



**HAL**  
open science

# EPISTASIS, PLEIOTROPY, AND THE MUTATION LOAD IN SEXUAL AND ASEXUAL POPULATIONS

Denis Roze, Alexandre Blanckaert

► **To cite this version:**

Denis Roze, Alexandre Blanckaert. EPISTASIS, PLEIOTROPY, AND THE MUTATION LOAD IN SEXUAL AND ASEXUAL POPULATIONS. *Evolution - International Journal of Organic Evolution*, 2014, 68, pp.137-149. 10.1111/evo.12232 . hal-01137848

**HAL Id: hal-01137848**

**<https://hal.science/hal-01137848>**

Submitted on 31 Mar 2015

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

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

Epistasis, pleiotropy and the mutation load in sexual and asexual  
populations

Denis Roze<sup>1,2</sup> and Alexandre Blanckaert<sup>3</sup>

1- CNRS, UMR 7144, Adaptation et Diversité en Milieu Marin, 29682 Roscoff, France

2- UPMC Université Paris VI, 29682 Roscoff, France

3- Mathematics and Biosciences Group, Department of Mathematics, University of  
Vienna, 1090 Vienna, Austria

*Corresponding author:* Denis Roze

Station Biologique de Roscoff

Place Georges Teissier, CS90074

29688 Roscoff Cedex, France

Ph: (+33) 2 98 29 23 20; fax: (+33) 2 98 29 23 24

roze@sb-roscoff.fr

*Running head:* Fitness landscapes and the mutation load

## ABSTRACT

1  
2 Mutation may impose a substantial load on populations, which varies accord-  
3 ing to the reproductive mode of organisms. Over the last years, different authors used  
4 adaptive landscape models to predict the long term effect of mutation on mean fitness;  
5 however, many of these studies assumed very weak mutation rates, so that at most  
6 one mutation segregates in the population. In this paper we derive several simple  
7 approximations (confirmed by simulations) for the mutation load at high mutation  
8 rate ( $U$ ), using a general model that allows us to play with the number of selected  
9 traits ( $n$ ), the degree of pleiotropy of mutations and the shape of the fitness function  
10 (which affects the average sign and magnitude of epistasis among mutations). When  
11 mutations have strong fitness effects, the equilibrium fitness  $\bar{W}$  of sexuals and asexu-  
12 als is close to  $e^{-U}$ ; under weaker mutational effects, sexuals reach a different regime  
13 where  $\bar{W}$  is a simple function of  $U$  and of a parameter describing the shape of the  
14 fitness function. Contrarily to weak-mutation results showing that  $\bar{W}$  is an increasing  
15 function of population size and a decreasing function of  $n$ , these parameters may have  
16 opposite effects in sexual populations at high mutation rate.

17

18 *Keywords:* adaptive landscape, epistasis, evolutionary quantitative genetics, multi-  
19 locus models, mutation load, stabilizing selection

21 Although mutation represents the ultimate fuel for adaptation, it is also the  
22 source of a fitness cost for populations due to the production of sub-optimal geno-  
23 types. This “mutation load” may in turn affect many important evolutionary pro-  
24 cesses, such as the evolution of sex and recombination (Kondrashov, 1988; Barton,  
25 1995; Keightley and Otto, 2006; Otto, 2009), inbreeding depression and the evolu-  
26 tion of mating systems (Lande and Schemske, 1985; Charlesworth and Charlesworth,  
27 1999; Charlesworth, 2006), mate choice (Rowe and Houle, 1996) or ploidy levels (Otto  
28 and Goldstein, 1992; Otto and Marks, 1996). Different types of models have explored  
29 the effect of recurrent mutation on the average fitness of populations. The simplest  
30 model assumes that each mutation decreases fitness by a fixed factor, independently  
31 of the genetic background (multiplicative model). In that case, the average fitness of a  
32 population at mutation-selection balance (relative to the maximal possible fitness) is  
33 approximately  $e^{-U}$ , where  $U$  is the genomic rate of deleterious mutation (e.g., Crow,  
34 1970). This results holds for both sexual (randomly mating) and asexual populations,  
35 as long as stochastic effects can be neglected. However, it seems unlikely that most  
36 mutations have independent effects: for example, direct measures of fitness effects of  
37 mutations (alone and in combination) in microorganisms usually show a wide distri-  
38 bution of epistatic interactions among pairs of mutations (e.g., Martin et al., 2007).  
39 Kimura and Maruyama (1966) explored the effects of epistasis among deleterious al-  
40 leles on mean fitness, assuming that epistasis is the same for all pairs of mutations.  
41 They showed that when interactions among mutations tend to reinforce their deleteri-  
42 ous effect (negative epistasis) the mean fitness of sexual populations increases, while it

43 decreases when deleterious alleles tend to compensate each other (positive epistasis).  
44 By contrast, the mean fitness of asexual populations is not affected by epistasis and  
45 remains approximately  $e^{-U}$ .

46 In a different class of models, mutations affect a given number of phenotypic  
47 traits which in turn influence fitness. Fisher (1930), Haldane (1932) and Wright (1935)  
48 considered different models of stabilizing selection acting on a single quantitative trait  
49 influenced by many genes. In particular, Haldane and Wright derived expressions for  
50 equilibrium allele frequencies, from which mean fitness can be deduced; further de-  
51 velopments (to which we will return below) were done in the sixties and seventies,  
52 in particular by Kimura (1965), Bulmer (1971, 1972) and Lande (1976, 1977). Inter-  
53 estingly, compensatory effects between mutations emerge naturally from this type of  
54 model (a mutation that displaces from the optimum can be compensated by another  
55 mutation bringing closer to the optimum), together with distributions of epistatic in-  
56 teractions among mutations. In his *Genetical Theory of Natural Selection*, Fisher also  
57 proposed a model involving multiple phenotypic traits in support of his idea that adap-  
58 tation is mainly due to mutations of small effect: this geometrical model represents  
59 an  $n$ -dimensional phenotypic space where each dimension corresponds to the value of  
60 a quantitative trait, and where mutations correspond to random vectors displacing  
61 individuals in phenotypic space. Fitness is assumed to decrease monotonously as the  
62 Euclidean distance from a point corresponding to the optimal phenotype increases.  
63 Although the initial goal of Fisher's geometrical model was to describe the dynamics  
64 of adaptation (Fisher, 1930; Orr, 1998, 2000; Welch and Waxman, 2003), it has also  
65 been used to explore the effects of various parameters (population size, number of  
66 phenotypic traits under selection, shape of the fitness peak...) on the mean fitness

67 of populations in equilibrium situations. Several authors considered a weak-mutation  
68 limit ( $NU \ll 1$ , where  $N$  is population size and  $U$  the total rate of mutations affecting  
69 the traits under selection) so that at most one mutation segregates in the popula-  
70 tion at a given time (Hartl and Taubes, 1998; Poon and Otto, 2000; Sella and Hirsh,  
71 2005; Tenailon et al., 2007; Sella, 2009): the population evolves away from the op-  
72 timum by fixing weakly deleterious alleles by random drift, and occasionally returns  
73 near the optimum by fixing a compensatory mutation. For example, using the fitness  
74 function  $W = \exp(-d^Q)$ , where  $d$  is the Euclidean distance from the optimum in  
75 phenotypic space and  $Q$  a parameter affecting the shape of the fitness peak, Tenail-  
76 lon et al. (2007) showed that the equilibrium mean fitness of a haploid population is  
77 approximately  $(1 - 1/(2N))^{n/Q}$ , where  $n$  is the number of traits affecting fitness (often  
78 called “complexity”). Note that this result does not depend on the reproductive mode  
79 of organisms (sexual or asexual) since at most one locus is polymorphic in this low  
80 mutation limit.

81         It seems likely that in many organisms, however,  $NU \gg 1$ , in which case many  
82 of the loci affecting the traits under selection may be polymorphic at the same time.  
83 In this regime, and assuming that the position of the optimum remains constant over  
84 time, the population can be represented by a collection of points centered around the  
85 optimum, while mean fitness can usually be expressed in terms of the variance of the  
86 different trait values in the population. Since the original works of Haldane (1932)  
87 and Wright (1935), numerous models have been used to calculate the genetic variance  
88 of a quantitative trait (or a set of traits) maintained at mutation-selection balance,  
89 and different regimes have been described (e.g., Bulmer, 1989; Bürger, 2000; Johnson  
90 and Barton, 2005; Zhang and Hill, 2005). When mutations tend to have strong fitness

91 effects, the distribution of phenotypic values is often highly leptokurtic (with a sharp  
92 peak at the optimum) and may present a singularity (Waxman and Peck, 1998). In  
93 the limit of very strong selection, the expected contribution to future generations  
94 of individuals located away from the optimum becomes negligible; in this case, one  
95 expects that mean fitness, relative to the fitness of individuals at the optimum ( $\bar{W}/W_0$ )  
96 should be approximately  $e^{-U}$  when the number of new mutations per individual is  
97 Poisson-distributed, independently of the exact shape of the fitness peak. Indeed,  
98 a recursion for the frequency of individuals at the optimum ( $x_0$ ) is given by  $x_0' =$   
99  $e^{-U}x_0W_0/\bar{W}$  ( $e^{-U}$  being the probability that no mutation occurs), which yields the  
100 desired result when  $x_0' = x_0$  (e.g., Kimura and Maruyama, 1966).

101         When mutations tend to have weak fitness effects, different results can be ob-  
102 tained by assuming that equilibrium distributions of phenotypic values are Gaussian.  
103 Most models exploring this regime considered a single trait under selection and a  
104 Gaussian (or quadratic) fitness function, although multivariate models have also been  
105 proposed (Lande, 1980; Turelli, 1985). Different assumptions regarding the genetic ar-  
106 chitecture of traits (in particular the number of alleles per locus) have been explored.  
107 *Continuum-of-alleles* models assume that an infinite number of alleles (having different  
108 phenotypic effects) segregate at each locus; expressions for the genetic variance and  
109 mean fitness under this scenario have been obtained by assuming that the equilibrium  
110 distribution of allelic effects in the population is Gaussian at each locus (e.g., Kimura,  
111 1965; Lande, 1976). Lande (1980) generalized these models to the case where selection  
112 acts on multiple traits, assuming a multivariate Gaussian fitness function. Although  
113 the general expression for equilibrium genetic variances (and mean fitness) is cumber-  
114 some, it takes a simpler form when selection acts independently on each trait, and in

115 the absence of correlation of mutational effects across traits (p.294 in Bürger, 2000).  
116 Neglecting linkage disequilibria among loci, mean fitness can then be written as a sim-  
117 ple function that depends on the number of selected traits, number of loci, strength  
118 of selection and mutational variance (see Results section below). Other models con-  
119 sidered stabilizing selection (Gaussian fitness function) acting on a single trait coded  
120 by multiple biallelic loci (e.g., Bulmer, 1972, 1985; Barton, 1986, 1989). As shown by  
121 Barton (1986), multiple possible equilibria exist for allele frequencies at the different  
122 loci. Although the population mean phenotype may deviate from the optimum at  
123 some of these equilibria, Barton (1989) argues that perturbations generated for exam-  
124 ple by random drift should keep the population mean near the optimum. Neglecting  
125 linkage disequilibria, and as long as the effects of mutation and drift at each locus are  
126 weaker than selection, one obtains from the expression of the equilibrium genetic vari-  
127 ance that mean fitness should be close to  $e^{-U}$ . More recently, Zhang and Hill (2003)  
128 investigated the maintenance of genetic variation in multiple-trait models (including  
129 correlations among traits), assuming that the fitness function is multivariate Gaussian,  
130 and that selection at each locus is sufficiently strong so that mutations remain rare in  
131 the population (only one mutant allele segregates at each locus). Results for the case  
132 of a changing optimum have been derived by Zhang (2012). Finally, effects of linkage  
133 disequilibria have been explored in both continuum-of-alleles and biallelic models, for  
134 the case of a single trait under Gaussian stabilizing selection (Bulmer, 1974; Lande,  
135 1976; Turelli and Barton, 1990), and were shown to be minor as long as the total rate  
136 of mutation on loci affecting the trait remains small, and linkage is not too tight.

137 In this article, we derive several simple approximations for the mutation load in  
138 both sexual and asexual populations at high mutation rate, using a general model that



139 allows us to play with the shape of the fitness peak (which in turn affects the average  
140 sign and magnitude of epistasis among mutations), the number of traits under selection  
141 and the degree of pleiotropy of mutations. These results are tested by individual-  
142 based simulations representing stabilizing selection on traits coded by large numbers  
143 of loci. As we will see, two different regimes are observed in asexual populations: when  
144 mutations tend to have large fitness effects, mean fitness only depends on the mutation  
145 rate and is close to  $e^{-U}$ , while under weaker mutational effects mean fitness increases  
146 as mutational effects decreases, and can be expressed in terms of the mutation rate,  
147 the dimensionality of the landscape, the shape of the fitness peak and the average  
148 deleterious effect of mutations. By contrast, three regimes are observed in the case  
149 of sexual populations: under strong mutational effects, mean fitness is again close to  
150  $e^{-U}$ , while it switches to a different value that only depends on the mutation rate  
151 and the shape of the fitness peak as mutational effects decrease. Interestingly, mean  
152 fitness is the same under these two regimes when the fitness function is Gaussian  
153 ( $\bar{W} \approx e^{-U}$ ), while  $\bar{W} > e^{-U}$  in the second regime when the fitness peak is flatter  
154 than a Gaussian around the optimum, and  $\bar{W} < e^{-U}$  when the fitness peak is sharper.  
155 Finally, as mutational effects continue decreasing a third regime is entered, where allele  
156 frequency dynamics are dominated by mutation and drift. We also obtain expressions  
157 for linkage disequilibria and find that they can have a substantial effect on fitness, in  
158 particular when the mutation rate is high and the number of traits under selection is  
159 low.

161 **Multivariate stabilizing selection.** Parameters and variables are summarized in  
 162 Table 1. The model represents an infinite haploid population with discrete generations,  
 163 under stabilizing selection on  $n$  phenotypic traits; for each individual, trait values are  
 164 given by the vector  $\mathbf{z} = (z_0, z_1, \dots, z_n)$ . Following others (Wilke and Adami, 2001;  
 165 Tenaillon et al., 2007; Gros et al., 2009), we assume for simplicity that the fitness  
 166 function is spherically symmetric, and is given by:

$$W = \exp \left[ -\frac{d^Q}{2V_S} \right] \quad (1)$$

167 where  $d = \sqrt{\sum_{i=1}^n z_i^2}$  is the Euclidean distance in phenotypic space from the optimal  
 168 phenotype (located at  $z_i = 0$  for each trait  $i$ ),  $V_S$  measures the strength of selection,  
 169 while the parameter  $Q$  determines the shape of the fitness peak:  $Q = 2$  corresponds to  
 170 a (multivariate) Gaussian fitness function, while  $Q < 2$  (resp.  $Q > 2$ ) leads to a sharper  
 171 (resp. flatter) fitness peak. Traits are coded by  $L$  biallelic loci, each locus affecting  
 172 a subset  $m \leq n$  of phenotypic traits (sampled randomly and independently for each  
 173 locus among the set of  $n$  traits). As in Lourenço et al. (2011), the parameter  $m$  thus  
 174 measures the number of traits affected by a single mutation (“mutation pleiotropy”),  
 175 while  $n$  is the total number of traits under selection (“complexity”):  $m = n$  corresponds  
 176 to maximal pleiotropy, where each mutation affects all selected traits (as in Fisher’s  
 177 model). The genomic mutation rate is denoted  $U$ , so that each locus mutates at rate  
 178  $u = U/L$ . For simplicity, we assume additive effects of the different loci on phenotypic  
 179 traits, and no environmental variance. Denoting 0 and 1 the two alleles at each locus,  
 180 and  $X_j$  an indicator variable that equals 0 if an individual carries allele 0 at locus  $j$ ,

181 and 1 otherwise, the value of trait  $i$  in a given individual is given by:

$$z_i = \sum_{j=1}^L r_{ij} X_j \quad (2)$$

182 where  $r_{ij}$  measures the effect of allele 1 at locus  $j$  on trait  $i$ : an individual carrying  
183 alleles 0 at all loci is thus at the phenotypic optimum. For each locus  $j$ , the effect  
184 of allele 1 on each of the  $m$  traits affected by the locus is sampled in a Gaussian  
185 distribution with mean 0 and variance  $a^2$  (therefore,  $r_{ij}$  can be positive or negative).  
186 Finally, we denote  $\bar{s}$  the mean deleterious effect of mutations (measured on log-fitness)  
187 in a population at the optimum, given by Gros et al. (2009):

$$\bar{s} = \frac{(2a^2)^{\frac{Q}{2}} \Gamma\left(\frac{m+Q}{2}\right)}{2V_S \Gamma\left(\frac{m}{2}\right)} \quad (3)$$

188 (where  $\Gamma$  is Euler's gamma function), which simplifies to  $\bar{s} = \frac{m}{2} (a^2/V_S)$  in the case of  
189 a Gaussian fitness function ( $Q = 2$ , in agreement with Martin and Lenormand, 2006).  
190 As shown by Gros et al. (2009), the average epistasis among pairs of mutations (at the  
191 optimum, and measured on log-fitness) is given by  $\bar{e} = (2 - 2^{Q/2}) \bar{s}$ , which equals zero  
192 when  $Q = 2$  (see also ref. Martin et al., 2007) but it positive for  $Q < 2$ , and negative  
193 for  $Q > 2$ .

194 Finally, we can note that the strength of selection  $V_S$  (equation 1) can be con-  
195 sidered as a scaling parameter: using the scaled variables  $z_{i,s} = z_i / (2V_S)^{\frac{1}{Q}}$  to measure  
196 phenotypic traits, fitness in terms of  $z_{i,s}$  variables becomes independent of  $V_S$ , while the  
197 variance of the distribution of mutational effects on each  $z_{i,s}$  equals  $a_s^2 = a^2 / (2V_S)^{\frac{2}{Q}}$ :  
198 therefore,  $a^2$  and  $V_S$  affect the results only through the compound parameter  $a_s^2$ . In  
199 the simulations (described next), we used  $V_S = 1/2$  so that  $a_s^2 = a^2$ . As shown by  
200 equation 3, large (resp. small) values of  $a_s^2$  imply that mutations have strong (resp.  
201 weak) fitness effects.

202 **Simulations.** Analytical predictions are tested using individual-based simulations.  
203 Our program (written in C++, and available upon request) represents a haploid pop-  
204 ulation of  $N$  individuals with discrete generations. The genome of each individual is  
205 represented by a sequence of  $L$  bits (0 or 1) corresponding to the different loci. Phe-  
206 notypic effects of mutation at each locus are sampled at the start of the simulation:  
207 the  $m$  traits affected by a given locus are sampled randomly (and independently for  
208 each locus) among the  $n$  selected traits, and the effect of mutation on each of these  $m$   
209 traits is sampled from a Gaussian distribution with mean 0 and variance  $a^2$ . At the  
210 start of a generation, phenotypic values are computed for each individual; from this,  
211 one obtains the fitness of the individual based on equation 1. For each individual of  
212 the next generation, two parents are sampled (the probability that an individual is  
213 sampled being proportional to its fitness); selfing is allowed if the same individual is  
214 sampled twice. To generate a recombinant chromosome, the number of cross-overs is  
215 sampled from a Poisson distribution with parameter  $R$  (genome map length), and the  
216 position of each cross-over along the chromosome is sampled from a uniform distribu-  
217 tion ( $R = 0$  corresponds to the case of an asexual population). Finally, the number  
218 of new mutations occurring in each genome is sampled from a Poisson distribution  
219 with parameter  $U$ , and the positions of mutant loci are sampled randomly; alleles are  
220 switched at mutant loci, from 0 to 1 or from 1 to 0 (mutation and back mutation thus  
221 occur at the same rate). In order to gain execution speed, the program is parallelized  
222 (using the MPI library) to run on several processors, each processor dealing with a  
223 given segment of the genome, for all individuals; execution speed can be considerably  
224 increased when the number of processors is sufficiently large so that the probability  
225 that an event (mutation or cross-over) occurs in a given segment per generation is

226 low (in which case most genome segments stay unchanged from one generation to the  
227 next).

228 At the start of the simulation each genome contains only “0” alleles, which cor-  
229 responds to the fitness optimum. Simulations run for 50000 generations, equilibrium  
230 being reached during the first 30000 generations for most parameter values tested.  
231 Every 1000 generations, the program records the mean fitness of the population and  
232 the first 6 moments of the distribution of each phenotypic trait in the population.  
233 The results shown in the different figures correspond to averages over the last 20000  
234 generations. Error bars were calculated using Hastings’ (1970) batching method, di-  
235 viding these 20000 generations into 4 batches of 5000 generations and calculating the  
236 standard error over these 4 batches; however, error bars were generally small relative  
237 to the size of symbols used in the figures and are thus not shown. All our simulation  
238 results (with the *Mathematica* commands used to generate the figures) can be found  
239 in the Supplementary Material.

240

## RESULTS

241 **Asexual population.** Figure 1 shows the equilibrium mean fitness of an asexual  
242 population as a function of the variance of mutational effects (on scaled traits)  $a_s^2$   
243 (left) or of the mean deleterious effect of mutations  $\bar{s}$  (right), and for different values  
244 of the shape parameter  $Q$ . As  $a_s^2$  becomes large, mean fitness converges to  $e^{-U}$  (which  
245 is close to 0.6 for  $U = 0.5$ , dashed line). For smaller values of the scaled mutational  
246 variance, however, the population reaches a different regime where  $\bar{W} > e^{-U}$ . In this  
247 case, an approximation for mean fitness can be obtained by assuming that distributions

248 of phenotypic traits  $z_i$  are Gaussian at equilibrium. As explained in Supplementary  
 249 File S1, Euclidean distances from the optimum ( $d$ ) then follow a  $\chi$ -distribution, which  
 250 yields:

$$\overline{\ln W} = -\frac{(2V_G)^{\frac{Q}{2}}}{2V_S} \frac{\Gamma\left(\frac{n+Q}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \quad (4)$$

251 where  $V_G$  is the genetic variance at equilibrium (the variance of  $z_i$ , which by symme-  
 252 try should be the same for all traits  $i$ ); for  $Q = 2$  (Gaussian fitness function), this  
 253 simplifies to  $\overline{\ln W} = -(n/2)(V_G/V_S)$ . An expression for  $V_G$  at equilibrium is derived  
 254 in Supplementary File S1. Assuming that  $\overline{\ln W} \approx \ln \bar{W}$ , one obtains that when  $m = n$   
 255 (that is, when each mutation affects all phenotypic traits):

$$\bar{W} \approx \exp \left[ -\bar{s} \left( \frac{nU}{\bar{s}Q} \right)^{\frac{Q}{2+Q}} \right] \quad (5)$$

256 When  $Q = 2$ , equation 5 simplifies to  $\exp \left[ -\sqrt{\frac{n}{2}} U \bar{s} \right]$ , in agreement with the result  
 257 obtained from Lande's analysis (1980) when selection acts independently on the dif-  
 258 ferent traits, and in the absence of mutational covariances among traits (that is, when  
 259 the variance-covariance matrices representing the effects of selection and mutation on  
 260 the traits are diagonal). Indeed, in an asexual population the whole genome can be  
 261 considered as a single locus with many alleles (provided that the number of loci is  
 262 sufficiently large), in which case Lande's analysis of the continuum-of-alleles model  
 263 yields  $V_G = \sqrt{U a^2 V_S}$  when mutation and selection covariance matrices are diagonal  
 264 (the equilibrium genetic variance for each trait is not affected by the other traits under  
 265 selection, see also p. 294 in ref. Bürger, 2000). When  $m < n$ , equation 5 still holds  
 266 when the fitness function is Gaussian ( $Q = 2$ ), while  $\bar{W}$  is given by a slightly more  
 267 complicated expression when  $Q \neq 2$  (Supplementary File S1). However, this expres-  
 268 sion only depends weakly on  $m$  (for fixed  $\bar{s}$ ) and remains close to equation 5 in most

269 cases (unless  $Q$  is high and  $m$  is small). As shown by Figure 1, equation 5 provides  
 270 accurate predictions for  $\bar{W}$  as long as  $a_s^2$  is not too high. Supplementary Figure S1  
 271 shows the genetic variance  $V_G$  at equilibrium, for the same parameter values as in  
 272 Figure 1.

273

274 **Sexual population.** Figure 2 shows the equilibrium mean fitness of a sexual popu-  
 275 lation as a function of  $a_s^2$  and  $\bar{s}$ . When mutations have strong effects on fitness, one  
 276 obtains again that  $\bar{W} \approx e^{-U}$  (right-most points on Figure 2). As  $a_s^2$  decreases, a sec-  
 277 ond regime is entered where  $\bar{W}$  reaches a new value which is independent of  $a_s^2$ , but  
 278 depends on the shape parameter  $Q$  (roughly for  $-4 < \log_{10}(a_s^2) < -2$ ). Finally, as  $a_s^2$   
 279 continues decreasing a third regime is entered where  $\bar{W}$  increases up to  $\bar{W} \approx 1$  as  $a_s^2$   
 280 decreases. In Supplementary File S2, we show that approximations for mean fitness  
 281 under these last two regimes ( $a_s^2$  small) can be obtained by assuming that distributions  
 282 of phenotypic values in the population are Gaussian. As in previous models (e.g., Bul-  
 283 mer, 1972; Barton, 1986) one obtains that the equilibrium frequency of allele 1 at locus  
 284  $j$  ( $p_j$ ) is either (i) at  $p_j = 1/2$  (when selection at locus  $j$  is weak relative to mutation)  
 285 or (ii) at one of the two symmetric equilibria where  $p_j(1 - p_j) = X < 1/4$  (where  $X$  is  
 286 a function of the model's parameters), the locus being closer to fixation (for allele 0 or  
 287 1) as selection increases. Assuming that all loci are at equilibrium (ii) and neglecting  
 288 linkage disequilibria among loci, one obtains the following approximation for mean  
 289 fitness:

$$\bar{W} \approx \exp \left[ -\frac{2U}{Q} \right] \quad (6)$$

290 Figure 2 shows that this expression does indeed provide a correct prediction of  $\bar{W}$  for  
 291 intermediate values of  $a_s^2$  (dashed lines). A better approximation can be obtained by

292 taking into account effects of the linkage disequilibria, which is done in Supplementary  
 293 File S2 using the methods developed by Turelli and Barton (1990). As described before  
 294 (Bulmer, 1971, 1974; Lande, 1976; Turelli and Barton, 1990), stabilizing selection tends  
 295 to generate negative covariances among loci, which affect  $\bar{W}$  through two effects: a  
 296 reduction of the genetic variance  $V_G$  which directly increases  $\bar{W}$ , and an effect on  
 297 equilibrium allele frequencies and thus on the *genic* variance  $V_g$  (see Appendix A)  
 298 which has the opposite effect on  $\bar{W}$ . As long as linkage disequilibria remain small  
 299 the second effect predominates, and one obtains the following approximation for  $\bar{W}$   
 300 (assuming that the number of loci  $L$  is large):

$$\bar{W} \approx \exp \left[ -\frac{2U}{Q} \left( 1 + \frac{2U}{n} \left( \frac{1}{r_H} - 1 \right) \right) \right] \quad (7)$$

301 where  $r_H$  is the harmonic mean recombination rate between all pairs of loci (derived in  
 302 Appendix B under our simulated genetic architecture). Figure 2 shows that equation  
 303 7 does indeed provide a slightly better approximation than equation 6 for intermediate  
 304 values of  $a_s^2$  (horizontal solid lines). Interestingly, Supplementary Figure S1 shows that  
 305  $Q$  has only a weak effect on the equilibrium genetic variance  $V_G$  in this intermediate  
 306 regime, which remains close to the value obtained for  $Q = 2$ , neglecting linkage dise-  
 307 quilibria:  $V_G = 2UV_S/n$  (which is also obtained from *house-of-cards* models assuming  
 308 a Gaussian fitness function, e.g. Turelli, 1985; Bürger, 2000). Finally, when  $a_s^2$  is very  
 309 small (so that mutations have very weak fitness effects) one predicts that most loci  
 310 should be at the equilibrium where  $p_j = 1/2$ , in which case  $\bar{W} \approx \exp \left[ -\bar{s} (L/4)^{\frac{Q}{2}} \right]$   
 311 where  $L$  is the total number of loci. However, in this regime we expect that the  
 312 relative effect of genetic drift on allele frequency dynamics could be important. In  
 313 Supplementary File S4, we show that when selection at each locus becomes weaker



314 than mutation and drift, an approximation for mean fitness is given by:

$$\bar{W} \approx \exp \left[ -\bar{s} \left( \frac{NU}{1 + 4Nu} \right)^{\frac{Q}{2}} \right] \quad (8)$$

315 where  $N$  is population size. Figure 2 shows that equation 8 provides accurate pre-  
316 dictions when  $a_s^2$  is very small (left part of the figures). Furthermore, combining  
317 equations 7 and 8 (solid curves) gives good predictions for all values of  $a_s^2$  as long  
318 as they remain small (roughly,  $< 10^{-2}$  for the parameter values of Figure 2) so that  
319 the Gaussian approximation holds. For higher values of  $a_s^2$ , simulations indicate that  
320 distributions of phenotypic values depart from Gaussian distributions: in particular,  
321 the fourth and sixth cumulants  $K_4$  and  $K_6$  are positive and increase as  $a^2$  increases  
322 (while the third and fifth cumulants stay close to zero), for all values of  $Q$  (results not  
323 shown). In Supplementary File S3, we show that using a Gram-Charlier expansion to  
324 approximate the distribution of phenotypic effects yields an expression for  $\bar{W}$  in terms  
325 of the genetic variance  $V_G$  and of higher cumulants ( $K_4$ ,  $K_6$ ). Furthermore, these  
326 higher cumulants can in turn be expressed in terms of the genetic variance, assuming  
327 that  $p_j q_j$  is small at each locus (*rare-alleles approximation*, e.g. Barton and Turelli,  
328 1987; Turelli and Barton, 1990): one obtains in particular  $K_4 \approx a^2 V_G [3m / (2 + m)]$ ,  
329  $K_6 \approx a^4 V_G [15m / (4 + m)]$ , which fits well with the simulation results (see Figure 1 in  
330 Supplementary File S3). Finally, the methods developed by Turelli and Barton (1994)  
331 can be used to calculate  $V_G$  at equilibrium in this non-Gaussian regime. As shown in  
332 Supplementary File S3, the result (which has to be obtained numerically) fits well with  
333 the simulation results as long as  $a_s^2$  is not too large, and  $Q \geq 2$ ; however, we could not  
334 find any simple analytical expression for  $\bar{W}$  in this non-Gaussian regime. From the  
335 expressions of  $K_4$  and  $K_6$  above, one predicts that the Gaussian regime should be left

336 more rapidly (as  $a^2$  increases) for larger  $m$ , since  $K_4$  and  $K_6$  are increasing functions  
 337 of  $m$ ; Figure 3 (left) confirms that this is indeed the case. However, Figure 3 right  
 338 indicates that for a fixed  $\bar{s}$ ,  $\bar{W}$  is relatively insensitive to  $m$ .

339 Figure 4 explores the effect of linkage on mean fitness: for  $U = 0.5$  and  $n = 50$ ,  
 340 the effect of linkage disequilibria on  $\bar{W}$  remains small as long as the harmonic mean  
 341 recombination rate  $r_H$  is not too small (with  $L = 10000$  loci,  $r_H \approx 0.053$  for a map  
 342 length of 1 Morgan, while  $r_H \approx 0.3$  for a map length of 10 Morgans — see Appendix  
 343 B). In this regime, increasing linkage tends to reduce mean fitness (through an in-  
 344 crease in genic variance  $V_g$ ), and equation 7 matches well with the simulation results  
 345 (except for  $Q = 1$  where the fit is less good). For tighter linkage, however (roughly,  
 346  $R < 1$ ) simulations results depart from the prediction from equation 7 (which assumes  
 347 that the contribution of linkage disequilibria to the genetic variance remains small)  
 348 and show that  $\bar{W}$  increases as recombination decreases: overall, linkage thus has a  
 349 non-monotonic effect on mean fitness. Interestingly, equation 7 indicates that the con-  
 350 tribution of linkage disequilibria becomes more important when the number of traits  
 351 under selection ( $n$ ) is small. Figure 5 confirms this result, showing that lower values  
 352 of  $n$  lead to a stronger decrease in  $\bar{W}$  for both  $\bar{s} = 0.01$  and  $\bar{s} = 0.1$ , while the effect  
 353 is less marked (and may even be reversed) when  $\bar{s} = 0.001$ . This effect of  $n$  can be  
 354 understood as follows: mean fitness decreases when the sum of genetic variances for  
 355 the different traits  $\sum_{i=1}^n V_{G,i}$  increases. Due to the symmetry of our model, the two  
 356 components of the genetic variance (the genic variance  $V_{g,i}$ , and the contribution of  
 357 linkage disequilibria  $d_i$ , see Appendix A) should be the same for all traits. As shown  
 358 in Supplementary File S2 (see also Turelli, 1985; Bürger, 2000), when the effects of  
 359 linkage disequilibria are neglected the genetic variance for each trait at equilibrium is

360 proportional to  $1/n$  (roughly,  $p_j q_j$  at each locus scales with  $1/m$ , since the strength  
361 of selection at each locus increases with  $m$ ; however, a proportion  $m/n$  only of loci  
362 contribute to trait  $i$ : therefore, the effect of  $m$  cancels and  $V_{g,i}$  scales with  $1/n$ ). As  
363 a consequence  $\overline{W}$  (which depends on the product  $nV_G$ ) is independent of  $n$ . The  
364 contribution of associations between loci depends on all pairwise linkage disequilibria  
365 and is proportional to  $V_g^2$  (see Supplementary File S2) and therefore to  $1/n^2$ . As a  
366 consequence, the overall contribution of linkage disequilibria to mean fitness is pro-  
367 portional to  $1/n$ . However, pairwise linkage disequilibria may not be the only cause  
368 of the strong reduction of  $\overline{W}$  shown on Figure 5 at low values of  $n$  (for  $\bar{s} = 0.01$  and  
369 0.1): indeed, simulations indicate that for these low values of  $n$ , distributions of phe-  
370 notypic traits in the population depart from Gaussian distributions (fourth and sixth  
371 cumulants become significantly different from zero), in which case the results derived  
372 in Supplementary File S2 (which assume Gaussian distributions of phenotypic traits)  
373 do not hold.

374

375 **Effects of population size and number of loci.** Many of the results shown above  
376 assume a large population size and large number of loci. Figure 6 explores the effects  
377 of varying population size: in the case of a sexual population (Figure 6A), smaller  
378 population sizes tend to decrease  $\overline{W}$  when the average fitness effect of mutations is  
379 sufficiently strong (right part of the figure). However, for lower values of  $\bar{s}$  one observes  
380 the opposite effect, with higher values of mean fitness for lower population sizes. With  
381 a Gaussian fitness function ( $Q = 2$ , as in Figure 6), a diffusion model can be used to  
382 predict the equilibrium genetic variance under selection, mutation and drift, assuming  
383 that mean phenotypes stay at the optimum (Bulmer, 1972). In Supplementary File S4,

384 we obtain an approximation for mean fitness (in terms of hypergeometric functions)  
 385 valid for  $Nu \ll 1$ , which fits well with simulation results for  $N = 500$  (dotted curve in  
 386 Figure 6A). In contrast to the sexual case, the effects of population size are less marked  
 387 in the case of asexual organisms (Figure 6B), and  $\overline{W}$  always decreases as  $N$  decreases.  
 388 These effects of finite population size can be understood as follows. Genetic drift affects  
 389 mean fitness through two different effects: an effect on the genetic variance maintained  
 390 at equilibrium, which is usually (but not always) lower in smaller populations, and an  
 391 effect on mean phenotypes, which may be displaced from the optimum due to drift.  
 392 In the case of a Gaussian fitness function ( $Q = 2$ ), these two effects can easily be  
 393 separated: indeed, log-fitness equals  $-\sum_{i=1}^n (z_i - z_i^*)^2 / (2V_S)$ , where  $z_i$  is the value of  
 394 phenotype  $i$  and  $z_i^*$  the optimal value for this phenotype (fixed to zero in our model).  
 395 Therefore, the average log-fitness is given by  $-\sum_{i=1}^n (V_{G,i} + (\overline{z}_i - z_i^*)^2) / (2V_S)$ , where  
 396  $V_{G,i}$  is the genetic variance for trait  $i$  and  $\overline{z}_i$  the average value of trait  $i$  (see also  
 397 Charlesworth, 2013). Assuming that  $\overline{\ln W} \approx \ln \overline{W}$ , mean fitness can thus be written  
 398 as a product of two terms:

$$\overline{W} \approx \exp \left[ -\frac{\sum_{i=1}^n V_{G,i}}{2V_S} \right] \exp \left[ -\frac{\sum_{i=1}^n (\overline{z}_i - z_i^*)^2}{2V_S} \right]. \quad (9)$$

399 The first term (denoted hereafter  $\overline{W}_{V_G}$ ) shows that increasing the variance of pheno-  
 400 typic traits tends to decrease mean fitness, while the second term (denoted hereafter  
 401  $\overline{W}_{\overline{z}}$ ) shows that displacing mean phenotypes from the optimum also decreases  $\overline{W}$ .  
 402 Simulations show that injecting equilibrium values of genetic variances  $V_{G,i}$  and mean  
 403 phenotypes  $\overline{z}_i$  (measured in the simulations) into equation 9 accurately predict  $\overline{W}$  for  
 404 all parameters tried with  $Q = 2$ , as long as  $\log_{10}(\overline{s}) \leq -1$  (results not shown). Note  
 405 that all the mathematical derivations performed in the different Supplementary Files

406 assume that mean phenotypes are at the optimum, and therefore that  $\overline{W}_{\bar{z}} = 1$ . Sup-  
 407 plementary Figure S2 shows measures of  $\overline{W}_{V_G}$  and  $\overline{W}_{\bar{z}}$  for the different values of  $N$   
 408 and  $\bar{s}$  explored in Figure 6. These results can be summarized as follows: in a sexual  
 409 population (Figure 6A),  $\overline{W}_{\bar{z}}$  stays close to 1 in most cases, but may be slightly less  
 410 than 1 (around 0.97 - 0.98) for  $N = 500$ : mean phenotypes thus generally stay close  
 411 to their optimal values, and the mutation load is mainly generated by the term  $\overline{W}_{V_G}$ .  
 412 When mutations tend to have strong fitness effects ( $\log_{10}(\bar{s}) \geq -2$ ), drift tends to  
 413 increase genetic variances  $V_{G,i}$  (by increasing  $\overline{p_j q_j}$  at each locus  $j$  above its value at  
 414 mutation-selection balance), and thus decreases  $\overline{W}_{V_G}$  — note that this effect is taken  
 415 into account by the diffusion model (equation 6 in Supplementary File S4). When  
 416 mutations tend to have weak fitness effects ( $\log_{10}(\bar{s}) \leq -3$ ) however, drift tends to  
 417 reduce  $V_{G,i}$  by bringing the different loci closer to fixation, thereby increasing  $\overline{W}_{V_G}$   
 418 (and hence  $\overline{W}$ ). By contrast, in an asexual population drift always reduce the genetic  
 419 variance ( $\overline{W}_{V_G}$  increases as  $N$  decreases), but also has a substantial effect on mean  
 420 phenotypes ( $\overline{W}_{\bar{z}}$  decreases as  $N$  decreases). Because the second effect is stronger than  
 421 the first, the overall effect of drift is to reduce the mean fitness of asexuals.

422 Our model is somehow artificial, however, as we assumed that the fitness opti-  
 423 mum is located at the origin ( $z_i^* = 0$  for all  $i$ ). Due to the symmetry of the model, one  
 424 expects that mean phenotypes should be at the origin (and thus at the optimum) in  
 425 the regime where dynamics are mainly driven by mutation and drift (low  $\bar{s}$ ), when the  
 426 number of loci is large: for example if a large, random proportion of loci are fixed for  
 427 allele 1 and the other loci fixed for allele 0, mean phenotypes should be close to zero  
 428 since the sum of all mutational effects (corresponding to the effects of fixed alleles 1)  
 429 converges to the average effect of mutations, which is zero. It seems artificial, however,

430 to assume that optimal trait values are precisely the values towards which the pop-  
431 ulation should converge at mutation-drift equilibrium, and it seems more likely that  
432 in many cases, mean phenotypes at mutation-drift equilibrium should be far from the  
433 optimal values (decreasing the value of  $\overline{W}_{\bar{z}}$  in the mutation-drift regime). In order to  
434 explore that, we modified our simulation program so that the fitness optimum is not  
435 located at the origin of the phenotypic space ( $z_i = 0$  for all  $i$ ). To do this, we select a  
436 genotype sufficiently distant from the origin (in phenotypic space) so that when setting  
437 optimal phenotypic values  $z_i^*$  to the values coded by this genotype, fitness at the origin  
438 is less than 0.1. At the start of the simulation, the population is fixed for the genotype  
439 corresponding to the fitness optimum. At mutation-drift balance, the population is  
440 still expected to converge to the origin (due to the symmetry of the distribution of  
441 mutational effects), where fitness is lower than 0.1. Results are shown on Figures 6C  
442 and 6D in the case of a sexual and asexual population, respectively. As can be seen  
443 on these Figures, the position of the fitness optimum has little effect as long as selec-  
444 tion is sufficiently strong ( $\log_{10}(\bar{s}) \geq -3$ ), since in this case selection maintains the  
445 population near the optimum. For lower  $\bar{s}$ , however, mutation and drift tend to bring  
446 populations away from the optimum (and closer to the origin), which decreases  $\overline{W}_{\bar{z}}$   
447 and thus decreases mean fitness (see also Supplementary Figure S2), this effect being  
448 more pronounced in asexual than in sexual populations (in sexuals, this effect becomes  
449 important only for  $\bar{s} = 10^{-5}$ ). Note that we could not obtain results for  $\bar{s} = 10^{-6}$ ,  
450 because with such low mutational effects (and with 10000 loci) it was not possible to  
451 find a genotype sufficiently distant from the origin so that fitness at the origin is lower  
452 than 0.1 when setting the optimum to the phenotypic values coded by this genotype.  
453 Interestingly, in the sexual case (Figure 6C) one still observes an intermediate range

454 of  $\bar{s}$  ( $\bar{s} = 10^{-4}, 10^{-3}$ ) where decreasing population size increases mean fitness: in this  
455 regime, drift tends to bring each locus closer to fixation (thus reducing  $V_{G,i}$ ), but se-  
456 lection still maintains mean phenotypes  $\bar{z}_i$  close to the optimum. Decreasing  $\bar{s}$  (or  
457 decreasing  $N$  further, as shown in Supplementary Figure S2 for  $N = 100$ ) generates  
458 departures of mean phenotypes from the optimum, decreasing  $\overline{W}_{\bar{z}}$  (and thus  $\overline{W}$ ).

459 Finally, Supplementary Figure S3 shows that changing the number of loci  $L$   
460 (for a fixed  $U$ ) has generally stronger effects in the case of sexual populations, where  
461 in the low  $\bar{s}$  regime, decreasing the number of loci decreases the genetic variance (and  
462 thus increases  $\overline{W}$ ).

## 463 DISCUSSION

464 As we have seen in introduction, different forms of models have been proposed  
465 to predict the overall effect of recurrent mutation on the mean fitness of populations.  
466 Typically, models representing deleterious mutations (without including a phenotypic  
467 dimension) do not consider possible compensatory effects among mutations, and as-  
468 sume that epistasis is the same for all pairs of mutations (Kimura and Maruyama,  
469 1966; Charlesworth, 1990). Using this type of model, Kimura and Maruyama (1966)  
470 predicted that epistasis should have no effect on the mean fitness of asexuals (which  
471 should remain close to  $e^{-U}$ ), while it should increase the fitness of sexuals if it is  
472 negative, and decrease  $\overline{W}$  if it is positive. On the other hand, models representing sta-  
473 bilizing selection on a set of quantitative phenotypic traits (such as Fisher's geometric  
474 model) provide a natural way of incorporating compensatory effects among mutations  
475 and distributions of selection coefficients and epistatic effects. In this paper, we used a

476 general model that allows us to play with the number of traits under selection ( $n$ ), the  
477 number of traits affected by a given mutation ( $m$ ) and the shape of the fitness peak  
478 ( $Q$ ). Interestingly, this last parameter was already shown by others to affect the av-  
479 erage sign of epistasis among pairs of mutations: epistasis is positive on average when  
480 the fitness peak is sharper than a Gaussian function ( $Q < 2$ ), while it is negative when  
481 the fitness peak is flatter than a Gaussian ( $Q > 2$ ), and equal to zero when fitness is  
482 exactly Gaussian (Martin et al., 2007; Gros et al., 2009). Furthermore, the average  
483 deleterious effect of mutations at the optimum  $\bar{s}$  can be expressed as a simple function  
484 of the variance of scaled phenotypic effects of mutations  $a_s^2$  and the parameters  $m$  and  
485  $Q$  (equation 3).

486 In the case of an asexual population, we found that mean fitness stays indeed  
487 close to  $e^{-U}$  when mutations tend to have strong fitness effects; however,  $\bar{W} > e^{-U}$   
488 under weaker fitness effects of mutations (Gaussian regime), due to the possibility of  
489 compensatory effects among mutations at different loci. For a fixed  $\bar{s}$ , mean fitness is  
490 higher under positive epistasis ( $Q < 2$ ) than under negative epistasis ( $Q > 2$ ); this is  
491 probably due to the fact that selection for compensatory mutations is stronger when  
492 the fitness peak is sharp, keeping the population closer to the fitness optimum. In  
493 contrast, under negative epistasis the fitness landscape presents a plateau around the  
494 optimum, and at equilibrium many genotypes are located near the edge of this plateau,  
495 where the effect of deleterious mutations tends to be strong. Finally, we showed that  
496 under this regime mean fitness can be expressed as a simple function of  $U$ ,  $\bar{s}$ ,  $n$  and  $Q$   
497 (equation 5), which indicates that  $\bar{W}$  decreases as the number of traits under selection  
498  $n$  increases.

499 In a sexual population, mean fitness is also at  $e^{-U}$  when mutations have very



500 strong fitness effects: this is expected, as mutations stay at very low frequency and  
501 epistatic interactions among mutations have thus little effect. For weaker mutational  
502 effects, Kimura and Maruyama (1966)'s prediction is verified: positive epistasis de-  
503 creases  $\overline{W}$ , while negative epistasis increases it; furthermore, we showed that  $\overline{W}$  con-  
504 verges to a very simple expression ( $e^{-\frac{2U}{Q}}$ ) when the phenotypic effect of mutations  
505 is sufficiently weak, so that distributions of phenotypic values in the population are  
506 nearly Gaussian. Although this expression assumes that linkage disequilibria among  
507 loci can be neglected, effects of genetic associations can be computed, showing that  
508 linkage disequilibria tend to reduce  $\overline{W}$  as long as they remain weak, the importance of  
509 this effect increasing as the number of selected traits  $n$  decreases. Contrarily to results  
510 obtained previously assuming  $NU \ll 1$  (Hartl and Taubes, 1998; Poon and Otto, 2000;  
511 Tenaillon et al., 2007; Lourenço et al., 2011), our model predicts a higher mean fitness  
512 when the number of traits is large, through this effect on linkage disequilibrium. Figure  
513 5 indeed shows important reductions in fitness when the number of traits decreases  
514 (for moderate to strong fitness effects of mutations, of the order 0.01 to 0.1), which  
515 may also involve higher-order genetic associations (since distributions of phenotypic  
516 values depart from Gaussian distributions for very low values of  $n$ ). Rather than the  
517 “cost of complexity” often described at low mutation rate, we thus predict a “cost of  
518 simplicity” when the overall mutation rate is large (as shown by equation 7, this effect  
519 is proportional to  $U^2$  and should thus become negligible when the overall mutation  
520 rate is small). Note that this cost does not depend on the degree of pleiotropy ( $m$ ) of  
521 individual mutations, but only on the total number of traits under selection. Interest-  
522 ingly, these results do not depend on the precise shape of the distribution of mutational  
523 effects (which was Gaussian in the simulations), since the derivation of equations 6 and

524 7 in Supplementary File S2 makes no assumption on this distribution. It does assume  
525 uncorrelated mutational effects on the different traits, however, and it would be inter-  
526 esting to extend our analysis by incorporating mutational and/or selective correlations  
527 among traits — as shown by Martin and Lenormand (2006), the overall effect of such  
528 correlations is to reduce the effective number of phenotypic traits (see also Zhang and  
529 Hill, 2003 for an analysis of the effects of correlations among traits). Furthermore,  
530 different ways of incorporating restricted pleiotropy could be considered: in particu-  
531 lar, Welch and Waxman (2003) proposed a model in which different sets of traits are  
532 affected by different subsets of loci (modular pleiotropy), and it would be interesting  
533 to explore how this form of modularity would affect our results.

534       Regarding genetic architecture, an important assumption of our model is the  
535 fact that loci are biallelic. By contrast, Lande (1980)’s analysis considered an infinite  
536 number of alleles per locus (continuum-of-alleles model); in that case, assuming di-  
537 agonal mutation and selection matrices (no mutational or selective covariance among  
538 traits) and neglecting linkage disequilibria, one obtains that the mean log-fitness should  
539 be  $-\sqrt{\frac{n}{2}LU\bar{s}}$  (where  $L$  is the number of loci) in the case of a Gaussian fitness function  
540 (this stems from the fact that the contribution of each locus to the genetic variance  
541 is  $\sqrt{ua^2V_S}$ , where  $u$  is the per-locus mutation rate). This contrasts with the result  
542 that mean log-fitness should be  $\approx -U$  when  $Q = 2$  in biallelic models, and indicates  
543 that assumptions about the number of alleles per locus may have important effects.  
544 However, a key assumption of Lande’s analysis is the fact that distributions of allelic  
545 effects in the population are Gaussian *at each locus*, which is much stronger than the  
546 assumption that distributions of overall phenotypes are Gaussian, and may not hold  
547 for realistic values of per-locus mutation rates and mutational variances (e.g., Turelli,

548 1984). It would thus be interesting to explore how sensitive are our results to as-  
549 sumptions regarding the number of alleles per locus. In particular, our assumption of  
550 biallelic loci should not hold when  $Nu \sim 1$  (in which case many alleles may segregate  
551 at the same locus).

552 Finally, we found that the degree of pleiotropy of mutations ( $m$ ) affects mean  
553 fitness mostly through its effect on  $\bar{s}$ , in regimes where  $\bar{W}$  is affected by  $\bar{s}$ : in partic-  
554 ular the Gaussian regime in asexual populations (equation 5), and the regime where  
555 selection at each locus is weaker than mutation and drift in sexual populations (equa-  
556 tion 8). This contrasts with the result obtained by Lourenço et al. (2011) in the low  
557 mutation limit ( $NU \ll 1$ ), showing that the average fitness at equilibrium depends  
558 on the total number of selected traits but not on the degree of mutational pleiotropy.  
559 Another difference concerns the effect of population size: although the load always  
560 increases as population size decreases when  $NU \ll 1$  (Hartl and Taubes, 1998; Poon  
561 and Otto, 2000; Tenaillon et al., 2007; Lourenço et al., 2011), we found that in sexu-  
562 als, the mutation load may actually be lower at smaller population sizes in the weak  
563 selection regime (Figure 6). As we have seen, this occurs whenever the fitness effect  
564 of mutations is small ( $\bar{s} = 10^{-4}, 10^{-3}$  in the case of Figure 6), so that diversity at  
565 each locus ( $p_j q_j$ ) at mutation-selection balance is substantial, in which case drift tends  
566 to reduce this diversity and therefore reduces the genetic variance  $V_G$  — the same  
567 effect has also recently been discussed by Charlesworth (2013) in a 1-trait model. For  
568 very low fitness effects of mutations, however, mutation and drift may displace mean  
569 phenotypes away from the optimum and substantially reduce mean fitness (Figures  
570 6C and D). This last effect may only occur for a restricted range of parameters, since  
571 when  $\bar{s}$  becomes very low, very large numbers of loci are required for fitness to decrease

572 significantly (see also Charlesworth, 2013).

573         In general, we found that asexual populations have a higher mean fitness than  
574 sexual populations, in agreement with previous models of stabilizing selection with a  
575 constant optimum (e.g., Charlesworth, 1993): this is due to the fact that genotypes  
576 coding for optimal trait values are preserved by asexual reproduction, but broken by  
577 recombination. Nevertheless, sexuals may have a higher mean fitness than asexuals  
578 when epistasis among mutations is negative on average ( $Q > 2$ ), as long as  $\bar{s}$  is not too  
579 small (so that  $\bar{W}$  stays close to  $e^{-U}$  in asexuals, but reach higher values in sexuals);  
580 this occurs for example for  $Q = 6$ , when  $\bar{s} \geq 10^{-4}$  (for the parameter values used in  
581 Figures 1 and 2). Sexuals may also have a higher mean fitness when the effects of  
582 drift are important: in particular, Figure 6C and 6D show that deviations of mean  
583 phenotypes from their optimal values caused by drift are generally more important  
584 in asexual than in sexual populations (for low  $\bar{s}$ ). Although we only observed this  
585 last effect for restricted ranges of parameters, it may become more important in the  
586 case of populations subdivided into many small demes. Previous models explored the  
587 evolution of recombination modifiers under stabilizing selection (Charlesworth, 1993;  
588 Barton, 1995), but ignored the effects of genetic drift: it would thus be interesting to  
589 extend these models to the case of finite (or spatially structured) populations. Explor-  
590 ing the effects of diploidy and of the mating system would be other possible extensions  
591 of the present work.

592

593 **Acknowledgements.** We thank Nick Barton, Harold de Vladar, Pierre-Alexis Gros,  
594 Guillaume Martin, Sylvain Glémin and two anonymous reviewers for helpful discus-  
595 sions and comments, and the bioinformatics and computing service of Roscoff's Bio-

596 logical Station for computing time. This work was supported by the French Agence  
597 Nationale de la Recherche (ANR-11-BSV7-013-03).

599 We call  $V_{G,i}$  the variance of the value of trait  $i$  in the population:  $V_{G,i} =$   
 600  $\text{Var}[z_i]$ . Because we assume additive effects of the different loci on phenotypic traits  
 601 ( $z_i = \sum_j r_{ij} X_j$ , where the sum is over all loci  $j$ ),  $V_{G,i}$  can be decomposed into two  
 602 terms (e.g., Lynch and Walsh, 1998):  $V_{G,i} = V_{g,i} + d_i$ , where  $V_{g,i} = \sum_j r_{ij}^2 p_j (1 - p_j)$   
 603 is the genic variance ( $p_j$  being the frequency of allele 1 at locus  $j$  in the population),  
 604 and  $d_i = \sum_{j \neq k} r_{ij} r_{ik} D_{jk}$  the effect of linkage disequilibria on the genetic variance  
 605 ( $D_{jk} = \text{Cov}[X_j, X_k]$  being the linkage disequilibrium between loci  $j$  and  $k$ ). In the  
 606 case of an asexual population, the genetic variance at equilibrium can be calculated  
 607 by considering the whole genome as a single locus with a large number of alleles  
 608 (Supplementary File S1). In the sexual case,  $V_{g,i}$  and  $d_i$  at equilibrium can be computed  
 609 using the methods developed by Turelli and Barton (1990), extended to deal with  
 610 multiple traits under selection (Supplementary File S2). Due to the symmetry of the  
 611 model we expect that  $V_{G,i}$  and  $V_{g,i}$  should be the same for all traits at equilibrium,  
 612 and are thus simply denoted  $V_G$  and  $V_g$  in the text.

614 Under the genetic setting described above, the harmonic mean recombination  
615 rate among pairs of loci  $r_H$  (that appears in equation 7) can be computed as follows.  
616 The genetic distance between adjacent loci (in Morgans) is  $R/(L-1)$ , and therefore  
617 the distance between two loci separated by  $i$  between-locus intervals is  $iR/(L-1)$ .  
618 Furthermore, the number of different pairs of loci separated by  $i$  intervals is  $L-i$ .  
619 Finally, the rate of recombination between two loci at genetic distance  $x$  (probability  
620 that an odd number of cross-overs occurs within the interval) is given by Haldane's  
621 mapping function:  $r(x) = \frac{1}{2}(1 - e^{-2x})$ , yielding:

$$\frac{1}{r_H} = \frac{1}{\frac{1}{2}L(L-1)} \sum_{i=1}^{L-1} \frac{L-i}{\frac{1}{2} [1 - \exp(-2i\frac{R}{L-1})]} \quad (10)$$

622 which has to be evaluated numerically. For  $L = 10000$  and  $R = 10$ , one obtains  
623  $r_H \approx 0.3$ .

625 Barton, N. H. 1986. The maintenance of polygenic variation through a balance between  
626 mutation and stabilizing selection. *Genet. Res.* 47:209–216.

627 ———. 1989. The divergence of a polygenic system subject to stabilizing selection,  
628 mutation and drift. *Genet. Res.* 54:59–77.

629 ———. 1995. A general model for the evolution of recombination. *Genet. Res.* 65:123–  
630 144.

631 Barton, N. H. and M. Turelli. 1987. Adaptive landscapes, genetic distance and the  
632 evolution of quantitative characters. *Genet Res* 49:157–173.

633 Bulmer, M. G. 1971. The effect of selection on genetic variability. *Am. Nat.* 105:201–  
634 211.

635 ———. 1972. The genetic variability of polygenic characters under optimizing selec-  
636 tion, mutation and drift. *Genet. Res.* 19:17–25.

637 ———. 1974. Linkage disequilibrium and genetic variability. *Genet. Res.* 23:281–289.

638 ———. 1985. *The Mathematical Theory of Quantitative Genetics*, 2nd edition. Oxford  
639 University Press, Oxford.

640 ———. 1989. Maintenance of genetic variability by mutation-selection balance: a  
641 child’s guide through the jungle. *Genome* 31:761–767.

642 Bürger, R. 2000. *The Mathematical Theory of Selection, Recombination, and Muta-*  
643 *tion*. Wiley, Chichester, U.K.



- 644 Charlesworth, B. 1990. Mutation-selection balance and the evolutionary advantage of  
645 sex and recombination. *Genet. Res.* 55:199–221.
- 646 ———. 1993. Directional selection and the evolution of sex and recombination. *Genet.*  
647 *Res.* 61:205–224.
- 648 ———. 2013. Why we are not dead 100 times over. *Evolution* in press.
- 649 Charlesworth, B. and D. Charlesworth. 1999. The genetic basis of inbreeding depres-  
650 sion. *Genet. Res.* 74:329–40.
- 651 Charlesworth, D. 2006. Evolution of plant breeding systems. *Curr. Biol.* 16:R726–  
652 R735.
- 653 Crow, J. F. 1970. Genetic loads and the cost of natural selection. Pp. 128–177 *in* K. Ko-  
654 jima, ed. *Mathematical Topics in Population Genetics*. Springer-Verlag, Berlin.
- 655 Fisher, R. A. 1930. *The genetical theory of natural selection*. Clarendon Press, Oxford.
- 656 Gros, P.-A., H. Le Nagard, and O. Tenaillon. 2009. The evolution of epistasis and  
657 its links with genetic robustness, complexity and drift in a phenotypic model of  
658 adaptation. *Genetics* 182:277–293.
- 659 Haldane, J. B. S. 1932. *The causes of evolution*. Longmans, Green, London.
- 660 Hartl, D. L. and C. H. Taubes. 1998. *Towards a theory of evolutionary adaptation*.  
661 *Genetica* 103:525–33.
- 662 Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their  
663 applications. *Biometrika* 57:97–109.

- 664 Johnson, T. and N. H. Barton. 2005. Theoretical models of selection and mutation on  
665 quantitative traits. *Phil. Trans. Roy. Soc. (Lond.) B* 360:1411–1425.
- 666 Keightley, P. D. and S. P. Otto. 2006. Interference among deleterious mutations favours  
667 sex and recombination in finite populations. *Nature* 443:89–92.
- 668 Kimura, M. 1965. A stochastic model concerning the maintenance of genetic variability  
669 in quantitative characters. *Proc. Natl. Acad. Sci. U. S. A.* 54:731–736.
- 670 Kimura, M. and T. Maruyama. 1966. The mutational load with epistatic gene inter-  
671 actions in fitness. *Genetics* 54:1337–1351.
- 672 Kondrashov, A. S. 1988. Deleterious mutations and the evolution of sexual reproduc-  
673 tion. *Nature* 336:435–440.
- 674 Lande, R. 1976. The maintenance of genetic variability by mutation in a polygenic  
675 character with linked loci. *Genet. Res.* 26:221–235.
- 676 ———. 1977. The influence of the mating system on the maintenance of genetic  
677 variability in polygenic characters. *Genetics* 86:485–498.
- 678 ———. 1980. The genetic covariance between characters maintained by pleiotropic  
679 mutations. *Genetics* 94:203–215.
- 680 Lande, R. and D. W. Schemske. 1985. The evolution of self-fertilization and inbreeding  
681 depression in plants. I. Genetic models. *Evolution* 39:24–40.
- 682 Lourenço, J., N. Galtier, and S. Glémin. 2011. Complexity, pleiotropy and the fitness  
683 effect of mutations. *Evolution* 65:1559–1571.

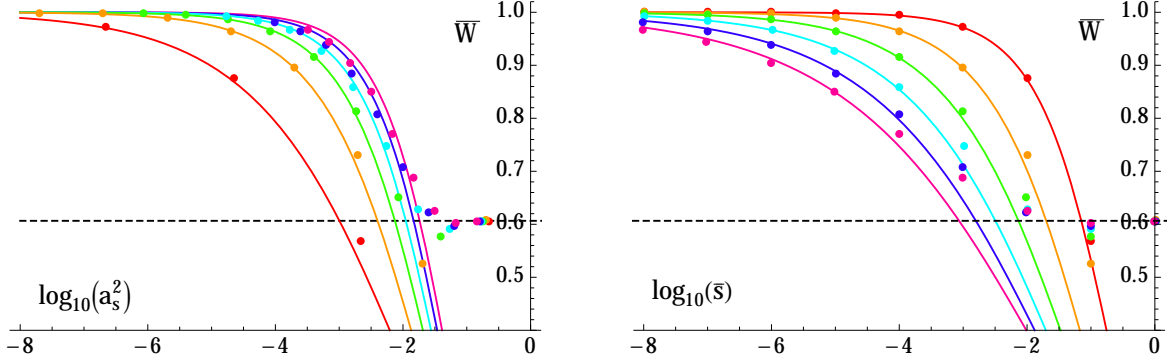
- 684 Lynch, M. and J. B. Walsh. 1998. *Genetics and Analysis of Quantitative Traits*.  
685 Sinauer Associates, Sunderland, MA.
- 686 Martin, G., S. F. Elena, and T. Lenormand. 2007. Distributions of epistasis in microbes  
687 fit predictions from a fitness landscape model. *Nat. Genet.* 39:555–560.
- 688 Martin, G. and T. Lenormand. 2006. A general multivariate extension of Fisher’s  
689 geometrical model and the distribution of mutation fitness effects across species.  
690 *Evolution* 60:893–907.
- 691 Orr, H. A. 1998. The population genetics of adaptation: the distribution of factors  
692 fixed during adaptive evolution. *Evolution* 52:935–949.
- 693 ———. 2000. Adaptation and the cost of complexity. *Evolution* 54:13–20.
- 694 Otto, S. P. 2009. The evolutionary enigma of sex. *Am. Nat.* 174:S1–S14.
- 695 Otto, S. P. and D. B. Goldstein. 1992. Recombination and the evolution of diploidy.  
696 *Genetics* 131:745–751.
- 697 Otto, S. P. and J. C. Marks. 1996. Mating systems and the evolutionary transition  
698 between haploidy and diploidy. *Biol. J. Linn. Soc.* 57:197–218.
- 699 Poon, A. and S. P. Otto. 2000. Compensating for our load of mutations: freezing the  
700 meltdown of small populations. *Evolution* 54:1467–1479.
- 701 Rowe, L. and D. Houle. 1996. The lek paradox and the capture of genetic variance by  
702 condition dependent traits. *Proc. Roy. Soc. (Lond.) B* 263:1415–1421.
- 703 Sella, G. 2009. An exact steady-state solution of Fisher’s geometric model and other  
704 models. *Theor. Popul. Biol.* 75:30–34.

- 705 Sella, G. and A. E. Hirsh. 2005. The application of statistical physics to evolutionary  
706 biology. *Proc. Natl. Acad. Sci. U. S. A.* 102:9541–9546.
- 707 Tenaillon, O., O. K. Silander, J.-P. Uzan, and L. Chao. 2007. Quantifying organismal  
708 complexity using a population genetic approach. *PLoS One* 2:e217.
- 709 Turelli, M. 1984. Heritable genetic variation via mutation-selection balance: Lerch’s  
710 zeta meets the abdominal bristle. *Theor. Popul. Biol.* 25:138–193.
- 711 ———. 1985. Effects of pleiotropy on predictions concerning mutation-selection bal-  
712 ance for polygenic traits. *Genetics* 11:165–195.
- 713 Turelli, M. and N. H. Barton. 1990. Dynamics of polygenic characters under selection.  
714 *Theor. Popul. Biol.* 38:1–57.
- 715 ———. 1994. Genetic and statistical analyses of strong selection on polygenic traits:  
716 what, me normal? *Genetics* 138:913–941.
- 717 Waxman, D. and J. R. Peck. 1998. Pleiotropy and the preservation of perfection.  
718 *Science* 279:1210–1213.
- 719 Welch, J. J. and D. Waxman. 2003. Modularity and the cost of complexity. *Evolution*  
720 57:1723–1734.
- 721 Wilke, C. O. and C. Adami. 2001. Interaction between directional epistasis and average  
722 mutational effects. *Proc. Roy. Soc. (Lond.) B* 268:1469–1474.
- 723 Wright, S. 1935. Evolution in populations in approximate equilibrium. *J. Genet.*  
724 30:257–266.

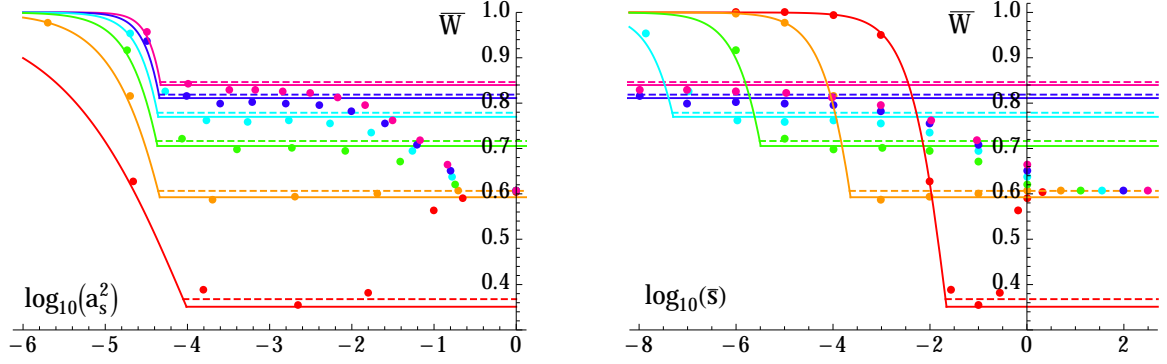
725 Zhang, X.-S. 2012. Fisher's geometrical model of fitness landscape and variance in  
726 fitness within a changing environment. *Evolution* 66:2350–2368.

727 Zhang, X.-S. and W. G. Hill. 2003. Multivariate stabilizing selection and pleiotropy  
728 in the maintenance of quantitative genetic variation. *Evolution* 57:1761–1775.

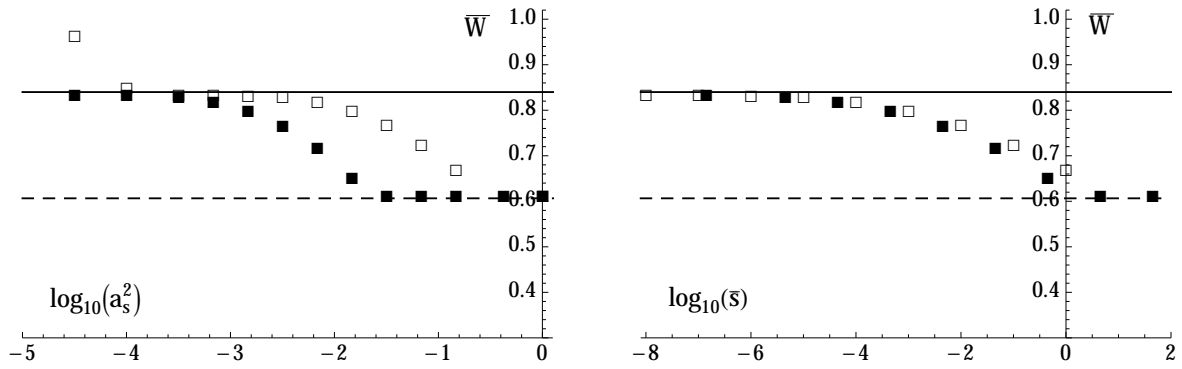
729 ———. 2005. Genetic variability under mutation selection balance. *Trends Ecol. Evol.*  
730 20:468–470.



**Figure 1.** Mean fitness of an asexual population as a function of  $a_s^2$  and  $\bar{s}$ . X-axes correspond to the mutational variance on scaled traits  $a_s^2 = a^2 / (2V_S)^{\frac{2}{Q}}$  (left) and the average deleterious effect of mutations (on log-fitness) at the optimum  $\bar{s}$  (right). The different colors correspond to different shapes of the fitness function, controlled by the parameter  $Q$ :  $Q = 1, 2, 3, 4, 5, 6$  from red to purple. The dashed horizontal line corresponds to  $e^{-U}$ , while colored curves are predictions from equation 5, and dots are simulation results. Parameter values are  $U = 0.5$ ,  $n = 50$ ,  $m = 5$ ; in the simulations  $V_S = 0.5$ , population size is set to  $N = 50000$  and the number of loci to  $L = 10000$ .

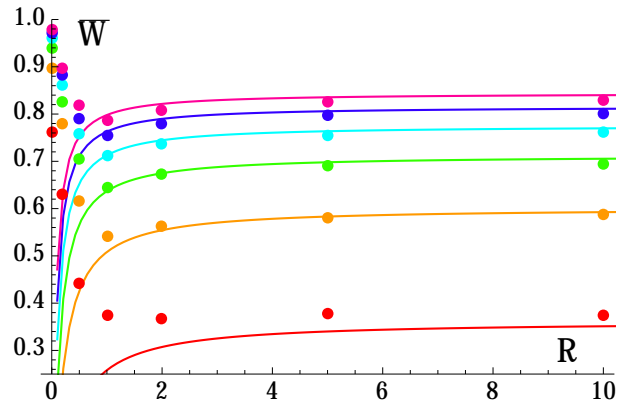


**Figure 2.** Mean fitness of a sexual population as a function of  $a_s^2$  and  $\bar{s}$  (defined as the average effect of mutations on log-fitness at the optimum). Parameter values are the same as in Figure 1, with genome map length  $R = 10$  Morgans (which leads to  $r_H \approx 0.3$ , see Appendix B). Dots correspond to simulation results, and curves to different approximations obtained assuming Gaussian distributions of phenotypic traits in the population: equations 6 (dashed lines), 7 (horizontal solid lines) and 8 (solid curves on the left). Note that in the left figure all points are superposed for  $\log_{10}(a_s^2) = 0$ .

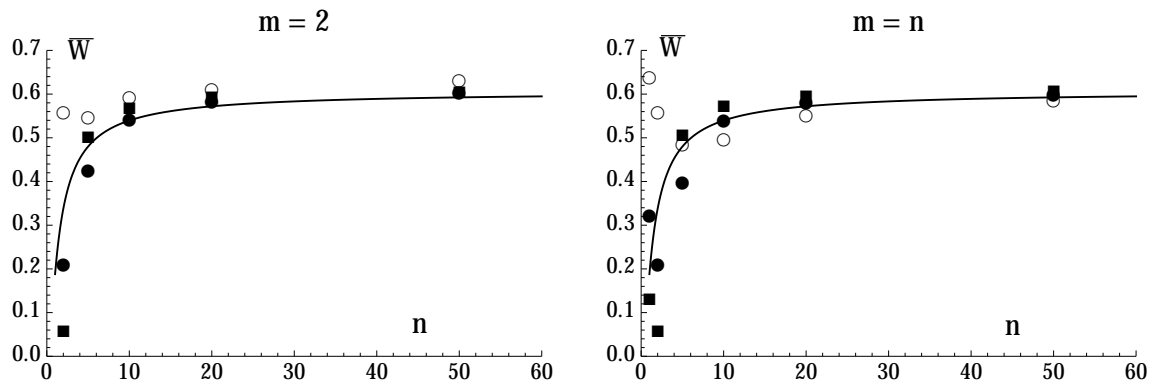


**Figure 3.** Mean fitness of a sexual population: effect of the degree of pleiotropy  $m$ . Squares: simulation results for  $Q = 6$ ,  $m = 5$  (white) and  $m = 50$  (black); other parameters are as in Figure 2. The solid line corresponds to the prediction from equation 7 (Gaussian regime), and the dashed line to  $e^{-U}$ .

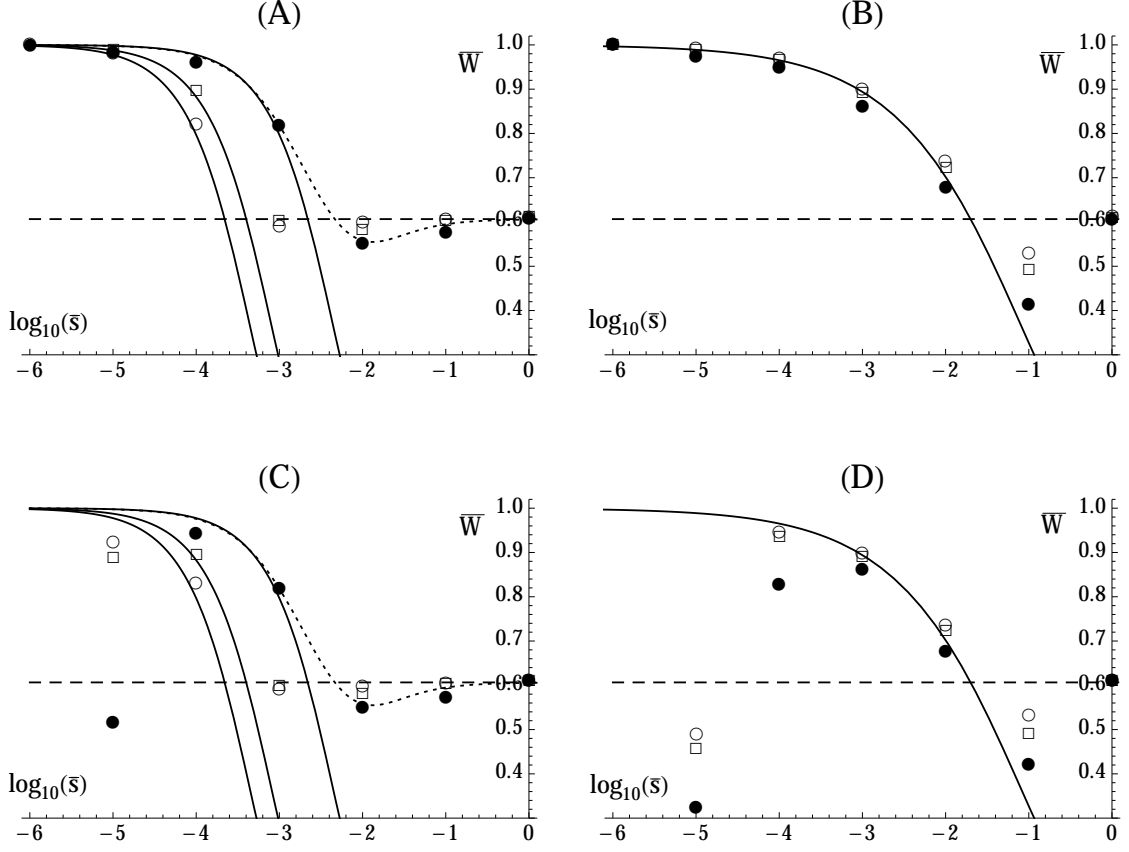




**Figure 4.** Mean fitness as a function of map length  $R$  in Morgans.  $R$  corresponds to the mean number of cross-overs per meiosis within the genome; note that  $R = 0$  corresponds to asexual reproduction. Parameter values are the same as in Figures 1 and 2, with  $a^2 = 0.0002$  in the simulations (so that  $\bar{s} = 0.001$  for  $Q = 2$ ). Dots correspond to simulation results, and curves to the prediction from equation 7.



**Figure 5.** Mean fitness of a sexual population as a function of the number selected traits  $n$ . The fitness function is Gaussian ( $Q = 2$ ). Curve: prediction from equation 7; dots: simulation results with  $\bar{s} = 0.001$  (empty circles),  $\bar{s} = 0.01$  (filled circles) and  $\bar{s} = 0.1$  (filled squares). Other parameters are as in Figure 2.



**Figure 6.** Mean fitness of a sexual (A, C) and asexual (B, D) populations: effects of population size. Dots correspond to simulation results for to  $N = 50000$  (empty circles), 5000 (empty squares), 500 (filled circles). Other parameters are as in Figure 1 and 2, with  $Q = 2$ . A, B: the optimum is located at the origin (in phenotypic space), that is, in  $z_i = 0$  for all traits  $i$ ; C, D: the optimum is located away from the origin, so that the fitness of an individual at the origin is less than 0.1 (see text for more explanations). Solid curves on the left figures (sexual populations) correspond to the predictions from equation 8 for  $N = 50000$ , 5000 and 500 from left to right, while the dotted curves corresponds to equation 6 in Supplementary File S4 (for  $N = 500$ ). Curves on the right figures (asexual populations) correspond to equation 5. Finally, the dashed horizontal lines correspond to  $e^{-U}$ .

**Table 1: Parameters and variables of the model.**

$n$	number of phenotypic traits under selection
$m$	number of traits affected by a single mutation
$Q$	shape of the fitness peak (see equation 1)
$V_S$	strength of selection (see equation 1)
$a^2$	variance of phenotypic effects of mutations
$a_s^2 = a^2 / (2V_S)^{\frac{2}{Q}}$	variance of mutational effects on scaled traits
$L$	number of loci affecting selected traits
$U$	total rate of mutation per generation on selected loci
$u = U/L$	mutation rate per locus
$\bar{s}$	mean deleterious effect of mutations (on log-fitness) at the fitness optimum (given by equation 3)
$R$	genome map length (in Morgans)
$r_H$	harmonic mean recombination rate among pairs of loci
$N$	population size
$z_i$	value of trait $i$ in a given individual
$r_{ij}$	effect of mutation at locus $j$ on trait $i$
$p_j$	frequency of allele 1 at locus $j$ in the population
$D_{jk}$	linkage disequilibrium between loci $j$ and $k$
$V_G$	genetic variance (variance of $z_i$ among individuals, which at equilibrium should be the same for all traits $i$ )
$V_g$	genic variance (see Appendix A)