Effects of Interference Between Selected Loci on the Mutation Load, Inbreeding Depression and Heterosis
Denis Roze

To cite this version:
Denis Roze. Effects of Interference Between Selected Loci on the Mutation Load, Inbreeding Depression and Heterosis. Genetics, Genetics Society of America, 2015, pp.115.178533.

HAL Id: hal-01187774
https://hal.archives-ouvertes.fr/hal-01187774
Submitted on 27 Aug 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.
Effects of interference between selected loci on the mutation load,
inbreeding depression and heterosis

Denis Roze*§

* CNRS, UMI 3614, Evolutionary Biology and Ecology of Algae, Roscoff, France
§ Sorbonne Universités, UPMC Université Paris VI, Roscoff, France
Running title: Selective interference in inbred populations

Keywords: deleterious mutation, multilocus population genetics, population structure, selective interference, self-fertilization

Address for correspondence:

Denis Roze  
Station Biologique de Roscoff  
Place Georges Teissier, CS90074  
29688 Roscoff Cedex  
France  
Phone: (+33) 2 98 29 23 20  
Fax: (+33) 2 98 29 23 24  
email: roze@sb-roscoff.fr
A classical prediction from single-locus models is that inbreeding increases the efficiency of selection against partially recessive deleterious alleles (purging), thereby decreasing the mutation load and level of inbreeding depression. However, previous multilocus simulation studies found that increasing the rate of self-fertilization of individuals may not lead to purging, and argued that selective interference among loci causes this effect. In this paper, I derive simple analytical approximations for the mutation load and inbreeding depression, taking into account the effects of interference between pairs of loci. I consider two classical scenarios of non-randomly mating populations: a single population undergoing partial selfing, and a subdivided population with limited dispersal. In the first case, correlations in homozygosity between loci tend to reduce mean fitness and increase inbreeding depression. These effects are stronger when deleterious alleles are more recessive, but only weakly depend on the strength of selection against deleterious alleles and on recombination rates. In subdivided populations, interference increases inbreeding depression within demes, but decreases heterosis between demes. Comparisons with multilocus, individual-based simulations show that these analytical approximations are accurate as long as the effects of interference stay moderate, but fail for high deleterious mutation rates and low dominance coefficients of deleterious alleles.
According to current estimates of spontaneous deleterious mutation rates in multicellular organisms (e.g., Baer et al., 2007; Haag-Liautard et al., 2007; Keightley, 2012) and estimated distributions of fitness effects of these mutations (e.g., Keightley and Eyre-Walker, 2007; Eyre-Walker and Keightley, 2007; Boyko et al., 2008; Haddrill et al., 2010), individuals may typically carry large numbers (possibly up to thousands) of deleterious alleles. Possible consequences of this load of deleterious mutations have been discussed since the early ages of theoretical population genetics (e.g., Haldane, 1937). In particular, it may reduce population mean rates of fecundity and viability, thereby increasing vulnerability to extinction (Lynch et al., 1995a,b). It may also affect a number of evolutionary processes, such as the evolution of sex or mating systems: for example, the fact that deleterious alleles are often partially recessive generates inbreeding depression, favoring outcrossing over self-fertilization (e.g., Lande and Schemske, 1985; Charlesworth and Charlesworth, 1987; Charlesworth, 2006).

In very large, panmictic populations, and in the absence of epistasis between mutations, genetic associations between deleterious alleles at different loci should remain weak, and may be neglected. In diploids, and assuming that the dominance coefficient of deleterious alleles is significantly greater than zero, the mutation load (reduction in mean fitness of the population due to deleterious alleles at mutation-selection balance) is approximately $1 - e^{-2U}$, where $U$ is the deleterious mutation rate per haploid genome (Crow, 1970; Agrawal and Whitlock, 2012). Furthermore, assuming for simplicity that all deleterious alleles have the same dominance coefficient $h$, inbreeding depression (defined here as the reduction in fitness of offspring produced...
by self-fertilization, relative to offspring produced by outcrossing) is approximately
\[ 1 - e^{-U(1-2h)/(2h)} \] (Charlesworth and Charlesworth, 2010). Analytical results on the
effects of genetic drift and non-random mating mainly stem from single-locus models.
Inbreeding increases the efficiency of selection against deleterious alleles, lowering the
mutation load and inbreeding depression (Lande and Schemske, 1985). Genetic drift
may also lead to better purging of partially recessive deleterious alleles (Kimura et al.,
1963), but this effect only causes a moderate reduction of the mutation load compared
to the effect of non-random mating, and only occurs when the effects of drift and
selection are of the same order of magnitude (Glémin, 2003). Drift has more notice-
able effects when it becomes stronger than selection and allows deleterious alleles to
reach fixation, which may increase the load by several orders of magnitude, and lowers
inbreeding depression (Bataillon and Kirkpatrick, 2000). Population subdivision has
similar consequences, due to the effects of drift within each local population (Whitlock,
2002; Glémin et al., 2003; Roze and Rousset, 2004).

These previous studies are based on single-locus models, and therefore do not
consider the effects of genetic associations between loci on the mutation load and
inbreeding depression. Between-locus associations are generated, however, as soon
as population size is finite or mating non-random (even in the absence of epistasis):
in particular, correlations in homozygosity described as “identity disequilibria” (Weir
and Cockerham, 1973; Vitalis and Couvet, 2001), and linkage disequilibria between
selected loci (Hill and Robertson, 1966; Roze and Lenormand, 2005; Kamran-Disfani
and Agrawal, 2014). Effects of deleterious mutations occurring at many loci have been
explored using simulation models of finite or infinite populations (e.g., Charlesworth et
al., 1990, 1991, 1992, 1993; Lande et al., 1994; Wang et al., 1999), sometimes showing
important deviations from single-locus predictions. In particular, using Kondrashov (1985)’s model to simulate recessive lethal mutations occurring at a very large (effectively infinite) number of unlinked loci in a partially selfing population, Lande et al. (1994) observed that contrarily to the predictions of single-locus models, recessive lethals cannot be purged by selfing unless the selfing rate exceeds a threshold value (see also Kelly, 2007). Lande et al. (1994) argued that this effect (called “selective interference”) is caused by identity disequilibria. Intuitively, selfing increases homozygosity at each locus and should thus purge recessive lethal mutations; however if many such mutations segregate in the population, any selfed offspring will almost certainly carry at least one mutation in the homozygous state, and will thus not survive. When this is the case, the population is effectively outcrossing, and purging does not occur.

To date, the effects of selective interference in partially inbred populations have only been explored numerically. How these effects scale with the strength of selection against deleterious alleles, dominance coefficients and recombination rates between loci thus remains unclear. In this paper, I derive analytical approximations describing the effect of interference between pairs of loci on the mean frequency of deleterious alleles, the mean and variance in fitness and the strength of inbreeding depression, assuming weak selection against deleterious alleles. I consider two classical scenarios of non-randomly mating populations: a single, large population in which individuals self-fertilize at a given rate, and a subdivided population with local mating followed by dispersal (island model of population structure). In the first case, interference between loci tends to reduce mean fitness and increase inbreeding depression. These effects are stronger when deleterious alleles are more recessive, but depend only weakly on the strength of selection against deleterious alleles and on recombination rates. In the
case of a subdivided population, I first show that combining two different approxi-
mations used in previous works (Glémin et al., 2003; Roze and Rousset, 2004) yields
more accurate expressions for the mutation load, inbreeding depression and heterosis
generated by a single deleterious allele. In a second step, I derive approximations for
the effects of interference between loci, and show that interference increases inbreeding
depression within demes, but decreases heterosis between demes. Comparisons with
individual-based, multilocus simulation results show that analytical approximations
incorporating the effects of associations between pairs of loci often provide accurate
predictions for the mutation load and inbreeding depression as long as the dominance
coefficient $h$ of deleterious alleles is not too low. These approximations fail when $h$ be-
comes close to zero and when the deleterious mutation rate is high, however, probably
due to the fact that higher-order interactions (involving three or more loci) become
important.

METHODS

I consider a diploid population with discrete generations, in which deleteri-
ous mutations occur at rate $U$ per haploid genome per generation. For simplicity, I
generally assume that all deleterious alleles have the same selection and dominance
coefficients ($s$, $h$), although distributions of $s$ and $h$ will be considered in the case of
a partially selfing population. Deleterious alleles at different loci have multiplicative
effects (no epistasis), so that the fitness of an organism carrying $j$ heterozygous and $k$
homozygous mutations is proportional to $(1 - hs)^j(1 - s)^k$. In the first model (par-
tial selfing), a parameter $\alpha$ measures the proportion of offspring produced by selfing,
while a proportion $1 - \alpha$ is produced by random union of gametes. The second model corresponds to the island model of population structure: the population is subdivided into a large number of demes, each containing $N$ adult individuals. These individuals produce large numbers of gametes (in proportion to their fitness), which fuse randomly within each deme to form juveniles. A proportion $m$ of these juveniles disperses, reaching any other deme with the same probability. Finally, $N$ individuals are sampled randomly within each deme to form the next adult generation. I assume soft selection, that is, all demes contribute equally to the migrant pool. In Supplementary Files A and B, I derive approximations for the mutation load and inbreeding depression that incorporate effects of pairwise associations between loci, assuming $s \ll U$ (so that individuals tend to carry many deleterious alleles) and that drift at the whole population level is negligible relative to selection. In the next sections, these analytical predictions are compared with individual-based, multilocus simulation results. The simulation programs (available from Dryad) are similar to those used in previous works (e.g., Roze and Rousset, 2009). Briefly, they represent a finite population of diploids, whose genome consists in a linear chromosome. Each generation, the number of new mutations per chromosome is drawn from a Poisson distribution with parameter $U$, the position of each mutation along the chromosome being drawn from a uniform distribution (in practice, a chromosome is represented by the positions of the deleterious alleles it carries). To form the next generation, a maternal parent is sampled for each offspring, either among all parents (in the case of a single population undergoing partial selfing) or among all parents from the offspring’s deme of origin (in the case of a subdivided population). In the first case, the parent self-fertilizes with probability $\alpha$, while with probability $1 - \alpha$ a second parent is sampled. In the second case (subdivided
population), a second parent is sampled from the same deme as the first. In all cases, the probability that a given parent is sampled is proportional to its fitness. Parents produce gametes by meiosis, a parameter $R$ measuring the genome map length: for each meiosis, the number of crossovers is sampled from a Poisson distribution with parameter $R$, the position of each crossover being drawn from a uniform distribution. Map length is fixed to 10 Morgans in most simulations, in order to mimic a whole genome with multiple chromosomes. The program runs for a large number of generations (generally $2 \times 10^5$), and measures the mean number of deleterious alleles per genome, mean fitness, variance in fitness, inbreeding depression and heterosis (in the case of a subdivided population) every 50 generations.

PARTIAL SELF-FERTILIZATION

In Supplementary File A, I derive approximate expressions for the mean and variance in log-fitness under weak selection (incorporating effects of associations between pairs of loci) and show that, neglecting higher moments of log-fitness, the average fitness is approximately:

$$
\bar{W} \approx e^{\ln W} \left(1 + \frac{\text{Var} [\ln W]}{2}\right) \tag{1}
$$

where $\ln W$ and $\text{Var} [\ln W]$ are the average and variance in log-fitness, respectively. Alternatively, an approximation for $\bar{W}$ can be obtained by assuming that the number of heterozygous mutations per outcrossed offspring follows a Poisson distribution, while the number of homozygous and heterozygous mutations per selfed offspring follow a bivariate Gaussian distribution — a similar method was used by Charlesworth et al.
(1991) to compute inbreeding depression using numerical recursions. However, both methods yield very similar results and only the first will be presented here.

In the following, I will first assume that all deleterious alleles have the same selection and dominance coefficients, and then turn to the more realistic situation where \( s \) and \( h \) vary among loci. Throughout, I assume that deleterious alleles stay at a low frequency in the population. In that case, and assuming fixed \( s \) and \( h \), the average log-fitness is approximately:

\[
\ln W \approx - \sum_i s \left[ 2h + (1 - 2h)F_i \right] p_i
\]

where the sum is over all loci, \( p_i \) is the equilibrium frequency of the deleterious allele at locus \( i \), and \( F_i \) is the probability of identity-by-descent at locus \( i \) due to partial selfing (generating an excess of homozygosity at locus \( i \)). Note that under random mating, equation 2 holds only when the dominance coefficient of deleterious alleles \( (h) \) is significantly greater than zero (otherwise, terms in \( p_i^2 \) must be included in the equation); however, equation 2 holds for all values of \( h \) under partial selfing \( (F_i > 0) \), as long as deleterious alleles stay at a low frequency.

As shown in Supplementary File A, the variance in log-fitness is approximately:

\[
\text{Var} \left[ \ln W \right] \approx 2s^2 \sum_i p_i + s^2 (1 - 2h^2) \sum_i F_i p_i
\]

\[
+ s^2 (1 - 2h)^2 \sum_{i \neq j} G_{ij} p_i p_j
\]

where \( G_{ij} \) is the identity disequilibrium between loci \( i \) and \( j \) (covariance in identity-by-descent, generating a correlation in homozygosity across loci). As explained in Supplementary File A, the terms on the first line of equation 3 are proportional to \( sU \), while the term in the second line is proportional to \( U^2 \). Therefore, assuming \( s \ll U \) and \( h \neq 0.5 \), the terms on the first line of equation 3 are relatively weak when
the population is partially selfing. Neglecting those terms, we have:

$$\text{Var} [\ln W] \approx s^2 (1 - 2h)^2 \sum_{i \neq j} G_{ij} p_i p_j .$$  \hspace{1cm} (4)

Identity disequilibria thus affect mean fitness through the term in $\text{Var} [\ln W]$ in equation 1. However, they also affect allele frequencies $p_i$ and excesses of homozygotes $F_i$, that appear in equation 2. Indeed, we have (see Supplementary File A):

$$F_i \approx \frac{\alpha}{2 - \alpha} \left[ 1 - s (1 - 2h) \sum_{j \neq i} G_{ij} p_j \right]$$  \hspace{1cm} (5)

while changes in allele frequencies due to selection are approximately:

$$\Delta_s p_i \approx -s \left[ h + (1 - h) F_i - s (1 - h) (1 - 2h) \left( 1 + \frac{\alpha}{2 - \alpha} \right) \sum_{j \neq i} G_{ij} p_j \right] p_i .$$  \hspace{1cm} (6)

Intuitively, homozygosity at locus $i$ (measured by $F_i$) is decreased by the fact that homozygotes at locus $i$ (either for the wild-type or for the deleterious allele) tend to be also homozygous at other loci, and that homozygotes at these loci have a lower fitness than heterozygotes when deleterious alleles are partially recessive (equation 5). Note that homozygosity at locus $i$ is also affected by selection acting at this locus, but this effect is negligible relative to the effects of all other loci when the number of segregating loci is large (i.e., when $s \ll U$). This decrease in homozygosity reduces the efficiency of selection against deleterious alleles, through the term in $F_i$ in equation 6. However, identity disequilibria further decrease the strength of selection against partially recessive deleterious alleles through two additional effects (explained below):

1. they reduce the “effective” dominance coefficient of deleterious alleles, and
2. they generate a relative excess of heterozygosity at locus $j$ among individuals carrying a deleterious allele at locus $i$ (measured by the association $D_{ij,j}$ in Supplementary File A). These two effects generate the last term within the brackets of equation 6 (see Supplementary File A for derivation).
The first effect stems from the fact that the fitness of mutant and wildtype homozygotes at locus $i$ are decreased by the same factor from associations with homozygotes at other selected loci; however, the fitness of heterozygotes at locus $i$ is decreased by a smaller factor, since these tend to be associated with heterozygotes at other loci, which have a higher fitness than homozygotes (provided $h < 0.5$). Therefore, identity disequilibria have a stronger impact on the fitness of homozygotes than on heterozygotes, decreasing the “effective” dominance coefficient of deleterious alleles, and thereby reducing the efficiency of selection against those alleles.

The second effect (deleterious alleles tend to be associated with more heterozygous backgrounds) stems from the fact that because heterozygotes at locus $i$ tend to be heterozygous at locus $j$ (while homozygotes at locus $i$ tend to be homozygous at locus $j$), and because selection is more efficient among homozygotes than among heterozygotes, selection against the deleterious allele at locus $i$ is less efficient among heterozygotes at locus $j$ than among homozygotes. This effect causes the deleterious allele at locus $i$ to be more frequent among heterozygotes than among homozygotes at locus $j$, in turn decreasing the efficiency of selection at locus $i$, since heterozygous backgrounds are fitter than homozygous ones when $h < 0.5$.

In the following, expressions for mean fitness $\bar{W}$ and inbreeding depression $\delta$ are obtained by replacing identity disequilibria $G_{ij}$ by their equilibrium values under neutrality. Because allele frequencies $p_i$ are of order $u/s$ (where $u$ is the deleterious mutation rate per locus), this will generate terms of order $U^2$ in the expressions for $\bar{W}$ and $\delta$ below. Taking into account the effect of selection acting at loci $i$ and $j$ on $G_{ij}$ would generate terms of order $sU^2$, which should be negligible relative to terms in $U$ and $U^2$ as long as selection is weak ($s$ small). However, $G_{ij}$ is also affected by selection
acting at other loci, due to three-locus identity disequilibria. Taking into account
the effects of these three-locus associations would introduce terms of order $U^3$ in the
expressions for $\bar{W}$ and $\delta$, which may become important when $U$ is sufficiently large.
As we will see, some discrepancies are observed between the analytical predictions and
the simulation results for high $U$ and low $h$, probably due to the fact that these higher-
order genetic associations (between three or more loci) are not taken into account in
the analysis.

Because the identity disequilibrium $G_{ij}$ depends on the recombination rate $r_{ij}$
between loci $i$ and $j$ (see Supplementary File A), $F_i$ and $p_i$ may depend on the posi-
tion of locus $i$ within the genome. However, the expression for $G_{ij}$ under neutrality
only weakly depends on $r_{ij}$, and is often close to the expression obtained for freely
recombining loci:

$$G_{ij} = \frac{4\alpha (1 - \alpha)}{(4 - \alpha) (2 - \alpha)^2}. \quad (7)$$

Injecting this expression into equations 5 and 6 yields the following approximation
for the average number of deleterious alleles per haplotype ($n = \sum_i p_i$) at mutation-
selection balance (to the second order in $U$):

$$n \approx U \frac{(2 - \alpha)}{s [2h + \alpha (1 - 2h)]} (1 + I_1) \quad (8)$$

where

$$I_1 = 2U (1 - h) (1 - 2h) \frac{2 + \alpha}{2 - \alpha} T, \quad (9)$$

$$T = \frac{2\alpha (1 - \alpha)}{(4 - \alpha) [2h + \alpha (1 - 2h)]^2} \geq 0. \quad (10)$$

The term $I_1$ in equation 8 represents the effect of identity disequilibria, increasing
the mean number of deleterious alleles when $h < 0.5$ (due to the three effects de-
scribed above). From this, and neglecting terms in $o(U^2)$, one obtains the following
approximation for mean fitness:

\[
\overline{W} \approx (1 + I_2) \exp \left[ -U \frac{4h + \alpha (1 - 4h)}{2h + \alpha (1 - 2h)} (1 + I_1) + \frac{2\alpha}{2 - \alpha} I_2 \right]
\]  

(11)

with:

\[I_2 = U^2 (1 - 2h)^2 T.\]  

(12)

As shown by equation 11 and the previous equations, identity disequilibria have three different effects on mean fitness (represented by the term in \(I_1\) and the two terms in \(I_2\) in equation 11), which can be interpreted as follows. (1) Correlations in homozygosity directly increase mean fitness when \(h \neq 0.5\), because double homozygotes and double heterozygotes have a higher fitness (on average) than genotypes that are homozygous at one locus and heterozygous at the other (e.g., Roze, 2009): this effect is represented by the term in \(\text{Var} \ln W\) is equation 1 (approximated by equation 4), corresponding to the factor \(1 + I_2\) in equation 11. (2) Identity disequilibria tend to decrease the excess of homozygosity \(F_i\) at each locus when \(h < 0.5\) (equation 5), increasing mean fitness since homozygotes have a lower fitness than heterozygotes when \(h < 0.5\) (term in \(e^{\ln \overline{W}}\) in equation 1, which increases as \(F_i\) decreases if \(h < 0.5\), as shown by equation 2). If \(h > 0.5\), \(F_i\) is now increased by identity disequilibria, but this again increases mean fitness since homozygotes have a higher fitness than heterozygotes. This second effect corresponds to the term \(2\alpha I_2/ (2 - \alpha)\) in equation 11. (3) Finally, identity disequilibria increase the frequency of deleterious alleles at mutation-selection balance when \(h < 0.5\) (as explained above), which decreases mean fitness: this corresponds to the factor \(1 + I_1\) in equation 11. One can show that effect (3) is stronger than effects (1) and (2) when \(h < 0.5\), causing identity disequilibria to decrease mean fitness (while when \(h > 0.5\), all three effects increase mean fitness). An approximation for the variance in fitness at
equilibrium is provided in Supplementary File A (equation A46); from this expression, it is possible to show that identity disequilibria generally increase the variance in fitness (unless $h = 0.5$, in which case their effect vanishes).

Finally, the effect of identity disequilibria on inbreeding depression is obtained as follows. Inbreeding depression is classically defined as:

$$\delta = 1 - \frac{W_{\text{self}}}{W_{\text{out}}}$$  \hspace{1cm} (13)

where $W_{\text{self}}$ and $W_{\text{out}}$ are the average fitnesses of individuals produced by selfing and by outcrossing, respectively (Charlesworth and Charlesworth, 1987). These quantities can be calculated as above, using expressions for $F_i$ and $G_{ij}$ in selfed individuals (for $W_{\text{self}}$) and in outcrossed individuals (for $W_{\text{out}}$). Because the last quantities equal zero, we have $W_{\text{out}} \approx e^{-2sh\sum p_i}$. Furthermore, denoting $F_{i,\text{self}}$ and $G_{ij,\text{self}}$ the excess of homozygosity and the identity disequilibrium among offspring produced by selfing, we have $F_{i,\text{self}} = (1 + F_i)/2$, while at the neutral equilibrium and under free recombination $G_{ij,\text{self}} = G_{ij}/4$. From this, one obtains:

$$\delta \approx 1 - \left(1 + \frac{I_2}{4}\right) \exp \left[-U \frac{1 - 2h}{2h + \alpha (1 - 2h)} (1 + I_1) + \frac{\alpha}{2 - \alpha} I_2\right]$$  \hspace{1cm} (14)

where $I_1$ and $I_2$ are given by equations 9 and 12. The three terms generated by identity disequilibria in equation 14 correspond to the three effects affecting mean fitness described above: (1) correlations in homozygosity tend to increase the fitness of inbred offspring whenever $h \neq 0.5$, thereby reducing inbreeding depression ($1 + I_2/4$ factor); (2) identity disequilibria reduce the excess homozygosity of inbred offspring, which also reduces inbreeding depression (term $\alpha I_2/(2 - \alpha)$) and (3) identity disequilibria increase the equilibrium frequency of partially recessive deleterious alleles, which increases inbreeding depression ($1 + I_1$ factor). Here again, the third effect is stronger.
than the first two, and the overall effect of identity disequilibria is thus to increase $\delta$.

Figure 1 shows that equation 11 provides accurate predictions for mean fitness when $U = 0.5$ and $h \geq 0.2$, while discrepancies are observed for $h = 0.1$. By contrast, ignoring effects of identity disequilibria overestimates mean fitness, in particular when $h$ is low. Figure 1 also shows that $\bar{W}$ is systematically lower than predicted when the selfing rate approaches 1; this effect is likely due to the fact that in the simulations, the effective population size is greatly reduced by background selection effects when outcrossing is very rare, in which case deleterious alleles may increase in frequency due to drift. As shown by Supplementary Figure S1, reducing the mutation rate from $U = 0.5$ to $U = 0.1$ reduces the effects of identity disequilibria, and leads to a better match between predictions from equation 11 and simulation results for $h = 0.1$. Supplementary Figures S2 and S3 show that changing the selection coefficient of deleterious alleles to $s = 0.01$ or $s = 0.1$ leads to very similar results (indeed, equation 11 does not depend on $s$), except that the effects of drift at high $\alpha$ are stronger for lower values of $s$. Genomic map length ($R$) was set to 10 Morgans in these simulations; additional simulations were run for the case of freely recombining loci, but yielded undistinguishable results unless $\alpha$ is close to 1 (in which case free recombination lowers the effects of drift — results not shown). The variance in fitness in the population at equilibrium is showed on Figure 2: when $h$ is low, the variance in fitness is maximised for intermediate values of the selfing rate $\alpha$, mainly due to the effects of identity disequilibria (which are maximised for intermediate values of $\alpha$).

Figure 3 compares the value of inbreeding depression measured in simulations with predictions from equation 14, also showing that taking into account the effects of identity disequilibria leads to more accurate predictions (although discrepancies appear
for $h = 0.1$). Results for the case of fully recessive mutations ($h = 0$) are shown in Figure 4: in agreement with Lande et al. (1994), for high mutation rates ($U = 0.25$ or 0.5) purging only occurs when the selfing rate exceeds a threshold value. Below this threshold, the population is effectively outcrossing, which is confirmed by the fact that mean fitness stays very close to the average fitness of a panmictic population ($\bar{W} \approx e^{-U}$ when $h = 0$) multiplied by the outcrossing rate (see Supplementary Figure S4). Figure 4 also shows that while equation 14 provides better predictions than the equivalent expression ignoring identity disequilibria, it does not fully capture the effect of selective interference for intermediate selfing rates and high values of $U$, indicating that higher-order genetic associations (in particular, joint homozygosity at multiple loci) must have important effects for these parameter values.

The previous results assume that all deleterious alleles have the same selection and dominance coefficients. However, Supplementary File A shows that they are easily extended to the more realistic situation where $s$ and $h$ vary among loci, as long as we can assume that selection is much stronger than drift at most loci. In that case, mean fitness and inbreeding depression at equilibrium do not depend on the strength of selection against deleterious alleles, and can be obtained by integrating terms appearing in the equations above over the distribution of dominance coefficients of these alleles (see equations A56 and A57 in Supplementary File A). In order to test these results, I modified the simulation program so that the distribution of selection coefficients of deleterious alleles is log-normal, with density function $\phi(s) = \exp \left[ - (\ln s - \mu)^2 / (2\sigma^2) \right] / (s\sigma\sqrt{2\pi})$ (where $\mu$ and $\sigma^2$ are the mean and variance of $\ln s$), truncated at $s = 1$ (this has a negligible effect for the parameter values considered here). Available data on fitness effects of deleterious alleles point to an ab-
sence of correlation between homozygous and heterozygous effects of deleterious
mutations (at least for mutations having sufficiently large homozygous effect, e.g., Manna
et al., 2012), the distribution of heterozygous effects (sh) being much less variable
than the distribution of homozygous effects (s). Here, I assume for simplicity that all
deleterious alleles have the same heterozygous effect \( \theta \): as a consequence, \( s \) and \( h \) are
negatively correlated, and the distribution of dominance coefficients (\( h = \theta/s \)) is given
by \( \psi(h) = (\theta/h^2) \phi(\theta/h) \). Figure 5 shows the distributions of \( s \) and \( h \) for \( \sigma = 0.8, \)
setting \( \mu \) and \( \theta \) so that \( \bar{s} = \exp[\mu + \sigma^2/2] = 0.05 \) and \( \bar{h} = \theta/\exp[\mu - \sigma^2/2] = 0.25 \)
(that is, \( \mu \approx -3.316 \) and \( \theta \approx 0.00659 \)); Supplementary Figure S5 shows \( h \) as a func-
tion of \( s \) for these parameter values. As shown by Figure 5, equations A56 and A57
provide accurate predictions for mean fitness and inbreeding depression when \( s \) and \( h \)
vary across loci (as before, discrepancies appear when \( \alpha \) approaches one, due to finite
population size effects). It also shows that introducing a variance in \( h \) has little effect
on mean fitness (its value being well predicted by the expression assuming fixed \( h \)),
while it strongly increases inbreeding depression, in particular when the selfing rate
is small. This may be understood from single-locus results: inbreeding depression in-
creases faster than linearly as \( h \) decreases (the effect of \( h \) on \( \delta \) being more marked when
\( \alpha \) is small), causing inbreeding depression to increase as the variance of \( h \) increases.
By contrast, the effect of \( h \) on mean fitness is weaker, and vanishes when \( \alpha = 0 \). Fi-
ally, Supplementary Figure S6 shows that when \( \bar{h} = 0.5 \), the variance of \( h \) generates
positive inbreeding depression, which is slightly increased by identity disequilibria.
The mutation load $L$, inbreeding depression $\delta$ and heterosis $H$ in a subdivided population may be defined as (e.g., Theodorou and Couvet, 2002; Whitlock, 2002; Glémin et al., 2003; Roze and Rousset, 2004):

\[
L = 1 - \frac{\overline{W}}{W_{\text{max}}}, \quad \delta = 1 - E_x \left[ \frac{W_{\text{self}, x}}{W_{\text{out}, x}} \right], \quad H = 1 - \frac{E_x [W_{\text{out}, x}]}{W_{\text{between}}} \quad (15)
\]

where $\overline{W}$ is the average fitness over the whole metapopulation, $W_{\text{max}}$ the maximal possible fitness, $W_{\text{self}, x}$ and $W_{\text{out}, x}$ the average fitnesses of individuals produced by selfing and by outcrossing in deme $x$ (respectively), $W_{\text{between}}$ the average fitness of offspring produced by crosses between parents from two different demes, while $E_x$ stands for the average over all demes $x$. In the present model $W_{\text{max}} = 1$, while the assumption of random mating within demes yields $E_x [W_{\text{out}, x}] = \overline{W}$. The definition of inbreeding depression given by equation 15 is equivalent to the “within-deme inbreeding depression” $\delta_{IS}$ in Roze and Rousset (2004) (or $\delta_1$ in Whitlock, 2002). Note that Theodorou and Couvet (2002) use a slightly different definition of within-deme inbreeding depression:

\[
\delta = 1 - E_x [W_{\text{self}, x}] / E_x [W_{\text{out}, x}]; \quad \text{however, we will see that both expressions often yield very similar results.}
\]

Supplementary File B shows how approximations for $L$, $\delta$ and $H$ can be derived, assuming that deme size $N$ is large, while the migration rate $m$ and strength of selection $s$ are small. As in the previous section, the total population size is supposed very large (large number of demes), so that the effects of drift at the whole population level can be neglected. In a first step, I show that improved approximations for $L$, $\delta$ and $H$ generated by mutation at a single locus can be obtained by combining previous results (Glémin et al., 2003; Roze and Rousset, 2004). Then, I extend these results to
the case of deleterious alleles occurring at a large number of loci, incorporating effects of pairwise associations among loci.

**Single-locus results.** As shown in Supplementary File B (see also Whitlock, 2002; Glémin et al., 2003; Roze and Rousset, 2004) the mutation load, inbreeding depression and heterosis generated by a single deleterious allele in a subdivided population (with random mating within demes) are approximately:

\[ L \approx 2shp + s(1 - 2h)F_{ST}p \]  

(16)

\[ \delta \approx \frac{1}{2}s(1 - 2h)(1 - F_{ST})p \]  

(17)

\[ H \approx s(1 - 2h)F_{ST}p \]  

(18)

where \( p \) is the frequency of the deleterious allele in the whole population, and \( F_{ST} \) measures the average genetic diversity within demes, relative to the genetic diversity in the whole metapopulation (Wright, 1969). As the number of demes tends to infinity, \( F_{ST} \) becomes equivalent to the probability that two genes sampled from the same deme are identical by descent (e.g., Rousset, 2002), that is, that their ancestral lineages coalesce in a finite number of generations — which is possible only if these lineages stay in the same deme until coalescence occurs, since it takes an infinite time for lineages present in different demes to coalesce.

Assuming \( N \) is large while \( s \) and \( m \) are small, the change in frequency of the deleterious allele due to selection is approximately (see Supplementary File B):

\[ \Delta_s p \approx -shp - s(1 - 3h)F_{ST}p + s(1 - 2h)\gamma p \]  

(19)

where \( \gamma \) is the probability that three genes sampled from the same deme are identical by descent (*i.e.*, that their ancestral lineages coalesce before migrating to different
In order to compute $\Delta s p$ in terms of the model parameters $(s, h, N, m)$, one may then assume that under weak selection $F_{ST}$ and $\gamma$ remain close to their equilibrium values under neutrality, and replace $F_{ST}$ and $\gamma$ by these values in equation 19 (Whitlock, 2002, 2003; Wakeley, 2003; Roze and Rousset, 2003, 2004). While this approximation yields accurate results as long as $s \ll m$, it generally fails when $s \geq m$, as the effect of selection on $F_{ST}$ and $\gamma$ cannot be neglected (Roze and Rousset, 2003, 2004). However, Supplementary File B shows that when $N$ is sufficiently large, $F_{ST}$ and $\gamma$ can be approximated by:

$$F_{ST} \approx \frac{1}{1 + 4N (m + sh)}, \quad \gamma \approx \frac{1}{[1 + 2N (m + sh)] [1 + 4N (m + sh)]}.$$

(20)

Replacing $F_{ST}$ and $\gamma$ by these expressions in equation 19 yields, at mutation-selection equilibrium:

$$p \approx \frac{(1 + 2\Gamma) (1 + 4\Gamma) u}{2\Gamma (1 + 4\Gamma h) s}$$

(21)

with $\Gamma = N (m + sh)$, and where $u$ is the mutation rate towards the deleterious allele.

From equations 16, 17 and 18, one then obtains:

$$L \approx \frac{(1 + 2\Gamma) (1 + 8\Gamma h)}{2\Gamma (1 + 4\Gamma h)} u$$

(22)

$$\delta \approx \frac{(1 - 2h) (1 + 2\Gamma)}{1 + 4\Gamma h} u$$

(23)

$$H \approx \frac{(1 - 2h) (1 + 2\Gamma)}{2\Gamma (1 + 4\Gamma h)} u.$$  

(24)

When $s \ll m$ (so that $\Gamma \approx Nm$), equations 21 - 24 become equivalent to the results obtained using expressions for $F_{ST}$ and $\gamma$ under neutrality (e.g., equations 35-39 in Roze and Rousset, 2004). As shown be Figure 6, however, taking into account the effect of selection on $F_{ST}$ and $\gamma$ (by using equation 20) greatly improves analytical predictions when $m \leq s$. Interestingly, the expression for $F_{ST}$ given by equation 20 was already
obtained by Glémin et al. (2003) using a method developed by Ohta and Kimura (1969, 1971) to compute moments of allele frequencies in finite populations (equation 11a in Glémin et al., 2003). However, Glémin et al. (2003) neglected the effect of population structure on the mean allele frequency $p$ (assuming that selection is strong relative to local drift) and thus replaced $p$ by $u/(sh)$ in equations 16 - 18. In effect, equations 21 - 24 thus combines the results of Glémin et al. (2003) — that take into account the effect of selection on $F_{ST}$, but neglect the effect of population structure on mean allele frequency — and the results of Roze and Rousset (2004), that take into account the effect of population structure on mean allele frequency, but neglect the effect of selection on $F_{ST}$. Supplementary Figure S7 compares these different approximations, and shows that equations 21 - 24 lead to significant improvement over these previous results.

Finally, we can note that when the migration rate $m$ is set to zero, the model represents an infinite number of replicates of a single population of size $N$. The above results thus predict that the variance in frequency of a deleterious allele due to drift in a single finite population should be approximately $\bar{p}\bar{q}/(1 + 4Nh)$ as long as the average frequency $\bar{p}$ of the deleterious allele remains small (from equation 20, with $\bar{q} = 1 - \bar{p}$). Furthermore, expressions for the average allele frequency, mutation load and inbreeding depression are obtained by setting $m = 0$ in equations 21-23. Figure 7 shows that these approximations are indeed accurate as long as $N$ is not too small (so that the deleterious allele stays rare in the population).

**Effects of interference between selected loci.** In the multilocus case, population structure generates different types of associations between alleles at different loci,
either from the same individual or from different individuals from the same deme.

As shown in Supplementary File B, selection against deleterious alleles is affected by these associations, through extra terms that appear in equation 19 (see equation B33 in Supplementary File B), and also through the fact that $F_{ST}$ and $\gamma$ at each locus are affected by interactions between loci. Assuming large deme size and weak selection and migration (so that $1/N$, $m$ and $s$ are of order $\epsilon$, where $\epsilon$ is a small term), fixed $s$ and $h$ and freely recombining loci, one obtains:

$$F_{ST} \approx \frac{1}{1 + 4N (m + sh)} \left[ 1 - s (1 - 2h) \frac{8Nm}{[1 + 4N (m + sh)]^2} \sum_j p_j \right]$$  \hspace{1cm} (25)$$

which is equivalent to equation 79 in Roze and Rousset (2008) when $sh \ll m$, while:

$$\gamma \approx \frac{1}{[1 + 2N (m + sh)] [1 + 4N (m + sh)]} \times \left[ 1 - s (1 - 2h) \frac{4Nm [3 + 8N (m + sh)]}{[1 + 2N (m + sh)] [1 + 4N (m + sh)]^2} \sum_j p_j \right]$$  \hspace{1cm} (26)$$

(where $p_j$ is the frequency of the deleterious allele at locus $j$ in the metapopulation).

Equations 25 and 26 show that $F_{ST}$ and $\gamma$ at a given locus are decreased by partially recessive deleterious alleles segregating at other loci: this effect stems from the fact that offspring from migrant individuals tend to be more heterozygous, and thus have higher fitness than offspring from philopatric individuals when deleterious alleles are partially recessive (heterosis). This increases the “effective” migration rate, and thus reduces genetic correlations between individuals within demes (e.g., Ingvarsson and Whitlock, 2000). As shown by equation 19, a lower $F_{ST}$ decreases selection against deleterious alleles when $h < 1/3$ (and increases selection otherwise), while a lower $\gamma$ increases selection against deleterious alleles when $h < 1/2$, and increases it otherwise. As a result, the effects of between-locus interactions on $F_{ST}$ and $\gamma$ may either increase or decrease the efficiency of selection against deleterious alleles, depending on parameter
Furthermore, Supplementary File B shows that all other effects of between-locus interactions should be negligible when $1/N$, $s$ and $m$ are small, $h \neq 0.5$ and assuming each deleterious allele remains rare in the metapopulation ($p_j$ small). From equations 19, 25 and 26, one obtains for the mean number of deleterious alleles per haplotype at equilibrium (to the second order in $U$):

$$n \approx (1 - I_3) \frac{(1 + 2\Gamma)(1 + 4\Gamma)U}{2\Gamma(1 + 4\Gamma h)}$$

(27)

where $I_3$ represents the effect of interactions between loci:

$$I_3 = (1 - 2h) \left( \frac{Nm}{\Gamma} \right) \frac{1 + 8\Gamma[h - (1 - 3h)\Gamma]}{\Gamma(1 + 4\Gamma)(1 + 4\Gamma h)^2} U.$$  

(28)

Note that the sign of $I_3$ depends on parameter values: while $I_3$ is always positive when $1/3 < h < 1/2$, it may become negative when $h < 1/3$, in particular if $\Gamma$ is large: therefore, interference between loci may either increase or decrease the frequency of deleterious alleles. Furthermore, one obtains for the mutation load:

$$L \approx 1 - \exp \left[ - (1 - I_4) \frac{(1 + 2\Gamma)(1 + 8\Gamma h)}{2\Gamma(1 + 4\Gamma h)} U \right]$$

(29)

with:

$$I_4 = (1 - 2h) \left( \frac{Nm}{\Gamma} \right) \frac{1 + 8\Gamma h [1 - (1 - 4h)\Gamma]}{\Gamma(1 + 4\Gamma h)^2(1 + 8\Gamma h)} U.$$  

(30)

Again, the sign of $I_4$ (representing the effect of interactions between loci) depends on parameter values: $I_4$ is always positive if $1/4 < h < 1/2$ (in which case interactions reduce the load), but becomes negative if $h < 1/4$ and $\Gamma$ is sufficiently large.

By contrast, the sign of the expressions obtained for the effects of interactions between loci on heterosis and inbreeding depression stays constant when $h < 1/2$.

Indeed, one obtains for heterosis (see Supplementary File B for derivation):

$$H \approx 1 - \exp \left[ - (1 - I_5) \frac{(1 - 2h)(1 + 2\Gamma)}{2\Gamma [1 + 4\Gamma h]} U \right]$$

(31)
\[ I_5 = (1 - 2h) \left( \frac{Nm}{\Gamma} \right) \frac{1 + 8\Gamma h (1 + \Gamma)}{\Gamma (1 + 4\Gamma h)^2} U. \]  
\[ (32) \]

showing that interactions between loci always decrease heterosis when \( h < 1/2 \). Finally, inbreeding depression is given by:

\[ \delta \approx 1 - \exp \left[ - (1 + I_6) \frac{(1 - 2h) (1 + 2\Gamma)}{1 + 4\Gamma h} U \right] \]
\[ (33) \]

with:

\[ I_6 = 2 (1 - 2h)^2 \left( \frac{Nm}{\Gamma} \right) \frac{1}{\Gamma (1 + 4\Gamma h)^2} U \]
\[ (34) \]

showing that interactions between loci always increase inbreeding depression within demes. Indeed, heterosis and inbreeding depression scale with \( F_{ST} n \) and \((1 - F_{ST}) n\), respectively (from equations 17 and 18), and one obtains from equations 25 and 27 that the effect of between-locus interactions on these products stays constant as long as \( h < 1/2 \) (to the second order in \( U \)). As shown by Figure 8, simulation results confirm that interactions between loci tend to increase inbreeding depression and decrease heterosis, fitting reasonably well with predictions from equations 31 and 33 (although discrepancies appear when \( m \) is very small). The effects of interactions between loci on inbreeding depression stays rather small for the parameter values used in Figure 8A, but become more important for lower values of \( s \) and \( h \) or higher values of \( U \), as shown by Figure 8C and 8D. As an aside, Supplementary File B also shows that defining inbreeding depression as \( 1 - E_x [W_{\text{self},x}] / E_x [W_{\text{out},x}] \) or as \( 1 - E_x [W_{\text{self},x}/W_{\text{out},x}] \) (where again \( E_x \) stands for the average over all demes \( x \), while \( W_{\text{self},x} \) and \( W_{\text{out},x} \) are the mean fitnesses of offspring produced by selfing and by outcrossing in deme \( x \)) should yield very similar results under our assumptions (\( N \) large, \( s, m \) small, \( p_j \) small), since the variance of \( W_{\text{out},x} \) and the covariance between \( W_{\text{self},x} \) and \( W_{\text{out},x} \) across demes remain
small under these conditions. Indeed, both measures were used in the simulations and gave nearly undistinguishable results (not shown).

DISCUSSION

Theoretical predictions regarding the effect of the mating system of organisms on the mutation load and inbreeding depression are often based on single-locus models. However, as previously shown by Lande et al. (1994), some of these predictions may not hold when considering more realistic situations involving multiple selected loci. In particular, when the genomic mutation rate towards recessive deleterious alleles is sufficiently high, inbreeding depression is maintained at high levels irrespective of the selfing rate of individuals (contrarily to the predictions of single-locus models), unless selfing exceeds a threshold value. This “selective interference” effect has been invoked by Scofield and Schultz (2006) and by Winn et al. (2011) to explain the lack of evidence of purging in meta-analyses comparing species with intermediate selfing rates to species with low selfing rate (while species with high selfing rates show reduced inbreeding depression): for example, Winn et al. (2011) observed that species with intermediate selfing rates (between 0.2 and 0.8) present similar levels of inbreeding depression as species with lower selfing rates (less than 0.2). Furthermore, it has been proposed that this effect may allow the stable maintenance of mixed mating systems (involving both selfing and outcrossing), since the classical prediction that only complete selfing or complete outcrossing should be evolutionarily stable (Lande and Schemske, 1985) is based on the assumption that inbreeding depression is a decreasing function of the selfing rate.
Most previous studies of selective interference were based on Kondrashov (1985)'s simulation model, representing deleterious alleles occurring at an infinite number of unlinked loci, in an infinite population. Lande et al. (1994) considered the case of fully (or almost fully, i.e., $h = 0.02$) recessive lethal mutations ($s = 1$), and found that selective interference becomes important when the genomic deleterious mutation rate is sufficiently high (0.2 – 1). Kelly (2007) showed that strong homozygous effects of deleterious alleles are not necessarily needed for interference to occur (the effect being actually stronger with $s = 0.1$ than with $s = 1$), while $h$ has to be sufficiently low in order to observe interference. Winn et al. (2011) modelled transitions from outcrossing to partial selfing, and showed that increased selfing leads to lower levels of inbreeding depression (purging) when $s = 0.05$ and $h = 0.2$ and when $s = 0.001$ and $h = 0.4$, but not when $s = 1$ and $h = 0.02$ (for a genomic mutation rate equal to 1), inbreeding depression staying close to 1 in the last situation.

To date, no analytical model has explored the mechanisms of selective interference. In this paper, I showed that analytical approximations can be obtained in regimes where interference stays moderate, by considering the effects of pairwise interactions between selected loci and assuming weak selection. As we have seen, the mechanisms underlying interference in partially inbred populations depend of the form of inbreeding considered. In a single, large population undergoing partial selfing, interference between loci are mainly driven by identity disequilibria between those loci (as long as the fitness of heterozygotes departs from the average of both homozygotes at each locus, i.e., $h \neq 0.5$). However, identity disequilibria affect inbreeding depression through several mechanisms: correlations in homozygosity directly reduce $\delta$, but also indirectly decrease homozygosity at each locus (which also reduces $\delta$) and decrease
the efficiency of selection against deleterious alleles, allowing them to be maintained
at higher frequencies (thereby increasing $\delta$). This last effect (which predominates over
the first two) corresponds to the verbal explanation proposed previously to explain
selective interference (purging is prevented by identity disequilibria, e.g., Lande et al.,
1994; Winn et al., 2011). However, we have seen that this effect itself involves three
different mechanisms: reduction of the “effective” dominance coefficient of deleterious
alleles, decrease in homozygosity at each locus, and positive correlations between the
presence of a deleterious allele at a given locus and heterozygosity at other loci. The
results presented here also show that interference is little affected by the strength of
selection against deleterious alleles (at least as long as selection is weak to moderate)
or by linkage, as long as genome map length is sufficiently high — in agreement with
the simulations results obtained by Charlesworth et al. (1992), showing that the effect
of linkage on mean fitness and inbreeding depression in partially selfing populations
often remains slight.

When inbreeding results from limited dispersal (population structure), interference
effects are more complicated as they involve associations between loci as well as
between different individuals from the same spatial location. However, we have seen
that when selection and migration are weak while deme size is large, the main effect of
interference between loci (assuming partially recessive deleterious alleles) is to increase
the “effective” migration rate at each locus (Ingvarsson and Whitlock, 2000), thereby
reducing probabilities of identity between alleles present in different individuals from
the same deme. This may either increase or decrease the strength of selection against
deleterious alleles depending on parameter values, but always increases inbreeding de-
pression within demes, while reducing heterosis between demes. In contrast with the
case of partial selfing in a single population, this effect does not involve identity dis-
equilibria (correlations in homozygosity across loci), but other types of associations
between alleles present in different individuals from the same deme (moments of link-
age disequilibrium and allele frequencies, see equations B44 and B45). Furthermore,
an important difference between partial selfing and population structure is that the
mutation load and inbreeding depression in a structured population may be affected
by the strength of selection against deleterious alleles (in particular when migration
is weak, see Figure 6). The effects of interference between loci also depend on the
strength of selection, being more marked for lower values of $s$.

Is selective interference likely to have important consequences in natural popula-
tions? Confirming previous results, we have seen that interference leads to substantial
deviations from single-locus results for parameter values leading to strong inbreeding
depression (high $U$, low $h$), independently of the strength of selection against deleteri-
ous alleles. In particular, the total absence of purging as the selfing rate increases (up
to a threshold value) is only observed when inbreeding depression is close to 1 (while
for lower values of $\delta$, interference only dampens the decline of inbreeding depression
with selfing). As observed by Winn et al. (2011), this condition may be fulfilled in
gymnosperms, which show very high levels of inbreeding depression. In contrast, an-
giosperms show lower values of inbreeding depression (on average), for which selective
interference may not be sufficiently strong to prevent purging. According to the results
shown here, interference between deleterious alleles may thus not represent a sufficient
explanation for the lack of evidence for purging in Angiosperms in Winn et al. (2011)'s
meta-analysis (for selfing rates between 0 and 0.8). Other possible explanations may
be a lack a sufficient power to detect purging, or synergistic epistasis between deleteri-
uous alleles, which tends to flatten the relationship between inbreeding depression and
the selfing rate (Charlesworth et al., 1991). Note also that, as discussed by Winn et
al. (2011), most estimates of inbreeding depression compiled in their dataset were ob-
tained under greenhouse conditions, and may thus be biased downwards if inbreeding
depression tends to be stronger in harsher environments (Armbruster and Reed, 2005).
More empirical studies of inbreeding depression in different sets of conditions are thus
needed to assess the potential importance of interactions between loci on selection
against deleterious alleles.

Finally, because the suppression of purging due to interference only occurs when
inbreeding depression is maximal, this mechanism does not seem a likely explanation
for the evolutionary maintenance of mixed mating systems (as proposed in previous
papers), since selfing should be strongly disfavored when $\delta$ is close to 1. Nevertheless,
the effects of associations between loci on the evolution of mating systems remain little
explored (but see Kamran-Disfani and Agrawal, 2014). Besides affecting inbreeding
depression, between-locus associations may modulate the advantage of selfers due to
more efficient purging (e.g., Uyenoyama and Waller, 1991; Epinat and Lenormand,
2009), and possibly generate additional selective forces acting on a modifier locus af-
flecting the selfing rate. These effects are still waiting for analytical exploration.

Acknowledgements. I thank the bioinformatics and computing service of Roscoff’s
Biological Station for computing time, and Joachim Hermisson and two anonymous re-
viewers for helpful suggestions and comments. This work was supported by the French
Agence Nationale de la Recherche (project TRANS, ANR-11-BSV7-013).


Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1991. Multilocus models of
inbreeding depression with synergistic selection and partial self-fertilization. Genet.

R735.

Charlesworth, D. and B. Charlesworth. 1987. Inbreeding depression and its evolution-

and population size on inbreeding depression due to mutational load. Genet. Res.
59:49–61.

———. 1993. Mutation accumulation in finite outbreeding and inbreeding populations.


Harper and Row, New York.

Epinat, G. and T. Lenormand. 2009. The evolution of assortative mating and selfing


Glémin, S. 2003. How are deleterious mutations purged? Drift versus nonrandom


33


Uyenoyama, M. K. and D. M. Waller. 1991. Coevolution of self-fertilization and in-


Figure 1. Average fitness at equilibrium as a function of the rate of self-fertilization $\alpha$, for different values of the dominance coefficient of deleterious alleles ($h$), and deleterious mutation rate per haploid genome $U = 0.5$. Solid curves: analytical approximation including effects of identity disequilibria (equation 11); dashed curves: neglecting effects of identity disequilibria (obtained by setting $I_1 = I_2 = 0$ in equation 11); dots: simulation results (in this and the following figures, error bars are smaller than the size of dots). In the simulations, $s = 0.05$, $N = 20,000$ and $R = 10$. 
Figure 2. Variance in fitness in the population at equilibrium, as a function of the rate of self-fertilization $\alpha$ and for different values of the dominance coefficient of deleterious alleles. Curves correspond to predictions from equation A46 in Supplementary File A (dotted: $h = 0.2$, long-dashed: $h = 0.3$, solid: $h = 0.4$); short-dashed curve: adding the term given in equation A47 for $h = 0.2$. Dots: simulation results for $h = 0.2$ (empty circles), $h = 0.3$ (filled circles) and $h = 0.4$ (filled squares). Parameter values are the same as in Figure 1.
**Figure 3.** Inbreeding depression as a function of the rate of self-fertilization $\alpha$, for different values of the dominance coefficient of deleterious alleles ($h = 0.1, 0.2, 0.3$ and 0.4 from top to bottom), and deleterious mutation rate per haploid genome $U = 0.5$. Solid curves: analytical approximation including effects of identity disequilibria (equation 14); dashed curve: neglecting effects of identity disequilibria (setting $I_1 = I_2 = 0$ in equation 14); dots: simulation results (same parameter values as for Figure 1).
**Figure 4.** Inbreeding depression as a function of the selfing rate $\alpha$: same as Figure 3 with fully recessive deleterious alleles ($h = 0$), and different values of the deleterious mutation rate $U$. 
Figure 5. Top: distributions of $s$ and $h$ assuming a log-normal distribution of $s$ with $\mu \approx -3.316$ and $\sigma = 0.8$ (so that $\overline{s} = 0.05$) and fixed heterozygous effects of deleterious alleles $\theta \approx 0.00659$ (so that $\overline{h} = 0.25$); see text for more explanations. Bottom: mean fitness and inbreeding depression as a function of the selfing rate $\alpha$. Dots: simulations results, using the distributions of $s$ and $h$ shown on top. Black curves: analytical predictions for fixed $h$, set to $\overline{h}$ (from equations 11 and 14). Red curves: analytical predictions for varying $h$ (from equation A56 and A57 in Supplementary File A). Dashed/solid curves: neglecting/including the effects of identity disequilibria. The mutation rate is set to $U = 0.5$; in the simulations, $N = 20,000$ and $R = 10$. 
Figure 6. Equilibrium values of $F_{ST}$, mutation load $L$ (divided by its value in a panmictic population, $2u$), heterosis and inbreeding depression in a subdivided population, when selection acts at a single locus. The x-axes show the migration rate between demes (on a log scale), and the different colours correspond to different values of $s$: 0.005 (orange), 0.01 (green), 0.05 (blue) and 0.1 (red). Coloured curves: predictions from equations 20 and 22-24. Dots: one-locus simulation results (30 replicates of $10^7$ generations; error bars are smaller than the size of dots). Black curves: predictions from Roze and Rousset, 2004 (obtained by replacing $\Gamma$ by $Nm$ in equations 20 and 22-24). Other parameter values: $h = 0.2$, $N = 100$, $u = 10^{-5}$; in the simulations the number of demes is set to 200, and back mutations occur at rate $10^{-7}$.
Figure 7. Variance of deleterious allele frequency (scaled by $\bar{p}\bar{q}$) and inbreeding depression in a single finite population, as a function of population size $N$ (on a log-scale). Solid curves correspond to predictions obtained from numerical integration over the standard diffusion result for the distribution of allele frequency (e.g., equation 9.3.4 in Crow and Kimura, 1970, see also Bataillon and Kirkpatrick, 2000), while dashed curves correspond to $1/(1 + 4Nh) \Gamma$ (left) and to the expression obtained by replacing $\Gamma$ by $Nsh$ in equation 23 (right). Dots: one-locus simulation results (averages over 30 replicates of $10^8$ to $10^9$ generations). Parameter values: $s = 0.005, 0.01, 0.05, 0.1$ (from right to left), $h = 0.3$, $u = 10^{-5}$; back mutation rate: $v = 10^{-7}$. 
Figure 8. Inbreeding depression (A, C, D) and heterosis (B) when deleterious mutations occur at a large number of loci, as a function of the migration rate between demes (on a log scale). Dots: multilocus simulation results; solid curves: predictions from equations 31 and 33; dotted curves: predictions ignoring effects of interactions between loci (setting $I_5$ and $I_6$ to zero in equations 31 and 33). Parameter values: A, B: $U = 0.5, h = 0.2, s = 0.05$ (squares, top curves in A, bottom curves in B), $s = 0.01$ (circles, bottom curves in A, top curves in B); C: $U = 0.5, h = 0.1, s = 0.01$; D: $U = 1, h = 0.2, s = 0.01$. Deme size: $N = 100$. In the simulations the number of demes is set to 200, and genome map length to $R = 20$ Morgans.