

Genetic Load in Sexual and Asexual Diploids: Segregation, Dominance and Genetic Drift

Christoph R. Haag¹ and Denis Roze²

University of Edinburgh, Institute of Evolutionary Biology, Edinburgh EH9 3JT, United Kingdom

Manuscript received March 8, 2007

Accepted for publication April 19, 2007

ABSTRACT

In diploid organisms, sexual reproduction rearranges allelic combinations between loci (recombination) as well as within loci (segregation). Several studies have analyzed the effect of segregation on the genetic load due to recurrent deleterious mutations, but considered infinite populations, thus neglecting the effects of genetic drift. Here, we use single-locus models to explore the combined effects of segregation, selection, and drift. We find that, for partly recessive deleterious alleles, segregation affects both the deterministic component of the change in allele frequencies and the stochastic component due to drift. As a result, we find that the mutation load may be far greater in asexuals than in sexuals in finite and/or subdivided populations. In finite populations, this effect arises primarily because, in the absence of segregation, heterozygotes may reach high frequencies due to drift, while homozygotes are still efficiently selected against; this is not possible with segregation, as matings between heterozygotes constantly produce new homozygotes. If deleterious alleles are partly, but not fully recessive, this causes an excess load in asexuals at intermediate population sizes. In subdivided populations without extinction, drift mostly occurs locally, which reduces the efficiency of selection in both sexuals and asexuals, but does not lead to global fixation. Yet, local drift is stronger in asexuals than in sexuals, leading to a higher mutation load in asexuals. In metapopulations with turnover, global drift becomes again important, leading to similar results as in finite, unstructured populations. Overall, the mutation load that arises through the absence of segregation in asexuals may greatly exceed previous predictions that ignored genetic drift.

MOST eukaryotes engage in sexual reproduction despite potentially high costs, such as the famous twofold cost of sex (MAYNARD SMITH 1978; BARTON and CHARLESWORTH 1998). Genetically, the key components of sexual reproduction are recombination and, in diploid organisms, segregation. Both are absent under pure asexual reproduction. Recombination and segregation rearrange the genotypic composition of offspring from sexual matings, by bringing together novel allelic combinations at a locus (segregation) or at a set of different loci (recombination). Hence, these processes may affect the distribution of fitness values within populations and may therefore generate indirect selective pressure for sexual reproduction (BARTON and CHARLESWORTH 1998; OTTO and LENORMAND 2002; OTTO 2003; AGRAWAL 2006; DE VISSER and ELENA 2007).

One possible advantage of recombination and segregation is that they allow sexual populations to reduce their genetic load through an improved efficiency of selection against deleterious alleles (KIMURA and MARUYAMA 1966; CROW 1970; CROW and KIMURA 1970). This requires

the existence of negative disequilibria such as when beneficial and deleterious alleles (within or between loci) occur more often in the same individual than expected by chance. Recombination and segregation bring together favorable alleles within the same individuals (and unfavorable alleles in others) and hence improve the efficiency of natural selection. Selection against recurrent deleterious mutation can create negative disequilibria between loci ("negative linkage disequilibrium") if deleterious alleles at different loci interact synergistically (KONDRASHOV 1982; CHARLESWORTH 1990). Equivalently, selection can create negative disequilibria within loci ("heterozygote excess") if deleterious alleles are fully or partially recessive. This is because (with partially recessive deleterious alleles) the fitness of heterozygotes is higher than the average fitness of the homozygotes, and hence heterozygote excess develops during selection. Once a heterozygote excess is established, sexual reproduction leads to improved selection and therefore to reduced genetic load, because segregation eliminates the heterozygote excess, resulting in an increased variance in fitness (CHASNOV 2000).

Arguments based on the genetic load are, however, not sufficient to predict how a modifier gene affecting the balance between sexual and asexual reproduction will evolve (*e.g.*, BARTON 1995; OTTO 2003). Indeed, there is always a cost (in terms of mean fitness of offspring) of breaking genetic associations that have been generated

¹Corresponding author: University of Edinburgh, Institute of Evolutionary Biology, Ashworth Lab. 2, W. Mains Rd., Edinburgh EH9 3JT, United Kingdom. E-mail: christoph.haag@ed.ac.uk

²Present address: Station Biologique de Roscoff, CNRS UMR 7144, 29682 Roscoff Cedex, France.

by selection. This cost is termed “recombination load” or “segregation load” (depending on whether negative linkage disequilibrium or heterozygote excess is broken). Analyses of modifier models have shown that, in infinite, randomly mating populations, sexual reproduction may be favored only when dominance and/or epistasis are sufficiently weak relative to the strength of selection, so that the recombination load and/or segregation load is not too high (BARTON 1995; OTTO 2003). These models have also shown that even a low rate of inbreeding may allow sex and recombination to be favored under less restrictive conditions than with random mating (OTTO 2003; ROZE and LENORMAND 2005). Whereas there is little empirical support for widespread weak synergistic epistasis (RICE 2002), there is ample evidence that new deleterious mutations are, on average, partly recessive (MULLER 1950; SIMMONS and CROW 1977; LYNCH and WALSH 1998; SZAFRANIEC *et al.* 2003). In diploid populations, genetic associations generated by dominance may thus play a greater role in the evolution of sex than genetic associations generated by epistasis (OTTO 2003).

Another factor that may contribute to the creation of negative linkage disequilibria is genetic drift in conjunction with directional selection. This is because genetic drift randomly creates positive and negative associations, but positive associations are rapidly consumed by selection (because they represent the most extreme fitness values), while negative associations tend to last longer (HILL and ROBERTSON 1966; FELSENSTEIN 1974). Genetic drift together with directional selection can lead to an advantage of recombination without the requirement of synergistic epistasis (OTTO and BARTON 1997, 2001; ILES *et al.* 2003; BARTON and OTTO 2005; KEIGHTLEY and OTTO 2006; ROZE and BARTON 2006), especially in subdivided populations (MARTIN *et al.* 2006; SALATHÉ *et al.* 2006).

Whether genetic drift can also lead to an advantage of segregation is less clear. Genetic drift has two important effects: first, in sexual populations, it may increase the average strength of selection against recessive deleterious alleles, an effect that has been termed “purging by drift” (GLÉMIN 2003); it is unclear whether this effect can also occur in asexuals. Second, it leads to random changes in allele frequencies, which renders selection less efficient: if drift is too strong compared to selection, frequency changes of deleterious alleles may be similar to those of neutral alleles (KIMURA *et al.* 1963). However, the strength of this effect may differ between sexual and asexual populations; indeed, due to the absence of segregation, asexuals inherit genotypes rather than alleles, which increases the sampling variance of genotype frequencies in asexual populations and thus reduces their variance effective size relative to sexual populations (BALLOUX *et al.* 2003).

Here, we analyze both of the effects of genetic drift explicitly by using equilibrium models to investigate the expected genetic load due to recurrent deleterious mutation in sexual and asexual populations subject to

drift. These models do not directly study the evolution of sex, because we fix the rate of sexual reproduction to either zero or one. Rather, they aim, as a first step, at comparing the relative effects of drift and selection between sexual and asexual diploids subject to recurrent deleterious mutation. To concentrate only on effects that are due to segregation, we use simple one-locus two-allele models, starting with a single population of varying effective size. We then extend this to metapopulations with finite, but large numbers of demes. This extension is important because it is likely to represent the natural situation, as most populations are subdivided to some extent, and because single small populations are unlikely to persist over long periods of time. AGRAWAL and CHASNOV (2001) derived the mutation load in diploid, infinite, and spatially structured sexual and asexual populations. In their model, population regulation occurs at the level of the whole population, and the only effect of population structure is to increase homozygosity in sexuals. However, it is likely that, in subdivided populations, most competition occurs locally, decreasing the efficiency of selection by increasing competition among related individuals (“local drift”). Population structure may also affect the load through effects on genetic drift at the total population level and on demography. We investigate these different effects using a finite-island model with extinction and recolonization. Overall, we show that in single undivided populations, as well as in metapopulations, the cumulative effects of genetic drift and segregation across a realistic number of loci may lead to an equilibrium fitness in sexuals that is many times higher than that in asexual populations.

THE MODEL

Throughout, we calculate the genetic load in sexual and asexual populations due to a single locus that mutates with rate u from a wild-type allele, A , to a mutant allele, a . Back mutation from a to A occurs at rate v , $v \ll u$. Relative genotypic fitness values for AA , Aa , and aa are 1, $1 - hs$, and $1 - s$, respectively, where h is the dominance coefficient and s is the selection coefficient. The genetic load L is defined as $L = 1 - W$, where W is the mean fitness of a population. Following CHASNOV (2000) and AGRAWAL and CHASNOV (2001), we extrapolate our results to many loci by assuming that each locus contributes independently (multiplicatively) to the genetic load; that is, $L_{\text{tot}} = 1 - W_{\text{tot}} = 1 - W^n$, where n is the number of loci. As is discussed later, this single-locus load underestimates the total load of asexuals, as interference between loci may greatly reduce the efficiency of selection at each locus.

Single sexual population: The mean fitness W_{sex} of a randomly mating population is determined by the frequency \bar{p} of the deleterious allele: $W_{\text{sex}} = 1 - 2hs\bar{p} - s(1 - 2h)\bar{p}^2$, and thus $L_{\text{sex}} = 2hs\bar{p} + s(1 - 2h)\bar{p}^2$. The expected allele frequency \bar{p} (and the expected squared

frequency, $\overline{p^2}$) in a population of arbitrary size N , subject to mutation, selection, and genetic drift, can be obtained by numerical integration of Wright's distribution (WRIGHT 1937; KIMURA *et al.* 1963; see also CABALLERO and HILL 1992; BATAILLON and KIRKPATRICK 2000; GLÉMIN 2003). All numerical calculations were done with Mathematica (WOLFRAM 2003), and we checked the approximations against simulation results, obtained by averaging the observed load over 10^8 generations, after the mutation–selection–drift equilibrium had been reached (which can easily be checked by visual inspection of the results).

Single asexual population: Due to the lack of segregation in obligate asexual diploids, their two haploid genomes will acquire mutations independently. Thus, a new mutation that arises in one of the two homologous chromosomes of an asexual will be restricted to that chromosome unless an independent mutation occurs at the same locus in the second chromosome (see also CHARLESWORTH and CHARLESWORTH 1997). Calculating the mutation–selection–drift balance for a diploid asexual population hence requires solving a two-dimensional stochastic model representing the change in frequency of genotypes Aa and aa . However, this can be simplified by noting that, when mutations are (partially) recessive, only two genotypes will usually segregate in the population. When N_e is large, the population is at mutation–selection balance and mutant homozygotes can be neglected (provided that $hs \gg \mu$, $N_e > 1/hs$). As N_e decreases, selection against Aa individuals becomes inefficient (roughly when $N_e < 1/hs$), and Aa goes to fixation. However, selection against aa individuals remains efficient, and the frequency of these individuals remains small, until, when N_e decreases to $\sim < 1/s$, selection against aa also becomes inefficient, and aa will eventually fix. Each of these two processes can be analyzed separately by standard diffusion models for haploid populations with only two genotypes with different fitnesses (CROW and KIMURA 1970).

The first process is represented by a diffusion in a population composed of Aa and AA individuals, corresponding to a standard haploid diffusion where Aa individuals have relative fitness $1 - hs$. The mutation rate from AA to Aa is $2u$ (because mutation in either of the two homologous chromosomes will form an Aa individual; CHARLESWORTH and CHARLESWORTH 1997), and the back-mutation rate (from Aa to AA) is v . Integration of the haploid diffusion described by these parameters yields Q , the expected frequency of Aa individuals in a population of Aa and AA individuals. The second process is represented by a diffusion in a population composed only of aa and Aa individuals (assuming the back-mutation rate is sufficiently small), that is, a standard haploid diffusion, where aa individuals have relative fitness $(1 - s)/(1 - hs)$, which equals $1 - (1 - h)s$ to the first order in s . The mutation rate from Aa to aa is u , and the back-mutation rate (from aa to Aa) is $2v$. This yields

R , the expected frequency of aa individuals in a population of aa and Aa . To combine these two processes, we approximate the expected frequencies of mutant homozygotes and heterozygotes, $\overline{p_{aa}}$ and $\overline{p_{Aa}}$, by QR and $Q(1 - R)$, respectively. Although mathematically not strictly correct, this gives good results (compared to simulations) here because it is only when the genotype Aa is close to fixation in the first diffusion (Q close to 1) that the frequency of aa is not negligibly small in the second diffusion. The load is then given by $L_{\text{asex}} = hs\overline{p_{Aa}} + s\overline{p_{aa}}$.

Large population approximation: The expected genetic load in sexual and asexual populations of infinite size is between u and $2u$, but stays close to $2u$ for most biologically realistic parameter values. Only when h is quite small ($h < \sqrt{u/s}$) is the load significantly reduced in sexuals compared to asexuals, because, as h decreases, L tends to u more quickly in sexuals than in asexuals (CHASNOV 2000; for the sexual case, see also KIMURA *et al.* 1963).

Small population approximation: When N_e is very small, the population is fixed for one genotype most of the time, and selection has little effect on the fixation probabilities. Neglecting the effect of selection, a simple calculation shows that asexual populations are fixed for aa , Aa , or AA with probabilities $u^2/(u+v)^2$, $2uv/(u+v)^2$, and $v^2/(u+v)^2$, respectively. On average, the load is thus given by $L_{\text{asex}} \approx us(u+2hv)/(u+v)^2$. Sexual populations are fixed for aa or AA with probabilities $u/(u+v)$ and $v/(u+v)$, respectively, and the load is given by $L_{\text{sex}} \approx us/(u+v)$. When $v \ll u$, L_{sex} and L_{asex} are both close to s ; however, the sexual load will be slightly higher than the asexual load as long as $h < \frac{1}{2}$: $L_{\text{sex}}/L_{\text{asex}} \approx (u+v)/(u+2hv)$.

Subdivided population: We next use the island model to consider the effects of population structure. The population is subdivided into n demes, each containing N diploid adults. These adults produce a large but, depending on their fitness, variable number of gametes (in the sexual case) or diploid juveniles (in the asexual case) and then die. In the sexual case, gamete fusion is random within each deme. Each juvenile then disperses with probability m . Each deme thus contributes to the pool of migrants in proportion to its average fecundity. Each migrant can reach any other deme with the same probability. Finally, N individuals are sampled randomly among all the juveniles present in each deme, to form the next adult generation. We also consider the effect of local extinctions of demes; for this, we use Slatkin's extinction–recolonization model (SLATKIN 1977). At the beginning of a generation, each deme goes extinct with probability e . During the dispersal phase, juveniles reaching an extinct deme do not survive. Then (after the dispersal phase), each extinct deme is recolonized by k juveniles, either sampled randomly from the whole population of juveniles (migrant pool model) or derived from the same deme (propagule pool model). In both cases, each juvenile has an equal probability to become a recolonizer.

It is assumed that recolonizers reproduce immediately, so that deme size goes back to N in all demes.

Our model corresponds, for instance, to a population subdivided into discrete patches in which the number of breeding sites is fixed (N adults per patch). This introduces local population regulation, but regulation is not completely local (as long as some migration occurs) because more fertile patches will produce more migrants. Therefore, our model may be seen as intermediate between complete local regulation (“soft selection”) and complete global regulation (sometimes called “hard selection”). Under complete global regulation, each deme contributes to the next generation (and not just to the migrant pool) in proportion to its mean fecundity. However, it is difficult to imagine a simple biological scenario that would correspond to complete global regulation in a spatially structured population. This can easily be seen by considering the limit when migration tends to zero, in which case one would have to assume that deme sizes can grow indefinitely, at rates depending on their mean fecundity. Also at intermediate levels of migration, it is difficult to imagine a life cycle that would make the contribution of each deme to the next generation exactly proportional to its mean fecundity. Therefore, rather than using a parameter that measures the “degree of local competition” as is sometimes done to scale between soft and hard selection, we preferred to investigate the effects of local competition in a simple life cycle where all parameters have immediate biological meanings (deme size, migration rate, extinction rate). Still, we can note that our model is equivalent to soft selection when $m = 0$ and to hard selection when $m = 1$.

We use the method of ROZE and ROUSSET (2003), which is sketched in APPENDIX A, to derive expressions for the expectation and the variance of the change in frequency of the deleterious allele, over one generation. Importantly, this method uses a separation-of-timescales argument that works best when selection is weak relative to migration ($s < m$). In sexuals, the expected change in frequency p of the mutant allele a , $E[\Delta p]$, is given by

$$E[\Delta p] = S_1 pq + S_2 p^2 q + uq - vp, \quad (1)$$

with $q = 1 - p$ and

$$S_1 = -s \left[1 - \frac{1}{N} (1 - m_b)^2 (1 - e) \right] [h + (1 - h)r_0] + s(1 - m_b)^2 (1 - e) \left(1 - \frac{1}{N} \right) [2hr_1 + (1 - 2h)\alpha] \quad (2)$$

$$S_2 = -s(1 - 2h) \left[1 - \frac{1}{N} (1 - m_b)^2 (1 - e) \right] (1 - r_0) + 2s(1 - 2h)(1 - m_b)^2 (1 - e) \left(1 - \frac{1}{N} \right) (r_1 - \alpha). \quad (3)$$

In the expressions above, r_0 , r_1 , and α are probabilities of coalescence within demes under neutrality, which are

functions of N , m , e , and k , and are given in APPENDIX A; m_b is the “backward” migration rate (the probability that, after dispersal, a juvenile comes from another deme), given by $m_b = m(1 - e)/(1 - me)$.

For asexuals, we calculate the expected genotype frequencies using two diffusions, as described previously for the single population. In the first case, where aa individuals are very rare, and genotypes Aa and AA segregate in the population with frequencies p and q , respectively, the expected change in the frequency of Aa individuals over one generation is given by

$$E[\Delta p] = S_{a1} pq + 2uq - vp \quad (4)$$

with

$$S_{a1} = -hs[1 - (1 - e)(1 - m_b)^2 r_1^R], \quad (5)$$

where r_1^R is the probability that the ancestral lineages of two genes sampled with replacement from the same deme stay in the same deme and coalesce, in a haploid model under neutrality (see APPENDIX A). In the second case, where AA is very rare, and aa and Aa segregate in the population (with p now being defined as the frequency of aa and q as the frequency of Aa), the expected change in frequency of aa individuals over one generation is given by

$$E[\Delta p] = S_{a2} pq + uq - 2vp \quad (6)$$

with

$$S_{a2} = -(1 - h)s[1 - (1 - e)(1 - m_b)^2 r_1^R], \quad (7)$$

where r_1^R is the same as above.

We then have for both the sexual and the asexual cases

$$\text{Var}[\Delta p] = \frac{r_1^R [1 - (1 - e)(1 - m_b)^2]}{n(1 - e)} pq \quad (8)$$

(e.g., ROZE and ROUSSET 2003), where r_1^R is again the neutral probability of coalescence for two genes sampled with replacement from a deme. The expression for r_1^R differs between the sexuals and the asexuals (see APPENDIX A), and, as above, p is the frequency of a in the sexual case and the frequency of Aa and aa in the first and second asexual models, respectively, and $q = 1 - p$.

The equations above take the same form as in a single finite population, the selection coefficients and the effective population size now depending on N , m , e , and k . As for a single population, these diffusion equations can be integrated numerically (see APPENDIX A) to obtain the load at equilibrium, the only difference being that, in the sexual case, the frequency of homozygotes is affected by population structure, and the load is now given by

$$L_{\text{sex}} = s[2h - (1 - 2h)r_0] \bar{p} + s(1 - 2h)(1 - r_0) \bar{p}^2, \quad (9)$$

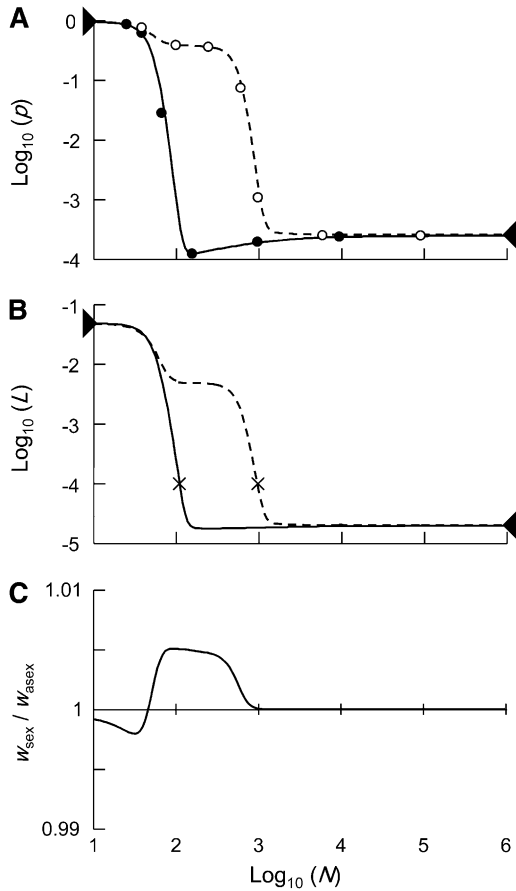


FIGURE 1.—Mean frequency p (A) of a partly recessive deleterious allele ($h = 0.1$, $s = 0.05$), mean load L (B), and relative fitness $w_{\text{sex}}/w_{\text{aseex}}$ (C). Solid lines in A and B represent sexual populations and dashed lines represent asexual populations. Triangles on the left and right (A and B) indicate expected values for very small and infinite populations, respectively, and \times 's (B) indicate N_{lim} . In A, solid circles (sexuals) and open circles (asexuals) are simulation results for selected values of N . Mutation parameters: $u = 10^{-5}$, $v = 10^{-7}$.

where \bar{p} and $\overline{p^2}$ are averages over the probability distribution of p , the frequency of allele a (e.g., ROZE and ROUSSET 2003). For asexuals, we combine the two diffusions as described above for the single population and express the load as

$$L_{\text{aseex}} = h\overline{p_{Aa}} + s\overline{p_{aa}}. \quad (10)$$

RESULTS

Single population: Figure 1 shows some results obtained by numerical integration for a partially recessive deleterious allele ($h = 0.1$, $s = 0.05$) in sexual and asexual populations. For sexual populations, Figure 1 illustrates findings already obtained by others (KIMURA *et al.* 1963; BATAILLON and KIRKPATRICK 2000; GLÉMIN 2003): in large populations, the mean mutant allele frequency and load are very close to mutation–selection balance.

As population size decreases, genetic drift increases, leading to partial purging by drift (GLÉMIN 2003), that is, a reduction in the mean frequency of the deleterious allele. The load also decreases slightly, although the effect is rather small for the parameter values used in Figure 1. This effect was first described by KIMURA *et al.* (1963, p. 1306), who noted that “Here there is the paradox that a finite population has a smaller load than an infinite population, which would seem to imply that a random process produces a higher average fitness than a deterministic one.” As shown by GLÉMIN (2003), this can be interpreted by considering the effect of averaging over a distribution of allele frequencies. When selection is weak, the change in frequency of the deleterious allele due to selection is given (to the first order in s) by

$$\begin{aligned} \Delta_s p &= -s[h + (1 - 2h)p]p(1 - p) \\ &= -shp - s(1 - 3h)p^2 + s(1 - 2h)p^3. \end{aligned} \quad (11)$$

Assuming that p follows a frequency distribution, we have

$$\overline{\Delta_s p} = -sh\bar{p} - s(1 - 3h)\overline{p^2} + s(1 - 2h)\overline{p^3}, \quad (12)$$

where the overbar stands for the average of the distribution. Assuming that most of the distribution stands at low values of p (selection remains efficient relative to drift), we may neglect the third moment $\overline{p^3}$ and obtain

$$\overline{\Delta_s p} = -sh\bar{p} - s(1 - 3h)(\sigma_p^2 + \bar{p}^2), \quad (13)$$

where σ_p^2 is the variance of the distribution of p . Equation 13 thus shows that when $h < \frac{1}{3}$, the efficiency of selection (as measured by $|\overline{\Delta_s p}|$) increases as the variance σ_p^2 increases (as long as drift is not too strong).

Whereas this effect may lead to a reduced load at intermediate population sizes in sexuals, decreasing population size also increases the stochastic component of allele frequency change. Hence, below a certain population size, selection is overwhelmed by drift, causing the mean frequency of the deleterious allele to rise until close to its neutral expectation, that is, close to fixation for $u \gg v$. This transition occurs over a narrow range of population sizes, leading also to a sharp increase in genetic load once population sizes decrease below a certain point. To quantify the population sizes at which this transition occurs, we arbitrarily define a limiting population size, N_{lim} , so that $L = 10u$ at N_{lim} (\times 's in Figure 1B).

In large asexual populations, mean allele frequency and load are very similar to values for large sexual populations (unless h is very small; CHASNOV 2000). However, a decrease in population size does not initially lead to purging, which can be understood from a similar argument as above. As long as selection against Aa remains efficient, the population consists essentially of

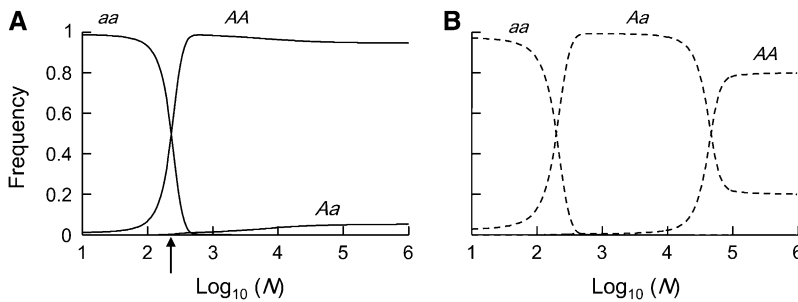


FIGURE 2.—Expected average genotype frequencies in sexual (A) and asexual (B) populations. a is a deleterious allele with $h = 0.01$, $s = 0.01$, $u = 10^{-5}$, and $v = 10^{-7}$. Note that the plotted frequencies are averages over many populations of a given size. For instance, sexual populations of the size indicated with an arrow in A have equal frequencies of aa and AA genotypes, on average, in the near absence of Aa . However, this does not imply heterozygote deficit within populations, but rather that most populations are fixed for AA or aa (with equal probability).

Aa and AA individuals. Calling now p the frequency of Aa , we have (to the first order in s) $\Delta_s p = -shp(1-p)$, and thus

$$\overline{\Delta_s p} = -sh\bar{p} + sh(\sigma_p^2 + \bar{p}^2), \quad (14)$$

showing that the variance in the distribution of p now decreases the efficiency of selection $|\overline{\Delta_s p}|$, for all $h > 0$ (Equations 13 and 14 become equivalent for $h = \frac{1}{2}$). Again, decreasing population size also increases the stochastic component of allele frequency change, and below a certain size selection against heterozygotes becomes inefficient, causing their frequency to increase to almost 1 (Figure 2). As a result, the load increases and N_{lim} is higher than in sexuals (Figure 1B). At the same time, homozygous mutants appear more frequently (because mutations arise in heterozygotes), but selection against these homozygotes remains efficient and hence their frequency remains low (Figure 2). However, as the population size decreases again, drift eventually overwhelms selection against deleterious homozygotes, and hence the mean mutant allele frequency and the load increase sharply a second time (Figure 1B). For the range of intermediate population sizes, in which selection is efficient against homozygous mutants, but not heterozygotes, the mutation load is substantially higher in asexuals than in sexuals, and hence the fitness of sexuals relative to asexuals peaks at these population sizes (Figure 1C). This is a direct consequence of the absence of segregation, because, with segregation, matings between heterozygotes constantly produce new homozygotes. Hence, in sexuals, heterozygotes cannot reach high frequencies due to drift while mutant homozygotes are still efficiently selected against (see also

Figure 2). Finally, Figure 1 also shows that our approximations, which are based on the assumption that only two genotypes segregate in asexual populations most of the time, work very well when compared with simulation results.

The stepwise increase in genetic load with decreasing population size in asexual populations is due to dominance. This can be seen in Figure 3, which shows the mean frequency of a deleterious mutation under multiplicative selection, that is, for $h = 1/(1 + \sqrt{1-s})$. In this case, genotype frequencies in infinite populations at mutation–selection balance are in Hardy–Weinberg proportions in both sexuals and asexuals (CHASNOV 2000; OTTO 2003). Figure 3A shows that under multiplicative selection, selection is overwhelmed by drift at larger population sizes in asexuals than in sexuals. This may be explained by the fact that drift has stronger effects in asexuals; indeed, the variance effective population size of asexuals is only half the variance effective population size of sexuals, because drift in asexuals occurs through random sampling of genotypes, whereas in sexuals it occurs through random sampling of alleles (BALLOUX *et al.* 2003). Indeed, Figure 3A shows that under multiplicative selection, the equilibrium frequency of the deleterious allele in an asexual population is the same as in a sexual population of half its size. As with the Hill–Robertson effect that occurs between selected loci (HILL and ROBERTSON 1966), an alternative interpretation of this process is through its effect on genetic associations: in the same way as drift and selection combine to generate negative linkage disequilibria between selected loci, they also combine to generate a negative intralocus association (excess of heterozygotes), which

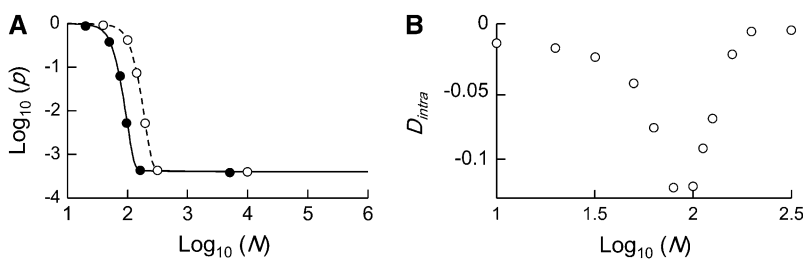


FIGURE 3.—(A) Mean frequency, p , of a deleterious allele under multiplicative selection ($s = 0.05$, $h = 1/(1 + \sqrt{1-s}) \approx 0.506$) in sexual (solid line) and asexual (dashed line) populations. The circles (solid for sexuals, open for asexuals) represent simulation results, and in each pair of horizontally adjacent circles, $N_{\text{asex}} = 2N_{\text{sex}}$. Mutation parameters: $u = 10^{-5}$, $v = 10^{-7}$. (B) Simulation results showing the mean intralocus association $D_{\text{intra}} = \bar{p}_{aa} - \bar{p}_a^2$ as a function of $\log_{10} N$, in asexuals (open circles) and sexuals (solid circles). For comparison, at $\log_{10} N = 2$, D_{intra} at a neutral locus with the same mutation parameters is -0.005 in asexuals and -2×10^{-7} in sexuals.

ual populations (same parameter values as in A). For comparison, at $\log_{10} N = 2$, D_{intra} at a neutral locus with the same mutation parameters is -0.005 in asexuals and -2×10^{-7} in sexuals.

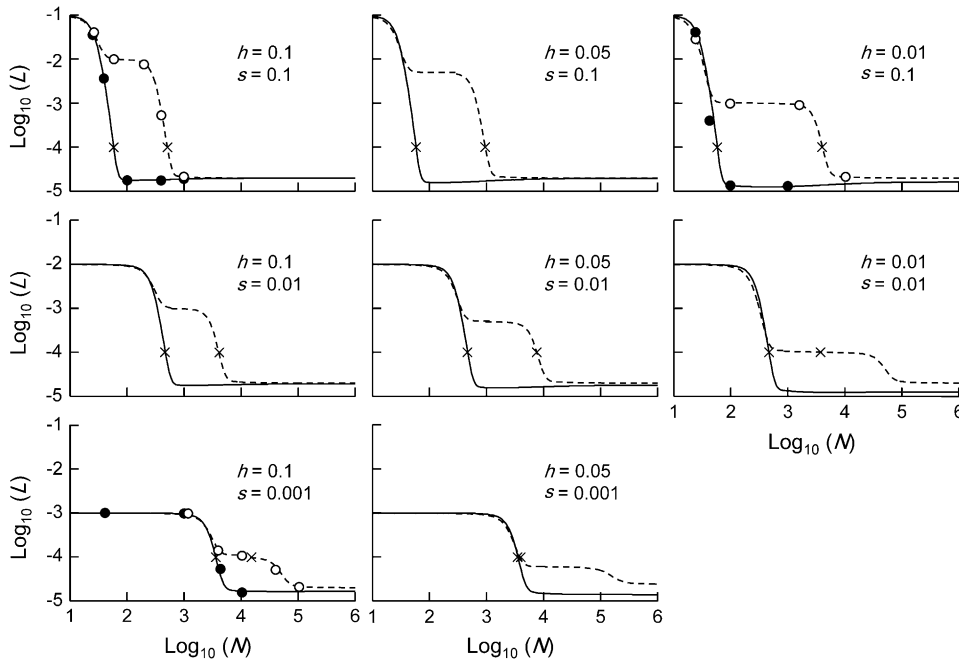


FIGURE 4.—Mean genetic load, L , in sexuals (solid lines) and asexuals (dashed lines) for mutations of different h and s . Simulation results (solid circles for sexuals, open circles for asexuals) are given for three selected parameter combinations of h and s . No graph is shown for $h = 0.01$, $s = 0.001$, because our model assumes $hs > u$. \times 's indicate N_{lim} . Mutation parameters: $u = 10^{-5}$, $v = 10^{-7}$.

may reach high values when segregation is reduced or absent. This is due to the fact that drift generates a variance in genotype frequencies, leading to situations where heterozygotes are in excess and situations where they are in deficit. Because the variance in fitness is lower in the first situation, negative associations tend to last longer and may accumulate over time if they are not broken down every generation by segregation (an extreme case being the situation where Aa is fixed in the population). This is illustrated in Figure 3B, which shows the mean intralocus association (the equivalent of the linkage disequilibrium between loci), defined as $D_{\text{intra}} = \overline{p_{aa}} - \overline{p_a}^2$, for different values of N in asexuals. Note that an equivalent definition is $D_{\text{intra}} = \overline{Fp_Ap_a}$, where F is the inbreeding coefficient (we use D_{intra} rather than F to emphasize the parallel to linkage disequilibrium between loci and because F is not defined for monomorphic populations). As shown by Figure 3B, D_{intra} is negative and peaks at intermediate values of N (where Aa often reaches high frequencies). However, it is important to note that this intralocus effect is not exactly equivalent to the Hill–Robertson effect, because asexual reproduction generates a negative D_{intra} , on average, even in the absence of selection (under mutation and drift alone), while the mean linkage disequilibrium between loci is zero under neutrality. In APPENDIX B, we show that in an asexual population, and under neutrality, D_{intra} at equilibrium is given by

$$D_{\text{intra}} = -\frac{\theta}{2(1+\theta)^2[1+2(N-1)u(1+\theta)]}, \quad (15)$$

where $\theta = v/u$ (for small u and v). Thus, in the neutral case, D_{intra} is close to -0.005 for the parameter values

used in Figure 3B ($\theta = 0.01$, $N < 10^3$); this is confirmed by simulations (not shown). Finally, one can note that in the sexual case, genetic drift also generates a negative D_{intra} in the neutral case (e.g., pp. 39–40 in GALE 1990) or under multiplicative selection, but this D_{intra} is much smaller in magnitude than in the asexual case. Under neutrality, it is given by

$$D_{\text{intra}} = -\frac{2\theta u}{(1+\theta)[1+2(2N-1)u(1+\theta)]} \quad (16)$$

at equilibrium (APPENDIX B), that is, for the parameter values in Figure 3B roughly four orders of magnitude lower than in the neutral case for asexuals.

Figure 4 shows the genetic load in sexual and asexual populations for deleterious alleles with a range of biologically realistic values of h and s (SIMMONS and CROW 1977; LYNCH and WALSH 1998; SZAFRANIEC *et al.* 2003). N_{lim} , the population sizes at which the load starts to increase above equilibrium values for infinite populations, is always higher for asexuals than for sexuals. There is thus always a range of intermediate population sizes where the load in asexual populations is considerably higher than in sexual ones. Figure 5 shows how N_{lim} depends on s for a range of different h . In the parameter range considered, N_{lim} in sexuals is almost independent of h and is $\sim 4/s$. In contrast, in asexuals, N_{lim} depends strongly on both h and s and is $\sim 4/hs$ ($N_{\text{lim}} = 4/s$ and $N_{\text{lim}} = 4/hs$ are numerical fits used to illustrate that N_{lim} depends on s in sexuals and on hs in asexuals).

Figure 4 also shows that for large populations, the genetic load is very similar in sexual and asexual populations ($\sim 2u$), except when h and s are small (e.g., $h = 0.001$, $s = 0.01$), when the load in sexuals is somewhat

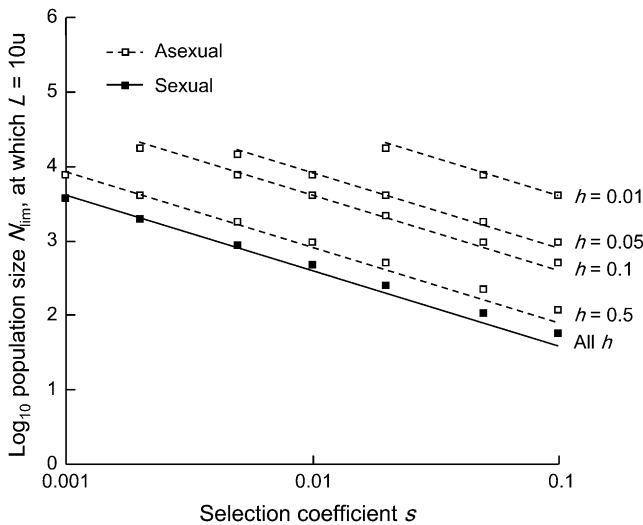


FIGURE 5.—Effect of selection coefficients s and dominance coefficients h on the limiting population size N_{lim} , below which genetic load $L > 0.0001$. Open (asexuals) and solid (sexuals) squares indicate different values of N_{lim} obtained from our model. The lines indicate the approximations, $N_{\text{lim}} = 4/s$ (sexuals, solid line) and $N_{\text{lim}} = 4/hs$ (asexuals, dashed lines). In sexuals, N_{lim} depends only minimally on h (Figure 4). Mutation parameters: $u = 10^{-5}$, $v = 10^{-7}$.

smaller (tending to u as h becomes 0). This is the effect studied by CHASNOV (2000). Multiplied across loci, it can compensate for the twofold cost of sex, but only for very small values of h . In contrast, in small populations we find a somewhat larger load in sexuals than in asexuals (see also Figure 1C). This is in agreement with the *small population approximation* given above and is due to the fact that small sexual populations are fixed either for AA or for aa most of the time, while asexual populations can be fixed for the Aa genotype.

It is difficult to extrapolate single-locus equilibrium results to a multilocus setting, but we follow previous work (CHASNOV 2000; AGRAWAL and CHASNOV 2001) in assuming complete independence of different loci and multiplicative fitness. This focuses only on the effects of segregation, neglecting any potential effects of linkage among loci. Our results thus underestimate the load in asexuals, as interference among loci can greatly reduce the effective population size of asexuals (HILL and ROBERTSON 1966; FELSENSTEIN 1974; CHARLESWORTH *et al.* 1993a; KEIGHTLEY and OTTO 2006; PALAND and LYNCH 2006). Figure 6 shows that the total fitness of sexuals relative to asexuals at equilibrium can greatly

exceed 2, even for conservative estimates of the genome-wide deleterious mutation rate U (HAAG-LIAUTARD *et al.* 2007). As already noted, the fitness of sexuals exceeds that of asexuals mainly for intermediate population sizes (and to a lower extent with large population sizes if h and s are small, see CHASNOV 2000). Conversely, for low population sizes, the fitness of sexuals is lower than that of asexuals (Figure 6). Therefore, these results indicate that sexuals may have a stronger advantage over asexuals at intermediate (rather than small or high) values of N_e . It would thus be interesting to study the effect of population size in a more dynamic model where an asexual mutant would spread within a sexual population, in which case the “population sizes” of sexuals and asexuals would vary over time.

Subdivided population: One might expect that increasing the degree of spatial structure of a large metapopulation (for example, by decreasing the size of local demes) has the same qualitative effect as decreasing N has in a single finite population. Indeed, population structure generates a variance in allele frequency between demes, which may lead to a similar effect as purging by drift in sexuals, but may also lead to local fixation of deleterious alleles if local drift is too strong. However, population structure increases the effective size of the total population (in the absence of any local demographic effects, *i.e.*, assuming constant deme sizes) and thus decreases the effect of drift on the change in allele frequencies in the whole metapopulation (WANG and CABALLERO 1999; ROUSSET 2003); therefore, the expected effect of increasing spatial structure will not be strictly equivalent to the effect of decreasing N_e in the single-population model. Figure 7 shows results for the genetic load, obtained using the island model of population structure without extinction. Here, we fixed the total population size at $N_T = nN = 10^4$ (where n is the number of demes and N is the number of adults per deme) and vary n and N . The x -axis in Figure 7 shows $\log_{10} N$, so one moves from 10^4 demes, each with a single individual, at the left, to a single population of 10^4 individuals at the right (*i.e.*, decreasing population structure). In sexuals, an effect similar to the purging by drift described in the single-population case occurs: a moderate degree of population structure increases the efficiency of selection (roughly when $h < \frac{1}{3}$), either under soft selection or under the present life cycle (WHITLOCK 2002; ROZE and ROUSSET 2004; THEODOROU and COUVET 2006). However, spatial structure also increases the

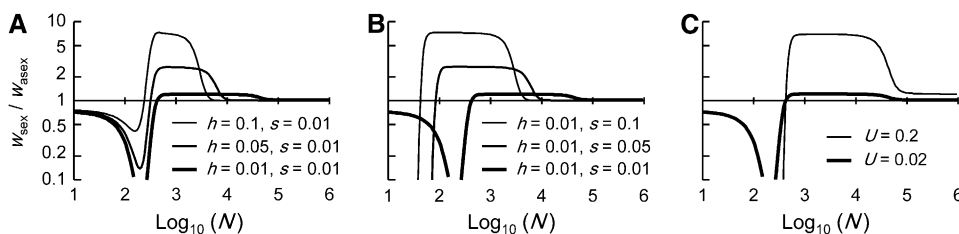


FIGURE 6.—Fitness of sexuals relative to asexuals for different values of h (A), s (B), and U (C). (A and B) $U = 0.02$; (C) $h = s = 0.01$. Per-locus mutation rates, $u = 10^{-5}$ and $v = 10^{-7}$; the number of loci, $n = 2000$ for $U = nu = 0.02$ and $n = 20,000$ for $U = 0.2$.

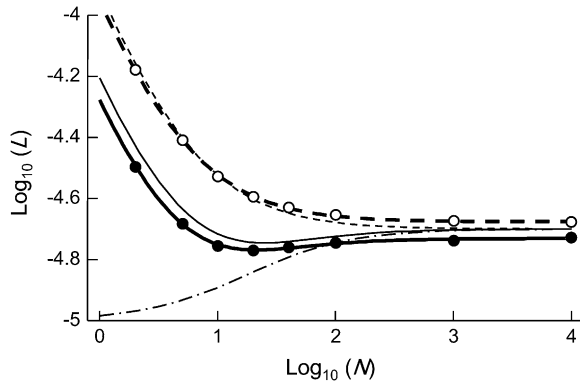


FIGURE 7.—Genetic load L as a function of deme size N in sexual (solid lines) and asexual (dashed lines) metapopulations without turnover ($e = 0$). Thick lines show diffusion results for sexuals (solid line) and asexuals (dashed line), while the open (asexuals) and solid (sexuals) circles show simulation results for $N = 2, 5, 10, 20, 40, 100, 1000$, and $10,000$ (average load over 10^8 generations). The thin lines correspond to approximations (10) and (11), while the dashed-dotted line shows the sexual load in AGRAWAL and CHASNOV'S (2001) model, with no local competition. Parameters: $n = 10^4/N$, $s = 0.01$, $h = 0.1$, $m = 0.1$, $u = 10^{-5}$, $v = 10^{-6}$.

degree of competition among related individuals (which tend to carry the same alleles), thereby reducing the effect of selection. As a result, population structure has a nonmonotonic effect on the mutation load in sexuals: the load first decreases slightly (with increasing population structure) and then increases with stronger population structure. Previous results have shown that, in general, population structure involving local competition does not greatly reduce the mutation load; rather, its main effect is that strong structure increases the load (GLÉMIN *et al.* 2003; ROZE and ROUSSET 2004; GLÉMIN 2005; THEODOROU and COUVET 2006). In asexuals, population structure does not improve the efficiency of selection; its only effect is to increase local competition, which increases the load. Although population structure increases the effective size of the whole metapopulation, this has little effect in the absence of extinction, for the parameter values used in Figure 7: N_e is always large, and the population is at mutation–selection balance over the whole range of deme sizes. Indeed, a deterministic solution (for an infinite population size N_T) gives a good approximation of our diffusion results. In the deterministic limit, and neglecting back mutation, the equilibrium frequency of allele a is $\sim u/S_1$ in the sexual case, where S_1 is given by Equation 2, while the frequency of the genotype Aa is $\sim 2u/S_{a1}$ in the asexual case, where S_{a1} is given by Equation 5. From this, one can obtain simple approximations for the load, assuming small m and large N ,

$$L_{\text{sex}} \approx u \frac{(1 + 2Nm)(1 + 8hNm)}{2Nm(1 + 4hNm)} \quad (17)$$

$$L_{\text{asex}} \approx 2u \frac{1 + 2Nm}{2Nm} \quad (18)$$

and thus

$$\frac{L_{\text{sex}}}{L_{\text{asex}}} \approx \frac{1 + 8hNm}{2 + 8hNm}. \quad (19)$$

Approximations (17) and (18) are shown in Figure 7 and exhibit a reasonable match with the diffusion and simulation results. Equation 19 indicates that, even when $h = \frac{1}{2}$ and when the effective size of the total population is infinite, sexuals may have a significantly lower load than asexuals. This contrasts with load in an infinite panmictic population, which, for both sexuals and asexuals, is very close to $2u$ when $h = \frac{1}{2}$. When $h = \frac{1}{2}$, S_1 simplifies to $(1 + r_0 - 2r_1)/2$ (while $S_2 = 0$), and S_{a1} simplifies to $1 - r_1$. The probabilities of identity by descent r_0 and r_1 can be shown to be equivalent to Wright's F -statistics F_{IT} and F_{ST} in a neutral infinite-island model (HUDSON 1998; ROUSSET 2002). Although our model includes selection, it is sufficient to use expressions for these F -statistics in a neutral model; indeed, taking into account the effect of selection on F_{IT} and F_{ST} would generate terms of order s^2 in the expressions of $E[\Delta p]$ and $\text{Var}[\Delta p]$ that would disappear in the diffusion limit. Using the relation $1 - F_{IT} = (1 - F_{IS})(1 - F_{ST})$, one arrives at the deterministic expressions (still for $h = \frac{1}{2}$):

$$L_{\text{sex}} = \frac{2u}{(1 - F_{ST})(1 + F_{IS})}, \quad L_{\text{asex}} = \frac{2u}{1 - F_{ST}}. \quad (20)$$

The factor $1 - F_{ST}$ comes from the fact that population structure decreases the efficiency of selection by increasing kin competition, while the factor $1 + F_{IS}$ corresponds to the fact that, in sexuals, two deleterious alleles can be eliminated at the same time when they occur in the same individual. Note that although we assume random mating within demes, F_{IS} is positive because it is measured after migration, which occurs in the diploid stage (and thus the two homologous genes of an individual are more likely to share a common ancestor than two genes in two different individuals from the same deme). Here, F_{IS} simply equals $1/(2N - 1)$: thus, this effect occurs mainly when deme size is small. More importantly, F_{ST} is not the same in the sexual and the asexual cases: in the asexual case, the F_{ST} that enters into Equation 20 corresponds to the F_{ST} of a *haploid* population ($\sim 1/[1 + 2Nm]$), which is higher than the F_{ST} of a diploid population with the same N and m ($\sim 1/[1 + 4Nm]$). For large N , and small m , we have:

$$\frac{1 - F_{ST}(\text{sex})}{1 - F_{ST}(\text{asex})} \approx 1 + \frac{1}{1 + 4Nm}. \quad (21)$$

This effect is equivalent to the result that, in a single population, asexuals have half the variance effective size of sexuals. Although here there is no drift at the whole-population level (as we are at the deterministic limit), local drift due to population structure (which reduces the efficiency of selection) is stronger in asexuals than in sexuals, generating a higher load in asexuals.

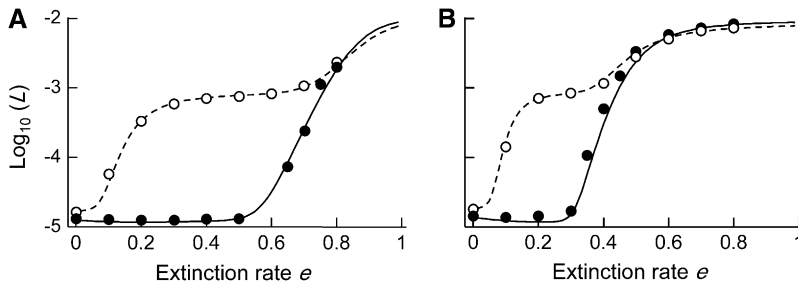


FIGURE 8.—Genetic load in sexual (solid lines) and asexual (dashed lines) metapopulations, as a function of the extinction rate e , for the migrant pool model of recolonization (A) and the propagule model (B). Solid (sexuals) and open (asexuals) circles show simulation results. Parameters: $n = 100$, $N = 100$, $s = 0.01$, $h = 0.1$, $m = 0.1$, $u = 10^{-5}$, $v = 10^{-6}$, $k = 10$.

Importantly, including local competition leads to a different effect of population structure than using a model of global competition as in AGRAWAL and CHASNOV (2001). The dashed-dotted curve in Figure 7 corresponds to Equation 5 in AGRAWAL and CHASNOV (2001) (load in sexuals), replacing f by $1/(1 + 4Nm)$ (note that this is not strictly correct, as it is difficult to imagine both global regulation and deme size staying constant over time, but it serves to illustrate their result). Here, the only effect of population structure is to increase homozygosity in sexuals (because individuals tend to mate with relatives), which leads to better purging and lowers the load. Conversely, population structure has no effect on the asexual load, which remains equal to $2u$ [that is, $\log_{10}(L) \approx -4.7$ for the parameters in Figure 7]. With local competition, the overall effect of population structure is to increase the load of both sexuals and asexuals (although a moderate degree of structure can slightly reduce the sexual load, as shown in Figure 7), the asexual load increasing faster than the sexual load.

Although population structure increases the effective size of metapopulations under the hypothesis of a constant deme size over time, adding local demography may reverse this relation (e.g., WHITLOCK and BARTON 1997). In particular, local extinctions may greatly reduce the effective size of a metapopulation. We investigated the effect of extinctions on the mutation load in sexual and asexual metapopulations using SLATKIN'S (1977) extinction–recolonization model (described above). Increasing the rate of deme extinctions (e) has similar effects to decreasing N_e in the case of a single population: as e increases, the importance of drift in the whole metapopulation increases, which may lead to the fixation (or quasi-fixation) of the deleterious allele. For intermediate values of e , selection against Aa individuals becomes ineffective in the case of asexuals, while selection remains efficient in the case of sexuals. For higher values of e , drift becomes stronger than selection in both sexuals and asexuals. This leads to stepwise increases of the genetic load, as shown in Figure 8, A (for the migrant pool model) and B (propagule pool model). In particular, Figure 8A shows that the asexual load may be far greater than the sexual load over a wide range of extinction rates, due to the lack of segregation (and a lower effective size) in asexuals.

DISCUSSION

In this article, we investigated the effects of population size and spatial structure on the mutation load in sexuals and asexuals. It is important to note that, in finite or structured populations, one cannot simply predict the outcome of competition between sexuals and asexuals by comparing their mutation loads at equilibrium: in a finite population, a mutation causing a transition to asexuality may reach fixation (if it occurs in a good genetic background) before mutation–selection balance is reached. In addition, in subdivided populations, one would have to account for the fact that asexuals compete more often with other asexuals (due to limited dispersal and local competition) than if the population was well mixed. Bearing this in mind, our results show that drift and population structure have different effects on the equilibrium mean fitness of sexuals and asexuals.

Single population: We found that, at intermediate population sizes, mutation load may be far greater in asexual populations than in sexual populations. Which population sizes are “intermediate” depends on the values of the parameters h and s , but, as a rule of thumb, we found that a 10-fold increase in mutation load compared to u was reached when $N < 4/s$ in sexuals and when $N < 4/hs$ in asexuals. In *Saccharomyces* and *Drosophila*, estimates indicate that $h > 0.01$ (SIMMONS and CROW 1977; LYNCH and WALSH 1998; SZAFRANIEC *et al.* 2003), so the parameter range at which drift overwhelms selection against heterozygotes in asexuals but is still effective in sexuals should span less than two orders of magnitude in N . However, as we discuss below, our estimate of mutation load in asexuals may be a gross underestimate when deleterious mutations occur at many loci, because it neglects the effects of clonal interference on the effective population size. Hence, it is possible that selection against heterozygotes in asexuals is ineffective even in populations $\gg N = 4/hs$.

Calculations of genetic loads are useful to assess the effects of segregation on the efficiency of selection in equilibrium situations. In contrast, to study the gradual evolution of sexual reproduction, the immediate costs (in terms of mean offspring fitness) of breaking genetic associations that have been generated by selection need to be taken into account (see review by AGRAWAL 2006).

This can be done, for instance, by studying the evolution of a modifier locus with different alleles that increase or decrease the proportion of offspring produced sexually. Using such a modifier model, OTTO (2003) found that, in infinite populations, conditions under which an advantage of segregation leads to an increase in frequency of a modifier promoting sexual reproduction are quite restricted, unless there is some inbreeding, which causes an excess of homozygotes. It may be argued that genetic drift has similar effects to inbreeding because both processes may lead to purging in sexuals. However, drift also leads to random changes in allele frequencies, which counteract the effectiveness of selection and may therefore generate a greater advantage for sex at intermediate population sizes, when deleterious alleles are fixed in the heterozygous state in asexuals, but not in sexuals. It would be interesting to see how this effect translates into an advantage for sex in a modifier model representing a finite diploid population.

A potential shortcoming of our model is that it neglects mitotic recombination in asexuals (BARBERA and PETES 2006; OMILIAN *et al.* 2006). Mitotic recombination leads to formation of homozygotes from heterozygous asexuals and is equivalent to increasing mutation rates from heterozygotes to homozygotes. High rates of mitotic recombination (of the order of magnitude of mutation rates or higher) may decrease the effects observed here because fixation in the heterozygous state becomes less common, if heterozygotes often “back mutate” to homozygotes. However, the significance of mitotic recombination in nature is unknown. For instance, if it involves long stretches of DNA containing many loci (OMILIAN *et al.* 2006), most cases of mitotic recombination may be deleterious because mutant homozygotes will be created at some loci.

Subdivided population: In large metapopulations without extinction and recolonization, drift at the global level can be neglected. In such metapopulations, the effect of increasing population structure (by decreasing deme size or migration rate) on the mutation load depends critically on the assumptions concerning population regulation. Here we found that the realistic assumption of local population regulation (*i.e.*, competition within local populations) has a strong effect on mutation load in both sexuals and asexuals. Local competition reduces the efficiency of selection, as individuals who tend to have similar genotypes compete with each other. Hence, in our model, the overall effect of population structure is to increase the load in both sexuals and asexuals (although a moderate degree of structure can lead to a small reduction of the load in sexuals). This effect is stronger in asexuals than in sexuals, because local drift has more effect in asexuals. Indeed, within a deme, the variance effective size of asexuals is half the effective size of sexuals. In the deterministic limit (very large total population size), we found that $L_{\text{sex}}/L_{\text{asex}} \approx (1 + 8hNm)/(2 + 8hNm)$. As for the case of a single population, it would be

interesting to model the evolution of a modifier locus affecting the proportion of sexually produced offspring in a metapopulation. This could be done using the method developed in ROZE and ROUSSET (2005).

Local demography can have important effects on the effective size of metapopulations (WHITLOCK and BARTON 1997). Here we used a very simple demographic model, where local extinctions occur at a rate e and where deme size immediately goes back to N after recolonization (SLATKIN 1977). Increasing the extinction rate reduces the effective size of the total population. Hence we found that, for a wide range of extinction rates, deleterious mutations may be fixed in the heterozygous state in asexuals, while still being efficiently eliminated in sexuals. More realistic demographic models of subdivided populations could be explored, for example, by using the method of ROUSSET and RONCE (2004).

Mutation load in small populations: In contrast to populations of intermediate size, sexuals have a similar or even a higher genetic load than asexuals in small populations (single-population model) and also in metapopulations with very high turnover rates (metapopulation model, large e). This occurs for parameter ranges at which selection is ineffective against all genotypes in both sexuals and asexuals and so that populations are fixed for the mutant genotype aa most of the time (when $u \gg v$).

The result that the improved effectiveness of selection in sexuals disappears, or is even reversed compared to that in asexuals at low effective population size, may be consistent with the well-documented pattern of geographic parthenogenesis (VANDEL 1928): asexuals often occur in more marginal (as opposed to core) habitats, such as at high latitudes or altitudes, than their closely related sexual counterparts (*e.g.*, BELL 1982; BIERZYCHUDEK 1985). Because marginal habitats may be environmentally less predictable and more patchily distributed than core habitats, populations in these habitats may tend to be smaller, subject to wider stochastic fluctuations in density, and hence subject to stronger drift (HAAG and EBERT 2004). If drift in marginal populations is indeed so strong that it overwhelms selection in asexuals and in sexuals, our model predicts that, in marginal habitats, asexuals have a slightly lower load than sexuals.

It is, however, unclear whether equilibrium results (such as those described here) apply to situations in which genetic drift is so strong that a majority of deleterious mutations are effectively neutral. If this were the case at many loci, deleterious mutations may accumulate through fixation and Muller’s ratchet in both sexual and asexual populations (although the effects differ between sexuals and asexuals, *e.g.*, PAMILO *et al.* 1987; CHARLESWORTH *et al.* 1993b; CHARLESWORTH and CHARLESWORTH 1997). This may eventually drive the populations to extinction (LYNCH *et al.* 1995), and, thus, equilibrium conditions would not be met. Nonetheless, if sexual populations fix mutations in a homozygous

state, whereas asexual ones fix them first as heterozygotes, the fitness decline may be faster in sexuals than in asexuals.

Multilocus simulation studies suggest that the advantage of recombination may also increase with population size (even without leveling off, at least for the studied parameter range; ILES *et al.* 2003; KEIGHTLEY and OTTO 2006; SALATHÉ *et al.* 2006), due to the fact that a greater number of selected mutations segregate in larger populations. Hence these studies, as well as ours, are consistent with the idea that the fitness of sexuals relative to asexuals increases with population size (in our study only up to a certain point). Again, this may be a possible explanation for geographical parthenogenesis, if marginal populations have a lower effective size than core populations.

Extrapolation to multilocus situations: The models presented here are single-locus models with recurrent deleterious mutations, but in real organisms mutations occur at many loci throughout the genome. It is difficult to extrapolate from single-locus models to multilocus situations. Nonetheless, our model suggests that, if deleterious alleles are at least partly recessive, asexuals suffer an increased load compared to sexuals at intermediate population sizes. Multiplied across loci, this could substantially reduce genetic load in sexual populations even for conservative estimates of U and in the absence of any effect of recombination. Neglecting potential effects of recombination (or its absence in asexuals) in a multilocus setting is, of course, unrealistic, but was done here to isolate the effects of segregation from those of recombination. In asexuals, the effectiveness of selection can be greatly reduced below levels predicted by single-locus models because of strong selective interference among nonrecombining loci (the “Hill–Robertson effect”; HILL and ROBERTSON 1966). Hence single-locus models underestimate the force of drift, and hence load, in asexuals. Selective interference also occurs in sexuals, but is much weaker, due to recombination (HILL and ROBERTSON 1966; FELSENSTEIN 1974; CHARLESWORTH *et al.* 1993a; KEIGHTLEY and OTTO 2006). Thus, it seems likely that combining the effects of segregation and recombination would lead to an increased parameter range in which asexuals have an increased load compared to sexuals.

We thank S. Glémin for discussion and for sharing his algorithm and B. Charlesworth, D. Charlesworth, P. Keightley, S. Otto, and an anonymous reviewer for discussion and comments on earlier versions of the manuscript. We are also grateful to S. Otto for pointing out the full extent of the analogy to the Hill–Robertson effect. This work was supported by the Swiss National Science Foundation, a Marie Curie Fellowship (C.R.H.), and a European Molecular Biology Organization Fellowship (D.R.).

LITERATURE CITED

- AGRAWAL, A. F., 2006 Evolution of sex: why do organisms shuffle their genotypes. *Curr. Biol.* **16**: R696–R704.
- AGRAWAL, A. F., and J. R. CHASNOV, 2001 Recessive mutations and the maintenance of sex in structured populations. *Genetics* **158**: 913–917.
- BALLOUX, F., L. LEHMANN and T. DE MEEÛS, 2003 The population genetics of clonal and partially clonal diploids. *Genetics* **164**: 1635–1644.
- BARBERA, M. A., and T. D. PETES, 2006 Selection and analysis of spontaneous reciprocal mitotic cross-overs in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **103**: 12819–12824.
- BARTON, N. H., 1995 Linkage and the limits to natural selection. *Genetics* **140**: 821–841.
- BARTON, N. H., and B. CHARLESWORTH, 1998 Why sex and recombination. *Science* **281**: 1986–1990.
- BARTON, N. H., and S. P. OTTO, 2005 Evolution of recombination due to genetic drift. *Genetics* **169**: 2353–2370.
- BATAILLON, T., and M. KIRKPATRICK, 2000 Inbreeding depression due to mildly deleterious mutations in finite populations: size does matter. *Genet. Res.* **75**: 75–81.
- BELL, G., 1982 *The Masterpiece of Nature: The Evolution and Genetics of Sexuality*. University of California Press, Berkeley, CA.
- BIERZYCHUDEK, P., 1985 Patterns in plant parthenogenesis. *Experientia* **41**: 1255–1264.
- CABALLERO, A., and W. G. HILL, 1992 Effects of partial inbreeding on fixation rates and variation of mutant genes. *Genetics* **131**: 493–507.
- CHARLESWORTH, B., 1990 Mutation–selection balance and the evolutionary advantage of sex and recombination. *Genet. Res.* **55**: 199–221.
- CHARLESWORTH, B., and D. CHARLESWORTH, 1997 Rapid fixation of deleterious alleles can be caused by Muller’s ratchet. *Genet. Res.* **70**: 63–73.
- CHARLESWORTH, B., M. T. MORGAN and D. CHARLESWORTH, 1993a The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**: 1289–1303.
- CHARLESWORTH, D., M. T. MORGAN and B. CHARLESWORTH, 1993b Mutation accumulation in finite populations. *J. Hered.* **84**: 321–325.
- CHASNOV, J. R., 2000 Mutation selection balance, dominance and the maintenance of sex. *Genetics* **156**: 1419–1425.
- CHERRY, J. L., and J. WAKELEY, 2003 A diffusion approximation for selection and drift in a subdivided population. *Genetics* **163**: 421–428.
- CROW, J. F., 1970 Genetic loads and the cost of natural selection, pp. 128–177 in *Mathematical Topics in Population Genetics*, edited by K. KOJIMA. Springer-Verlag, New York.
- CROW, J. F., and M. KIMURA, 1970 *An Introduction to Population Genetics Theory*. Harper & Row, New York.
- DE VISSER, J. A. G. M., and S. F. ELENA, 2007 The evolution of sex: empirical insights into the roles of epistasis and drift. *Nat. Rev. Genet.* **8**: 139–149.
- FELSENSTEIN, J., 1974 The evolutionary advantage of recombination. *Genetics* **78**: 737–756.
- GALE, J. S., 1990 *Theoretical Population Genetics*. Unwin Hyman, London.
- GLÉMIN, S., 2003 How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution* **57**: 2678–2687.
- GLÉMIN, S., 2005 Lethals in subdivided populations. *Genet. Res.* **86**: 41–51.
- GLÉMIN, S., J. RONFORT and T. BATAILLON, 2003 Patterns of inbreeding depression and architecture of the load in subdivided populations. *Genetics* **165**: 2193–2212.
- HAAG, C. R., and D. EBERT, 2004 A new hypothesis to explain geographic parthenogenesis. *Ann. Zool. Fenn.* **41**: 539–544.
- HAAG-LIAUTARD, C., M. DORRIS, X. MASIDE, S. MACASKILL, D. L. HALLIGAN *et al.*, 2007 Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*. *Nature* **445**: 82–85.
- HILL, W. G., and A. ROBERTSON, 1966 The effect of linkage on limits to artificial selection. *Genet. Res.* **8**: 269–294.
- HUDSON, R. R., 1998 Island models and the coalescent process. *Mol. Ecol.* **7**: 413–418.
- ILES, M. M., K. WALTERS and C. CANNINGS, 2003 Recombination can evolve in large finite populations given selection on sufficient loci. *Genetics* **165**: 2249–2258.
- KEIGHTLEY, P. D., and S. P. OTTO, 2006 Interference among deleterious mutations favours sex and recombination in finite populations. *Nature* **443**: 89–92.

- KIMURA, M., and T. MARUYAMA, 1966 The mutation load with epistatic gene interactions in fitness. *Genetics* **54**: 1337–1351.
- KIMURA, M., T. MARUYAMA and J. F. CROW, 1963 The mutation load in small populations. *Genetics* **48**: 1303–1312.
- KONDRASHOV, A. S., 1982 Selection against harmful mutations in large sexual and asexual populations. *Genet. Res.* **40**: 325–332.
- LYNCH, M., and B. WALSH, 1998 *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.
- LYNCH, M., J. CORNER and R. BÜRGER, 1995 Mutation accumulation and the extinction of small populations. *Am. Nat.* **146**: 489–518.
- MARTIN, G., S. P. OTTO and T. LENORMAND, 2006 Selection for recombination in structured populations. *Genetics* **172**: 593–609.
- MAYNARD SMITH, J., 1978 *The Evolution of Sex*. Cambridge University Press, Cambridge.
- MULLER, H. J., 1950 Our load of mutations. *Am. J. Hum. Genet.* **2**: 111–176.
- OMILIAN, A. R., M. E. A. CRISTESCU, J. L. DUDYCHA and M. LYNCH, 2006 Ameiotic recombination in asexual lineages of *Daphnia*. *Proc. Natl. Acad. Sci. USA* **103**: 18638–18643.
- OTTO, S. P., 2003 The advantages of segregation and the evolution of sex. *Genetics* **164**: 1099–1118.
- OTTO, S. P., and N. BARTON, 1997 The evolution of recombination: removing the limits to natural selection. *Genetics* **147**: 879–906.
- OTTO, S. P., and N. BARTON, 2001 Selection for recombination in small populations. *Evolution* **55**: 1921–1931.
- OTTO, S. P., and T. LENORMAND, 2002 Resolving the paradox of sex and recombination. *Nat. Rev. Genet.* **3**: 252–261.
- PALAND, S., and M. LYNCH, 2006 Transitions to asexuality result in excess amino acid substitutions. *Science* **311**: 990–992.
- PAMILO, P., M. NEI and W. LI, 1987 Accumulation of mutations in sexual and asexual populations. *Genet. Res.* **49**: 135–146.
- RICE, W. R., 2002 Experimental tests of the adaptive significance of sexual reproduction. *Nat. Rev. Genet.* **3**: 241–251.
- ROUSSET, F., 2002 Inbreeding and relatedness coefficients: What do they measure? *Heredity* **88**: 371–380.
- ROUSSET, F., 2003 Effective size in simple metapopulation models. *Heredity* **91**: 107–111.
- ROUSSET, F., and O. RONCE, 2004 Inclusive fitness for traits affecting metapopulation demography. *Theor. Popul. Biol.* **65**: 127–141.
- ROZE, D., and N. BARTON, 2006 The Hill–Robertson effect and the evolution of recombination. *Genetics* **173**: 1793–1811.
- ROZE, D., and T. LENORMAND, 2005 Self-fertilization and the evolution of recombination. *Genetics* **170**: 841–857.
- ROZE, D., and F. ROUSSET, 2003 Selection and drift in subdivided populations: a straightforward method for deriving diffusion approximations and applications involving dominance, selfing and local extinctions. *Genetics* **165**: 2153–2166.
- ROZE, D., and F. ROUSSET, 2004 Joint effects of self-fertilization and population structure on mutation load, inbreeding depression and heterosis. *Genetics* **167**: 1001–1015.
- ROZE, D., and F. ROUSSET, 2005 Inbreeding depression and the evolution of dispersal rates: a multilocus model. *Am. Nat.* **166**: 708–721.
- SALATHÉ, M., R. SALATHÉ, P. SCHMID-HEMPEL and S. BONHOEFFER, 2006 Mutation accumulation in space and the maintenance of sexual reproduction. *Ecol. Lett.* **9**: 941–946.
- SIMMONS, M. J., and J. F. CROW, 1977 Mutations affecting fitness in *Drosophila* populations. *Annu. Rev. Genet.* **11**: 49–78.
- SLATKIN, M., 1977 Gene flow and genetic drift in a species subject to frequent local extinctions. *Theor. Popul. Biol.* **12**: 253–262.
- SZAFRANIEC, K., D. M. WLOCH, P. SLIWA, R. H. BORTS and R. KORONA, 2003 Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous loci of the yeast *Saccharomyces cerevisiae*. *Genet. Res.* **82**: 19–31.
- THEODOROU, K., and D. COUVET, 2006 Genetic load in subdivided populations: interactions between the migration rate, the size and the number of subpopulations. *Heredity* **96**: 69–78.
- VANDEL, A., 1928 La parthénogénèse géographique. Contribution à l'étude biologique et cytologique de la parthénogénèse naturelle. *Bull. Biol. Fr. Belg.* **62**: 164–281.
- WANG, J. L., and A. CABALLERO, 1999 Developments in predicting the effective size of subdivided populations. *Heredity* **82**: 212–226.
- WHITLOCK, M. C., 2002 Selection, load and inbreeding depression in a large metapopulation. *Genetics* **160**: 1191–1202.
- WHITLOCK, M. C., and N. H. BARTON, 1997 The effective size of a subdivided population. *Genetics* **146**: 427–441.
- WOLFRAM, S., 2003 *The Mathematica Book*. Cambridge University Press, Cambridge.
- WRIGHT, S., 1937 The distribution of gene frequencies in populations. *Proc. Natl. Acad. Sci. USA* **23**: 307–320.

Communicating editor: H. G. SPENCER

APPENDIX A

We derive here expressions for the expected change in frequency of the deleterious allele (a) over one generation, in the island model of population structure, with extinctions and recolonizations. Note that to use the diffusion method, it is sufficient to express this expected change in the limit as the number of demes n tends to infinity (e.g., ROZE and ROUSSET 2003). Our life cycle assumes that at each generation, each deme may go extinct with probability e . In nonextinct demes, individuals produce a large but, depending on their fitness, variable number of gametes, which fuse immediately and at random within the deme (in the sexual case), or juveniles (in the asexual case). Each juvenile then has a probability m of entering the migrant pool. Therefore, demes contribute to the migrant pool on proportion of their mean fecundity. Migrants reach any other deme with the same probability, but following SLATKIN'S (1977) extinction–recolonization model we assume that migrants arriving in an extinct deme do not survive, each extinct deme being recolonized later by k juveniles. In the absence of extinction, the expected contribution of adult j in deme i to the next adult generation is given by

$$\frac{(1-m)w_{ij}}{(1-m)w_i + mw} + \frac{mw_{ij}}{w}, \quad (\text{A1})$$

where w_{ij} is the fecundity of individual ij , w_i is the average fecundity in deme i , and w is the average fecundity in the whole metapopulation; w_i is given by

$$\begin{aligned} w_i &= p_{AA,i} + (1-hs)p_{Aa,i} + (1-s)p_{aa,i} \\ &= 1 - 2hsp_{a,i} - s(1-2h)p_{aa,i}, \end{aligned} \quad (\text{A2})$$

where $p_{Aa,i}$, $p_{aa,i}$, and $p_{a,i}$ are the frequencies of Aa individuals, of aa individuals, and of the a allele in deme i , before selection. When extinctions occur, the backward migration rate (probability that, after dispersal, a juvenile comes

from another deme) is different from m and is given in the neutral case by $m_b = (1 - e)m/(1 - me)$. Although selection affects m_b , this effect generates a term of order s^2 in the expression of the change in frequency of the deleterious allele, which disappears in the diffusion limit. In the sexual case, the frequency of a in deme i (given that deme i does not go extinct) after selection (and before dispersal) is given by

$$p_{a,i}^s = \frac{(1 - hs)p_{Aa,i}/2 + (1 - s)p_{aa,i}}{w_i} = \frac{(1 - hs)p_{a,i} + (1 - h)sp_{aa,i}}{w_i}, \quad (\text{A3})$$

where w_i is given by equation (A2) above. The expected frequency of a in deme i , after dispersal and recolonization of extinct demes is then given by

$$E[p_{a,i}'] = ep_a^s + (1 - e) \frac{(1 - m_b)w_i p_{a,i}^s + m_b w(1 - e)p_a^s}{(1 - m_b)w_i + m_b w(1 - e)}, \quad (\text{A4})$$

where w is again the average fecundity in the whole metapopulation (the average over i of w_i), and where p_a^s is the frequency of a in the whole population, after selection, given by

$$p_a^s = \frac{\overline{w_i}}{w} p_{a,i}^s, \quad (\text{A5})$$

where the overbar means the average over all demes i . The first and second terms of Equation A4 represent the case where deme i went extinct and the case where it did not go extinct (respectively). In this second case, $(1 - m_b)w_i$ juveniles come from deme i , while $m_b w(1 - e)$ juveniles are migrants coming from other nonextinct demes. After expressing Equation A4 to the first order in s and averaging over all i , one obtains an expression for Δp_a (the change in frequency of a over one generation) as a function of the moments $\overline{(p_{a,i})^2}$ and $\overline{p_{a,i}p_{aa,i}}$ (where again the overbar means the average over all i). When selection is weaker than migration, we can use a separation-of-timescales argument and replace these moments by their equilibrium expressions under neutrality, for the current allele frequencies (e.g., WHITLOCK 2002; CHERRY and WAKELEY 2003; ROZE and ROUSSET 2003). This gives

$$\overline{(p_{a,i})^2} = r_1^R p_a + (1 - r_1^R) p_a^2 \quad (\text{A6})$$

$$\overline{p_{a,i}p_{aa,i}} = \alpha^R p_a + \beta^R p_a^2 + \gamma^R p_a^3, \quad (\text{A7})$$

where r_1^R is the probability that the ancestral lineages of two genes sampled with replacement from the same deme stay in the same deme and coalesce (in the neutral case and for an infinite number of demes), while α^R , β^R , and γ^R are the probabilities that the ancestral lineages of the two homologous genes of an individual, plus a third gene sampled with replacement from the same deme, all stay in the same deme and coalesce (α^R), that only two of them coalesce (β^R), or that no coalescence occurs ($\gamma^R = 1 - \alpha^R - \beta^R$). These can be expressed in terms of probabilities of coalescence of genes sampled *without* replacement,

$$r_1^R = \frac{1 + r_0}{2N} + \left(1 - \frac{1}{N}\right) r_1, \quad (\text{A8})$$

where r_0 and r_1 are the probabilities that the two homologous genes of an individual, and two genes sampled from two different individuals from the same deme (respectively), coalesce within the deme. Similarly, one has

$$\alpha^R = \frac{r_0}{N} + \left(1 - \frac{1}{N}\right) \alpha \quad (\text{A9})$$

$$\gamma^R = \left(1 - \frac{1}{N}\right) \gamma, \quad (\text{A10})$$

where α and γ are the probabilities that the two homologous genes of an individual, plus a gene from a different individual from the same deme, all coalesce (α) or that no coalescence occurs (γ). After some simplifications, and using the fact that $\gamma = 1 + 2\alpha - r_0 - 2r_1$, one arrives at Equation 1 in the text.

The coalescence probabilities r_0 , r_1 , and α are obtained by solving recurrence equations. For this, one defines ϕ (and ψ) as the probability that two (three) individuals colonizing the same extinct deme come from the same deme (before

dispersal). In the migrant pool model $\varphi = \psi = 0$, while in the propagule pool model $\varphi = (1 - m_b)^2$ and $\psi = (1 - m_b)^3$. Recursions for r_0 and r_1 are given by

$$r'_0 = \left[1 - e + e \left(\frac{1}{2k} + \left(1 - \frac{1}{k} \right) \varphi \right) \right] \left[\frac{1 + r_0}{2N} + \left(1 - \frac{1}{N} \right) r_1 \right] + \frac{e}{2k} \quad (\text{A11})$$

$$r'_1 = \left[(1 - e)(1 - m_b)^2 + e \left(\frac{1}{2k} + \left(1 - \frac{1}{k} \right) \varphi \right) \right] \left[\frac{1 + r_0}{2N} + \left(1 - \frac{1}{N} \right) r_1 \right] + \frac{e}{2k}. \quad (\text{A12})$$

Defining δ as the probability that the ancestral lineages of three genes sampled from three different individuals from the same deme stay in the same deme and coalesce, one then has

$$\begin{aligned} \alpha' &= \left[(1 - e)(1 - m_b)^2 + e \left(1 - \frac{1}{k} \right) \left(\frac{3}{2k} \varphi + \left(1 - \frac{2}{k} \right) \psi \right) \right] \\ &\times \left[\frac{1 + 3r_0}{4N^2} + \frac{3}{2N} \left(1 - \frac{1}{N} \right) (r_1 + \alpha) + \left(1 - \frac{1}{N} \right) \left(1 - \frac{2}{N} \right) \delta \right] \\ &+ \frac{3e}{2k} \left[\frac{1}{2k} + \left(1 - \frac{1}{k} \right) \varphi \right] \left[\frac{1 + r_0}{2N} + \left(1 - \frac{1}{N} \right) r_1 \right] + \frac{e}{4k^2}. \end{aligned} \quad (\text{A13})$$

Finally, δ' is given by the same equation, replacing $(1 - m_b)^2$ by $(1 - m_b)^3$. Solving these equations gives expressions for r_0 , r_1 , and α in terms of N , m , e , and k .

Using the same method for the asexual case yields an expression that depends on r_1^R , the probability that the ancestral lineages of two genes sampled with replacement from the same deme stay in the same deme and coalesce, this time in a *haploid* model. This is given by

$$r_1^R = \frac{1}{N} + \left(1 - \frac{1}{N} \right) r_1, \quad (\text{A14})$$

where r_1 is obtained by solving the recurrence equation

$$r'_1 = \left[(1 - e)(1 - m_b)^2 + e \left(1 - \frac{1}{k} \right) \varphi \right] \left[\frac{1}{N} + \left(1 - \frac{1}{N} \right) r_1 \right] + \frac{e}{k}. \quad (\text{A15})$$

This method leads to Equations 1–8 in the text, describing the expectation and variance of the change in allele frequency in sexual and asexual metapopulations. These expressions take the same form as in a single, nonsubdivided population, and standard diffusion methods (e.g., CROW and KIMURA 1970) yield

$$\phi_s(p) = K_s \exp[4N_c S_1 p + 2N_c S_2 p^2] p^{4N_c u - 1} (1 - p)^{4N_c v - 1}, \quad (\text{A16})$$

where $\phi_s(p)$ is the probability that allele a is present in frequency p in the metapopulation at equilibrium, in the sexual case, K_s is a constant, S_1 and S_2 are given by Equations 2 and 3, and $N_c = p(1 - p)/(2 \text{Var}[\Delta p])$, where $\text{Var}[\Delta p]$ is given by Equation 8 (see also ROZE and ROUSSET 2004). In the first asexual model (describing a population consisting of Aa and AA individuals), the equilibrium probability distribution of the frequency of Aa individuals is given by

$$\phi_{a1}(p) = K_{a1} \exp[2N_c S_{a1}] p^{4N_c u - 1} (1 - p)^{2N_c v - 1}, \quad (\text{A17})$$

where S_{a1} is given by Equation 5, and $N_c = p(1 - p)/(\text{Var}[\Delta p])$. In the second asexual model (describing a population consisting of aa and Aa individuals), the equilibrium probability distribution of the frequency of aa individuals is given by

$$\phi_{a2}(p) = K_{a2} \exp[2N_c S_{a2}] p^{2N_c u - 1} (1 - p)^{4N_c v - 1}, \quad (\text{A18})$$

where S_{a2} is given by Equation 7, and again $N_c = p(1 - p)/(\text{Var}[\Delta p])$. We then used Mathematica to integrate numerically over these distributions, to obtain \bar{p} and \bar{p}^2 in the sexual case and $\overline{p_{Aa}}$, $\overline{p_{aa}}$ in the asexual case.

APPENDIX B

We calculate here the equilibrium value of the intralocus association D_{intra} at the mutation–drift equilibrium (no selection), in the sexual and asexual cases. We have

$$\begin{aligned}
 D_{\text{intra}} &= \overline{p_{aa}} - \overline{p_a^2} \\
 &= \overline{p_{aa}} - \left(\overline{p_{aa}^2} + \overline{p_{Aa}p_{Aa}} + \frac{\overline{p_{Aa}^2}}{4} \right).
 \end{aligned}
 \tag{B1}$$

D_{intra} can thus be expressed by calculating $\overline{p_{aa}}$, $\overline{p_{aa}^2}$, $\overline{p_{Aa}p_{Aa}}$, and $\overline{p_{Aa}^2}$ at equilibrium. In the asexual case, recursions for $\overline{p_{aa}}$ and $\overline{p_{Aa}}$ are given by (to the first order in u and v)

$$\overline{p_{aa}'} = (1 - 2v)\overline{p_{aa}} + u\overline{p_{Aa}} \tag{B2}$$

$$\overline{p_{Aa}'} = 2u + (1 - 3u - v)\overline{p_{Aa}} + 2(v - u)\overline{p_{aa}}, \tag{B3}$$

giving at equilibrium $\overline{p_{aa}} = u^2/(u + v)^2$, $\overline{p_{Aa}} = 2uv/(u + v)^2$. A recursion for $\overline{p_{aa}^2}$ is obtained as follows: $\overline{p_{aa}^2}'$ is the probability that, at the next generation, two individuals sampled with replacement from the whole population are both aa . With probability $1/N$, the same individual is sampled twice, and this individual is aa with probability $\overline{p_{aa}'}$. With probability $1 - 1/N$, two different individuals are sampled, in which case they are both aa with probability $(\overline{p_{aa}^u})^2$, where $\overline{p_{aa}^u}$ is the frequency of aa after mutation, and before random sampling ($\overline{p_{aa}^u} = (1 - 2v)p_{aa} + up_{Aa}$). This gives

$$\overline{p_{aa}^2}' = \frac{\overline{p_{aa}'}}{N} + \left(1 - \frac{1}{N}\right) \left[(1 - 4v)\overline{p_{aa}^2} + 2u\overline{p_{Aa}p_{aa}} \right]. \tag{B4}$$

Similarly, one obtains the following recursions for $\overline{p_{Aa}^2}$ and $\overline{p_{Aa}p_{aa}}$:

$$\overline{p_{Aa}^2}' = \frac{\overline{p_{Aa}'}}{N} + \left(1 - \frac{1}{N}\right) \left[(1 - 6u - 2v)\overline{p_{Aa}^2} + 4u\overline{p_{Aa}} + 4(v - u)\overline{p_{Aa}p_{aa}} \right] \tag{B5}$$

$$\overline{p_{Aa}p_{aa}'} = \left(1 - \frac{1}{N}\right) \left[2u\overline{p_{aa}} + (1 - 3u - 3v)\overline{p_{Aa}p_{aa}} + u\overline{p_{Aa}^2} + 2(v - u)\overline{p_{aa}^2} \right]. \tag{B6}$$

Solving these recursions gives, at equilibrium,

$$D_{\text{intra}} = -\frac{\theta}{2(1 + \theta)^2[1 + 2(N - 1)u(1 + \theta)]}, \tag{B7}$$

where $\theta = v/u$. In the sexual case, one can write recursions for $\overline{p_{aa}}$ and $\overline{p_a^2}$,

$$\overline{p_a^2}' = \frac{\overline{p_a'}}{2N} + \left(1 - \frac{1}{2N}\right) \overline{(p_a^u)^2}, \tag{B8}$$

where $\overline{p_a^u}$ is the frequency of a after mutation, given by $\overline{p_a^u} = u + (1 - u - v)p_a$. Therefore,

$$\overline{p_a^2}' = \frac{\overline{p_a'}}{2N} + \left(1 - \frac{1}{2N}\right) \left[2u\overline{p_a} + (1 - 2u - 2v)\overline{p_a^2} \right] \tag{B9}$$

and

$$\overline{p_{aa}'} = \overline{(p_a^u)^2} = 2u\overline{p_a} + (1 - 2u - 2v)\overline{p_a^2}. \tag{B10}$$

Given that $p_a = u/(u + v)$, one obtains at equilibrium

$$D_{\text{intra}} = -\frac{2\theta u}{(1 + \theta)[1 + 2(2N - 1)u(1 + \theta)]}. \tag{B11}$$