, indirect selection vanishes when R tends to zero (since sex becomes genetically equivalent to asexual reproduction), in which case the rate of sex evolves towards zero when sex is costly — Figure 9 shows that low rates of sex may be maintained in the population, probably due to hitchhiking effects between loci affecting investment into sex and loci affecting selected traits. Finally, Figure 8 shows that higher rates of sex are maintained in the absence of a direct cost of sex ($c = 1$), although the rate of sex still evolves towards zero when mutational bias is absent ($\theta = 0$).

Our simulation program was modified in order to test the validity of the QLE approximations shown above (equations 32 – 34) for different values of $\overline{\sigma}$. In this modified version, we introduce genetic variation for investment in sex but constrain

$\overline{\sigma}$ to stay in a given range by sampling the value of alleles at loci affecting sex after mutation from a uniform distribution with variance a_s^2, without adding the value of the allele before mutation. The short and long-term selection gradients were estimated from equations 30 and 31 (divided by $V_{\mathrm{g},\sigma}$), using equations A2 and A3 and measuring the moments $\overline{z_\alpha}$, $V_{\mathrm{g},\alpha}$, $M_{\mathrm{g},\sigma\alpha\alpha}$ and $C_{\mathrm{g},\sigma\alpha}$ for all traits α. For this, the value of g_σ was estimated for each individual from the average rate of sex σ of 100 clonally produced offspring (all with different environmental components of investment in sex e_s), given by equation 13. The terms Δ_1 and Δ_2 of equation 34 were also measured every 100 generations by producing a pool of offspring by sexual reproduction and another pool by asexual reproduction, and measuring the mean fitness and additive variance in log fitness within each pool of offspring. The additive variance in log fitness was estimated from the covariance in log fitness $\mathrm{Cov}(\ln W)$ between sexually (or asexually) produced offspring and their own sexually produced offspring, using $\mathrm{Var}_{\mathrm{A}}(\ln W) = 4\mathrm{Cov}(\ln W) - \mathrm{Var}(\ln W)$ (Lynch and Walsh, 1998, Supplementary File S3). Figure 10 shows that the QLE approximation provides correct predictions of the indirect selection gradients when selection is sufficiently weak ($\overline{s_d} = 10^{-4}$, for the parameter values used in Figure 10) and for intermediate rates of sex (while the QLE expressions diverge as $\overline{\sigma}$ approaches zero). Discrepancies appear for $\overline{s_d} = 10^{-3}$, however, and become more important for $\overline{s_d} = 10^{-2}$. These discrepancies are probably due to a breakdown of the different assumptions used to derive equations 32 – 36 (e.g., weak genetic associations, negligible effect of associations involving more than 2 or 3 loci, distribution of breeding values close to a Gaussian distribution), and possibly also to the effect of drift on genetic associations (through the Hill-Robertson effect), which is not taken into account in our analysis.

DISCUSSION

Epistasis and drift are the two major sources of genetic associations that have been considered in theoretical studies on the benefits of sex and recombination. Epistasis may favor recombination when it is negative on average, that is, when the fitness effect of a deleterious allele is increased by the presence of other deleterious alleles at other loci, or conversely when the fitness effect of a beneficial allele is decreased by the presence of other beneficial alleles in the genome. However, epistatic interactions also generate a short-term cost for recombination (since recombinant offspring tend to have a lower mean fitness than their parents in a constant environment), so that high rates of recombination can only be favored when epistasis is weak relative to the strength of selection, and not too variable across loci (Barton, 1995; Otto and Feldman, 1997). Epistatic interactions (on fitness) arise naturally in models of selection acting on quantitative phenotypic traits. In agreement with the results mentioned above, Gaussian (or quadratic) stabilizing selection around a fixed optimum in an infinite population is expected to disfavor recombination in the absence of mutational bias (Charlesworth, 1993). Indeed, at equilibrium the mean phenotype of the population is centered on the optimum, in which case epistasis between deleterious alleles is zero on average, with a given variance (Martin et al., 2007) — epistasis between two alleles displacing the phenotype in the same direction is negative (due to the negative curvature of the fitness function), while epistasis between alleles having opposite (compensatory) effects on the phenotype is positive. Away from the optimum, epistasis between deleterious alleles is negative on average (while epistasis between beneficial alleles is also negative, e.g., Martin et al., 2007), generating a deterministic advantage for recombination

(Charlesworth, 1993).

Our simulation results confirm that, in the absence of mutational bias on phenotypic traits, populations evolve towards obligate asexuality when the phenotypic optimum remains constant over time, even when population size is finite. This stands in contrast with previous simulation results assuming fixed epistasis across loci (always negative or always positive), that found only minor effects of epistasis compared with the stochastic (Hill-Robertson) effects that favor recombination in initially asexual (or non recombining) populations (Keightley and Otto, 2006). When mutational bias is included in the model, however, positive rates of sex are maintained in the population at equilibrium. Indeed, mutational bias tends to displace mean phenotypes away from the optimum (thereby increasing the mutation load), this effect being stronger in asexual populations in which the variance in fitness may be greatly lowered by negative associations between loci, reducing their ability to respond to directional selection. Extending Barton's (1995) QLE analysis to our model, we obtained deterministic approximations for the short and long-term indirect selection gradients acting on sex in terms of mean trait values and genetic variances, and showed that these approximations provide reasonable predictions when selection acting at the different loci is sufficiently weak and when the rate of sex is not too low (Figure 10). This implies that, in this parameter range, selection for sex is mainly driven by negative linkage disequilibria caused by epistasis (although drift may play a significant role by increasing the distance between the mean phenotype and the optimum, and therefore the magnitude of directional selection). The Hill-Robertson effect may become more important in parameter ranges where the QLE approximation fails (strong selection and/or low rate of sex); however, the lack of suitable analytical method to cover such

regimes makes it difficult to assess its relative effect.

Recent experimental evolution studies showed that higher rates of sex or outcrossing may evolve in populations adapting to a new environment or coevolving with a pathogen, possibly through the generation of advantageous genotypes by recombination and segregation (Becks and Agrawal, 2010, 2012; Morran et al., 2011; Luijckx et al., 2017). In adapting populations of monogonont rotifers, Becks and Agrawal (2012) showed that sexually produced offspring tend to have a lower mean fitness and a higher variance in fitness than asexually produced offspring, in agreement with predictions from models with concave fitness functions such as the one used in this paper. However, how to relate the effect of sex on the mean and variance in fitness of offspring with the strength of indirect selection for sex is not immediately obvious. Transposing Barton's (1995) and Charlesworth and Barton's (1996) analysis of recombination modifier models to our sex modifier model, we showed that simple relations exist between the short and long-term selection gradients for sex and the effect of sex on the fitness of offspring (equations 34 – 36). However, several important caveats must be noted: (i) these relations only hold in the QLE regime, and thus break down when the rate of sex in the population is low; (ii) they depend on average recombination rates between loci affecting fitness and loci affecting the rate of sex (through $r_{\mathrm{h},1}$ and $r_{\mathrm{h},2}$ in equation 34), which are generally unknown (although lower bounds for selection gradients can be obtained by replacing these terms by their values under free recombination, *i.e.* $r_{\mathrm{h},1} = 3/4$ and $r_{\mathrm{h},2} = 3/8$); (iii) the long-term selection gradient is expressed in terms of the effect of sex on the *additive* variance in fitness among offspring, which will generally be more difficult to measure than the variance in fitness. Nevertheless, estimations of the effect of sex on the mean and variance in fitness among offspring

still convey important information on the existence and sign of short and long-term selection gradients on sex (e.g., Peters and Otto, 2003; Sharp and Otto, 2016).

For a given genomic mutation rate U, our model predicts that increasing the dimensionality of the fitness landscape n increases selection for sex (Figure 8). Indeed, the variance of epistasis between mutations decreases as n increases (Martin et al., 2007), epistasis vanishing as n tends to infinity, since mutations become orthogonal in this limit (without any possible compensatory effect). In other words, strong epistatic interactions (in particular, compensatory effects between deleterious alleles) are more likely to occur when the dimensionality of the fitness landscape is low, and these strong interactions tend to favor asexual reproduction (that can maintain coadapted multilocus genotypes). However, we can note that our model assumes that all loci have the same probability of affecting any trait: under a more modular genetic architecture where different sets of loci affect different sets of traits (modular pleiotropy, e.g., Welch and Waxman, 2003; Chevin et al., 2010; Chebib and Guillaume, 2017), the magnitude of epistatic interactions may be more dependent on the average number of traits coded by a given module than on the total number of selected traits, which may lead to different results. In general, the range of realistic values for the dimensionality of fitness landscapes remains difficult to assess: while a large number of traits in an organism may be under selection, many of those traits are probably correlated, reducing the effective dimensionality of the landscape (Martin and Lenormand, 2006). In VSV and ϕX174 viruses, the effective number of selected traits was estimated to be around 10 and 45 (respectively) based on predictions from Fisher's geometric model on the relation between N_e and population mean fitness (Tenaillon et al., 2007; Lourenço et al., 2011), but this number may be much higher in multicellular eukaryotes.

As we have seen, mutational bias is required for sex to be favored in a constant environment. Some evidence for mutational bias on quantitative traits has been obtained from *Drosophila* and *Caenorhabditis elegans* (e.g., Santiago et al., 1992; Lyman et al., 1996; Keightley and Ohnishi, 1998; Ostrow et al., 1997; García-Dorado et al., 1999); however, how to relate these data with the parameter θ measuring bias in our model is not immediately obvious. In particular, a downward mutational bias is often observed on traits that may be seen as fitness components, but such a bias is expected in our model at the optimum even when $\theta = 0$ (since fitness can only decrease at an optimum). Traits that have a less direct relation with fitness sometimes show mutational bias (e.g., metabolite pool size, Davies et al., 2016), sometimes not (e.g., mitotic spindle traits, Farhadifar et al., 2016) but it is again difficult to relate such measures to θ, since the relation between these traits and fitness is generally poorly known. Information on θ may rather be obtained from the distribution of fitness effects of mutations. Indeed, bias causes mutation to push phenotypic traits in a given direction away from the optimum, so that the proportion of beneficial mutations should always stay below 0.5, even for small-effect mutations occurring in a non-optimal genotype. By contrast, in the absence of bias the proportion of beneficial mutations tends to 0.5 as one moves away from the optimum, the convergence to 0.5 being faster for smaller-effect mutations. As a consequence of this high rate of compensatory mutations, drift load generally stays mild in the absence of bias unless population size is very small (the load being roughly proportional to n/N, e.g., Lande, 1980b; Hartl and Taubes, 1998; Poon and Otto, 2000), while it may reach much higher values when mutational bias is present, as shown in the present paper. Compensatory mutations has been best studied in model organisms such as bacteriophages, bacteria, nematodes and yeasts where

they were shown to be common (e.g., Levin et al., 2000; Poon and Chao, 2005; Estes et al., 2011; Szamecz et al., 2014). However, more work is needed to better understand how the rate of compensatory mutations changes with the degree of maladaptation of individuals, in order to gain more insights on realistic levels of mutational bias (as modeled here).

Finally, we can note that the equilibrium rate of sex in the population generally stays small when the cost of sex is moderate to strong (Figures 8, 9), the highest rates of sex being always achieved under complete bias ($\theta = 1$), that is, when compensatory mutations are not possible. Similarly, low levels of costly sex are also maintained in most cases in models on the evolution of sex due to deleterious mutations without epistasis (Roze and Michod, 2010; Roze and Otto, 2012; Roze, 2014). Exploring to what extent higher levels of sex may be maintained in models including environmental change would thus be of interest, and will be the subject of future work.

665 LITERATURE CITED

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

668 ———. 2009. Spatial heterogeneity and the evolution of sex in diploids. Am. Nat.
669 174:S54–S70.

670 Anton, H. 2005. Elementary Linear Algebra, 9th edition. John Wiley & Sons, Hoboken,
671 NJ.

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

674 Barton, N. H. 1989. The divergence of a polygenic system subject to stabilizing selec-
675 tion, mutation and drift. Genet. Res. 54:59–77.

676 ———. 1995. A general model for the evolution of recombination. Genet. Res. 65:123–
677 144.

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

680 Becks, L. and A. F. Agrawal. 2010. Higher rates of sex evolve in spatially heterogeneous
681 environments. Nature 468:89–93.

682 ———. 2012. The evolution of sex is favoured during adaptation to new environments.
683 PLoS Biology 5:e1001317.

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* R. E. Michod and B. R. Levin, eds. The evolution of sex. An examination of current ideas. Sinauer, Sunderland.

Bürger, R. 1998. Mathematical properties of mutation-selection models. Genetica 102/103:279–298.

Campbell, R. B. 1984. The manifestation of phenotypic selection at constituent loci. I. Stabilizing selection. Evolution 38:1033–1038.

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

———. 2013a. Stabilizing selection, purifying selection, and mutational bias in finite populations. Genetics 194:955–971.

———. 2013b. Why we are not dead one hundred times over. Evolution 67:3354–3361.

Charlesworth, B. and N. H. Barton. 1996. Recombination load associated with selection for increased recombination. Genet. Res. 67:27–41.

Chebib, J. and F. Guillaume. 2017. What affects the predictability of evolutionary constraints using a G-matrix? The relative effects of modular pleiotropy and mutational correlation. Evolution 71:2298–2312.

Chevin, L.-M., G. Martin, and T. Lenormand. 2010. Fisher's model and the genomics of adaptation: restricted pleiotropy, heterogenous mutation, and parallel evolution. Evolution 64:3213–3221.

Davies, S. K., A. Leroi, A. Burt, G. J. Bundy, and C. F. Baer. 2016. The mutational structure of metabolism in *Caenorhabditis elegans*. Evolution 70:2239–2246.

de Visser, J. A. G. M. and S. F. Elena. 2007. The evolution of sex: empirical insights into the roles of epistasis and drift. Nat. Rev. Genet. 8:139–149.

Estes, S., P. C. Phillips, and D. R. Denver. 2011. Fitness recovery and compensatory evolution in natural mutant lines of *C. elegans*. Evolution 65:2335–2344.

Farhadifar, R., J. Miguel Ponciano, E. C. Andersen, D. J. Needleman, and C. F. Baer. 2016. Mutation is a sufficient and robust predictor of genetic variation for mitotic spindle traits in *Caenorhabditis elegans*. Genetics 203:1859–1870.

Felsenstein, J. 1974. The evolutionary advantage of recombination. Genetics 78:737–756.

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., C. López-Fanjul, and A. Caballero. 1999. Properties of spontaneous mutations affecting quantitative traits. Genet. Res. 74:341–350.

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

724 Hartfield, M. and P. Keightley. 2012. Current hypotheses regarding the evolution of
725 sex and recombination. Integr. Zool. 7:192–209.

726 Hartfield, M., S. P. Otto, and P. D. Keightley. 2010. The role of advantageous muta-
727 tions in enhancing the evolution of a recombination modifier. Genetics 184:1153–
728 1164.

729 Hartl, D. L. and C. H. Taubes. 1998. Towards a theory of evolutionary adaptation.
730 Genetica 103:525–33.

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

733 Keightley, P. D. and O. Ohnishi. 1998. EMS-induced polygenic mutation rates for nine
734 quantitative characters in *Drosophila melanogaster*. Genetics 148:753–766.

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

737 Kimura, M. and T. Maruyama. 1966. The mutational load with epistatic gene inter-
738 actions in fitness. Genetics 54:1337–1351.

739 Kondrashov, A. S. and J. F. Crow. 1988. King's formula for the mutation load with
740 epistasis. Genetics 120:853–856.

741 Kondrashov, A. S. and L. Y. Yampolsky. 1996. Evolution of amphimixis and recombi-
742 nation under fluctuating selection in one and many traits. Genet. Res. 68:165–173.

743 Lande, R. 1976a. The maintenance of genetic variability by mutation in a polygenic
744 character with linked loci. Genet. Res. 26:221–235.

———. 1976b. Natural selection and random genetic drift in phenotypic evolution. Evolution 30:314–334.

———. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33:402–416.

———. 1980a. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215.

———. 1980b. Genetic variation and phenotypic evolution during allopatric speciation. Am. Nat. 116:463–479.

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

Levin, B. R., V. Perrot, and N. Walker. 2000. Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. Genetics 154:959–997.

Lourenço, J., N. Galtier, and S. Glémin. 2011. Complexity, pleiotropy and the fitness effect of mutations. Evolution 65:1559–1571.

Luijckx, P., E. K. H. Ho, M. Gasim, S. Chen, A. Stanic, C. Yanchus, Y. S. Kim, and A. F. Agrawal. 2017. Higher rates of sex evolve during adaptation to more complex environments. Proc. Natl. Acad. Sci. U. S. A. 114:534–539.

Lyman, R. F., F. Lawrence, S. V. Nuzhdin, and T. F. C. Mackay. 1996. Effects of single *P*-element insertions on bristle number and viability in *Drosophila melangaster*. Genetics 143:277–292.

766 Lynch, M. and J. B. Walsh. 1998. Genetics and Analysis of Quantitative Traits.
767 Sinauer Associates, Sunderland, MA.

768 Manna, F., R. Gallet, G. Martin, and T. Lenormand. 2012. The high-throughput yeast
769 deletion fitness data and the theories of dominance. J. Evol. Biol. 25:892–903.

770 Martin, G., S. F. Elena, and T. Lenormand. 2007. Distributions of epistasis in microbes
771 fit predictions from a fitness landscape model. Nat. Genet. 39:555–560.

772 Martin, G. and T. Lenormand. 2006. A general multivariate extension of Fisher's
773 geometrical model and the distribution of mutation fitness effects across species.
774 Evolution 60:893–907.

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

777 Maynard Smith, J. 1980. Selection for recombination in a polygenic model. Genet.
778 Res. 35:269–277.

779 ———. 1988. Selection for recombination in a polygenic model — the mechanism.
780 Genet. Res. 51:59–63.

781 Morran, L. T., O. G. Schmidt, I. A. Gelarden, R. C. I. Parrish, and C. M. Lively. 2011.
782 Running with the Red Queen: host-parasite coevolution selects for biparental sex.
783 Science 333:216–218.

784 Ostrow, D., N. Phillips, A. Avalos, D. Blanton, A. Boggs, T. Keller, L. Levy, J. Rosen-
785 bloom, and C. F. Baer. 1997. Mutational bias for body size in rhabditid nematodes.
786 Genetics 176:1653–1661.

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

———. 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.

———. 2001. Selection for recombination in small populations. Evolution 55:1921–1931.

Otto, S. P. and M. W. Feldman. 1997. Deleterious mutations, variable epistatic interactions, and the evolution of recombination. Theor. Popul. Biol. 51:134–47.

Peters, A. D. and S. P. Otto. 2003. Liberating genetic variance through sex. BioEssays 25:533–537.

Phillips, P. C., S. P. Otto, and M. C. Whitlock. 2000. Beyond the average: the evolutionary importance of gene interactions and variability of epistatic effects. Pp. 20–38 *in* J. B. Wolf, E. D. Brodie, and M. J. Wade, eds. Epistasis and the Evolutionary Process. Oxford University Press, New York.

Poon, A. and L. Chao. 2005. The rate of compensatory mutation in the DNA bacteriophage ϕX174. Genetics 170:898–899.

Poon, A. and S. P. Otto. 2000. Compensating for our load of mutations: freezing the meltdown of small populations. Evolution 54:1467–1479.

Robertson, A. 1960. A theory of limits in artificial selection. Proc. Roy. Soc. (Lond.) B 153:234–249.

808 Roze, D. 2014. Selection for sex in finite populations. J. Evol. Biol. 27:1304–1322.

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

811 Roze, D. and A. Blanckaert. 2014. Epistasis, pleiotropy and the mutation load in
812 sexual and asexual populations. Evolution 68:137–149.

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

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

817 Roze, D. and S. P. Otto. 2012. Differential selection between the sexes and selection
818 for sex. Evolution 66:558–574.

819 Santiago, E., J. Albornoz, A. Domínguez, M. A. Toro, and C. López-Fanjul. 1992.
820 The distribution of spontaneous mutations on quantitative traits and fitness in
821 *Drosophila melanogaster*. Genetics 132:771–781.

822 Sharp, N. P. and S. P. Otto. 2016. Evolution of sex: using experimental genomics to
823 select among competing theories. BioEssays 38:751–757.

824 Szamecz, B., G. Boross, D. Kalapis, K. Kovács, G. Fekete, Z. Farkas, V. Lázár,
825 M. Hrtyan, P. Kemmeren, M. J. A. Groot Koerkamp, E. Rutkai, F. C. P. Holstege,
826 B. Papp, and C. Pál. 2014. The genomic landscape of compensatory evolution. PLoS
827 Biology 12:e1001935.

828 Tenaillon, O., O. K. Silander, J.-P. Uzan, and L. Chao. 2007. Quantifying organismal
829 complexity using a population genetic approach. PLoS One 2:e217.

830 Waxman, D. and J. R. Peck. 2003. The anomalous effects of biased mutation. Genetics
831 164:1615–1626.

832 Welch, J. J. and D. Waxman. 2003. Modularity and the cost of complexity. Evolution
833 57:1723–1734.

834 Zhang, X.-S. and W. G. Hill. 2008. The anomalous effects of biased mutation revis-
835 ited: mean-optimum deviation and apparent directional selection under stabilizing
836 selection. Genetics 179:1135–1141.

APPENDIX: GENERAL QLE RESULTS

Assuming that the distribution of phenotypic traits affecting fecundity in the population is approximately Gaussian and that selection is weak, a general expression for indirect selection on the rate of sex is given by (see Supplementary File S3 for derivation):

$$\Delta_{\text{ind}}\,\overline{g_\sigma} \approx \sum_{\alpha=1}^{n} \frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}}\, C_{\text{g},\sigma\alpha} + \sum_{\alpha\leq\beta} \frac{\partial \ln \overline{W}}{\partial C_{\text{g},\alpha\beta}}\, M_{\text{g},\sigma\alpha\beta} \tag{A1}$$

where the second sum is over all possible pairs of selected traits, including $\alpha = \beta$. Equation A1 is equivalent to Charlesworth's (1993) decomposition of the selection gradient for a recombination modifier allele into two terms (equation A10 in Charlesworth, 1993, see also Appendix 2 of Barton, 1995). The first term of equation A1 (equivalent to the term in $\delta\overline{z}$ in Charlesworth, 1993) represents indirect selection caused by the effect of sex on mean phenotypes. With our Gaussian, isotropic fitness function (equation 3), we have:

$$\frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}} = -\frac{\overline{z_\alpha}}{V_{\text{g},\alpha} + V_{\text{s}}}, \tag{A2}$$

which is approximately $-\overline{z_\alpha}/V_{\text{s}}$ when selection is weak ($V_{\text{g},\alpha} \ll V_{\text{s}}$). In our model, directional selection occurs along the axis corresponding to the direction of the mutational bias, and therefore only the first term of the sum (for $\alpha = 1$) will contribute when phenotypes are measured in the basis defined by equations 9 and 10. The second term of equation A1 (equivalent to the term in δV_g in Charlesworth, 1993) represents indirect selection caused by the effect of sex on the genetic variance-covariance matrix: $\partial \ln \overline{W}/\partial C_{\text{g},\alpha\beta}$ describes how mean fitness is affected by the genetic covariance between traits α and β, while the third moment $M_{\text{g},\sigma\alpha\beta} = \text{E}\left[(g_\sigma - \overline{g_\sigma})\,(g_\alpha - \overline{g_\alpha})\,(g_\beta - \overline{g_\beta})\right]$ (where E stands for the average over all individuals) describes to what extent the

genetic covariance between traits α and β differs between subsets of the populations with different rates of sex. As shown in Supplementary File S3, under an isotropic, Gaussian fitness function and measuring phenotypes in the basis defined by equations 9 and 10, only the terms with $\alpha = \beta$ differ from zero, and the second term of equation A1 thus becomes $\sum_{\alpha=1}^{n} \left(\partial \ln \overline{W}/\partial V_{\mathrm{g},\alpha}\right) M_{\mathrm{g},\sigma\alpha\alpha}$. The selection gradient $\partial \ln \overline{W}/\partial V_{\mathrm{g},\alpha}$ measures the strength of stabilizing selection on trait α, and is given by:

$$\begin{aligned} \partial \ln \overline{W}/\partial V_{\mathrm{g},\alpha} &= -\frac{1}{2\left(V_{\mathrm{g},\alpha} + V_{\mathrm{s}}\right)} + \frac{1}{2}\left(\frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}}\right)^2 \\ &\approx -\frac{1}{2V_{\mathrm{s}}}\left(1 - \frac{\overline{z_\alpha}^2}{V_{\mathrm{s}}}\right) \end{aligned} \tag{A3}$$

where again the term $\overline{z_\alpha}^2$ will differ from zero only for the first phenotypic trait in the basis defined by equations 9 and 10. Note that the second term of equation A1 does not appear in classic expressions describing the effect of selection on correlated characters (Lande, 1979), as these assume a multivariate Gaussian distribution of phenotypic traits. Here we cannot assume that the joint distribution of the rate of sex σ and of the traits affecting fecundity is multivariate Gaussian: in particular, sex tends to increase $V_{\mathrm{g},\alpha}$ by breaking negative genetic associations (linkage disequilibria between alleles with compensatory effects on trait α), generating a positive third moment $M_{\mathrm{g},\sigma\alpha\alpha}$.

Following Charlesworth (1993) and Barton (1995), the moments $C_{\mathrm{g},\sigma\alpha}$ and $M_{\mathrm{g},\sigma\alpha\beta}$ that appear in equation A1 may be expressed in terms of the genetic variance for the rate of sex $V_{\mathrm{g},\sigma}$ and genetic variances (and covariances) for selected traits using a QLE argument. The derivation (shown in Supplementary File S3) supposes that selection is weak relative to effective recombination rates between loci (and thus that the rate of sex is not too low), so that linkage disequilibria remain small. Furthermore, it neglects the effects of genetic associations involving more than three loci.

Under these assumptions, one obtains for $M_{\mathrm{g},\sigma\alpha\beta}$:

$$M_{\mathrm{g},\sigma\alpha\beta} \approx -\frac{\Delta_{\mathrm{sel}}\mathcal{D}_{\alpha\beta}}{r_{\mathrm{h},1}\,\overline{\sigma}^2}\,V_{\mathrm{g},\sigma} \tag{A4}$$

where $r_{\mathrm{h},1}$ is defined in the main text. The term $\Delta_{\mathrm{sel}}\mathcal{D}_{\alpha\beta}$ in the numerator of equation A4 measures the change in $\mathcal{D}_{\alpha\beta}$ (per generation) due to selection, where $\mathcal{D}_{\alpha\beta}$ is the contribution of linkage disequilibria to the genetic covariance between traits α and β. As shown in Supplementary File S3, when phenotypes are measured in a basis that eliminates covariances between traits, we have (assuming that the number of loci affecting selected traits is large):

$$\Delta_{\mathrm{sel}}\mathcal{D}_{\alpha\beta} \approx \left[(1+I_{\alpha\beta})\,\frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\alpha\beta}} - \frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}}\frac{\partial \ln \overline{W}}{\partial \overline{z_\beta}}\right] V_{\mathrm{g},\alpha}V_{\mathrm{g},\beta} \tag{A5}$$

where $I_{\alpha\beta}$ equals 1 if $\alpha = \beta$, and 0 otherwise. Under an isotropic, Gaussian fitness function, it is possible to show that the term between brackets in equation A5 equals 0 when $\alpha \neq \beta$, while it is approximately $-1/V_{\mathrm{s}}$ when $\alpha = \beta$ (Supplementary File S3, equation A3). In this case, equations A3 – A5 yield equation 32 in the main text.

The QLE expression for the genetic covariance $C_{\mathrm{g},\sigma\alpha}$ that appears in the first term of equation A1 writes (see Supplementary File S3 for derivation):

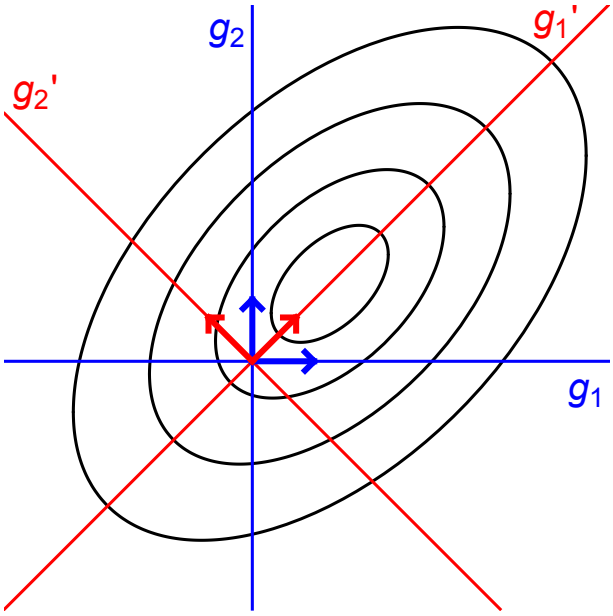
$$C_{\mathrm{g},\sigma\alpha} \approx -\left(\frac{1}{r_{\mathrm{h},2}\,\overline{\sigma}} - \frac{1}{r_{\mathrm{h},1}}\right)\frac{1}{\overline{\sigma}^2}\sum_{\beta=1}^{n}\frac{\partial \ln \overline{W}}{\partial \overline{z_\beta}}\left(\Delta_{\mathrm{sel}}\mathcal{D}_{\alpha\beta}\right)V_{\mathrm{g},\sigma} \tag{A6}$$

where $r_{\mathrm{h},2}$ is defined in the main text. Under an isotropic, Gaussian fitness function, equations A2 – A5 yield equation 33 in the main text.

Table 1: Parameters and variables of the model.

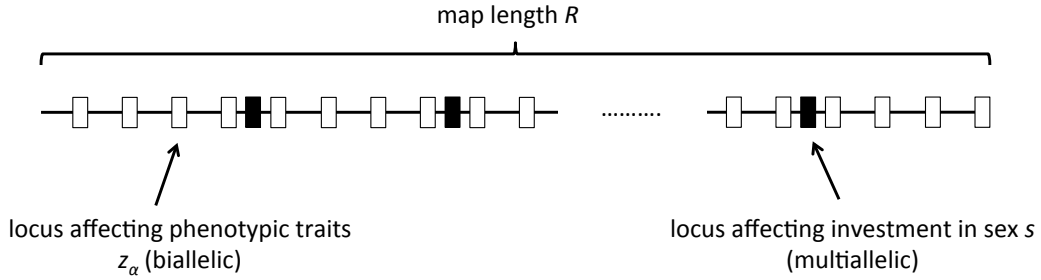
N	Population size
n	Number of selected traits
m	Degree of pleiotropy of mutations
$\rho = m/n$	Scaled pleiotropy
V_{e}	Environmental variance (on selected traits)
ω^2	Strength of stabilizing selection on phenotypic traits
$V_{\mathrm{s}} = \omega^2 + V_{\mathrm{e}}$	Strength of stabilizing selection on breeding values g_α
$W_{\mathrm{g,max}} = \left(\omega^2/V_{\mathrm{s}}\right)^{n/2}$	Mean fitness of an optimal genotype
ℓ	Number of loci affecting selected traits
u	Mutation rate per locus per generation
$U = u\,\ell$	Overall mutation rate on loci affecting selected traits
R	Genome map length
a^2	Variance of mutational effects on selected traits
b	Mutational bias on selected traits
$\theta = b^2/\left(a^2 + b^2\right)$	Scaled mutational bias
$\langle X \rangle$	Expected value of X at mutation-selection-drift equilibrium
$\overline{s_d}$	Average deleterious effect of mutations on log fitness (in an optimal genotype)
z_α	Value of phenotypic trait α (in a given individual)
g_α, e_α	Genetic and environmental components of trait α

$g_{\alpha j}$	Effect of the allele present at locus j on trait α
$\overline{z_\alpha}$	Average value of trait α (in the population)
$V_{\mathrm{g},\alpha}$	Genetic variance for trait α (variance of g_α)
$C_{\mathrm{g},\alpha\beta}$	Genetic covariance between traits α and β
$\mathcal{D}_{\alpha\alpha}$, $\mathcal{D}_{\alpha\beta}$	Effect of linkage disequilibria on $V_{\mathrm{g},\alpha}$ and $C_{\mathrm{g},\alpha\beta}$
$r_{\alpha j}$	Effect of allele 1 at locus j on trait α
p_j, q_j	Frequencies of allele 1 and allele 0 at locus j
s	Investment into sexual reproduction
c	Cost of sex
$\sigma = \frac{s}{c(1-s)+s}$	Rate of sex (proportion of sexually produced offspring among maternally produced offspring)
$\overline{\sigma}$	Mean rate of sex in the population
$V_{\mathrm{g},\sigma}$	Genetic variance for the rate of sex σ
ℓ_s	Number of loci affecting s
U_s	Mutation rate per generation on loci affecting s
a_s^2	Variance of mutational effects on s
$V_{\mathrm{e},s}$	Environmental variance on s



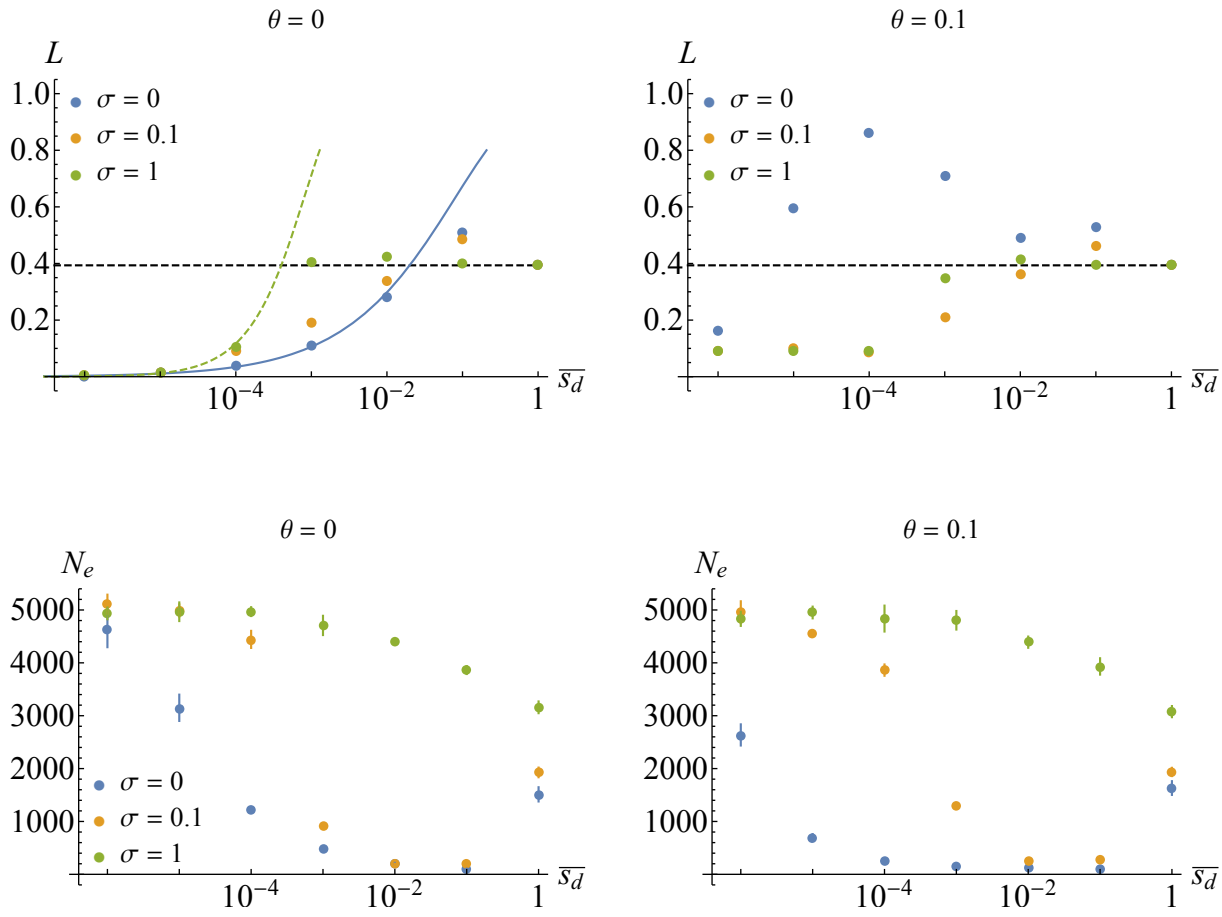
899

900 **Figure 1.** Effect of mutational bias, illustrated for $n = 2$ (the fitness optimum
901 corresponds to the axes' origin). The black curves show the shape of the frequency
902 distribution of individuals with different values of traits g_1 and g_2 (blue axes). Due
903 to the symmetry of our model, mutational bias tends to displace $\overline{g_1}$ and $\overline{g_2}$ from their
904 optimal values by the same amount, and generates a positive covariance among traits.
905 Traits g_1' and g_2' are defined by rotating the phenotypic basis (equations 9 and 10, red
906 axes) so that the covariance between g_1' and g_2' is zero, while mutational bias displaces
907 phenotypes along the g_1' axis.



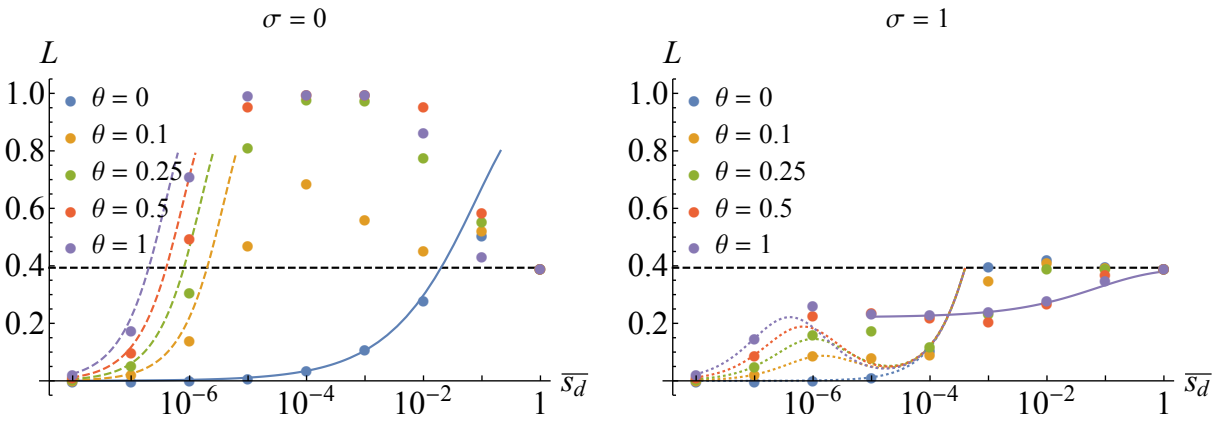
908

909 **Figure 2.** Simulated genetic architecture. Traits affecting fecundity are coded by ℓ
910 biallelic loci uniformly distributed along a chromosome with map length R Morgans.
911 Investment in sex is coded by ℓ_s multiallelic loci, which are also regularly spaced along
912 the chromosome. When $\ell_s = 1$, the locus affecting investment in sex is located at the
913 mid-point of the chromosome.



914

915 **Figure 3.** Top: average mutation load as a function of the mean fitness effect of
916 mutations $\overline{s_d}$, for different rates of sex σ and different degrees of mutational bias θ.
917 Dots: simulation results (note that all points are superposed for $\overline{s_d} = 1$). In this
918 and the following figures, error bars (computed by splitting the last generations of the
919 simulation into 6 batches of 10^4 generations and calculating the standard error over
920 batches) are smaller than the size of symbols in most cases. The horizontal dashed line
921 correspond to equation 24 $(1 - e^{-U})$, the green dashed curve to equation 22 and the
922 solid blue curve to equation 25. Bottom: estimated effective population size N_e (see
923 Methods) for the same parameter values. Parameter values are $N = 5000$, $U = 0.5$,
924 $\ell = 10^4$, $n = 50$, $m = 5$, $R = 10$.



925

926 **Figure 4.** Average mutation load in asexual (left) and sexual (right) populations as a
927 function of the mean fitness effect of mutations $\overline{s_d}$, for different degrees of mutational
928 bias θ. The horizontal dashed lines correspond to equation 24 ($1 - e^{-U}$). Left: the
929 colored dashed curves correspond to equation 22, and the solid blue curve to equation
930 25. Right: the dotted curves correspond to equation 23, and the solid curve to equation
931 26. Parameter values are as in Figure 3.

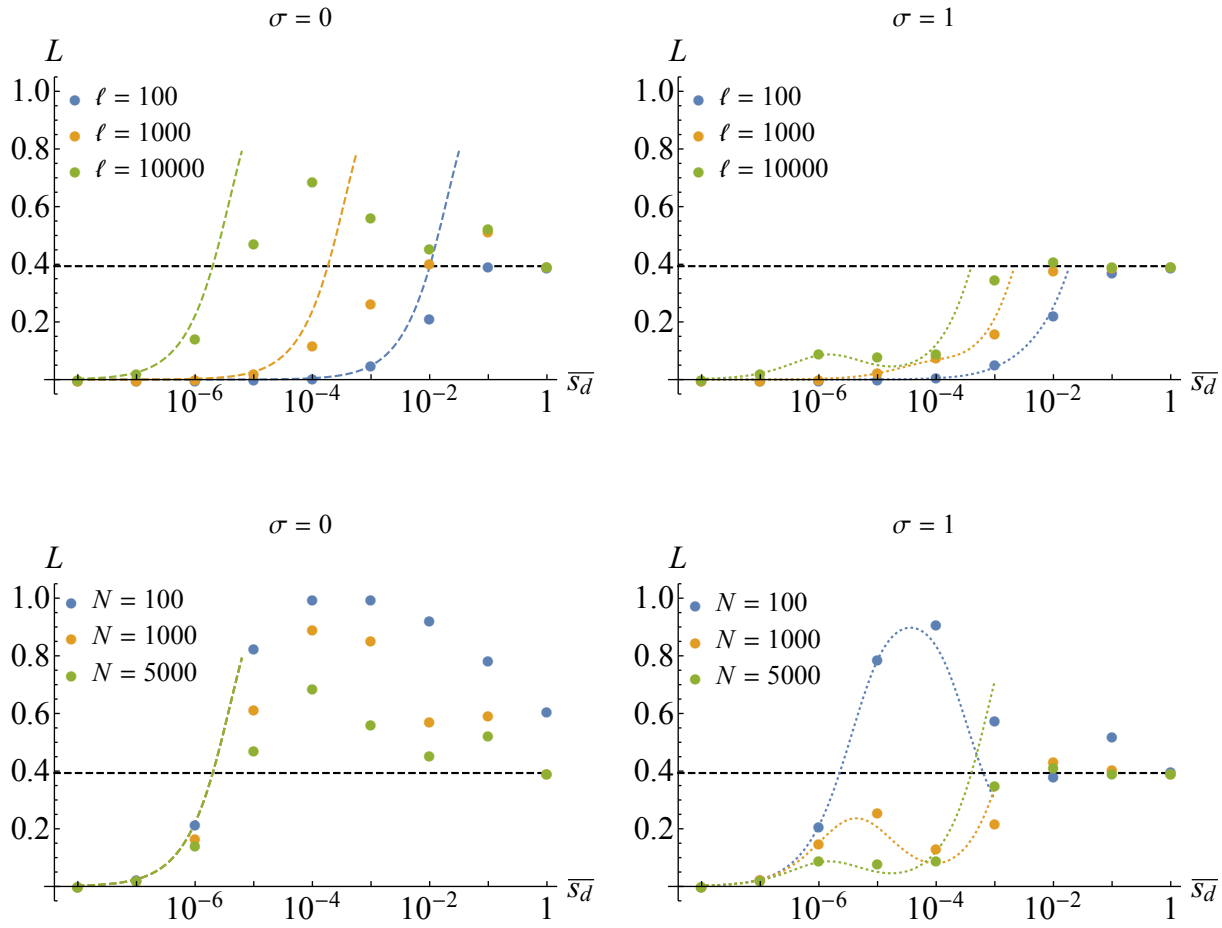
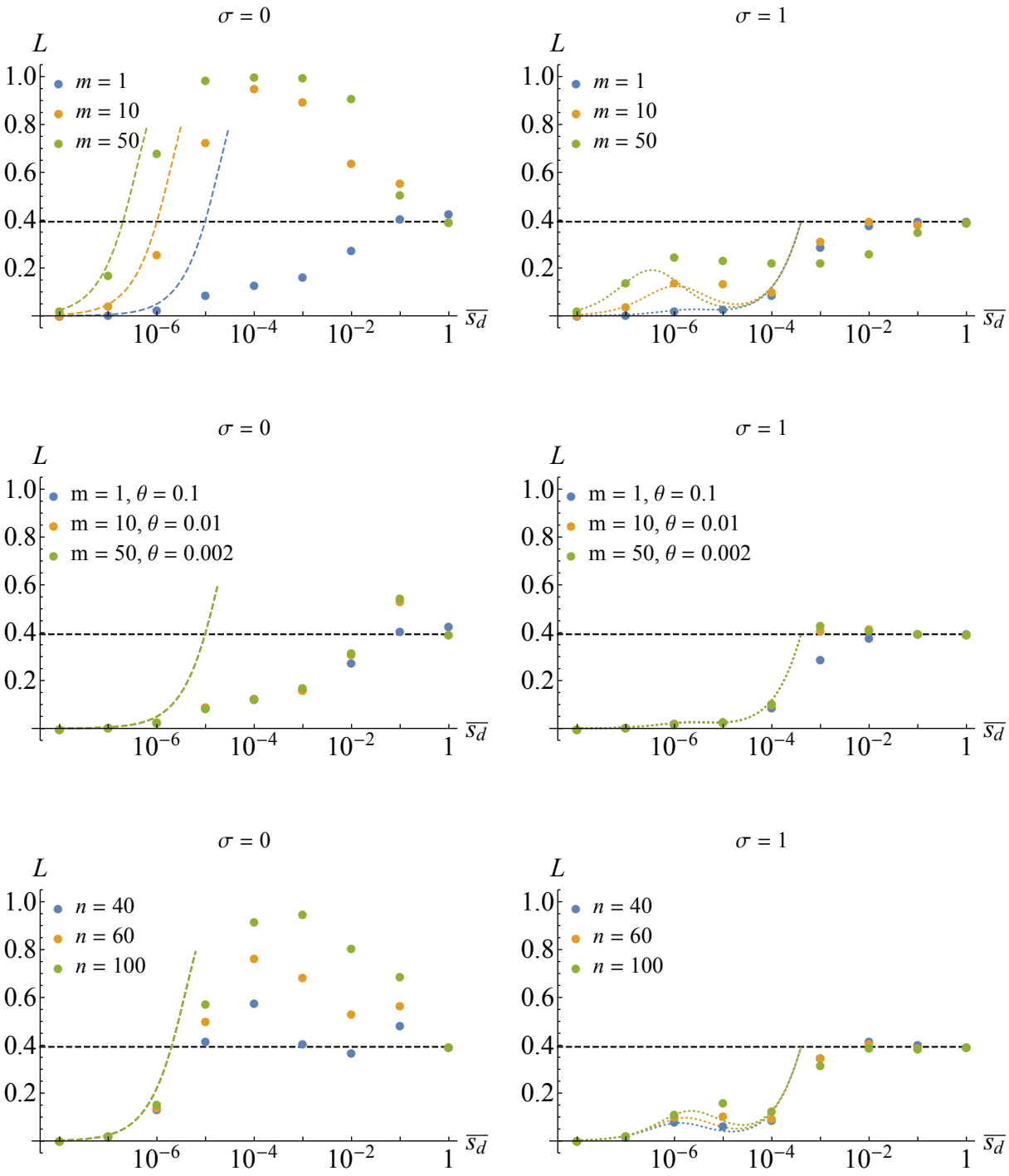


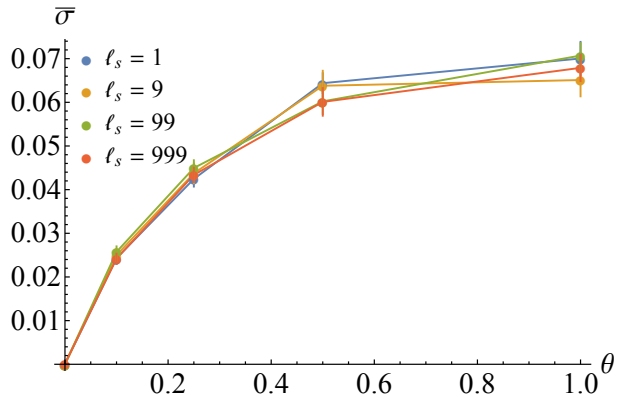
Figure 5. Average mutation load in asexual (left) and sexual (right) populations as a function of the mean fitness effect of mutations $\overline{s_d}$, for different numbers of loci ℓ affecting selected traits (top) and different values of population size N (bottom). The horizontal dashed lines correspond to equation 24 ($1 - e^{-U}$), the dashed curves to equation 22, and the dotted curves to equation 23. Parameter values are as in Figure 3, with $\theta = 0.1$.



939

940 **Figure 6.** Average mutation load in asexual (left) and sexual (right) populations as
941 a function of the mean fitness effect of mutations $\overline{s_d}$, for different degrees of pleiotropy
942 of mutations m and numbers of selected traits n. In the middle panels, $m\,\theta$ is kept

constant by decreasing θ as m increases, while m/n is kept constant in the bottom panels by increasing m as n increases (*i.e.*, $m = 4$, 6 and 10 when $n = 40$, 60 and 100, respectively). The horizontal dashed lines correspond to equation 24 ($1 - e^{-U}$), the dashed curves to equation 22, and the dotted curves to equation 23. Parameter values are as in Figure 3 with $\theta = 0.1$ and $n = 50$ unless specified otherwise.



948

949 **Figure 7.** Mean rate of sex in the population at equilibrium as a function of the
950 degree of mutational bias θ, for different values of the number of loci ℓ_s affecting
951 investment in sex. Parameter values: $N = 5000$, $\overline{s_d} = 10^{-3}$, $n = 50$, $m = 5$, $\ell = 10^4$,
952 $U = 0.5$, $R = 10$, $c = 1.2$, $U_s = 10^{-3}$, $a_s^2 = V_{\mathrm{e},s} = 5 \times 10^{-5}$, initial investment in sex:
953 $s_{\mathrm{init}} = 0.05$. In this and the following figures, error bars were computed by splitting the
954 last generations of the simulation into 15 batches of 10^5 generations and calculating
955 the standard error over batches.

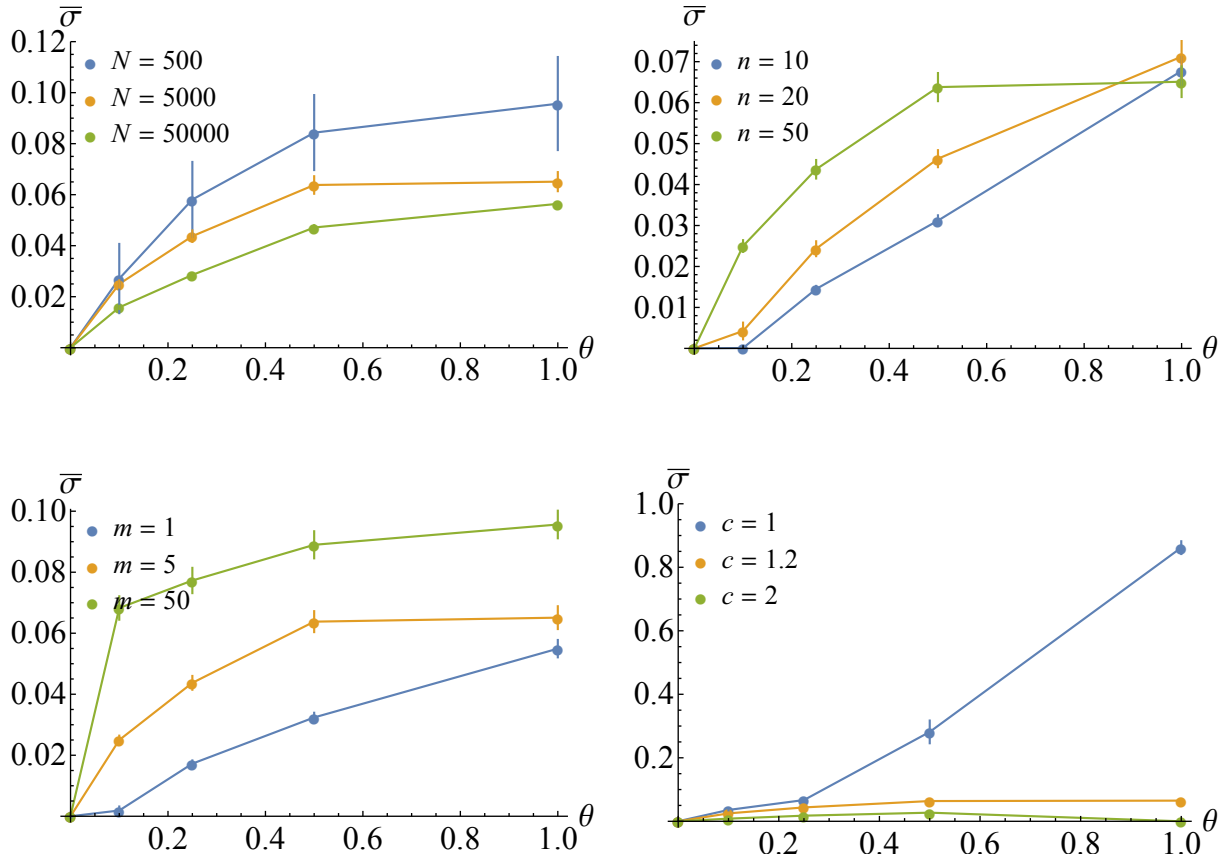
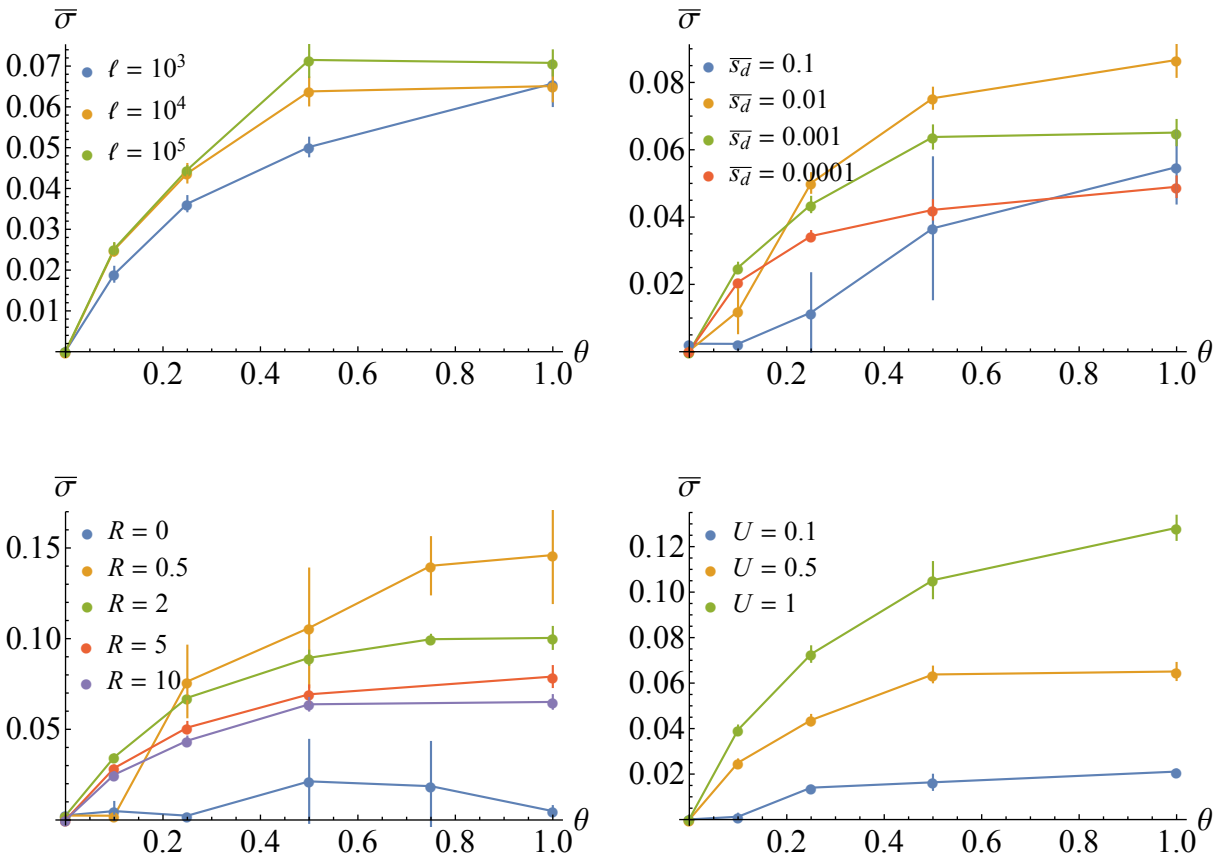


Figure 8. Mean rate of sex at equilibrium as a function of the degree of mutational bias θ, for different values of population size N, number of selected traits n, degree of pleiotropy m and cost of sex c. Parameter values are as in Figure 7 with $\ell_s = 9$ unless specified otherwise. The m/n ratio is kept constant (and equal to 0.1) in the panel showing results for different values of n (top right panel), *i.e.*, $m = 1$, 2 and 5 for $n = 10$, 20 and 50, respectively.



963

964 **Figure 9.** Mean rate of sex at equilibrium as a function of the degree of mutational
965 bias θ, for different values of the number of selected loci ℓ, average deleterious effect
966 of mutations $\overline{s_d}$, genome map length R and overall mutation rate on selected traits U.
967 Parameter values are as in Figure 7 with $\ell_s = 9$ unless specified otherwise.

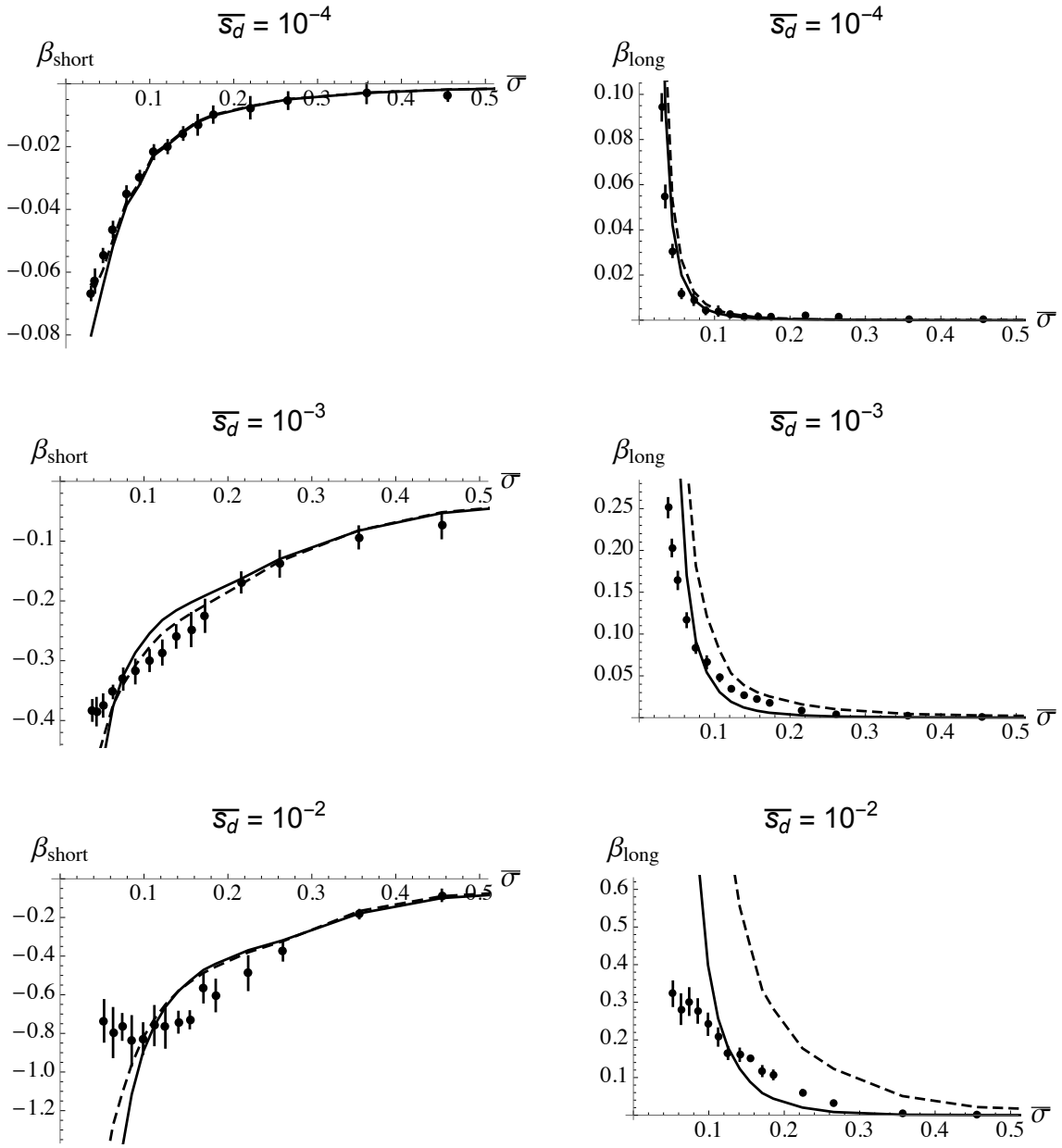


Figure 10. Short and long-term selection gradients for sex as a function of the mean rate of sex in the population, for different values of $\overline{s_d}$. The dots show β_{short} and β_{long} estimated using equations 30 – 31 (divided by $V_{\text{g},\sigma}$) and equations A2 – A3. Solid curves correspond to equations 32 – 33 (using the values of $\overline{\sigma}$, $\overline{z_\alpha}$ and $V_{\text{g},\alpha}$ measured in the simulations), and dashed curves to equations 34 – 36 (where Δ_1 and Δ_2 are measured in the simulations as explained in the main text). Parameter values are as in Figure 7 with $\ell_s = 9$ and $\theta = 0.1$, leading to $r_{\text{h},1} \approx 0.66$ and $r_{\text{h},2} \approx 0.13$.

SUPPLEMENTARY FIGURES

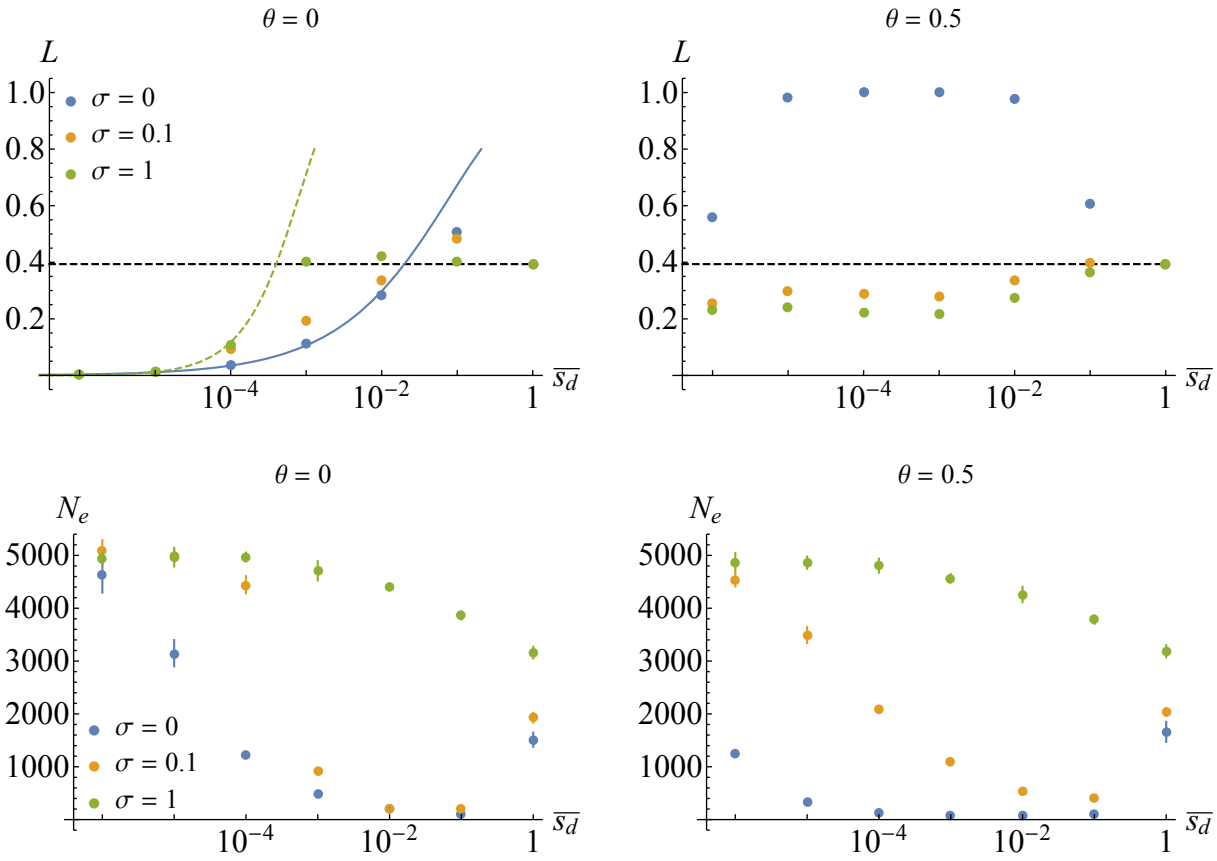


Figure S1. Same as Figure 3 in the main text, comparing $\theta = 0$ and $\theta = 0.5$.

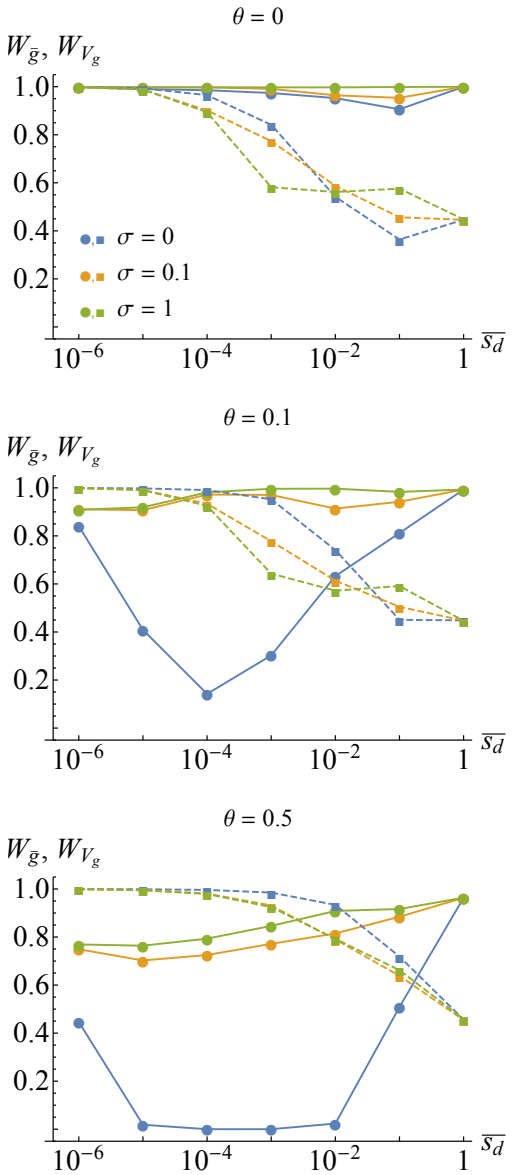


Figure S2. Terms $W_{\bar{g}} = \exp\left[-\sum_{\alpha=1}^{n} \langle \overline{g_\alpha}^2 \rangle / (2V_s)\right]$, $W_{V_g} = \exp\left[-\sum_{\alpha=1}^{n} \langle V_{g,\alpha} \rangle / (2V_s)\right]$ representing the effect of departures of mean phenotypes from the optimum ($W_{\bar{g}}$, circles, solid lines) and the effect of genetic variance (W_{V_g}, squares, dashed lines) on the mutation load ($L \approx 1 - W_{\bar{g}} W_{V_g}$, see Supplementary File S2), for different values of $\overline{s_d}$ and θ. Parameter values are as in Figures 3 and S1. Note that the lines simply connect simulation results and do not correspond to analytical approximations. Mutational bias causes an increase in load through $W_{\bar{g}}$.

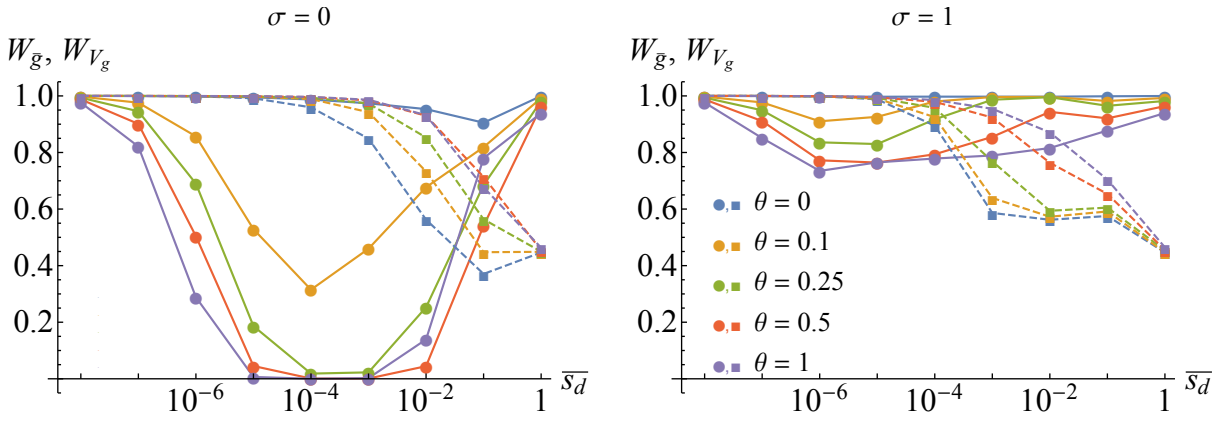


Figure S3. Same as Figure 4 in the main text, showing $W_{\bar{g}} = \exp\left[-\sum_{\alpha=1}^{n} \langle \overline{g_\alpha}^2 \rangle \,/\, (2V_\mathrm{s})\right]$ (circles, solid lines) and $W_{V_\mathrm{g}} = \exp\left[-\sum_{\alpha=1}^{n} \langle V_{\mathrm{g},\alpha} \rangle \,/\, (2V_\mathrm{s})\right]$ (squares, dashed lines).

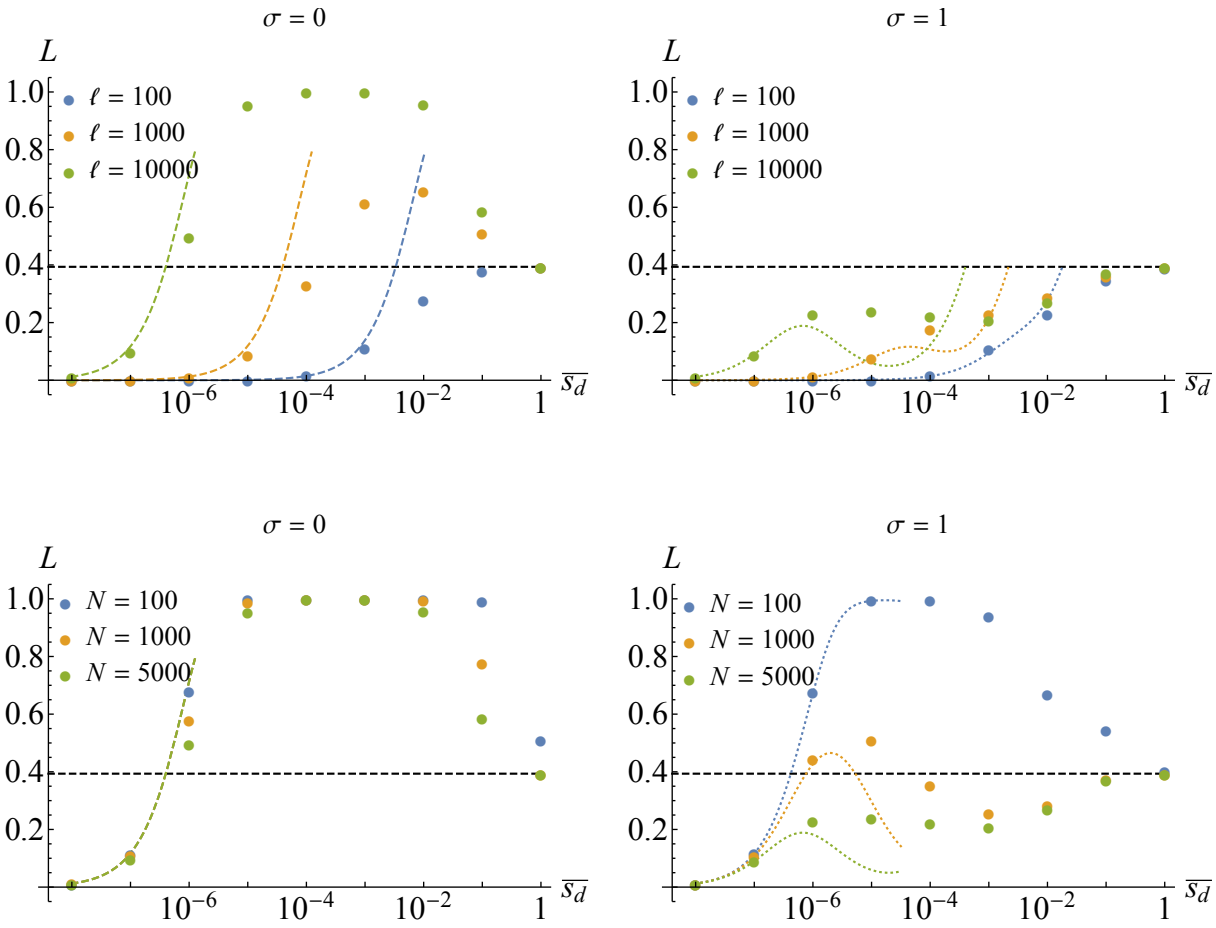


Figure S4. Same as Figure 5 in the main text, with $\theta = 0.5$.

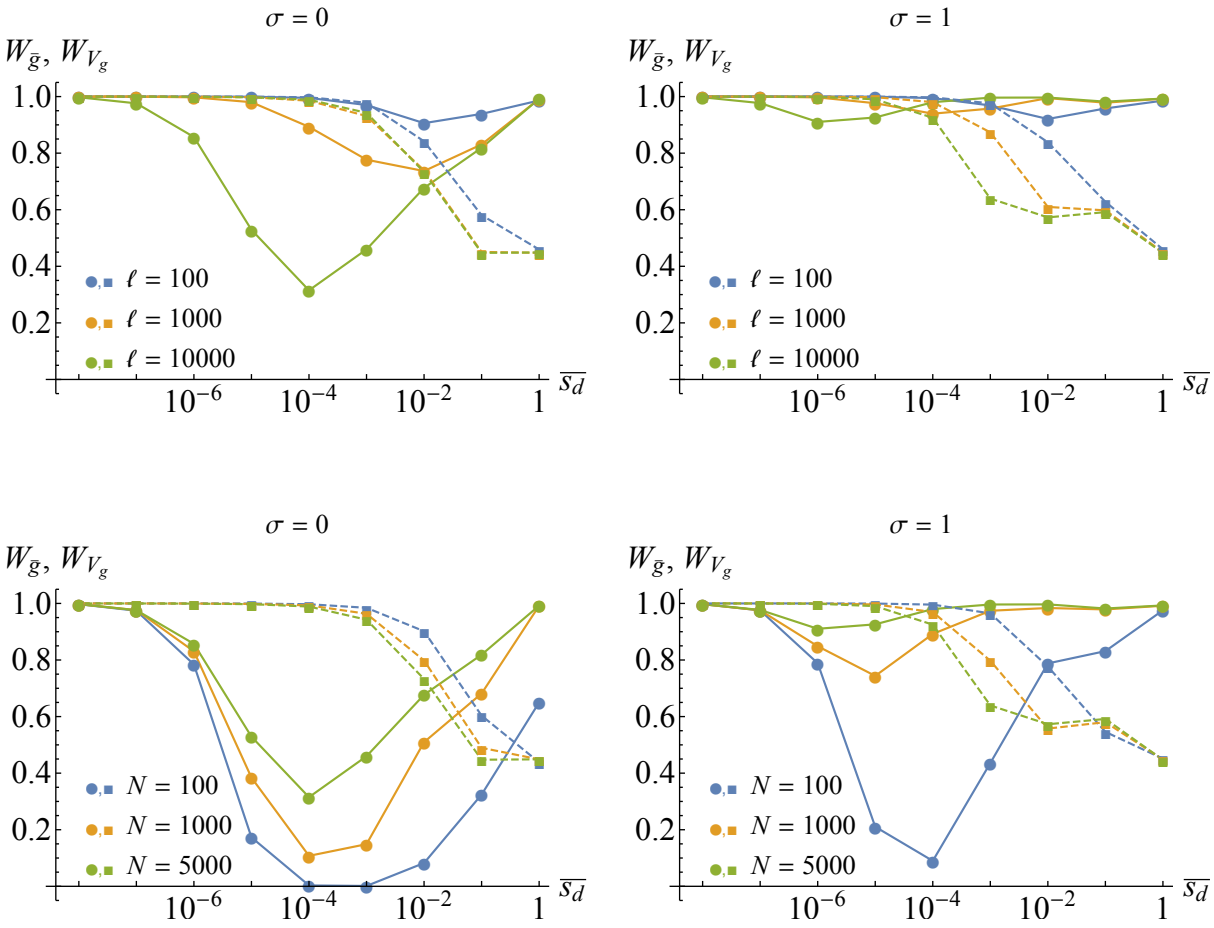


Figure S5. Same as Figure 5 in the main text, showing $W_{\bar{g}} = \exp\left[-\sum_{\alpha=1}^{n} \langle \overline{g_\alpha}^2 \rangle / (2V_\mathrm{s})\right]$ (circles, solid lines) and $W_{V_\mathrm{g}} = \exp\left[-\sum_{\alpha=1}^{n} \langle V_{\mathrm{g},\alpha} \rangle / (2V_\mathrm{s})\right]$ (squares, dashed lines).

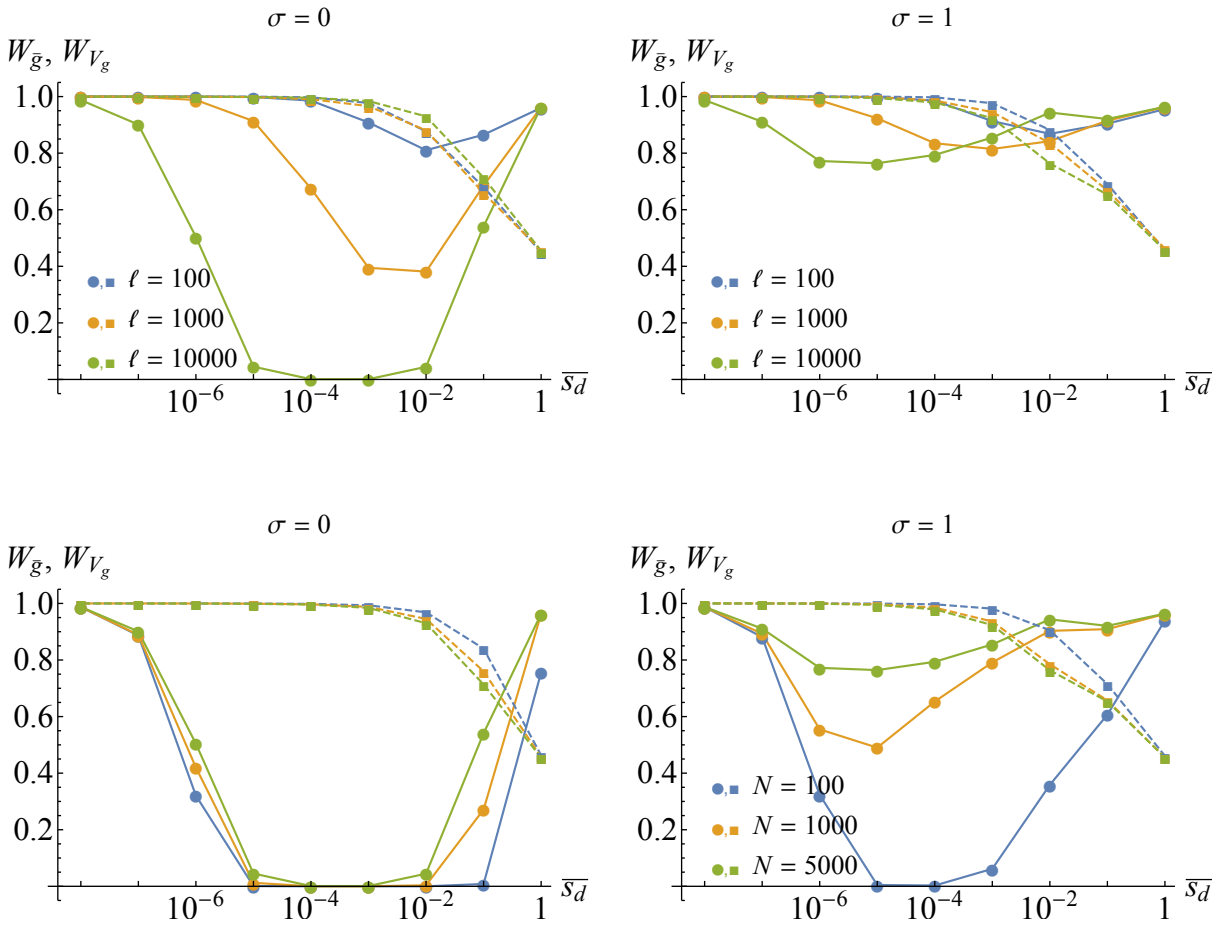


Figure S6. Same as Figure S4, showing $W_{\bar{g}} = \exp\left[-\sum_{\alpha=1}^{n} \langle \overline{g_\alpha}^2 \rangle \,/\, (2V_\mathrm{s})\right]$ (circles, solid lines) and $W_{V_\mathrm{g}} = \exp\left[-\sum_{\alpha=1}^{n} \langle V_{\mathrm{g},\alpha} \rangle \,/\, (2V_\mathrm{s})\right]$ (squares, dashed lines).

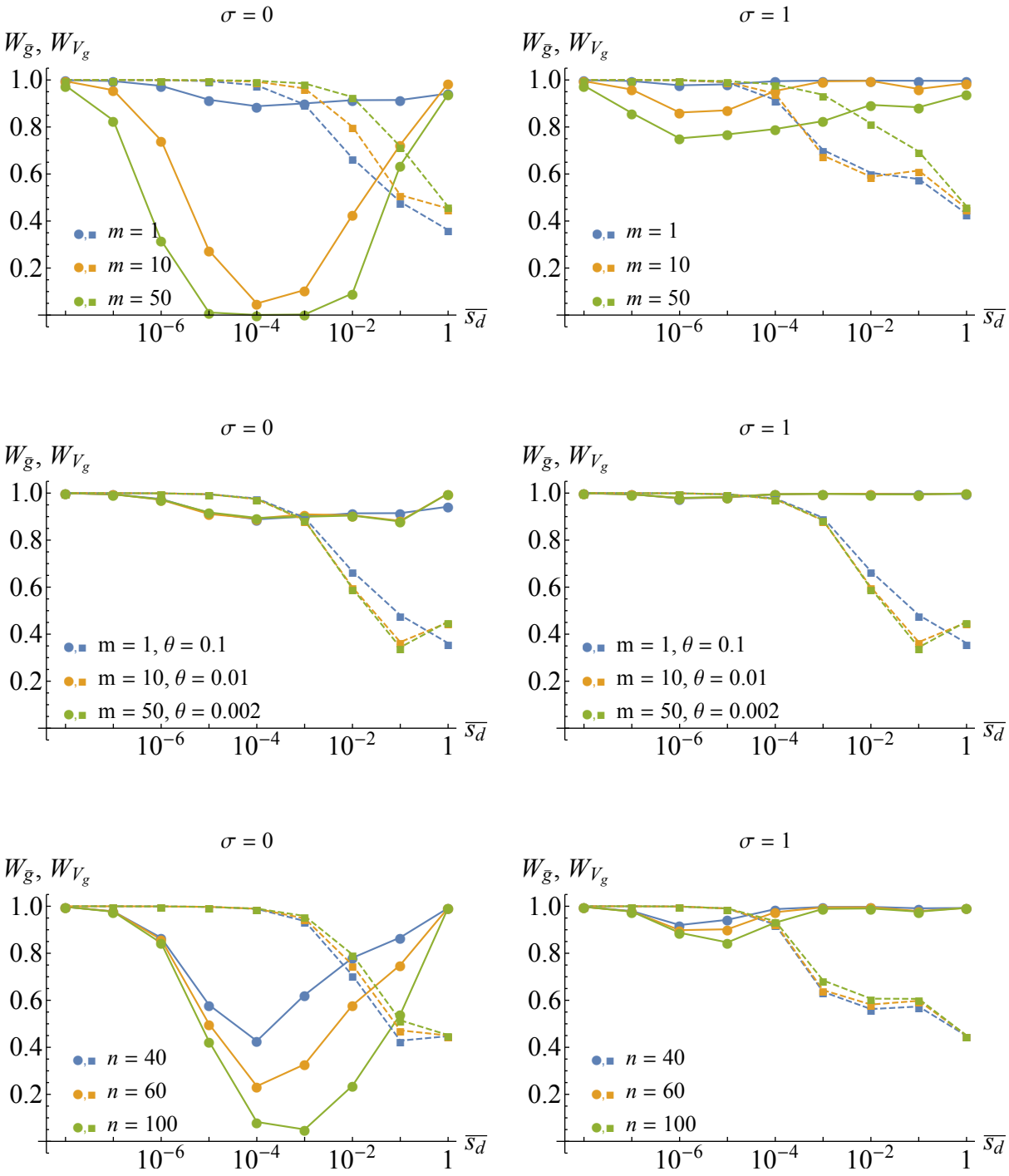


Figure S7. Same as Figure 6 in the main text, showing $W_{\bar{g}} = \exp\left[-\sum_{\alpha=1}^{n} \langle \overline{g_\alpha}^2 \rangle / (2V_s)\right]$ (circles, solid lines) and $W_{V_g} = \exp\left[-\sum_{\alpha=1}^{n} \langle V_{g,\alpha} \rangle / (2V_s)\right]$ (squares, dashed lines).

FILE S1: DESCRIPTION OF SIMULATION PROGRAMS

The genome of each individual consists in a single linear chromosome with map length R (average number of cross-overs at meiosis). The ℓ loci affecting the n traits under stabilizing selection are biallelic and equally spaced along the chromosome, the genome of an individual at these loci being represented by a set of bits (0 or 1). At the beginning of the simulation, the effects of allele 1 at each locus on the different phenotypes are drawn and stored in a table: as explained above, each locus only affects a subset of m randomly sampled traits, the effect on each of these traits being drawn from a Gaussian distribution with standard deviation $a = \sqrt{2V_\mathrm{s}\,(1-\theta)\,\overline{s_d}/m}$ and average $b = \sqrt{2V_\mathrm{s}\,\theta\,\overline{s_d}/m}$. At the start of each generation, genetic components g_α are computed for each individual given its genotype, and environmental components e_α are drawn from a Gaussian distribution with mean 0 and variance V_e, fixed to $1/n$ to avoid that fitness reaches very low values when the number of selected traits is large. The fitness of each individual is then computed according to equation 3 in the main text, where ω^2 is fixed to 10; however, as noted above, the values of ω^2 and V_e should have little effect on the results (for given values of $\overline{s_d}$ and θ), since $V_\mathrm{s} = \omega^2 + V_\mathrm{e}$ may be considered as a scaling factor.

Investment in sexual reproduction s is coded by ℓ_s loci, which are also equally spaced along the chromosome. These loci are multiallelic, investment in sex being given by:

$$s = s_\mathrm{init} + \sum_{i=1}^{\ell_s} g_{si} + e_s \tag{1}$$

where s_init is the initial investment in sex, g_{si} the effect of locus i on s, and e_s an environmental component drawn from a centered Gaussian distribution with variance

$V_{\mathrm{e},s}$. If the value of s obtained from equation 1 is lower than 0 or higher than 1, it is then set to 0 or 1 (respectively). During a number of preliminary generations, g_{si} is fixed to zero for all loci affecting the rate of sex. Then, mutation occurs at a rate U_s per generation on the whole set of ℓ_s loci. When a mutation occurs at locus i, a quantity drawn from a centered Gaussian distribution with variance a_s^2 is added to g_{si}.

For each of the N individuals of the next generation, a maternal parent is sampled with a probability proportional to $W\left(1-s+\frac{s}{c}\right)$, where W is its fitness and s its investment into sex. With probability $1-\sigma$ (where σ is given by equation 13 in the main text), the new individual is produced asexually and carries the same genotype as its mother, except for new mutations (the number of mutations on biallelic loci affecting the traits under stabilizing selection is drawn from a Poisson distribution with parameter U). With the complementary probability, the new individual is produced sexually; in this case a paternal parent is sampled with a probability proportional to Ws, and a recombinant offspring is produced (the number of cross-overs occurring at meiosis is sampled from a Poisson distribution with parameter R, and the position of each cross-over is drawn from a uniform distribution along the chromosome). Every 100 generations, the mean investment in sex, mean rate of sex, mean fitness, mean trait values, genetic variances and covariances among traits and some higher moments of phenotypic distributions are recorded by the program.

FILE S2: APPROXIMATIONS FOR THE EFFECT OF MUTATIONAL BIAS ON MEAN FITNESS

We explain here the derivation of the approximations given in the main text for the effect of mutational bias on the load, assuming a Gaussian fitness function and biallelic loci. Throughout the following, the notation $\overline{X}$ (also denoted $\mathrm{E}[X]$) stands for the average of the quantity X over all individuals of the population, while the notation $\langle Y \rangle$ stands for the expected value of quantity Y in the population at mutation-selection-drift equilibrium. In particular, $\langle \overline{W} \rangle$ is the expected value of the population mean fitness. Assuming that the variance in log-fitness among individuals remains small, we have $\overline{W} \approx e^{\overline{\ln W}}$; furthermore, assuming that the variance in $\overline{\ln W}$ due to drift is small yields:

$$\langle \overline{W} \rangle \approx \left\langle e^{\overline{\ln W}} \right\rangle \approx e^{\langle \overline{\ln W} \rangle}. \tag{1}$$

From equations 4 and 18 in the main text, this yields:

$$\langle L \rangle \approx 1 - \exp\left[-\frac{1}{2V_\mathrm{s}} \sum_{\alpha=1}^{n} \left\langle \overline{g_\alpha}^2 \right\rangle \right] = 1 - \exp\left[-\frac{1}{2V_\mathrm{s}} \sum_{\alpha=1}^{n} \left(\langle V_{\mathrm{g},\alpha} \rangle + \langle \overline{g_\alpha}^2 \rangle \right) \right] \tag{2}$$

Equation 1 shows that the load can be decomposed into the two terms, $W_{V_\mathrm{g}} = \exp\left[-\sum_{\alpha=1}^{n} \langle V_{\mathrm{g},\alpha} \rangle / (2V_\mathrm{s}) \right]$ and $W_{\overline{g}} = \exp\left[-\sum_{\alpha=1}^{n} \langle \overline{g_\alpha}^2 \rangle / (2V_\mathrm{s}) \right]$ representing the decrease in mean fitness due to genetic variance maintained in the population, and to deviations of the mean phenotype from the optimum, respectively. If population size is sufficiently large, the variance of mean phenotypes due to drift should remain small (Lande, 1976; Charlesworth, 2013b), so that $\langle \overline{g_\alpha}^2 \rangle \approx \langle \overline{g_\alpha} \rangle^2$; this is confirmed by simulations (results not shown). In the following, we thus derive approximations for $\langle L \rangle$ by computing expressions for $\langle \overline{g_\alpha} \rangle$ and $\langle V_{\mathrm{g},\alpha} \rangle$ in different limit cases.

Genetic associations and decomposition of the genetic variance. Using the notation of Barton and Turelli (1991) and Kirkpatrick et al. (2002), we denote $p_i = \mathrm{E}[X_i]$ the frequency of allele 1 at locus i and define ζ_i as:

$$\zeta_i = X_i - p_i. \tag{3}$$

Furthermore, products of ζ_i variables are denoted:

$$\zeta_{\mathbb{U}} = \prod_{i \in \mathbb{U}} \zeta_i \tag{4}$$

where $\mathbb{U}$ represents a set of loci. For example, for $\mathbb{U} = \{i, j\}$, we have:

$$\zeta_{ij} = (X_i - p_i)(X_j - p_j). \tag{5}$$

Finally, genetic associations $D_{\mathbb{U}}$ are defined as averages of $\zeta_{\mathbb{U}}$ variables over all individuals:

$$D_{\mathbb{U}} = \mathrm{E}[\zeta_{\mathbb{U}}] \tag{6}$$

In particular, D_{ij} is the linkage disequilibrium between loci i and j. As we will see, associations involving repeated indices (such as $D_{iij} = \mathrm{E}\Big[(X_i - p_i)^2 (X_j - p_j)\Big]$) sometimes appear in the computations. Using the fact that ${X_i}^2 = X_i$ (since X_i equals 0 or 1), repeated indices can be eliminated using the relation:

$$D_{\mathbb{U}ii} = p_i q_i D_{\mathbb{U}} + (1 - 2p_i) D_{\mathbb{U}i} \tag{7}$$

with $q_i = 1 - p_i$ (e.g., equation 5 in Kirkpatrick et al., 2002). In particular (and because $D_j = \mathrm{E}[X_j - p_j] = 0$), we have $D_{iij} = (1 - 2p_i) D_{ij}$. Similarly, $D_{iijj} = p_i q_i p_j q_j + (1 - 2p_i)(1 - 2p_j) D_{ij}$, while $D_{ii} = p_i q_i$.

The genetic variance for trait α in the population is given by:

$$V_{\mathrm{g},\alpha} = \mathrm{E}\left[(g_\alpha - \overline{g_\alpha})^2\right] \tag{8}$$

From equations 5 and 6 in the main text:

$$\overline{g_\alpha} = \mathrm{E}\left[\sum_{i=1}^{\ell} r_{\alpha i}\, X_i\right] = \sum_{i=1}^{\ell} r_{\alpha i}\, p_i \tag{9}$$

so that $\langle \overline{g_\alpha} \rangle = \sum_i r_{\alpha i} \langle p_i \rangle$. Using the definitions above, we have:

$$\begin{aligned} V_{\mathrm{g},\alpha} &= \mathrm{E}\left[\left(\sum_{i=1}^{\ell} r_{\alpha i}\,(X_i - p_i)\right)^2\right] \\ &= \mathrm{E}\left[\left(\sum_{i=1}^{\ell} r_{\alpha i}\,\zeta_i\right)^2\right] = \mathrm{E}\left[\sum_{i,j} r_{\alpha i}\, r_{\alpha j}\,\zeta_i\,\zeta_j\right] \end{aligned} \tag{10}$$

where the last sum is over all i and j (including $i = j$). Using equations 4 and 6, one obtains:

$$V_{\mathrm{g},\alpha} = \sum_{i=1}^{\ell} {r_{\alpha i}}^2\, p_i q_i + \sum_{i \neq j} r_{\alpha i}\, r_{\alpha j}\, D_{ij}. \tag{11}$$

In the following, we assume that linkage disequilibria remain negligible, so that $\langle V_{\mathrm{g},\alpha} \rangle \approx \sum_{i=1}^{\ell} {r_{\alpha i}}^2 \langle p_i q_i \rangle$.

Neglecting the effects of selection on mean trait values and genetic variance. Simple approximations for $\langle \overline{g_\alpha} \rangle$ and $\langle V_{\mathrm{g},\alpha} \rangle$ are obtained for the regime where $\overline{s_d}$ is so low that selection has negligible effects on $\langle p_i \rangle$ and $\langle p_i q_i \rangle$, compared with the effects of mutation and drift. Because drift does not change expected allele frequencies, the change in $\langle p_i \rangle$ over one generation is given by (neglecting the effect of selection):

$$\langle p_i \rangle_{t+1} = u + (1 - 2u) \langle p_i \rangle_t \,. \tag{12}$$

yielding $\langle p_i \rangle = 1/2$ at equilibrium. Using the change in phenotypic basis given by equations 9 and 10 in the main text, we have:

$$\frac{1}{2V_{\mathrm{s}}} \sum_{\alpha=1}^{n} \langle \overline{g_\alpha} \rangle^2 = \frac{1}{2V_{\mathrm{s}}} \left\langle \overline{g_{1'}} \right\rangle^2 \tag{13}$$

since only the first trait in the new basis (along which the mutational bias occurs) should differ from zero, on average, at equilibrium. Equation 9 and equation 11 in the main text then yield:

$$\frac{1}{2V_{\mathrm{s}}}\sum_{\alpha=1}^{n}\langle\overline{g_\alpha}\rangle^2=\frac{1}{4}\left(\ell\,\tilde{b}_1{}'\right)^2 \tag{14}$$

where $\tilde{b}_1{}'$ is the scaled magnitude of mutational bias.

Neglecting the effects of selection yields the following recursion for $\langle p_i q_i\rangle$:

$$\langle p_i q_i\rangle_{t+1}\approx\left(1-\frac{1}{N}\right)\left[u+(1-4u)\,\langle p_i q_i\rangle_t\right] \tag{15}$$

so that $\langle p_i q_i\rangle\approx Nu/\left(1+4Nu\right)$ at equilibrium (assuming large N and small u). Noting that $\sum_\alpha\sum_i {r_{\alpha i}}^2=2V_{\mathrm{s}}\,\overline{s_d}\,\ell$ (see equation 7 in the main text), one obtains:

$$\frac{1}{2V_{\mathrm{s}}}\sum_{\alpha=1}^{n}\langle V_{\mathrm{g},\alpha}\rangle\approx\overline{s_d}\,\frac{NU}{1+4Nu}. \tag{16}$$

Equations 2, 14 and 16 yield the following approximation for the load (assuming $\langle\overline{g_\alpha}^2\rangle\approx\langle\overline{g_\alpha}\rangle^2$):

$$\langle L\rangle\approx 1-\exp\left[-\overline{s_d}\,\frac{NU}{1+4Nu}-\frac{1}{4}\left(\ell\,\tilde{b}_1{}'\right)^2\right], \tag{17}$$

equivalent to equation 8 in Roze and Blanckaert (2014) in the absence of mutational bias ($\tilde{b}_1{}'=0$). Comparisons with individual-based simulations show that equation 17 does indeed provide correct predictions when $\overline{s_d}$ is very low (see Figures 2-4 in the main text). As $\overline{s_d}$ increases, $\langle\overline{g_\alpha}\rangle$ and $\langle V_{\mathrm{g},\alpha}\rangle$ depart more and more from equations 14 and 16; however, simulations indicate that equation 16 stays valid over a wider range of values of s than equation 14, in agreement with previous observations that selection may have significant effects on mean trait values even when $\langle p_i q_i\rangle$ at each locus is mainly controlled by mutation and drift (Campbell, 1984; Barton, 1989; Charlesworth, 2013a). Based on this, we can derive a better approximation for low $\overline{s_d}$ by taking the

effect of selection on $\langle\overline{g_\alpha}\rangle$ into account, but still neglect the effect of selection on $\langle V_{\mathrm{g},\alpha}\rangle$, as shown in the next subsection.

Effect of selection on mean trait values in the low $\overline{s_d}$ regime. From equation 12, we have:

$$\langle\overline{g_\alpha}\rangle_{t+1} = u\sum_i r_{\alpha i} + (1-2u)\left(\langle\overline{g_\alpha}\rangle_t + \langle\Delta_{\mathrm{sel}}\overline{g_\alpha}\rangle_t\right) \tag{18}$$

where $\langle\Delta_{\mathrm{sel}}\overline{g_\alpha}\rangle$ is the change in $\langle\overline{g_\alpha}\rangle$ due to selection, given by:

$$\langle\Delta_{\mathrm{sel}}\overline{g_\alpha}\rangle = \left\langle \mathrm{E}\left[\frac{W_\mathrm{g}}{\overline{W}}\,g_\alpha\right]\right\rangle - \langle\overline{g_\alpha}\rangle = \left\langle \mathrm{E}\left[\frac{W_\mathrm{g}}{\overline{W}}\,(g_\alpha-\overline{g_\alpha})\right]\right\rangle. \tag{19}$$

Assuming weak selection ($V_{\mathrm{g},\alpha}/V_\mathrm{s}$ small, of order ϵ), we have from equation 4 in the main text (to the first order in ϵ):

$$\frac{W_\mathrm{g}}{W_{\mathrm{g,max}}} \approx 1 - \frac{1}{2V_\mathrm{s}}\sum_{\alpha=1}^{n} {g_\alpha}^2 = 1 - \frac{1}{2V_\mathrm{s}}\sum_{\alpha=1}^{n}\left[\overline{g_\alpha}^2 + 2\,(g_\alpha-\overline{g_\alpha})\,\overline{g_\alpha} + (g_\alpha-\overline{g_\alpha})^2\right], \tag{20}$$

yielding:

$$\frac{\overline{W}}{W_{\mathrm{g,max}}} \approx 1 - \frac{1}{2V_\mathrm{s}}\sum_{\alpha=1}^{n}\left[\overline{g_\alpha}^2 + V_{\mathrm{g},\alpha}\right], \tag{21}$$

and thus:

$$\frac{W_\mathrm{g}}{\overline{W}} \approx 1 - \frac{1}{V_\mathrm{s}}\sum_{\alpha=1}^{n}\overline{g_\alpha}\,(g_\alpha-\overline{g_\alpha}) - \frac{1}{2V_\mathrm{s}}\sum_{\alpha=1}^{n}\left[(g_\alpha-\overline{g_\alpha})^2 - V_{\mathrm{g},\alpha}\right]. \tag{22}$$

From equations 19 and 22, one obtains:

$$\langle\Delta_{\mathrm{sel}}\overline{g_\alpha}\rangle \approx -\frac{1}{V_\mathrm{s}}\sum_{\beta=1}^{n}\langle\overline{g_\beta}\,C_{\mathrm{g},\alpha\beta}\rangle - \frac{1}{2V_\mathrm{s}}\sum_{\beta=1}^{n}\langle M_{\mathrm{g},\alpha\beta\beta}\rangle \tag{23}$$

where $M_{\mathrm{g},\alpha\beta\beta}$ is the third moment $\mathrm{E}\left[(g_\alpha-\overline{g_\alpha})\,(g_\beta-\overline{g_\beta})^2\right]$. Assuming that the distribution of phenotypes in the population stays close to a Gaussian distribution, $M_{\mathrm{g},\alpha\beta\beta}$ should be close to zero. Furthermore, assuming that fluctuations in $\overline{g_\beta}$ and $C_{\mathrm{g},\alpha\beta}$ due

to drift remain small, $\langle \overline{g_\beta}\, C_{\mathrm{g},\alpha\beta} \rangle \approx \langle \overline{g_\beta} \rangle \langle C_{\mathrm{g},\alpha\beta} \rangle$. Measuring traits in the phenotypic basis defined by equations 9 and 10 in the main text (so that $C_{\mathrm{g},\alpha\beta} = 0$ for $\alpha \neq \beta$), one obtains the following expression for $\langle \overline{g_1{}'} \rangle$ at equilibrium (from equations 18 and 23):

$$\langle \overline{g_1{}'} \rangle \approx \frac{U\, b_1{}'}{1 - (1-2u)\left(1 - \frac{\langle V_{\mathrm{g},1}{}' \rangle}{V_\mathrm{s}}\right)}\,. \tag{24}$$

Neglecting the effects of selection on $\langle V_{\mathrm{g},1}{}' \rangle \approx \sum_{i=1}^{\ell} (r_{1i}{}')^2 \langle p_i q_i \rangle$ and noting that $\sum_{i=1}^{\ell} (r_{1i}{}')^2 = \frac{m}{n}\,(a^2 + m\,b^2)\,\ell = 2V_\mathrm{s}\,\overline{s_d}\,\ell\,\frac{1}{n}\,[1 + \theta\,(m-1)]$ (using equation 9 in the main text), we have:

$$\langle V_{\mathrm{g},1}{}' \rangle \approx 2V_\mathrm{s}\,\overline{s_d}\,\frac{1}{n}\,[1 + \theta\,(m-1)]\,\frac{NU}{1+4Nu}. \tag{25}$$

From equations 24 and 25, one obtains (assuming u and $\overline{s_d}$ are small):

$$\langle \overline{g_1{}'} \rangle \approx \frac{\ell\, b_1{}'}{2\left[1 + \frac{\overline{s_d}}{n}\,[1 + \theta\,(m-1)]\,\frac{N\ell}{1+4Nu}\right]} \tag{26}$$

Equations 2, 16 and 26 yield (assuming $\langle \overline{g_\alpha}^2 \rangle \approx \langle \overline{g_\alpha} \rangle^2$):

$$\langle L \rangle \approx 1 - \exp\left[-\overline{s_d}\,\frac{NU}{1+4Nu} - \frac{\left(\ell\,\tilde{b}_1{}'\right)^2}{4\left[1 + \frac{\overline{s_d}}{n}\,[1 + \theta\,(m-1)]\,\frac{N\ell}{1+4Nu}\right]^2}\right]. \tag{27}$$

Comparisons with individual-based simulations confirm that equation 27 provides better predictions than equation 17 in the case of sexual populations, as long as $\overline{s_d}$ is sufficiently small (see Figures 3-5 in the main text). Equation 27 fails when $\overline{s_d}$ is not very small, however, as selection affects genetic variances at equilibrium. Unfortunately, we could not obtain any simple expression for the genetic variance (and mean fitness) in this regime for arbitrary θ, although an approximation can be obtained for $\theta = 1$, as shown in the next subsection.

Effect of selection on genetic variance and approximations for the mutation-selection regime. Neglecting linkage disequilibria, genetic variances can be expressed in terms of the genetic diversities $p_i q_i$ at the different loci (equation 11). Extending equation 15 to include selection yields:

$$\langle p_i q_i \rangle_{t+1} \approx \left(1 - \frac{1}{N}\right) \left[u + (1-4u) \left\langle p_i^{\mathrm{sel}} q_i^{\mathrm{sel}} \right\rangle_t\right]. \tag{28}$$

Furthermore, noting that $p_i q_i = D_{ii} = \mathrm{E}\left[\zeta_{ii}\right]$, we have, to the first order in ϵ:

$$\left\langle p_i^{\mathrm{sel}} q_i^{\mathrm{sel}} \right\rangle = \left\langle \mathrm{E}\left[\frac{W}{\overline{W}}\,\zeta_{ii}\right]\right\rangle. \tag{29}$$

Decomposing g_α, $\overline{g_\alpha}$ and $V_{\mathrm{g},\alpha}$ as sums over loci (using equations 9 and 11) and introducing the centered variables $\zeta_i = X_i - p_i$, we have from equation 22:

$$\frac{W_{\mathrm{g}}}{\overline{W}} = 1 + \sum_{i=1}^{\ell} a_i \zeta_i + \sum_{i,j} a_{ij} \left(\zeta_{ij} - D_{ij}\right) \tag{30}$$

with $a_i = -\frac{1}{V_{\mathrm{s}}} \sum_{\alpha=1}^{n} \overline{g_\alpha}\, r_{\alpha i}$ and $a_{ij} = -\sum_{\alpha=1}^{n} \left(r_{\alpha i}\, r_{\alpha j}\right) / \left(2V_{\mathrm{s}}\right)$, both of order ϵ (Barton and Turelli, 1991). Using equations 29 and 30, and neglecting linkage disequilibria, one obtains:

$$\left\langle p_i^{\mathrm{sel}} q_i^{\mathrm{sel}} \right\rangle = \langle p_i q_i \rangle - \frac{1}{V_{\mathrm{s}}} \sum_{\alpha=1}^{n} r_{\alpha i} \left\langle \overline{z_\alpha} \left(1 - 2p_i\right) p_i q_i \right\rangle - \frac{1}{2V_{\mathrm{s}}} \sum_{\alpha=1}^{n} {r_{\alpha i}}^2 \left\langle \left(1 - 2p_i\right)^2 p_i q_i \right\rangle. \tag{31}$$

Equations 28 and 31 lead to the following recursion for the genetic variance:

$$\begin{aligned} \langle V_{\mathrm{g},\alpha} \rangle_{t+1} \approx \left(1 - \frac{1}{N}\right) \Bigg[& 2V_{\mathrm{s}}\, \overline{s_d}\, \frac{U}{n} + (1-4u) \Bigg(\langle V_{\mathrm{g},\alpha} \rangle_t - \frac{1}{V_{\mathrm{s}}} \sum_{\beta=1}^{n} \left\langle \overline{z_\beta}\, C_{\alpha\alpha\beta} \right\rangle_t \\ & - \frac{1}{2V_{\mathrm{s}}} \sum_{i=1}^{\ell} {r_{\alpha i}}^2 \sum_{\beta=1}^{n} {r_{\beta i}}^2 \left\langle \left(1 - 2p_i\right)^2 p_i q_i \right\rangle_t \Bigg) \Bigg]. \end{aligned} \tag{32}$$

It is not possible to derive an expression for $\langle V_{\mathrm{g},\alpha} \rangle$ at equilibrium under mutation, selection and drift from equation 32 — one may assume that $\langle \overline{z_\beta}\, C_{\alpha\alpha\beta} \rangle \approx \langle \overline{z_\beta} \rangle \langle C_{\alpha\alpha\beta} \rangle$ and that $\langle C_{\alpha\alpha\beta} \rangle \approx 0$, and assume that $p_i q_i$ is small at each locus so that $\left(1 - 2p_i\right)^2 \approx 1$

(and the second sum on the second line of equation 32 becomes $\sum_\beta \langle V_{\mathrm{g},\beta}\rangle$), but the resulting approximation does not work well when $\overline{s_d}$ is small (as $p_i q_i$ may not be small), nor when $\overline{s_d}$ is large and in the presence of mutational bias (as $\langle C_{\alpha\alpha\beta}\rangle \neq 0$, results not shown). Neglecting drift, and in the absence of mutational bias ($\theta = 0$), the change in $p_i q_i$ over one generation is (from equations 28 and 31):

$$\Delta\left(p_i q_i\right) \approx u\left(1-2p_i\right)^2 - \frac{1}{2V_\mathrm{s}} \sum_{\alpha=1}^{n} {r_{\alpha i}}^2 \left(1-2p_i\right)^2 p_i q_i \tag{33}$$

so that either $p_i = 1/2$ or $p_i q_i = 2V_\mathrm{s}\, u / \left(\sum_{\alpha=1}^{n} {r_{\alpha i}}^2\right)$ at equilibrium. When $\overline{s_d} \gg u$, most loci should be at the second equilibrium, in which case $V_{\mathrm{g},\alpha} \approx 2V_\mathrm{s} U/n$, and $L \approx 1 - e^{-U}$.

Another approximation can be obtained for the case where $\theta = 1$ (no variance of mutational effects) and when drift is negligible. Indeed, in this case alleles 1 are deleterious, and p_i should thus be small at equilibrium. To the first order in p_i, equations 28 and 31 give for the change in p_i over one generation:

$$\Delta p_i \approx u - \frac{1}{V_\mathrm{s}} \sum_{\alpha=1}^{n} r_{\alpha i}\, \overline{g_\alpha}\, p_i - \frac{1}{2V_\mathrm{s}} \sum_{\alpha=1}^{n} {r_{\alpha i}}^2 p_i \tag{34}$$

which may also be written as (using the change in phenotypic basis given by equations 9 and 10 in the main text):

$$\Delta p_i \approx u - \frac{1}{V_\mathrm{s}} {b_1}' \overline{g_1'}\, p_i - \overline{s_d}\, p_i \,. \tag{35}$$

From this, the change in $\overline{g_1'} \approx \sum_i r_{1i}' p_i$ is:

$$\Delta\overline{g_1'} \approx U\, {b_1}' - \frac{1}{V_\mathrm{s}} {b_1}' \left(\overline{g_1'}\right)^2 - \overline{s_d}\, \overline{g_1'} \tag{36}$$

yielding, at equilibrium:

$$\frac{\overline{g_1'}}{\sqrt{2V_\mathrm{s}}} \approx \frac{\sqrt{8\tilde{b}_1'^2 U + \overline{s_d}^2} - \overline{s_d}}{4\tilde{b}_1'} \,. \tag{37}$$

Finally, we have $\sum_\alpha V_{\mathrm{g},\alpha} \approx \sum_i \sum_\alpha {r_{\alpha i}}^2 p_i$, which is also $2V_\mathrm{s}\,\overline{s_d}\,\ell\,p_i$ (as p_i should be the same at all loci when $\theta = 1$). Noting that $\overline{g_1{}'} \approx {b_1}'\,\ell\,p_i$, we thus have:

$$\frac{\sum_{\alpha=1}^{n} V_{\mathrm{g},\alpha}}{2V_\mathrm{s}} \approx \frac{\overline{s_d}}{\tilde{b}_1{}'} \left(\frac{\overline{g_1{}'}}{\sqrt{2V_\mathrm{s}}} \right). \tag{38}$$

Equations 2, 37 and 38 finally lead to:

$$L \approx 1 - \exp\left[-\frac{4\tilde{b}_1{}'^2 U + \overline{s_d}\left(\sqrt{8\tilde{b}_1{}'^2 U + \overline{s_d}^2} - \overline{s_d} \right)}{8\tilde{b}_1{}'^2} \right] \tag{39}$$

or, in terms of $\overline{s_d}$, ρ and U:

$$L \approx 1 - \exp\left[-\frac{4\rho U - \overline{s_d} + \sqrt{\overline{s_d}\,(8\rho U + \overline{s_d})}}{8\rho} \right]. \tag{40}$$

Simulations confirm that equation 40 provides accurate predictions for $\theta = 1$ (in sexual populations), when $\overline{s_d}$ is sufficiently high (see Figure 4 in the main text).

LITERATURE CITED

Barton, N. H. 1989. The divergence of a polygenic system subject to stabilizing selection, mutation and drift. Genet. Res. 54:59–77.

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

Campbell, R. B. 1984. The manifestation of phenotypic selection at constituent loci. I. Stabilizing selection. Evolution 38:1033–1038.

Charlesworth, B. 2013a. Stabilizing selection, purifying selection, and mutational bias in finite populations. Genetics 194:955–971.

———. 2013b. Why we are not dead one hundred times over. Evolution 67:3354–3361.

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

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution 30:314–334.

Roze, D. and A. Blanckaert. 2014. Epistasis, pleiotropy and the mutation load in sexual and asexual populations. Evolution 68:137–149.

FILE S3: QLE MODEL FOR THE EVOLUTION OF SEX

We derive here expressions for the change in mean rate of sex in the limit of an infinite population, using a quasi-linkage equilibrium (QLE) argument. For this, we use Turelli and Barton's (1990) method (see also Barton, 1995) to express the effect of selection on genetic associations in terms of partial derivatives of $\ln \overline{W}$ with respect to mean trait values and genetic variances/covariances. Note that the derivations given below are in principle valid for any number of possible alleles at each locus (not necessarily biallelic loci) and any fitness function (not necessarily Gaussian), as long as the distribution of phenotypes affecting fecundity stays approximately Gaussian.

Definitions. Extending the notation of Turelli and Barton (1990) to multiple traits, we define the centered variable $\zeta_{\alpha j}$ as:

$$\zeta_{\alpha j} = g_{\alpha j} - \overline{g_{\alpha j}} \tag{1}$$

(where again $g_{\alpha j}$ is the effect of the allele present at locus j on trait α in a given individual, and $\overline{g_{\alpha j}}$ its average over all individuals). Genetic associations are defined as

$$C_{\mathbb{U}} = \mathrm{E}\left[\zeta_{\mathbb{U}}\right] \tag{2}$$

where E stands for the average over all individuals, and with $\zeta_{\mathbb{U}} = \prod_x \zeta_x$, each x bearing two elements, the trait α and the locus j. For example, $C_{\alpha j\, \alpha j} = \mathrm{E}\left[\left(g_{\alpha j} - \overline{g_{\alpha j}}\right)^2\right]$ while $C_{\alpha j\, \alpha k\, \beta k} = \mathrm{E}\left[\left(g_{\alpha j} - \overline{g_{\alpha j}}\right)\left(g_{\alpha k} - \overline{g_{\alpha k}}\right)\left(g_{\beta k} - \overline{g_{\beta k}}\right)\right]$. Using these definitions, the genetic

variance for trait α can be written as:

$$\begin{aligned} V_{\mathrm{g},\alpha} &= \mathrm{E}\left[\left(\sum_j \zeta_{\alpha j}\right)^2\right] = \sum_{j=1}^{\ell} C_{\alpha j\, \alpha j} + \sum_{j\neq k} C_{\alpha j\, \alpha k} \\ &= V^0_{\mathrm{g},\alpha} + \mathcal{D}_{\alpha\alpha} \end{aligned} \tag{3}$$

where $V^0_{\mathrm{g},\alpha} = \sum_j C_{\alpha j\, \alpha j}$ is the "genic variance" for trait α (genetic variance in a population with the same allele frequencies, at linkage equilibrium), and $\mathcal{D}_{\alpha\alpha} = \sum_{j\neq k} C_{\alpha j\, \alpha k}$ is the effect of linkage disequilibria on the variance. Similarly, the genetic covariance between traits α and β can be decomposed as:

$$\begin{aligned} C_{\mathrm{g},\alpha\beta} &= \sum_{j=1}^{\ell} C_{\alpha j\, \beta j} + \sum_{j\neq k} C_{\alpha j\, \beta k} \\ &= C^0_{\mathrm{g},\alpha\beta} + \mathcal{D}_{\alpha\beta}\,. \end{aligned} \tag{4}$$

As explained in the main text, we assume that investment in sexual reproduction s is also a polygenic trait with independent genetic and environmental contributions:

$$s = \overline{s} + g_s + e_s \tag{5}$$

where e_s is sampled in a Gaussian distribution with mean 0 and variance $V_{\mathrm{e},s}$, and assuming additive effects of loci affecting s:

$$g_s = \sum_j g_{sj} \tag{6}$$

where g_{sj} is the effect of the allele at locus j on investment in sex (we assume that loci affecting the rate of sex do not affect the other traits). Assuming that the variance in s in the population is sufficiently small (g_s, g_e small, of order η), the rate of sex $\sigma = s/\left[c\,(1-s)+s\right]$ of an individual can also be decomposed into an additive genetic and an environmental component:

$$\sigma = \overline{\sigma} + g_\sigma + e_\sigma \tag{7}$$

with (to leading order in η):

$$\overline{\sigma} \approx \frac{\overline{s}}{c\,(1-\overline{s})+\overline{s}}, \quad g_\sigma \approx \frac{c}{\left[c\,(1-\overline{s})+\overline{s}\right]^2}\, g_s, \quad V_{\mathrm{g},\sigma} \approx \frac{c^2}{\left[c\,(1-\overline{s})+\overline{s}\right]^4}\, V_{\mathrm{g},s}\,. \tag{8}$$

The expected change in $\overline{\sigma}$ over one generation (denoted $\Delta\overline{\sigma}$) corresponds to the change in $\overline{g_\sigma}$ in the parental generation due to differences in fecundities among individuals and to the cost of sex. In the following, we derive deterministic approximations for $\Delta\overline{\sigma}$, assuming that phenotypic traits affecting fecundity are normally distributed (this implies that the number of loci affecting each of these traits is sufficiently large, each locus having a sufficiently small effect on the trait). We will also use a quasi-linkage equilibrium approximation, assuming that rates of sex and recombination are not too small, so that genetic associations between loci are small and equilibrate fast relative to change in allele frequencies. Finally, we will assume that the genetic variance in the rate of sex in the population ($V_{\mathrm{g},\sigma}$) is small (however, we do not make any assumption on the number of loci affecting the rate of sex). For this, we will decompose a generation into two steps: the first ("selection") corresponds to the differential reproduction of individuals due to differences in fecundity (according to the values of their phenotypes $z_1, \ldots z_n$), while the second ("reproduction") corresponds to the effect of the cost of sex and of sexual recombination (strictly, this second step also involves selection when $c > 1$, since individuals investing more in sex are disfavored). In the next sections, we derive expressions for changes in mean breeding values during these two steps.

Effect of selection on mean breeding values. The effect of selection on $\overline{g_\alpha}$ can be written as:

$$\Delta_{\mathrm{sel}}\overline{g_\alpha} = \mathrm{E}\left[\frac{W_{\mathbf{g}}}{\overline{W}}\,(g_\alpha - \overline{g_\alpha})\right] \tag{9}$$

where $W_{\mathbf{g}}$ is the mean fecundity of individuals with breeding values $\mathbf{g} = (g_1, g_2, \ldots)$ and $\overline{W}$ the mean fecundity of the whole population. Following Barton (1995), we assume that selection is weak and approximate $W_{\mathbf{g}}/\overline{W}$ by:

$$\begin{aligned}\frac{W_{\mathbf{g}}}{\overline{W}} \approx 1 &+ \sum_{\alpha} (g_\alpha - \overline{g_\alpha}) \frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}} \\ &+ \sum_{\alpha \leq \beta} \left[(g_\alpha - \overline{g_\alpha}) (g_\beta - \overline{g_\beta}) - C_{\mathrm{g},\alpha\beta} \right] \frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\alpha\beta}} .\end{aligned} \tag{10}$$

(see Appendix A), where the last sum includes the terms for $\alpha = \beta$, which involve partial derivatives of $\ln \overline{W}$ with respect to $V_{\mathrm{g},\alpha}$. From equations 9 and 10, and assuming a Gaussian distribution of breeding values, we recover the classic expression:

$$\Delta_{\mathrm{sel}} \overline{g_\alpha} = \sum_{\beta} \frac{\partial \ln \overline{W}}{\partial \overline{z_\beta}} C_{\mathrm{g},\alpha\beta} \tag{11}$$

(Lande, 1979). The change in $\overline{g_\sigma}$ is obtained similarly:

$$\Delta_{\mathrm{sel}} \overline{g_\sigma} = \mathrm{E} \left[\frac{W_{\mathbf{g}}}{\overline{W}} (g_\sigma - \overline{g_\sigma}) \right] . \tag{12}$$

However, we can no longer assume that the joint distribution of investment into sex σ and of traits affecting fecundity is multivariate normal (indeed, genetic variances and covariances may differ between subgroups of individuals differing in their values of g_σ, due to the effect of sexual recombination on genetic associations). From equations 10 and 12, one obtains:

$$\Delta_{\mathrm{sel}} \overline{g_\sigma} \approx \sum_{\alpha} \frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}} C_{\mathrm{g},\sigma\alpha} + \sum_{\alpha \leq \beta} \frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\alpha\beta}} M_{\mathrm{g},\sigma\alpha\beta} \tag{13}$$

where $M_{\mathrm{g},\sigma\alpha\beta}$ is the moment $\mathrm{E}\left[(g_\sigma - \overline{g_\sigma}) (g_\alpha - \overline{g_\alpha}) (g_\beta - \overline{g_\beta})\right]$. Equation 13 is equivalent to Charlesworth (1993)'s decomposition of the selection gradient for a recombination modifier allele into two terms (equation A10 in Charlesworth, 1993, see also Appendix 2 of Barton, 1995). The first part of equation 13 shows that under directional selection acting on trait α, a covariance between g_α and g_σ generates indirect

selection on σ (this is equivalent to the term in $\delta\overline{z}$ in Charlesworth, 1993). The second part of equation 13 (equivalent to the term in δV_g in Charlesworth, 1993) corresponds to indirect selection on σ due to different genetic variances and covariances for selected traits among subgroups of individuals with different rates of sex. For example, $\partial \ln \overline{W}/\partial V_{\mathrm{g},\alpha} < 0$ under stabilizing selection acting on a single trait α (the immediate effect of increasing genetic variance is to decrease mean fitness) and in this situation we also expect that higher rates of sex tend to increase genetic variance, so that $\mathrm{E}\left[(g_\sigma - \overline{g_\sigma})\,(g_\alpha - \overline{g_\alpha})^2\right] > 0$, and the second term of equation 13 selects against sex. This term is equivalent to the "short-term effect" in models for the evolution of sex (or recombination) with epistasis (e.g., Agrawal, 2006). Now, if the population mean phenotype $\overline{\mathbf{z}} = (\overline{z_1}, \ldots \overline{z_n})$ is displaced from the optimum, the higher genetic variance associated with sex will increase the efficiency of selection, generating associations between higher values of g_σ and values of g_α closer to the optimum, that in turn favor sex. This effect is represented by the first term of equation 13 and corresponds to the "long term effect" (favoring sex due to an increased efficiency of selection).

Selection gradients $\partial \ln \overline{W}/\partial \overline{z_\alpha}$ and $\partial \ln \overline{W}/\partial C_{\mathrm{g},\alpha\beta}$ take simple forms in the case of a fully isotropic model with Gaussian stabilizing selection:

$$W = \exp\left[-\frac{\sum_\alpha (z_\alpha - \theta_\alpha)^2}{2\omega^2}\right] \tag{14}$$

where ω^2 represents the strength of selection (the same for all traits), and where the phenotypic optimum is located at $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots \theta_n)$. A general expression for mean fitness under Gaussian stabilizing selection (and when the maximal possible fitness is 1, as implied by equation 14) is given by:

$$\overline{W} = \sqrt{\det\left((\mathbf{S}+\mathbf{P})^{-1}\,\mathbf{S}\right)}\,\exp\left[-\frac{1}{2}\,(\overline{\mathbf{z}} - \boldsymbol{\theta})^T\,(\mathbf{S}+\mathbf{P})^{-1}\,(\overline{\mathbf{z}} - \boldsymbol{\theta})\right] \tag{15}$$

(Gomulkiewicz and Houle, 2009), where $\det(\mathbf{A})$ is the determinant of matrix $\mathbf{A}$, T stands for matrix/vector transpose, $\mathbf{S}$ is a matrix determining the pattern of multivariate stabilizing selection, and $\mathbf{P}$ is the phenotypic variance-covariance matrix. When fitness is given by equation 14, $\mathbf{S} = \omega^2 \mathbf{I}$ (where $\mathbf{I}$ is the identity matrix). Furthermore, our assumption of independent, identically distributed environmental effects yields $\mathbf{P} = \mathbf{G} + V_{\mathrm{e}} \mathbf{I}$, where $\mathbf{G}$ is the genetic variance-covariance matrix (whose diagonal elements are genetic variances $V_{\mathrm{g},\alpha}$, and off-diagonal elements genetic covariances $C_{\mathrm{g},\alpha\beta}$). It is always possible to find an orthonormal basis in which the $\mathbf{G}$ matrix is diagonal, that is, to define new phenotypic traits as linear combinations of the "true" phenotypic traits so that the new traits are independent, and fitness is still given by equation 14. Assuming that phenotypes are measured in this new basis, we show in Appendix B that:

$$\frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}} = -\frac{\overline{z_\alpha} - \theta_\alpha}{V_{\mathrm{g},\alpha} + V_{\mathrm{s}}} \tag{16}$$

where $V_{\mathrm{s}} = \omega^2 + V_{\mathrm{e}}$, while:

$$\frac{\partial \ln \overline{W}}{\partial V_{\mathrm{g},\alpha}} = -\frac{1}{2\left(V_{\mathrm{g},\alpha} + V_{\mathrm{s}}\right)} + \frac{1}{2}\left(\frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}}\right)^2 \tag{17}$$

$$\frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\alpha\beta}} = \left(\frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}}\right)\left(\frac{\partial \ln \overline{W}}{\partial \overline{z_\beta}}\right), \tag{18}$$

for $\alpha \neq \beta$.

Change in mean rate of sex during reproduction. To compute the change in $\overline{g_\sigma}$ during reproduction (due to the cost of sex), we first compute the change in $\overline{g_s}$. We have:

$$\begin{aligned} \Delta_{\mathrm{rep}} \overline{g_s} = {} & \mathrm{E}'\left[\frac{c\left(1-s\right)}{c\left(1-\overline{s}'\right)+\overline{s}'}\left(g_s - \overline{g_s}'\right)\right] \\ & + \mathrm{E}'\left[\frac{s_{\mathrm{♀}}\, s_{\mathrm{♂}}}{\overline{s}'\left[c\left(1-\overline{s}'\right)+\overline{s}'\right]}\,\frac{\left(g_{s,\mathrm{♀}} - \overline{g_s}'\right) + \left(g_{s,\mathrm{♂}} - \overline{g_s}'\right)}{2}\right] \end{aligned} \tag{19}$$

where the primes denote averages among individuals after selection (that is, weighting each individual by its relative fecundity), and where the average on the second line is over all possible pairs of female and male parents, $s_{\female}$ and $s_{\male}$ being the investments in sex of these parents, and $g_{s,\female}$, $g_{s,\male}$ their value of g_s. Equation 19 can be understood as follows. The term on the first line is the proportion of asexually produced offspring — which is $(1-\overline{s}')\,/\left(1-\overline{s}'+\frac{\overline{s}'}{c}\right)$ — multiplied by the change in the mean value of g_s among those offspring relative to the parents: for this, each parent is weighted by its relative contribution to the pool of asexually produced offspring, which is $(1-s)\,/\,(1-\overline{s}')$. The term on the second line is the proportion of sexually produced offspring — which is $\frac{\overline{s}'}{c}/\left(1-\overline{s}'+\frac{\overline{s}'}{c}\right)$ — multiplied by the change in the mean value of g_s among those offspring. On average, the mean value of g_s among the offspring of a given female and male is $(g_{s,\female}+g_{s,\male})\,/2$, where $g_{s,\female}$ and $g_{s,\male}$ are the values of g_s in the parents. Furthermore, the relative contributions of these parents to the pool of sexually produced offspring are $s_{\female}/\overline{s}'$ and $s_{\male}/\overline{s}'$. Replacing s by $\overline{s}+g_s-\overline{g_s}'+\overline{g_s}'+e_s$ (and similarly for $g_{s,\female}$, $g_{s,\male}$) in equation 19, and using $\overline{s}'=\overline{s}+\overline{g_s}'$ finally yields:

$$\Delta_{\mathrm{rep}}\overline{g_s}=-\frac{c-1}{c\,(1-\overline{s}')+\overline{s}'}\,{V_{\mathrm{g},s}}'\,. \tag{20}$$

Equation 20 represents the effect of direct selection against sex (whenever $c>1$), and is equivalent to the expression derived in Roze (2014) in the case of a single biallelic sex modifier locus. Strictly, $\overline{s}'$ and ${V_{\mathrm{g},s}}'$ in equation 20 are the mean and genetic variance for investment in sex after selection (weighting each individual by its relative fecundity). However, taking into account the change in $\overline{s}$ and $V_{\mathrm{g},s}$ due to selection would introduce terms in ${V_{\mathrm{g},s}}^2$ in equation 20; neglecting those terms, $\overline{s}'$ and ${V_{\mathrm{g},s}}'$ in equation 20 can thus be replaced by their values $\overline{s}$ and $V_{\mathrm{g},s}$ at the start of the

generation (before selection). From equations 8 and 20, one then obtains:

$$\Delta_{\mathrm{rep}}\overline{g_\sigma} \approx -\frac{c-1}{1+(c-1)\,\overline{\sigma}}\,V_{\mathrm{g},\sigma}\,. \tag{21}$$

Assuming no mutational bias on σ, the change in the mean rate of sex over one generation is given by:

$$\Delta\overline{\sigma} = \Delta_{\mathrm{sel}}\overline{g_\sigma} + \Delta_{\mathrm{rep}}\overline{g_\sigma}\,. \tag{22}$$

In the following, we derive approximate expressions for the moments $C_{\mathrm{g},s\alpha}$ and $M_{\mathrm{g},s\alpha\beta}$ that appear in the expression of $\Delta_{\mathrm{sel}}\overline{g_\sigma}$ (equation 13). However before that, we will compute an expression for the contribution of linkage disequilibria to the genetic variances and covariances between traits affecting fecundity ($\mathcal{D}_{\alpha\beta} = \sum_{j\neq k} C_{\alpha j\,\beta k}$, equations 3 and 4), at quasi-linkage equilibrium.

Genetic associations between selected loci. Neglecting genetic variance in the rate of sex, $C_{\alpha j\,\beta k}$ at the next generation is given by:

$$C''_{\alpha j\,\beta k} = (1-\overline{\sigma})\,C'_{\alpha j\,\beta k} + \overline{\sigma}\,(1-r_{jk})\,C'_{\alpha j\,\beta k} \tag{23}$$

where the double prime denotes variables measured at the next generation (after reproduction), and r_{jk} is the recombination rate between loci j and k. The first term of equation 23 is the proportion of asexually produced offspring, multiplied by the genetic association among those offspring, which is the same as among parents. The second term is the proportion of sexually produced offspring, in which $C_{\alpha j\,\beta k}$ is decreased by a factor $1-r_{jk}$ due to recombination. Equation 23 can be written under the simpler form:

$$C''_{\alpha j\,\beta k} = (1-\rho_{jk})\,C'_{\alpha j\,\beta k} \tag{24}$$

where $\rho_{jk} = \overline{\sigma}\, r_{jk}$ is the "effective" recombination rate between loci j and k.

The effect of selection on $C_{\alpha j\, \beta k}$ can be computed as follows (Turelli and Barton, 1990; Barton, 1995). We have:

$$C'_{\alpha j\, \beta k} = \mathrm{E}' \left[(g_{\alpha j} - \overline{g_{\alpha j}}') \, (g_{\beta k} - \overline{g_{\beta k}}') \right] \tag{25}$$

where again the prime denotes averages after selection (weighting each individual by its relative fecundity). Equation 25 can also be written:

$$C'_{\alpha j\, \beta k} = \mathrm{E}' \left[(g_{\alpha j} - \overline{g_{\alpha j}} - \Delta_{\mathrm{sel}} \overline{g_{\alpha j}}) \, (g_{\beta k} - \overline{g_{\beta k}} - \Delta_{\mathrm{sel}} \overline{g_{\beta k}}) \right]. \tag{26}$$

In the following, we use the notation $C^{\mathrm{sel}}_{\mathbb{U}}$ for genetic associations measured after selection, but using as "reference values" (the $\overline{g_{\alpha j}}$ in equation 1) allelic averages before selection: in particular, $C^{\mathrm{sel}}_{\alpha j\, \beta k} = \mathrm{E}' \left[(g_{\alpha j} - \overline{g_{\alpha j}}) \, (g_{\beta k} - \overline{g_{\beta k}}) \right]$. Expanding equation 26 and noting that $C^{\mathrm{sel}}_{\alpha j} = \mathrm{E}' \left[g_{\alpha j} - \overline{g_{\alpha j}} \right] = \Delta_{\mathrm{sel}} \overline{g_{\alpha j}}$, one obtains:

$$C'_{\alpha j\, \beta k} = C^{\mathrm{sel}}_{\alpha j\, \beta k} - (\Delta_{\mathrm{sel}} \overline{g_{\alpha j}}) \, (\Delta_{\mathrm{sel}} \overline{g_{\beta k}}) \, . \tag{27}$$

Furthermore, we have:

$$C^{\mathrm{sel}}_{\alpha j\, \beta k} = \mathrm{E} \left[\frac{W_{\mathbf{g}}}{\overline{W}} \, \zeta_{\alpha j\, \beta k} \right], \quad \Delta_{\mathrm{sel}} \overline{g_{\alpha j}} = \mathrm{E} \left[\frac{W_{\mathbf{g}}}{\overline{W}} \, \zeta_{\alpha j} \right]. \tag{28}$$

From equation 10, and noting that $g_\alpha - \overline{g_\alpha} = \sum_j \zeta_{\alpha j}$, while $(g_\alpha - \overline{g_\alpha}) \, (g_\beta - \overline{g_\beta}) - C_{\mathrm{g},\, \alpha\beta} = \sum_{j,k} (\zeta_{\alpha j\, \beta k} - C_{\alpha j\, \beta k})$ (where the last sum is over all pairs of loci j and k including $j = k$), we have:

$$\frac{W_{\mathbf{g}}}{\overline{W}} \approx 1 + \sum_\alpha \frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}} \sum_j \zeta_{\alpha j} + \sum_{\alpha \leq \beta} \frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\, \alpha\beta}} \sum_{j,k} (\zeta_{\alpha j\, \beta k} - C_{\alpha j\, \beta k}) \, . \tag{29}$$

Equations 28 and 29 yield:

$$\begin{aligned} C^{\mathrm{sel}}_{\alpha j\, \beta k} = C_{\alpha j\, \beta k} &+ \sum_\gamma \frac{\partial \ln \overline{W}}{\partial \overline{z_\gamma}} \sum_i C_{\gamma i\, \alpha j\, \beta k} \\ &+ \sum_{\gamma \leq \delta} \frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\gamma\delta}} \sum_{h,i} (C_{\gamma h\, \delta i\, \alpha j\, \beta k} - C_{\gamma h\, \delta i} \, C_{\alpha j\, \beta k}) \, . \end{aligned} \tag{30}$$

Equation 30 shows that $C_{\alpha j\,\alpha k}$ is affected by higher-order associations (involving 3 or 4 loci). These associations are in turn generated by the effect of selection, and eroded by recombination. In the following we assume that selection is sufficiently weak relative to recombination, so that between-locus associations remain small (Turelli and Barton, 1990). Ignoring terms involving between-locus associations in the sums on the right-hand-side of equation 30, only the terms for $h = j$, $i = k$ and $h = k$, $i = j$ in the last sum remain, giving:

$$C^{\mathrm{sel}}_{\alpha j\,\alpha k} \approx C_{\alpha j\,\alpha k} + \sum_{\gamma\leq\delta} \frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\gamma\delta}} \left(C_{\alpha j\,\gamma j}\, C_{\beta k\,\delta k} + C_{\alpha j\,\delta j}\, C_{\beta k\,\gamma k}\right). \tag{31}$$

Equations 28 and 29 also yield (neglecting between-locus associations):

$$\Delta_{\mathrm{sel}}\overline{g_{\alpha j}} = \sum_{\gamma} \frac{\partial \ln \overline{W}}{\partial \overline{z_\gamma}} \sum_{i} C_{\alpha j\,\gamma j} + \sum_{\gamma\leq\delta} \frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\gamma\delta}} \sum_{h,i} C_{\alpha j\,\gamma j\,\delta j}\,. \tag{32}$$

$C_{\alpha j\,\beta k}$ at QLE is obtained by solving $C''_{\alpha j\,\beta k} = C_{\alpha j\,\beta k}$. From equations 24, 27, 31 and 32, this yields:

$$C_{\alpha j\,\beta k} = \left(\frac{1}{\rho_{jk}} - 1\right) \Delta_{\mathrm{sel}} C_{\alpha j\,\beta k} \tag{33}$$

with

$$\Delta_{\mathrm{sel}} C_{\alpha j\,\beta k} = \sum_{\gamma,\delta} \left(1 + I_{\gamma\delta}\right) \frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\gamma\delta}}\, C_{\alpha j\,\gamma j}\, C_{\beta k\,\delta k} - \left(\Delta_{\mathrm{sel}}\overline{g_{\alpha j}}\right)\left(\Delta_{\mathrm{sel}}\overline{g_{\alpha k}}\right). \tag{34}$$

where $I_{\gamma\delta}$ equals 1 if $\gamma = \delta$, and 0 otherwise. Summing over all loci, one obtains for $\mathcal{D}_{\alpha\beta} = \sum_{j\neq k} C_{\alpha j\,\beta k}$:

$$\mathcal{D}_{\alpha\beta} \approx \left(\frac{1}{\rho_{\mathrm{h}}} - 1\right) \Delta_{\mathrm{sel}} \mathcal{D}_{\alpha\beta} \tag{35}$$

where ρ_{h} is the harmonic mean of ρ_{jk} over all pairs of loci affecting fecundity, and with

$$\Delta_{\mathrm{sel}} \mathcal{D}_{\alpha\beta} \approx \sum_{\gamma,\delta} \left[\left(1 + I_{\gamma\delta}\right) \frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\gamma\delta}} - \frac{\partial \ln \overline{W}}{\partial \overline{z_\gamma}} \frac{\partial \ln \overline{W}}{\partial \overline{z_\delta}}\right] C^0_{\mathrm{g},\alpha\gamma}\, C^0_{\mathrm{g},\beta\delta}\,. \tag{36}$$

Because $C^0_{\mathrm{g},\alpha\beta} \approx C_{\mathrm{g},\alpha\beta}$ in the QLE regime (weak linkage disequilibria), we may replace $C^0_{\mathrm{g},\alpha\gamma}$ and $C^0_{\mathrm{g},\beta\delta}$ in equation 35 by $C_{\mathrm{g},\alpha\gamma}$ and $C_{\mathrm{g},\beta\delta}$. If phenotypes are measured in a

basis that eliminates genetic covariances among traits, one obtains from equations 35 and 36:

$$\mathcal{D}_{\alpha\alpha} = \sum_{j\neq k} C_{\alpha j\, \alpha k} \approx \left(\frac{1}{\rho_{\mathrm{h}}} - 1\right) \left[2\,\frac{\partial \ln \overline{W}}{\partial V_{\mathrm{g},\,\alpha}} - \left(\frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}}\right)^2\right] {V_{\mathrm{g},\,\alpha}}^2 \,. \tag{37}$$

When the fitness function is given by equation 14, this simplifies to (using equations 16 and 17):

$$\mathcal{D}_{\alpha\alpha} \approx -\left(\frac{1}{\rho_{\mathrm{h}}} - 1\right) \frac{{V_{\mathrm{g},\,\alpha}}^2}{V_{\mathrm{g},\,\alpha} + V_{\mathrm{s}}} \tag{38}$$

corresponding to the result obtained by Bulmer (1985) under the assumption of exchangeable loci (equations A3c and A6a in Charlesworth, 1993).

Indirect selection for sex: "short-term effect". As discussed earlier, the "short-term effect" is represented by the term on the second line of equation 13, which depends on moments $M_{\mathrm{g},\,\sigma\alpha\beta} = \mathrm{E}\left[\left(g_\sigma - \overline{g_\sigma}\right)\left(g_\alpha - \overline{g_\alpha}\right)\left(g_\beta - \overline{g_\beta}\right)\right]$ (for all traits α, β affecting fecundity). From equation 8, we have:

$$M_{\mathrm{g},\,\sigma\alpha\beta} \approx \frac{c}{\left[c\,(1-\overline{s}) + \overline{s}\right]^2} M_{\mathrm{g},\,s\alpha\beta} \tag{39}$$

with $M_{\mathrm{g},\,s\alpha\beta} = \mathrm{E}\left[\left(g_s - \overline{g_s}\right)\left(g_\alpha - \overline{g_\alpha}\right)\left(g_\beta - \overline{g_\beta}\right)\right]$. Furthermore, using our definition of genetic associations, $M_{\mathrm{g},\,s\alpha\beta}$ can be decomposed as:

$$M_{\mathrm{g},\,s\alpha\beta} = \sum_{i,j,k} C_{si\,\alpha j\,\beta k} \tag{40}$$

where the sum is over all loci i affecting investment sex, and over all pairs of loci j and k affecting traits α, β. A QLE approximation for $C_{si\,\alpha j\,\beta k}$ can be obtained as follows. At the next generation, we have:

$$C''_{si\,\alpha j\,\beta k} = \mathrm{E}''\left[\left(g_{si} - \overline{g_{si}}''\right)\left(g_{\alpha j} - \overline{g_{\alpha j}}''\right)\left(g_{\beta k} - \overline{g_{\beta k}}''\right)\right] \tag{41}$$

where again the double primes denote averages over individuals of the next generation (after reproduction). Equation 41 can also be written:

$$C''_{si\,\alpha j\,\beta k} = \mathrm{E}''\Big[\left(g_{si} - \overline{g_{si}}' - \Delta_{\mathrm{rep}}\overline{g_{si}}\right)\left(g_{\alpha j} - \overline{g_{\alpha j}}' - \Delta_{\mathrm{rep}}\overline{g_{\alpha j}}\right) \times \left(g_{\beta k} - \overline{g_{\beta k}}' - \Delta_{\mathrm{rep}}\overline{g_{\beta k}}\right)\Big] \tag{42}$$

where $\overline{g_{si}}'$, $\overline{g_{\alpha j}}'$, $\overline{g_{\beta k}}'$ are the averages of g_{si}, $g_{\alpha j}$, $g_{\beta k}$ among selected parents (weighting each parent by its relative fecundity), and $\Delta_{\mathrm{rep}}\overline{g_{si}} = \overline{g_{si}}'' - \overline{g_{si}}'$ the change in $\overline{g_{si}}$ during reproduction, due to the cost of sex (and similarly for $\Delta_{\mathrm{rep}}\overline{g_{\alpha j}}$ and $\Delta_{\mathrm{rep}}\overline{g_{\beta k}}$). In the following, we use the notation $C^{\mathrm{rep}}_{\mathbb{U}}$ for genetic associations measured after reproduction, but using as "reference values" (the $\overline{g_{\alpha j}}$ in equation 1) allelic averages after selection ($\overline{g_{\alpha j}}'$): for example, $C^{\mathrm{rep}}_{\alpha j\,\beta k} = \mathrm{E}''\left[\left(g_{\alpha j} - \overline{g_{\alpha j}}'\right)\left(g_{\beta k} - \overline{g_{\beta k}}'\right)\right]$. Expanding equation 42 and noting that $\mathrm{E}''\left[g_{si} - \overline{g_{si}}'\right] = \Delta_{\mathrm{rep}}\overline{g_{si}}$, one obtains:

$$\begin{aligned} C''_{si\,\alpha j\,\beta k} = {} & C^{\mathrm{rep}}_{si\,\alpha j\,\beta k} - \left(\Delta_{\mathrm{rep}}\overline{g_{si}}\right) C^{\mathrm{rep}}_{\alpha j\,\beta k} - \left(\Delta_{\mathrm{rep}}\overline{g_{\alpha j}}\right) C^{\mathrm{rep}}_{si\,\beta k} \\ & - \left(\Delta_{\mathrm{rep}}\overline{g_{\beta k}}\right) C^{\mathrm{rep}}_{si\,\alpha j} + 2\left(\Delta_{\mathrm{rep}}\overline{g_{si}}\right)\left(\Delta_{\mathrm{rep}}\overline{g_{\alpha j}}\right)\left(\Delta_{\mathrm{rep}}\overline{g_{\beta k}}\right). \end{aligned} \tag{43}$$

The change in $\overline{g_{\alpha j}}$ during reproduction is generated by the cost of sex and by genetic associations between locus j and loci affecting investment in sex, and is thus proportional to $V_{\mathrm{g},s}$ (the same is true for $\Delta_{\mathrm{rep}}\overline{g_{\beta k}}$). Furthermore, the sum over all i and j of $C_{si\,\beta k}$ is the genetic covariance between trait β and investment in sex s, which is also proportional to $V_{\mathrm{g},s}$. As a consequence, the last three terms of equation 43 will generate terms in $O({V_{\mathrm{g},s}}^2)$, and will thus be ignored, so that:

$$C''_{si\,\alpha j\,\beta k} \approx C^{\mathrm{rep}}_{si\,\alpha j\,\beta k} - \left(\Delta_{\mathrm{rep}}\overline{g_{si}}\right) C^{\mathrm{rep}}_{\alpha j\,\beta k}. \tag{44}$$

Using a similar reasoning as when deriving equation 19, an expression for $C^{\mathrm{rep}}_{si\,\alpha j\,\beta k}$ is

given by (using $\overline{s}' \approx \overline{s}$, and for $j \neq k$):

$$\begin{aligned} C^{\mathrm{rep}}_{si\,\alpha j\,\beta k} = {} & \mathrm{E}'\left[\frac{c\,(1-s)}{c\,(1-\overline{s})+\overline{s}}\,\zeta_{si\,\alpha j\,\beta k}\right] \\ & + \mathrm{E}'\left[\frac{s_{\varphi}\,s_{\male}}{\overline{s}\,[c\,(1-\overline{s})+\overline{s}]}\Big(r_{ijk,\emptyset}\,\zeta_{si\,\alpha j\,\beta k,\varphi} + r_{\emptyset,ijk}\,\zeta_{si\,\alpha j\,\beta k,\male}\right. \\ & \qquad + r_{i,jk}\,\zeta_{si,\varphi}\,\zeta_{\alpha j\,\beta k,\male} + r_{jk,i}\,\zeta_{\alpha j\,\beta k,\varphi}\,\zeta_{si,\male} \\ & \qquad + r_{ij,k}\,\zeta_{si\,\alpha j,\varphi}\,\zeta_{\beta k,\male} + r_{k,ij}\,\zeta_{\beta k,\varphi}\,\zeta_{si\,\alpha j,\male} \\ & \qquad \left. + r_{ik,j}\,\zeta_{si\,\beta k,\varphi}\,\zeta_{\alpha j,\male} + r_{j,ik}\,\zeta_{\alpha j,\varphi}\,\zeta_{si\,\beta k,\male}\Big)\right]. \end{aligned} \tag{45}$$

In equation 45, $r_{\mathbb{S},\mathbb{T}}$ is the probability that a meiotic product inherits the set $\mathbb{S}$ of loci from the maternal genome, and the set $\mathbb{T}$ of loci from the paternal genome, while $\zeta_{\mathbb{S},\varphi}$, $\zeta_{\mathbb{S},\male}$ variables are measured in the maternal and paternal parent, respectively. Writing s on the first line of equation 45 under the form $\overline{s} + \sum_h \zeta_{sh} + e_s$, and s_{φ}, $s_{\male}$ on the second line as $\overline{s} + \sum_h \zeta_{sh,\varphi} + e_{s,\varphi}$ and $\overline{s} + \sum_h \zeta_{sh,\male} + e_{s,\male}$, one arrives at:

$$\begin{aligned} C^{\mathrm{rep}}_{si\,\alpha j\,\beta k} = {} & \left[1 - \frac{\overline{s}}{c\,(1-\overline{s})+\overline{s}}\,(1 - r_{ijk,\emptyset} - r_{\emptyset,ijk})\right] C'_{si\,\alpha j\,\beta k} \\ & + \frac{1}{c\,(1-\overline{s})+\overline{s}}\left[-\,(c - r_{ijk,\emptyset} - r_{\emptyset,ijk}) \sum_h C'_{sh\,si\,\alpha j\,\beta k}\right. \\ & \qquad + (r_{i,jk} + r_{jk,i})\left(C'_{\alpha j\,\beta k} \sum_h C'_{sh\,si} + \sum_l C'_{sl\,\alpha j\,\beta k} \sum_h C'_{sh\,si}\right) \\ & \qquad + (r_{ij,k} + r_{k,ij})\left(C'_{si\,\alpha j} \sum_h C'_{sh\,\beta k} + \sum_l C'_{sl\,si\,\alpha j} \sum_h C'_{sh\,\beta k}\right) \\ & \qquad \left. + (r_{ik,j} + r_{j,ik})\left(C'_{si\,\beta k} \sum_h C'_{sh\,\alpha j} + \sum_l C'_{sl\,si\,\beta k} \sum_h C'_{sh\,\alpha j}\right)\right]. \end{aligned} \tag{46}$$

Many of the terms of equation 46 may be neglected when $V_{\mathrm{g},\,s}$ is small, using the fact that sums over all loci of associations involving one or several "s" indices are proportional to $V_{\mathrm{g},\,s}$: therefore, the terms on the last two lines of equation 46 and the last term of the third line will generate terms in ${V_{\mathrm{g},\,s}}^2$. Furthermore, we will neglect linkage disequilibria between loci affecting investment in sex, so that only the terms

for $h = i$ remain in the sums above, and equation 46 simplifies to:

$$\begin{aligned} C^{\mathrm{rep}}_{si\,\alpha j\,\beta k} \approx (1-\rho_{ijk})\, C'_{si\,\alpha j\,\beta k} - \frac{c - r_{ijk,\emptyset} - r_{\emptyset,ijk}}{c\,(1-\overline{s}) + \overline{s}}\, C'_{si\,si\,\alpha j\,\beta k} \\ + \frac{r_{i,jk} + r_{jk,i}}{c\,(1-\overline{s}) + \overline{s}}\, C'_{si\,si}\, C'_{\alpha j\,\beta k} \end{aligned} \tag{47}$$

with $\rho_{ijk} = \overline{\sigma}\, r_{ijk}$, $r_{ijk} = 1 - r_{ijk,\emptyset} - r_{\emptyset,ijk}$ being the probability that at least one recombination event occurs between loci i, j and k. Because $C'_{si\,si\,\alpha j\,\beta k} \approx C'_{si\,si}\, C'_{\alpha j\,\beta k}$ to leading order, equation 47 further simplifies to:

$$C^{\mathrm{rep}}_{si\,\alpha j\,\beta k} \approx (1-\rho_{ijk})\, C'_{si\,\alpha j\,\beta k} - \frac{c - 1 + r_{jk}}{c\,(1-\overline{s}) + \overline{s}}\, C'_{si\,si}\, C'_{\alpha j\,\beta k} \tag{48}$$

The term $\Delta_{\mathrm{rep}}\overline{g_{si}}$ in equation 44 is given by:

$$\Delta_{\mathrm{rep}}\overline{g_{si}} = \mathrm{E}'\left[\frac{c\,(1-s)}{c\,(1-\overline{s}) + \overline{s}}\,\zeta_{si}\right] + \mathrm{E}'\left[\frac{s_{\mathrm{♀}}\, s_{\mathrm{♂}}}{\overline{s}\,[c\,(1-\overline{s}) + \overline{s}]}\left(\frac{\zeta_{si,\mathrm{♀}} + \zeta_{si,\mathrm{♂}}}{2}\right)\right]. \tag{49}$$

Neglecting linkage disequilibria between loci affecting the rate of sex, this yields:

$$\Delta_{\mathrm{rep}}\overline{g_{si}} \approx -\frac{c-1}{c\,(1-\overline{s}) + \overline{s}}\, C'_{si\,si}. \tag{50}$$

From equation 24, $C^{\mathrm{rep}}_{\alpha j\,\beta k} = (1-\rho_{jk})\, C'_{\alpha j\,\beta k}$ to leading order (that is, neglecting genetic variation for the rate of sex), so that:

$$(\Delta_{\mathrm{rep}}\overline{g_{si}})\, C^{\mathrm{rep}}_{\alpha j\,\beta k} \approx -\frac{c-1}{c\,(1-\overline{s}) + \overline{s}}\,(1-\rho_{jk})\, C'_{si\,si}\, C'_{\alpha j\,\beta k}\,. \tag{51}$$

Putting everything together, one obtains from equations 44, 48 and 51:

$$C''_{si\,\alpha j\,\beta k} \approx (1-\rho_{ijk})\, C'_{si\,\alpha j\,\beta k} - \frac{c\, r_{jk}}{[c\,(1-\overline{s}) + \overline{s}]^2}\, C'_{si\,si}\, C'_{\alpha j\,\beta k}\,. \tag{52}$$

The change in $C_{si\,si}$ due to selection may be neglected under our assumption that $V_{\mathrm{g},s}$ is small (as it would generate terms in ${V_{\mathrm{g},s}}^2$). Furthermore, the effect of selection on $C_{si\,\alpha j\,\beta k}$ can be neglected when selection is weak, as it involves higher-order associations

between loci i, j, k and other loci, which are themselves generated by the effect of selection at these loci. Using these approximations, equation 52 becomes:

$$C''_{si\,\alpha j\,\beta k} \approx (1-\rho_{ijk})\, C_{si\,\alpha j\,\beta k} - \frac{c\, r_{jk}}{\left[c\,(1-\overline{s})+\overline{s}\right]^2}\, C_{si\,si}\, C'_{\alpha j\,\beta k} \tag{53}$$

giving at QLE:

$$C_{si\,\alpha j\,\beta k} \approx -\frac{1}{\rho_{ijk}} \frac{c\, r_{jk}\, C'_{\alpha j\,\beta k}}{\left[c\,(1-\overline{s})+\overline{s}\right]^2}\, C_{si\,si}\,. \tag{54}$$

From the results of the preceding subsection, $C'_{\alpha j\,\beta k} \approx \left(\Delta_{\rm sel} C_{\alpha j\,\beta k}\right)/\rho_{jk}$ (where $\Delta_{\rm sel} C_{\alpha j\,\beta k}$ is given by equation 34), so that $r_{jk}\, C'_{\alpha j\,\beta k} \approx \Delta_{\rm sel} C_{\alpha j\,\beta k}/\overline{\sigma}$. Equation 54 thus simplifies to:

$$C_{si\,\alpha j\,\beta k} \approx -\frac{1}{r_{ijk}} \frac{c}{\overline{s}^2} \left(\Delta_{\rm sel} C_{\alpha j\,\beta k}\right) C_{si\,si}\,. \tag{55}$$

The same reasoning as above can be used to compute $C_{si\,\alpha j\,\beta j}$, which is found to be negligible. Summing over all loci, one thus obtains:

$$M_{{\rm g},\, s\alpha\beta} \approx -\frac{1}{r_{\rm h,1}} \frac{c}{\overline{s}^2} \left(\Delta_{\rm sel} \mathcal{D}_{\alpha\beta}\right) V_{{\rm g},\, s} \tag{56}$$

where $r_{\rm h,1}$ is the harmonic mean of r_{ijk} over all triplets of loci involving one locus affecting sex and two loci affecting fecundity. From this, we have (using equations 8 and 39):

$$M_{{\rm g},\, \sigma\alpha\beta} \approx -\frac{\Delta_{\rm sel} \mathcal{D}_{\alpha\beta}}{r_{\rm h,1}\, \overline{\sigma}^2}\, V_{{\rm g},\, \sigma}\,. \tag{57}$$

Indirect selection for sex due to the short-term effect (the second term of equation 13) can thus be written approximately as:

$$-\frac{1}{r_{\rm h,1}\, \overline{\sigma}^2} \left(\sum_{\alpha \le \beta} \frac{\partial \ln \overline{W}}{\partial C_{{\rm g},\, \alpha\beta}}\, \Delta_{\rm sel} \mathcal{D}_{\alpha\beta} \right) V_{{\rm g},\, \sigma}\,. \tag{58}$$

We will see later that the term between parentheses can be expressed in terms of the effect of sex on the mean fitness of offspring.

Indirect selection for sex: "long-term effect". The long-term effect depends on genetic covariances between the rate of sex and traits affecting fecundity ($C_{\mathrm{g},\sigma\alpha}$). From equation 8, we have:

$$C_{\mathrm{g},\sigma\alpha} \approx \frac{c}{\left[c\,(1-\overline{s})+\overline{s}\right]^2}\,C_{\mathrm{g},s\alpha} \tag{59}$$

with $C_{\mathrm{g},s\alpha} = \mathrm{E}\left[(g_s-\overline{g_s})\,(g_\alpha-\overline{g_\alpha})\right]$, which can be decomposed as:

$$C_{\mathrm{g},s\alpha} = \sum_{i,j} C_{si\,\alpha j}\,. \tag{60}$$

Using the same approach as above, one obtains for the effect of reproduction on $C_{si\,\alpha j}$:

$$C''_{si\,\alpha j} = C^{\mathrm{rep}}_{si\,\alpha j} - (\Delta_{\mathrm{rep}}\overline{g_{si}})\,(\Delta_{\mathrm{rep}}\overline{g_{\alpha j}}) \approx C^{\mathrm{rep}}_{si\,\alpha j} \tag{61}$$

since the term $(\Delta_{\mathrm{rep}}\overline{g_{si}})\,(\Delta_{\mathrm{rep}}\overline{g_{\alpha j}})$ will generate terms in ${V_{\mathrm{g},s}}^2$. Neglecting linkage disequilibria between loci affecting the rate of sex and other terms of order ${V_{\mathrm{g},s}}^2$, we have:

$$C^{\mathrm{rep}}_{si\,\alpha j} \approx (1-\rho_{ij})\,C'_{si\,\alpha j} - \frac{c-1+r_{ij}}{c\,(1-\overline{s})+\overline{s}}\,C'_{si\,si\,\alpha j}. \tag{62}$$

To the first order in $V_{\mathrm{g},s}$,

$$C'_{si\,si\,\alpha j} \approx C^{\mathrm{sel}}_{si\,si\,\alpha j} - (\Delta_{\mathrm{sel}}\overline{g_{\alpha j}})\,C_{si\,si} \tag{63}$$

while from equation 29:

$$\begin{aligned} C^{\mathrm{sel}}_{si\,si\,\alpha j} \approx C_{si\,si\,\alpha j} &+ \sum_\beta \frac{\partial \ln \overline{W}}{\partial \overline{z_\beta}} \sum_k C_{si\,si\,\alpha j\,\beta k} \\ &+ \sum_{\beta\le\gamma} \frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\beta\gamma}} \sum_{k,l} \left(C_{si\,si\,\alpha j\,\beta k\,\gamma l} - C_{si\,si\,\alpha j} C_{\beta k\,\gamma l}\right) \end{aligned} \tag{64}$$

$$\Delta_{\mathrm{sel}}\overline{g_{\alpha j}} \approx \sum_\beta \frac{\partial \ln \overline{W}}{\partial \overline{z_\beta}} \sum_k C_{\alpha j\,\beta k} + \sum_{\beta\le\gamma} \frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\beta\gamma}} \sum_{k,l} C_{\alpha j\,\beta k\,\gamma l}\,. \tag{65}$$

From equations 63 – 65, and using the fact that $C_{si\,si\,\alpha j\,\beta k} \approx C_{si\,si}\,C_{\alpha j\,\beta k}$, $C_{si\,si\,\alpha j\,\beta k\,\gamma l} \approx C_{si\,si}\,C_{\alpha j\,\beta k\,\gamma l}$ to the first order in $V_{\mathrm{g},s}$, one obtains that the effect of selection on $C_{si\,si\,\alpha j}$ is negligible, which finally leads to $C_{si\,si\,\alpha j} \approx C'_{si\,si\,\alpha j} \approx 0$ at QLE.

The effect of selection on $C_{si\,\alpha j}$ is given by:

$$C'_{si\,\alpha j} \approx C^{\mathrm{sel}}_{si\,\alpha j} - (\Delta_{\mathrm{sel}}\overline{g_{si}})\,(\Delta_{\mathrm{sel}}\overline{g_{\alpha j}})\,, \tag{66}$$

where

$$\begin{aligned} C^{\mathrm{sel}}_{si\,\alpha j} \approx C_{si\,\alpha j} &+ \sum_{\beta} \frac{\partial \ln \overline{W}}{\partial \overline{z_\beta}} \sum_{k} C_{si\,\alpha j\,\beta k} \\ &+ \sum_{\beta \leq \gamma} \frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\beta\gamma}} \sum_{k,l} \left(C_{si\,\alpha j\,\beta k\,\gamma l} - C_{si\,\alpha j} C_{\beta k\,\gamma l}\right), \end{aligned} \tag{67}$$

while the term $(\Delta_{\mathrm{sel}}\overline{g_{si}})\,(\Delta_{\mathrm{sel}}\overline{g_{\alpha j}})$ is of higher order in the strength of selection, and may thus be neglected. Finally, using the same method as in the previous subsection shows that associations $C_{si\,\alpha j\,\beta k\,\gamma l}$ (that appear on the second line of equation 67) are proportional to $C_{si\,si}\,C_{\alpha j\,\beta k\,\gamma l}$. However, 3-locus associations $C_{\alpha j\,\beta k\,\gamma l}$ are of higher order in the strength of selection than 2-locus associations, and we will assume that the sum over all loci of these associations is negligible relative to the sum of pairwise associations $C_{\alpha j\,\beta k}$. This leaves us with the following recursion for $C_{si\,\alpha j}$:

$$C''_{si\,\alpha j} \approx (1-\rho_{ij}) \left[C_{si\,\alpha j} + \sum_{\beta} \frac{\partial \ln \overline{W}}{\partial \overline{z_\beta}} \sum_{k} C_{si\,\alpha j\,\beta k} \right]. \tag{68}$$

At QLE, and using equation 55, we thus have:

$$C_{si\,\alpha j} \approx -\left(\frac{1}{\rho_{ij}} - 1\right) \frac{c}{\overline{s}^2} \sum_{k} \frac{1}{r_{ijk}} \sum_{\beta} \frac{\partial \ln \overline{W}}{\partial \overline{z_\beta}} \left(\Delta_{\mathrm{sel}} C_{\alpha j\,\beta k}\right) C_{si\,si}, \tag{69}$$

and summing over all loci:

$$C_{\mathrm{g},s\alpha} \approx -\left(\frac{1}{r_{\mathrm{h},2}\,\overline{\sigma}} - \frac{1}{r_{\mathrm{h},1}}\right) \frac{c}{\overline{s}^2} \sum_{\beta} \frac{\partial \ln \overline{W}}{\partial \overline{z_\beta}} \left(\Delta_{\mathrm{sel}} \mathcal{D}_{\alpha\beta}\right) V_{\mathrm{g},s} \tag{70}$$

where $r_{\mathrm{h},2}$ is the harmonic mean of $r_{ij}\,r_{ijk}$ over all triplets of loci i, j and k, where i affects investment in sex while j and k affect fecundity. Equations 8 and 59 then yield:

$$C_{\mathrm{g},\sigma\alpha} \approx -\left(\frac{1}{r_{\mathrm{h},2}\,\overline{\sigma}} - \frac{1}{r_{\mathrm{h},1}}\right) \frac{1}{\overline{\sigma}^2} \sum_{\beta} \frac{\partial \ln \overline{W}}{\partial \overline{z_\beta}} \left(\Delta_{\mathrm{sel}} \mathcal{D}_{\alpha\beta}\right) V_{\mathrm{g},\sigma} \tag{71}$$

and indirect selection for sex due to the long term effect (first term of equation 13) is thus approximately:

$$-\left(\frac{1}{r_{\mathrm{h},2}\,\overline{\sigma}}-\frac{1}{r_{\mathrm{h},1}}\right)\frac{1}{\overline{\sigma}^2}\left(\sum_{\alpha,\beta}\frac{\partial\ln\overline{W}}{\partial\overline{z_\alpha}}\frac{\partial\ln\overline{W}}{\partial\overline{z_\beta}}\,\Delta_{\mathrm{sel}}\mathcal{D}_{\alpha\beta}\right)V_{\mathrm{g},\sigma}\,. \tag{72}$$

Note that the term in the first parenthesis of equation 72 is positive, $1/r_{\mathrm{h},1}$ becoming negligible compared with $1/\left(r_{\mathrm{h},2}\,\overline{\sigma}\right)$ as the rate of sex decreases.

Expressing indirect selection in terms of the effect of sex on the fitness of offspring. The terms between parentheses in equation 58 and 72 (involving $\Delta_{\mathrm{sel}}\mathcal{D}_{\alpha\beta}$) provide intuitive understanding of the mechanisms generating indirect selection for sex, but would be difficult to measure in a real population. However, using our hypothesis of weak selection and Gaussian distribution of traits affecting fecundity, these can be expressed in terms of the effect of sex on the fecundity of offspring, that could (at least in principle) be measured in an experimental population. Indeed, a Taylor series on $\ln\overline{W}$ provides the following approximation for the effect of a change in mean breeding values and/or in the genetic variance-covariance matrix on $\ln\overline{W}$:

$$\Delta\ln\overline{W}\approx\sum_{\alpha}\Delta\overline{z_\alpha}\,\frac{\partial\ln\overline{W}}{\partial\overline{z_\alpha}}+\sum_{\alpha\leq\beta}\Delta C_{\mathrm{g},\alpha\beta}\,\frac{\partial\ln\overline{W}}{\partial C_{\mathrm{g},\alpha\beta}}\,. \tag{73}$$

If we now imagine an experiment where we sample a sufficiently large number of individuals from the population (so that genetic associations within this pool of individuals are representative of associations in the whole population) and let them produce a pool of offspring by sexual reproduction and another pool by asexual reproduction, both pools should have the same mean breeding values (on average), while genetic variances and covariances (measured separately within each pool of offspring) should differ by

an amount:

$$\Delta_{\text{sex/asex}} C_{\text{g},\,\alpha\beta} = C^{\text{sex}}_{\text{g},\,\alpha\beta} - C^{\text{asex}}_{\text{g},\,\alpha\beta} = -\sum_{j\neq k} r_{jk}\, C'_{\alpha j\,\beta k} \tag{74}$$

due to the effect of sexual recombination. From equation 33, we have $\Delta_{\text{sel}}\mathcal{D}_{\alpha\beta} \approx \sum_{j\neq k} \rho_{jk}\, C'_{\alpha j\,\beta k}$, so that:

$$\Delta_{\text{sex/asex}} C_{\text{g},\,\alpha\beta} \approx -\frac{1}{\overline{\sigma}}\, \Delta_{\text{sel}}\mathcal{D}_{\alpha\beta}. \tag{75}$$

Therefore, from equation 73, the difference in $\ln \overline{W}$ between sexually and asexually produced offspring is given by:

$$\Delta_1 = \ln \overline{W}_{\text{sex}} - \ln \overline{W}_{\text{asex}} \approx -\frac{1}{\overline{\sigma}} \sum_{\alpha\leq\beta} \frac{\partial \ln \overline{W}}{\partial C_{\text{g},\,\alpha\beta}}\, \Delta_{\text{sel}}\mathcal{D}_{\alpha\beta} \tag{76}$$

and indirect selection for sex due to the short-term effect (equation 58) thus becomes:

$$\frac{\Delta_1}{r_{\text{h},1}\,\overline{\sigma}}\, V_{\text{g},\,\sigma}\,. \tag{77}$$

Following Barton (1995) and Charlesworth and Barton (1996), selection for sex due to the long-term effect can be expressed in terms of the effect of sex on the variance in log-fitness among offspring. From equation 10 we have, to leading order in selection gradients:

$$\begin{aligned} \ln W_{\mathbf{g}} - \ln \overline{W} \approx & \sum_{\alpha} (g_\alpha - \overline{g_\alpha})\, \frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}} \\ & + \sum_{\alpha\leq\beta} \left[(g_\alpha - \overline{g_\alpha})\, (g_\beta - \overline{g_\beta}) - C_{\text{g},\,\alpha\beta} \right] \frac{\partial \ln \overline{W}}{\partial C_{\text{g},\,\alpha\beta}} \end{aligned} \tag{78}$$

so that the variance in $\ln W_{\mathbf{g}}$ among individuals is:

$$\begin{aligned} \text{Var}\left[\ln W_{\mathbf{g}}\right] \approx & \sum_{\alpha,\beta} C_{\text{g},\,\alpha\beta}\, \frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}} \frac{\partial \ln \overline{W}}{\partial \overline{z_\beta}} \\ & + \sum_{\alpha\leq\beta} \sum_{\gamma\leq\delta} \left(C_{\text{g},\,\alpha\gamma}\, C_{\text{g},\,\beta\delta} + C_{\text{g},\,\alpha\delta}\, C_{\text{g},\,\beta\gamma} \right) \frac{\partial \ln \overline{W}}{\partial C_{\text{g},\,\alpha\beta}} \frac{\partial \ln \overline{W}}{\partial C_{\text{g},\,\gamma\delta}}\,. \end{aligned} \tag{79}$$

Equation 79 is approximately equivalent to the first two lines of equation A3b in Charlesworth and Barton (1996), corresponding to the additive and epistatic components of the variance in log fitness (denoted hereafter V_{A} and V_{AA}). Using equations 75 and 79, the sum appearing in the expression for the strength of selection for sex due to the long term effect (equation 72) can be expressed as $-\overline{\sigma}\,(V_{\mathrm{A,\,sex}} - V_{\mathrm{A,\,asex}})$, where $V_{\mathrm{A,\,sex}}$ and $V_{\mathrm{A,\,asex}}$ are the additive components of the variance in log fitness (first term of equation 79) among offspring produced by sexual and asexual reproduction, respectively. Selection for sex due to the long term effect thus becomes:

$$\left(\frac{1}{r_{\mathrm{h},2}\,\overline{\sigma}} - \frac{1}{r_{\mathrm{h},1}}\right)\frac{\Delta_2}{\overline{\sigma}}\,V_{\mathrm{g},\sigma} \tag{80}$$

with $\Delta_2 = V_{\mathrm{A,\,sex}} - V_{\mathrm{A,\,asex}}$. Assuming that epistasis is weak relative to directional selection, Charlesworth and Barton (1996) show that the effect of recombination on V_{AA} may be neglected, in which case the long term effect can be expressed in terms of the effect of recombination on $\mathrm{Var}\,[\ln W_{\mathbf{g}}]$. However, in situations where epistatic interactions may be of the same order of magnitude as directional selection (as in the present model), the additive component of $\mathrm{Var}\,[\ln W_{\mathbf{g}}]$ should be estimated, for example from the covariance between parents and offspring (e.g., Lynch and Walsh, 1998). Indeed, under the assumption of a sufficiently large number of loci with weak effects so that the joint distribution of trait values in parents and offspring is approximately multivariate Gaussian, the covariance in log fitness between parents and offspring is:

$$\begin{aligned}\mathrm{Cov}^{\mathrm{PO}}\,[\ln W_{\mathbf{g}}] \approx & \sum_{\alpha,\beta} C^{\mathrm{PO}}_{\mathrm{g},\alpha\beta}\,\frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}}\frac{\partial \ln \overline{W}}{\partial \overline{z_\beta}} \\ & + \sum_{\alpha\le\beta}\sum_{\gamma\le\delta}\left(C^{\mathrm{PO}}_{\mathrm{g},\alpha\gamma}\,C^{\mathrm{PO}}_{\mathrm{g},\beta\delta} + C^{\mathrm{PO}}_{\mathrm{g},\alpha\delta}\,C^{\mathrm{PO}}_{\mathrm{g},\beta\gamma}\right)\frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\alpha\beta}}\frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\gamma\delta}}\end{aligned} \tag{81}$$

where $C^{\mathrm{PO}}_{\mathrm{g},\alpha\beta}$ is the covariance between g_α in the parents and g_β in their offspring. Using

$C^{\mathrm{PO}}_{\mathrm{g},\alpha\beta} = C_{\mathrm{g},\alpha\beta}/2$, equation 81 becomes:

$$\mathrm{Cov}^{\mathrm{PO}}\left[\ln W_{\mathbf{g}}\right] \approx \frac{V_{\mathrm{A}}}{2} + \frac{V_{\mathrm{AA}}}{4} \tag{82}$$

yielding

$$V_{\mathrm{A}} \approx 4\mathrm{Cov}^{\mathrm{PO}}\left[\ln W_{\mathbf{g}}\right] - \mathrm{Var}\left[\ln W_{\mathbf{g}}\right]. \tag{83}$$

APPENDIX A: APPROXIMATION FOR $W_{\mathbf{g}}/\overline{W}$

Assuming that selection is weak (meaning that the variance in $W_{\mathbf{g}}$ is small), we can approximate $W_{\mathbf{g}}$ by a Taylor series around $\overline{\mathbf{g}} = (\overline{g_1}, \overline{g_2}, \ldots)$:

$$\begin{aligned} W_{\mathbf{g}} \approx W_{\mathbf{g}}(\overline{\mathbf{g}}) &+ \sum_{\alpha} (g_\alpha - \overline{g_\alpha}) \frac{\partial W_{\mathbf{g}}}{\partial g_\alpha} \\ &+ \frac{1}{2} \sum_{\alpha,\beta} (g_\alpha - \overline{g_\alpha}) (g_\beta - \overline{g_\beta}) \frac{\partial^2 W_{\mathbf{g}}}{\partial g_\alpha \partial g_\beta} \end{aligned} \tag{A1}$$

where the partial derivatives are taken in $\overline{\mathbf{g}}$, and the last sum includes $\alpha = \beta$. Averaging over all individuals yields $\overline{W} \approx W_{\mathbf{g}}(\overline{\mathbf{g}}) + \frac{1}{2} \sum_{\alpha,\beta} C_{\mathrm{g},\alpha\beta}\, \partial^2 W_{\mathbf{g}} / (\partial g_\alpha \partial g_\beta)$, so that equation A1 can also be written as:

$$\begin{aligned} W_{\mathbf{g}} \approx \overline{W} &+ \sum_{\alpha} (g_\alpha - \overline{g_\alpha}) \frac{\partial W_{\mathbf{g}}}{\partial g_\alpha} \\ &+ \frac{1}{2} \sum_{\alpha,\beta} \left[(g_\alpha - \overline{g_\alpha}) (g_\beta - \overline{g_\beta}) - C_{\mathrm{g},\alpha\beta} \right] \frac{\partial^2 W_{\mathbf{g}}}{\partial g_\alpha \partial g_\beta} \,. \end{aligned} \tag{A2}$$

The derivatives of $W_{\mathbf{g}}$ in equation can be expressed in terms of derivatives of $\overline{W}$ (Barton and Turelli, 1991; Turelli and Barton, 1994). Consider the effect of a slight change in the distribution of breeding values $\mathbf{g}$ on mean fitness: $\overline{g_\alpha}$ and $C_{\mathrm{g},\alpha\beta}$ change to $\overline{g_\alpha}^*$ and $C^*_{\mathrm{g},\alpha\beta}$, causing mean fitness to change from $\overline{W}$ to $\overline{W}^*$. Replacing $g_\alpha - \overline{g_\alpha}$ by $g_\alpha - \overline{g_\alpha}^* + \overline{g_\alpha}^* - \overline{g_\alpha}$ in equation A2 and averaging over the new state of the population yields:

$$\overline{W}^* \approx \overline{W} + \sum_{\alpha} (\overline{g_\alpha}^* - \overline{g_\alpha}) \frac{\partial W_{\mathbf{g}}}{\partial g_\alpha} + \frac{1}{2} \sum_{\alpha,\beta} \left(C^*_{\mathrm{g},\alpha\beta} - C_{\mathrm{g},\alpha\beta} \right) \frac{\partial^2 W_{\mathbf{g}}}{\partial g_\alpha \partial g_\beta} \,. \tag{A3}$$

Note that terms $(\overline{g_\alpha}^* - \overline{g_\alpha})(\overline{g_\beta}^* - \overline{g_\beta})$ appearing in the second sum have been neglected, as we assume that $\overline{g_\alpha}^* - \overline{g_\alpha}$ is small for all α. Another expression for $\overline{W}^*$ can be obtained by developing $\overline{W}$ (which is a function of $\overline{g_\alpha} = \overline{z_\alpha}$ and $C_{\mathrm{g},\alpha\beta}$ for all α, β) as a Taylor

series:

$$\overline{W}^* \approx \overline{W} + \sum_{\alpha} \left(\overline{g_\alpha}^* - \overline{g_\alpha}\right) \frac{\partial \overline{W}}{\partial \overline{z_\alpha}} + \sum_{\alpha \leq \beta} \left(C^*_{\mathrm{g},\,\alpha\beta} - C_{\mathrm{g},\,\alpha\beta}\right) \frac{\partial \overline{W}}{\partial C_{\mathrm{g},\,\alpha\beta}} \tag{A4}$$

(note that each (α, β) pair is counted only once in the last sum). From equations A3 and A4, we have

$$\frac{\partial W_{\mathbf{g}}}{\partial g_\alpha} \approx \frac{\partial \overline{W}}{\partial \overline{z_\alpha}}, \quad \frac{\partial^2 W_{\mathbf{g}}}{\partial {g_\alpha}^2} \approx 2\,\frac{\partial \overline{W}}{\partial V_{\mathrm{g},\,\alpha}}, \quad \frac{\partial^2 W_{\mathbf{g}}}{\partial g_\alpha \partial g_\beta} \approx \frac{\partial \overline{W}}{\partial C_{\mathrm{g},\,\alpha\beta}} \;(\alpha \neq \beta) \tag{A5}$$

and equation A2 and A5 yield (after dividing both sides by $\overline{W}$):

$$\begin{aligned} \frac{W_{\mathbf{g}}}{\overline{W}} \approx 1 &+ \sum_{\alpha} \left(g_\alpha - \overline{g_\alpha}\right) \frac{\partial \ln \overline{W}}{\partial \overline{z_\alpha}} \\ &+ \sum_{\alpha \leq \beta} \left[\left(g_\alpha - \overline{g_\alpha}\right)\left(g_\beta - \overline{g_\beta}\right) - C_{\mathrm{g},\,\alpha\beta}\right] \frac{\partial \ln \overline{W}}{\partial C_{\mathrm{g},\,\alpha\beta}}\,. \end{aligned} \tag{A6}$$

APPENDIX B: SELECTION GRADIENTS WITH ISOTROPIC, GAUSSIAN FITNESS FUNCTION

From equation 15, we have:

$$\begin{aligned}\ln \overline{W} &= \frac{1}{2}\ln\left[\det\left(\left(\mathbf{S}+\mathbf{P}\right)^{-1}\mathbf{S}\right)\right] - \frac{1}{2}\left(\overline{\mathbf{z}}-\boldsymbol{\theta}\right)^{T}\left(\mathbf{S}+\mathbf{P}\right)^{-1}\left(\overline{\mathbf{z}}-\boldsymbol{\theta}\right) \\ &= \frac{1}{2}\ln\left[\det(\mathbf{S})\right] - \frac{1}{2}\ln\left[\det(\mathbf{S}+\mathbf{P})\right] - \frac{1}{2}\left(\overline{\mathbf{z}}-\boldsymbol{\theta}\right)^{T}\left(\mathbf{S}+\mathbf{P}\right)^{-1}\left(\overline{\mathbf{z}}-\boldsymbol{\theta}\right)\end{aligned} \tag{B1}$$

so that:

$$\frac{\partial \ln \overline{W}}{\partial \overline{\mathbf{z}}} = -\left(\mathbf{S}+\mathbf{P}\right)^{-1}\left(\overline{\mathbf{z}}-\boldsymbol{\theta}\right) \tag{B2}$$

and

$$\begin{aligned}\frac{\partial \ln \overline{W}}{\partial \mathbf{G}} &= -\frac{1}{2}\frac{\partial \ln\left[\det(\mathbf{S}+\mathbf{P})\right]}{\partial \mathbf{G}} - \frac{1}{2}\left(\overline{\mathbf{z}}-\boldsymbol{\theta}\right)^{T}\frac{\partial\left(\mathbf{S}+\mathbf{P}\right)^{-1}}{\partial \mathbf{G}}\left(\overline{\mathbf{z}}-\boldsymbol{\theta}\right) \\ &= -\frac{1}{2}\mathrm{Tr}\left(\left(\mathbf{S}+\mathbf{P}\right)^{-1}\frac{\partial\left(\mathbf{S}+\mathbf{P}\right)}{\partial \mathbf{G}}\right) \\ &\qquad + \frac{1}{2}\left(\overline{\mathbf{z}}-\boldsymbol{\theta}\right)^{T}\left(\mathbf{S}+\mathbf{P}\right)^{-1}\frac{\partial\left(\mathbf{S}+\mathbf{P}\right)}{\partial \mathbf{G}}\left(\mathbf{S}+\mathbf{P}\right)^{-1}\left(\overline{\mathbf{z}}-\boldsymbol{\theta}\right)\end{aligned} \tag{B3}$$

where Tr stands for the trace of a matrix. If phenotypes are measured in a basis that eliminates covariances among traits, $\left(\mathbf{S}+\mathbf{P}\right)^{-1}$ is a diagonal matrix, with elements $1/\left(V_{\mathrm{g},\alpha}+V_{\mathrm{s}}\right)$ on its diagonal (with $V_{\mathrm{s}}=\omega^2+V_{\mathrm{e}}$). In that case, equations B2 and B3 yield equations 16 – 18.

LITERATURE CITED

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

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

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

Bulmer, M. G. 1985. The Mathematical Theory of Quantitative Genetics, 2nd edition. Oxford University Press, Oxford.

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

Charlesworth, B. and N. H. Barton. 1996. Recombination load associated with selection for increased recombination. Genet. Res. 67:27–41.

Gomulkiewicz, R. and D. Houle. 2009. Demographic and genetic constraints on evolution. Am. Nat. 174:E218–E229.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33:402–416.

Lynch, M. and J. B. Walsh. 1998. Genetics and Analysis of Quantitative Traits. Sinauer Associates, Sunderland, MA.

Roze, D. 2014. Selection for sex in finite populations. J. Evol. Biol. 27:1304–1322.

Turelli, M. and N. H. Barton. 1990. Dynamics of polygenic characters under selection. Theor. Popul. Biol. 38:1–57.

———. 1994. Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? Genetics 138:913–941.