Accumulation of mutations in sexual and asexual populations

PEKKA PAMILO*, MASATOSHI NEI and WEN-HSIUNG LI

Center for Demographic and Population Genetics, University of Texas Health Science Center at Houston, P.O. Box 20334, Houston, Texas 77225, U.S.A.

Summary

The accumulation of beneficial and harmful mutations in a genome is studied by using analytical methods as well as computer simulation for different modes of reproduction. The modes of reproduction examined are biparental (bisexual, hermaphroditic), uniparental (selfing, automictic, asexual) and mixed (partial selfing, mixture of hermaphroditism and parthenogenesis). It is shown that the rates of accumulation of both beneficial and harmful mutations with weak selection depend on the within-population variance of the number of mutant genes per genome. Analytical formulae for this variance are derived for neutral mutant genes for hermaphroditic, selfing and asexual populations; the neutral variance is largest in a selfing population and smallest in an asexual population. Directional selection reduces the population variance in most cases, whereas recombination partially restores the reduced variance. Therefore, biparental organisms accumulate beneficial mutations at the highest rate and harmful mutations at the lowest rate. Selfing organisms are intermediate between biparental and asexual organisms. Even a limited amount of outcrossing in largely selfing and parthenogenetic organisms markedly affects the accumulation rates. The accumulation of mutations is likely to affect the mean population fitness only in long-term evolution.

1. Introduction

The main genetic feature that distinguishes asexual reproduction from sexual reproduction is the lack of recombination; the offspring duplicate the parental genotype. Therefore, different advantageous mutations cannot be incorporated into the population unless they arise in the same individual (Fisher, 1930; Muller, 1932; Crow & Kimura, 1965). Another consequence is the accumulation of harmful mutations, a phenomenon commonly referred to as Muller's ratchet (Muller, 1964; Felsenstein, 1974). In asexual organisms, the offspring will inherit all the deleterious mutant alleles from the parent unless back mutation occurs. Because back mutation has to occur at a specific locus whereas new harmful mutations can arise at any locus, harmful mutant genes keep accumulating at least in the initial stage (Maynard Smith, 1978). Harmful mutations can also accumulate in a sexual population, but the rate of accumulation is expected to be slower than that in an asexual population (Felsenstein, 1974).

The accumulation of beneficial mutations at a higher rate and the accumulation of harmful mutations at a lower rate have been suggested as the long-term advantages of sexual reproduction or recombination (Muller, 1932, 1964), and they have played an important role in the evolutionary theory of sex (Maynard Smith, 1978; Bell, 1982). In order to understand the real evolutionary significance of sex or recombination, however, we must quantify the effect of recombination on the rates of accumulation of mutant genes. At the present time no rigorous analytical solution exists, and the few quantitative studies available are based mainly on computer simulations (Felsenstein, 1974; Felsenstein & Yokoyama, 1976; Haigh, 1978; Takahata, 1982; Takahata & Slatkin, 1983). These studies focus on the comparison of populations with no recombination with those with free recombination between any pair of loci. In nature, however, there are a variety of reproductive systems with restricted recombination. Heller & Maynard Smith (1979) have suggested that Muller's ratchet also operates in organisms which reproduce uniparentally with recombination, though this view has been challenged by Templeton (1982, 1983).

* Corresponding author. Permanent address: Department of Genetics, University of Helsinki, Arkadiankatu 7, SF-00100 Helsinki 10, Finland.

It seems therefore necessary to conduct a thorough study on the relative merits of different modes of reproduction. In this paper we compare the rates of accumulation of both beneficial and harmful mutations in genomes of diploid organisms with biparental reproduction (bisexual, hermaphroditic), uniparental reproduction (selfing, parthenogenetic, asexual), and a mixed form of reproduction (partial selfing, a mixture of sexual and asexual reproduction). We first develop a theoretical framework for these comparisons and then compare the different modes of reproduction by means of computer simulation.

2. Theoretical framework

(i) *General formulation*

We consider a diploid organism with L loci per genome and assume that at each of these loci mutation occurs from the wildtype allele A to the mutant allele a with a frequency of v per gene per generation. Backward mutations are assumed to be negligible. We also assume that selection is genic and the relative fitness of allele a is $1-s$ compared with that of A. The mathematical model will be presented in terms of harmful mutations with $s > 0$, but the same model can be applied to beneficial mutations with $s < 0$. Fitness is multiplicative over loci, the fitness of an individual with k a genes in the genome being $w_k = (1-s)^k$. When s and the variance of k among the individuals in a population are small, the fitness difference between two individuals can be approximated with an additive model, $w_i - w_j = s(j-i)$. We use the additive approximation in this section, assuming that the mean $(\bar{k})$ of k is small.

We denote the mean number of mutant genes per diploid individual at a given time by $\bar{k} = 2L\bar{X}$, where $\bar{X}$ is the mean of the proportion (X) of mutant genes among $2L$ genes in an individual. The amount of increase of $\bar{X}$ due to mutation is $\Delta_m\bar{X} = v(1-\bar{X})$ per generation. This increment decreases (or increases) by selection acting against (or for) the mutant alleles. Since we assume multiplicative fitness among different loci, the average reduction in fitness by having one harmful mutation is s. Therefore, the mean change of fitness per generation is $\Delta\bar{w} = -s\Delta\bar{k}$, where $\Delta\bar{k}$ refers to the amount of change in $\bar{k}$ caused by selection. Obviously, $\Delta\bar{k} = -\Delta\bar{w}/s$. When the effects of mutant genes are additive, the change in fitness is proportional to the variance of fitness among individuals in the population, $\Delta\bar{w} = V_w/\bar{w}$ (where $\bar{w}$ is the mean fitness of the population) (e.g. Ewens, 1979). In the additive model the variance of fitness is given by $V_w = s^2 V_k = 4L^2s^2V_X$, where V_k and V_X are the variances of k and X in the population, respectively. Using these relationships and $\bar{k} = 2L\bar{X}$, the change in $\bar{X}$ due to selection becomes $\Delta_s\bar{X} = -2LsV_X/\bar{w}$. Note that the mean fitness is given here in reference to the best genotype present in the population. When the selective differences are small, we can assume $\bar{w} \approx 1$,

and the total change in a deterministic model is $\Delta_m\bar{X} + \Delta_s\bar{X}$ and can be approximately written as

$$\frac{d\bar{X}}{dt} = v(1-\bar{X}) - 2LsV_X = v(1-\bar{X}) - sV_k/(2L). \quad (1)$$

Haigh's (1978) equation (10) gives essentially the same result. Equation (1) also holds for beneficial mutations $(s < 0)$.

It is clear from (1) that when the within-population variance V_k is 0 (i.e. when the population size is very small), the accumulation of mutations depends on the mutation frequency alone, as expected. On the other hand, when the population size becomes infinite, (1) reduces to a form of a single-locus model (see equation (5) below). Equation (1) also shows that the accumulation stops when the within-population variance becomes $V_k = 2Lv(1-\bar{X})/s$. In an infinite population, the proportion of mutant genes per individual is expected to reach an equilibrium value of v/s with a binomial (or Poisson) variance (Kimura & Maruyama, 1966). This expectation is the same for all modes of reproduction.

To evaluate the value of V_k, let us consider a large number of replicate populations that were set up with no mutant genes initially. Mutations accumulate with time, and we can count the number of mutant genes per individual at a given time. The total variance (V_T) of this number among all individuals in all replicate populations can be divided into two components: the variance among the populations means $(V_{\bar{k}})$ and the variance among individuals within a single population (V_k) with

$$V_T = V_{\bar{k}} + V_k. \quad (2)$$

For neutral mutations, the expected proportion of mutant genes among $2L$ genes in an individual $(\bar{X})$ equals the mean $(\bar{x})$ of the gene frequency at a single locus (x), and the variance among the population means $(V_{\bar{X}})$ is related to the single-locus stochastic variance (V_S) of the gene frequency. Furthermore, the variance $V_{\bar{k}}$ is equal to $4L^2V_{\bar{X}}$. If we assume that the same properties hold approximately for slightly deleterious mutations $(1 \gg s > 0)$, we can derive $V_{\bar{k}}$ by using a single-locus model of weak selection with irreversible mutation. Our computer simulations have shown that selection must be extremely weak for this approximation to be satisfactory. At any rate, we derive the variances $V_{\bar{k}}$ and V_k for neutral genes and later examine the applicability of the formulae to the case of weak selection by computer simulation. These variances depend on the population size as well as on the mode of reproduction. We now derive these variances for three different modes of reproduction: biparental reproduction, selfing and asexual reproduction.

(ii) *Biparental reproduction*

It is not necessary to specify the exact mode of reproduction, because bisexual and hermaphroditic

populations should have the same variance if the effective population sizes are the same. The total (V_T) and within-population (V_k) variances of the number of mutant genes per individual have been derived in Appendix A under the assumption that all loci are independent of each other. They are

$$V_T = 2L\overline{X}(1-\overline{X}) + 2LV_S \qquad (3)$$

$$V_k = 2L[\overline{X}(1-\overline{X}) - V_S]. \qquad (4)$$

These variances depend on the single-locus stochastic variance, V_S, which is derived below. The change in the gene frequency (x) per generation under the deterministic single-locus model is given by

$$\Delta x = \frac{v(1-x) + (1-s)x - sx^2}{1 - 2sx} - x \approx sx^2$$
$$- (s+v)x + v. \qquad (5)$$

Approximating Δx by dx/dt and integrating this with the initial condition $x(0) = 0$, we obtain the expected frequency at time t

$$x(t) = (1 - e^{(v-s)t})/\left(\frac{s}{v} - e^{(v-s)t}\right) \quad \text{for } s \neq v, \qquad (6a)$$

$$x(t) = t/\left(t + \frac{1}{s}\right) \quad \text{for } s = v. \qquad (6b)$$

In order to derive the stochastic variance of $x(t)$ in a finite population of effective size N, we simplify (5) by ignoring the second-order term sx^2. This can be done when the frequency of the mutant allele is low and the selection coefficient s is small. With this simplification, (5) can be approximated by

$$\frac{dx}{dt} = (s+v)\left(\frac{v}{s+v} - x\right). \qquad (7)$$

This has the same mathematical form as that for the migration model studied by Crow & Kimura (1970, p. 390). Applying this model, we obtain the expected single-locus variance

$$V_S = \frac{B(A-B)}{A^2(A+1)} + \frac{2B(2B-A)}{A^2(A+2)}e^{-(s+v)t}$$
$$+ \frac{B(B+1)}{(A+1)(A+2)}e^{-2(s+v+1/4N)t} - \frac{B^2}{A^2}e^{-2(s+v)t}, \qquad (8)$$

where $A = 4N(s+v)$ and $B = 4Nv$. [If back-mutation occurs at frequency u, we can write $s+u$ instead of s in (7) and (8).]

Substitution of the variance (4) and (8) into (1) gives the expected rate of accumulation of slightly harmful mutations. It should be noted, however, that when there is selection the single-locus variance V_S and the multilocus variance do not necessarily behave in the same way. In practice, selection will create linkage disequilibrium between loci (Nei, 1963; Hill & Robertson, 1966), and (3) and (4) do not hold. Therefore, this approach gives only approximate results. We shall return to this problem when simulation results are presented.

Evaluating the means and variances at time t, we find from (4) that the multilocus within-population variance for neutral mutations ($s = 0$) is given by

$$V_k = \frac{4LNv}{2Nv+1}(e^{-vt} - e^{-(2v+1/2N)t}). \qquad (9)$$

(iii) *Selfing*

In the case of selfing, heterozygous genotypes are formed by mutation at a rate of $2v$ per locus per generation. The heterozygotes are unstable and quickly segregate into the two homozygotes with equal frequency. Hence, we have the total 'mutation' rate v from AA to aa. The population variances are derived in Appendix B, and from (A 13) and (A 14) we have

$$V_T = 4L\overline{X}(1-\overline{X}) \qquad (10)$$

$$V_k = 4L[\overline{X}(1-\overline{X}) - V_S]. \qquad (11)$$

To derive the single-locus stochastic variance, V_S, we use the same approach as that in the case of biparental production. Equations (6)–(8) are modified for selfing by replacing s with $2s$ and $2N$ with N. This is because selfing can be modelled as a haploid model with a selection coefficient $2s$ against the mutant type. Strictly speaking, the extension to multiple loci is valid only for neutral mutant genes, as in the case of biparental reproduction. Inserting the V_S obtained for neutral mutations into (11) gives

$$V_k = \frac{4LNv}{Nv+1}(e^{-vt} - e^{-(2v+1/N)t}). \qquad (12)$$

This selfing model also applies to automictic parthenogenesis. In fact, the model applies better to automixis with gamete duplication (i.e. diploidy is restored by duplication of a single haploid nucleus) or with terminal fusion (i.e. fusion of two haploid nuclei produced by the same second meiotic division) than to selfing. In automictic parthenogenesis with central fusion (i.e. fusion with haploid nuclei originating from two different secondary oocytes produced in the first meiotic division), the rate of segregation of heterozygotes depends on the recombination frequency within the chromosomes. In the vicinity of the centromeres the segregation takes place slowly, and the expected variance V_k is between the values given for selfing and asexual reproduction.

(iv) *Asexual reproduction*

In asexual reproduction (and apomictic parthenogenesis), new mutations appear almost exclusively as heterozygotes. Because there is no recombination, we can apply a haploid model with L loci, in which mutation occurs from the initial genotype (AA) to the mutant type (Aa) with a rate of $2v$ per locus per generation. We neglect the mutations from AA or Aa to aa, since the frequency is very low. The total and

within-population variances of the number of mutant alleles per individual are derived in Appendix C, and from (A 15) and (A 16) we have

$$V_T = 2L\overline{X}(1-\overline{X})$$ (13)

$$V_k = L[2\overline{X}(1-2\overline{X})-V_{S'}]$$ (14)

where $V_{S'}$ is the single-locus stochastic variance of heterozygote frequency. When deriving $V_{S'}$, equations (6)–(8) are modified by replacing v with $2v$ and $2N$ with N, because the asexual model can be treated as a haploid model with a mutation rate $2v$ from one type (AA) to another (Aa). When there is selection ($s > 0$ or $s < 0$), the variance in (14) gives an overestimate because different loci are not inherited independently. For neutral mutations ($s = 0$), we obtain from (14)

$$V_k = \frac{2LNv}{2Nv+1}(e^{-2vt}-e^{-(4v+1/N)t}).$$ (15)

It can be shown that for neutral mutations, the order of the magnitude of the within-population variance for the same t value is (12) > (9) > (15). We thus expect the largest variance for selfing (or automictic) populations and the smallest variance for asexual populations. Lynch & Gabriel (1983) reached a similar conclusion concerning the genetic variances of polygenic traits in asexual and sexual (biparental) populations.

3. Simulation

By using computer simulation, we studied the accumulation of mutations under seven different modes of reproduction (Table 1), considering a population of diploid organisms with $L = 5000$ loci.

Table 1. *Modes of reproduction examined*

Biparental
  1 Bisexual (dioecy)
  2 Hermaphroditism (monoecy)

Uniparental
  3 Selfing
  4 Automictic parthenogenesis with central fusion
  5 Asexual (apomixis)

Mixed
  6 Partial selfing
  7 Mixture of parthenogenesis and hermaphroditism

In *automictic* parthenogenesis, a new zygote is formed by the fusion of two meiotically reduced nuclei descending from the same primary oocyte. There are three types of fusion. In central fusion, the two uniting nuclei descend from two different secondary oocytes, and the fusion restores the parental genotype if no recombination has taken place in the first meiotic division. In terminal fusion, the two uniting nuclei descend from the same secondary oocyte, so that the fusion produces homozygotes if no recombination has taken place in the first meiotic division. In gamete duplication, a single haploid nucleus is duplicated, resulting in a homozygous diploid zygote. *Apomictic* parthenogenesis has no meiotic reduction and the new zygote is identical with the parental genotype. See Bell (1982) for detail.

We assume that all individuals are initially of the AA genotype for all loci. To save computer time, $N = 100$ was used for most of our simulations. We examined the accumulation of beneficial and harmful mutations separately. The rate of accumulation was calculated after the population variance stabilized. Somewhat different simulation procedures were used for different modes of reproduction, as mentioned below.

Asexual reproduction and apomictic parthenogenesis. In these modes of reproduction, no recombination takes place and the offspring are identical to the parent, barring new mutations. For our purpose, it is not necessary to know which loci have mutant alleles. We simply considered the distribution of the number of mutant genes per individual. This distribution can be presented as a vector $\bar{n}$, in which the element n_k is the number of individuals with k mutant genes. Initially, $n_0 = N$, and $n_k = 0$ for $k \geqslant 1$. The transition matrix between two consecutive generations has elements that depend on the mutation rate and selection coefficient against (or for) the mutant alleles. Mutation was assumed to occur following a Poisson distribution with a parameter $v(2L-k)$. If we denote $p_k = n_k/N$, the expected value of p_k in the next generation (indicated by a prime) is

$$p_k' = \sum_{j=0}^{k} p_{k-j} \frac{[v(2L-k+j)]^j}{j!} e^{-v(2L-k+j)}(1-s)^{k-j}/\overline{w},$$

where

$$\overline{w} = \sum_{i=0}^{2L} p_i(1-s)^i$$

is the mean population fitness compared with that of a population with no mutant alleles (i.e. the initial population). With a low mutation rate, the probability of having many mutations occurring in the same individual is very small, so that we ignored the possibility of having seven or more mutations. The probability of having any of the latter events was less than 10^{-8} in our simulations.

Using the probabilities p_k' we determined a new $\bar{n}$ vector following a method similar to that used by Li (1980). When $Np_k' \leqslant 10$ (or $\geqslant N-10$), n_k' was obtained by using a Poisson distribution with a parameter Np_k' [or $N(1-p_k')$]. When $10 < Np_k' < N-10$, it was determined by using the normal distribution with mean Np_k' and variance $Np_k'(1-p_k')$. After determining n_k' for the smallest value of k, i.e. $k = k_0$, the remaining p_k' values were standardized by $p_k' = p_k'/(1-p_{k_0}')$ to keep their sum equal to one. The population size used to calculate the expected value was also reduced by n_{k_0}' before sampling the class k_0+1. When 50, or fewer, individuals remain to be sampled, sampling was completed by drawing pseudorandom numbers and associating them directly with the relative probabilities of the remaining classes.

Automictic parthenogenesis. This mode of reproduction involves a fusion of two haploid nuclei generated

meiotically from the same primary oocyte. Several mechanisms exist to restore diploidy (Suomalainen, 1950; Asher, 1970); we study the case with central fusion. In this case, the fusing nuclei originate from different secondary oocytes. In other words, the centromeres of the two homologous chromosomes always originate from different parental chromosomes. This mechanism preserves the parental genotype at loci near the centromere, but if crossing-over takes place, some chromosomal segments can descend from the same parental chromosome and all loci within such segments immediately become homozygous. In our simulations, we assumed that the loci are arranged in five pairs of acrocentric chromosomes, 1000 loci each. Selection was assumed to occur deterministically, and N parents for the next generation were chosen by multinomial sampling after selection. In the determination of the offspring genotypes, we assumed that the number of crossovers follows a Poisson distribution with a parameter R for each pair of non-sister chromatids at the four-strand stage, and that the number of mutations follows a Poisson distribution with a parameter $vL(1-X)$ for a haploid set of chromosomes (X is the proportion of genes that are already of the mutant type). The numbers of crossovers and mutations are obtained by using Poisson pseudorandom numbers, and the locations of these events in the genome are determined by uniform pseudorandom numbers. Starting from the two parental centromeres of each chromosome pair and following one chromatid in each, we obtain the multilocus genotype of the offspring.

Selfing. Selfing differs from automixis in that haploid gametes produced by an individual can combine at random. Except for this aspect, our procedure of computer stimulation was the same as that for automixis. We also examined a version with

free recombination between each pair of heterozygous loci.

Hermaphroditism (monoecy). In this reproductive system, the gametes from all individuals are combined at random. Each individual of the parental generation contributes to the gamete pool according to its fitness, and the offspring are then formed by a random union of gametes. In our simulation, the gametes were formed as in the case of automixis. We examined the cases with linkage (five chromosome pairs) and free recombination.

Bisexual reproduction (dioecy). Simulation was conducted in the same way as that for hermaphroditism, except that the population was divided into $N/2$ males and $N/2$ females. No sex chromosomes or other sex-associated differences were considered.

Mixed models. We considered two types of mixed models of reproduction. One was the combination of parthenogenesis and hermaphroditism, in which the population reproduces with a cycle consisting of a fixed number of apomictic generations and one hermaphroditic generation. The other was partial selfing, in which a fixed proportion of individuals reproduce hermaphroditically, the remainder having self-fertilization.

4. Results

Table 2 shows the accumulation of harmful mutations in a hermaphroditic population with free recombination. The population is initially free of mutant genes. For the case of $s = 0.001$ equation (1) gives a quite accurate prediction for the mean number ($\bar{k}$) of mutant genes per individual (diploid genome). For the case of $s = 0.01$, equation (1) gives a value of $\bar{k}$ close to the observed value for $t < 200$ generations but gives an underestimate as t increases further. Equation (8) leads

Table 2. *Examples of temporal changes of the mean ($\bar{k}$) and variance (V_T, $V_{\bar{k}}$ and V_k) of the number of harmful mutant genes per individual in a hermaphroditic population with free recombination; 100 replications with $N = 100$, $L = 5000$, $v = 2 \times 10^{-5}$*

| s | Time | k | | | V_T | | $V_{\bar{k}}$ | | V_k | |
		$(6)^a$	$(1)^a$	Obs.	$(3)^b$	Obs.	$(8)^b$	Obs.	$(4)^b$	Obs.
0·001	100	19·0	19·2	19·0	22·9	22·6	7·8	7·1	15·1	15·7
	200	36·2	37·1	36·9	49·0	48·1	24·4	25·2	24·6	23·4
	300	51·7	54·2	54·1	75·9	79·9	44·2	52·1	31·6	28·6
	400	65·7	70·6	71·2	102·9	92·9	64·4	62·5	37·7	31·3
	500	78·4	86·3	87·0	128·1	135·0	83·8	103·0	43·6	33·3
0·01	100	12·6	13·3	13·8	15·6	17·1	3·6	5·7	11·4	11·5
	200	17·3	19·2	20·4	23·5	27·0	6·2	11·9	16·0	15·4
	300	19·0	21·9	24·0	27·6	32·1	7·4	15·6	18·2	16·8
	400	19·6	23·1	26·0	29·8	33·6	7·8	16·8	19·2	17·2
	500	19·9	23·6	28·5	32·3	38·4	7·8	20·8	19·6	18·0

[a] The predicted values of k are calculated using the single-locus equation (6), and by iterating equation (1).
[b] The variances predicted by applying the single-locus model are calculated V_T from equation (3), $V_{\bar{k}} = 4LV_S$ where V_S is obtained from equation (8), and V_k is from equation (4).

Table 3. *Simulated rates of accumulation and population variances*; 100 *replications with* $N = 100$, $L = 5000$, $v = 2 \times 10^{-5}$ (50 *replications in automixis*)

	Beneficial mutations ($s = -0.01$)		Harmful mutations ($s = 0.01$)	
	$\frac{d\bar{k}^a}{dt} \pm$ S.E.	V_k^b	$\frac{d\bar{k}^a}{dt} \pm$ S.E.	V_k^b
Biparental reproduction				
Bisexual				
$R = 0^c$	0.430 ± 0.007	24.0	0.057 ± 0.005	14.3
$R = 0.5^c$	0.679 ± 0.010	49.9	0.017 ± 0.006	18.2
Free recombination	0.743 ± 0.010	53.4	0.029 ± 0.006	18.3
Hermaphrodite				
$R = 0^c$	0.419 ± 0.007	23.8	0.058 ± 0.005	14.5
$R = 0.5^c$	0.692 ± 0.010	49.7	0.022 ± 0.006	18.2
Free recombination	0.712 ± 0.010	53.9	0.022 ± 0.006	17.9
Uniparental reproduction				
Selfing				
$R = 0^c$	0.359 ± 0.006	17.6	0.078 ± 0.005	12.1
$R = 0.5^c$	0.392 ± 0.006	19.6	0.080 ± 0.005	11.8
Free recombination	0.367 ± 0.006	17.5	0.083 ± 0.005	12.5
Automixis, $R = 0.5^c$	0.381 ± 0.009	18.6	0.072 ± 0.007	11.8
Asexual	0.296 ± 0.005	9.9	0.125 ± 0.004	7.9

$^a \dfrac{d\bar{k}}{dt}$ gives the average rate of change of the number of mutant genes per individual per generation for generations 450–500.

$^b V_k$ is the mean population variance in the same generations used for calculating the accumulation rate.

$^c R$ is the expected number of crossovers per one chromatid pair; the genome has five pairs of chromosomes

to serious underestimates of the variance among population means ($V_{\bar{k}}$), particularly when s and t are large. The discrepancy arises because the computer simulation allows a continuous accumulation of harmful mutations, whereas the approximations used to derive equation (8) lead to a prediction of a steady state. Thus, equation (8) should be used only for neutral mutations. Equation (4) gives fairly accurate values for the within-population variance (V_k) when $t < 400$ but tends to give overestimates when t is larger.

It is clear from Table 2 that, after passing the early generations, the rate of accumulation of new mutations and the within-population variance stabilize. In populations with $N = 100$, the accumulation rate becomes fairly steady when $t = 400$, and we therefore use the generations 450–500 to obtain the accumulation rate for a given model. We consider the mean and variance of the number of mutant genes per individual ($\bar{k}$ and V_k) because they are more conveniently expressed than those of the proportion. The standard errors of the accumulation rates are calculated over the last 50 generations of all replications. Because the successive generations of a replicate are not independent of each other, the standard errors given below would be underestimates.

The accumulation rates in a selfing population are not much affected by the number of crossovers, because most loci are homozygous and crossing over has little effect (Table 3). In biparental reproduction, crossing over has a clear effect on the population variances and accumulation rates (Table 3). However, there is not much difference in the accumulation rates between free recombination and limited recombination with $R = 0.5$ in the biparental models, between bisexual and hermaphroditic reproduction, or between selfing and automixis (central fusion with $R = 0.5$). We shall therefore concentrate on three models: asexual reproduction, selfing and hermaphroditism with free recombination.

First we compare the simulation results to the theoretical expectations outlined above. In equation (1) we derived the rate of increase of the mean proportion ($\bar{X}$) of mutant genes (loci) in a genome. Since the mean number of mutant genes per individual ($\bar{k}$) is related to $\bar{X}$ by $\bar{k} = 2L\bar{X}$, the rate of accumulation of mutant genes per individual is approximately given by

$$\frac{d\bar{k}}{dt} = 2Lv - sV_k, \tag{16}$$

when $\bar{X}$ is small. In our simulation, we assumed $v = 2 \times 10^{-5}$ and $L = 5000$ for most cases. Therefore, the rate of accumulation is given by $0.2 - sV_k$. Tables 3, 4 and 6 give the values of $d\bar{k}/dt$ and V_k obtained from computer simulation for various values of s and for different modes of reproduction. It is clear that equation (16) holds approximately for all cases. Therefore, if we know V_k we can predict the rate of accumulation of mutant genes fairly accurately by

Table 4. *Rates of accumulation of mutations with different selection intensities*[a]; 100 *replications with* $N = 100$, $L = 5000$, $v = 2 \times 10^{-5}$

	Hermaphrodite[b]		Selfing[b]		Asexual	
s	$\frac{d\bar{k}}{dt} \pm$ S.E.	$V_k{}^c$	$\frac{d\bar{k}}{dt} \pm$ S.E.	$V_k{}^c$	$\frac{d\bar{k}}{dt} \pm$ S.E.	$V_k{}^c$
-0.010	0.712 ± 0.010	53.9	0.367 ± 0.006	17.5	0.296 ± 0.005	9.9
-0.001	0.229 ± 0.009	37.2	0.241 ± 0.008	36.6	0.225 ± 0.006	18.5
0.000	0.201 ± 0.008	35.2	0.200 ± 0.009	37.8	0.189 ± 0.006	18.8
		(35.8)		(39.2)		(19.4)
0.001	0.153 ± 0.008	33.3	0.160 ± 0.008	32.9	0.183 ± 0.006	17.4
0.010	0.022 ± 0.006	17.9	0.083 ± 0.005	12.5	0.125 ± 0.004	7.9

[a] The entries in the table are mean values (mean $\pm$ S.E.) for generations 450–500.
[b] Free recombination.
[c] V_k is the mean variance of the number of mutant genes per genome; the values in parentheses are the expected neutral variances ($s = 0$) from equations (9), (12) and (15).

using equation (16) unless t is large ($t > 500$). When t is large, equation (16) is expected to give overestimates of $d\bar{k}/dt$. In this paper, however, we consider the relatively early stage of accumulation of mutant genes.

One problem with equation (16) is that it is difficult to determine V_k analytically except for the case of neutral genes. For neutral mutations, V_k is obtained from equations (9), (12) and (15). It is given in parentheses in Table 4. The theoretical variance agrees very well with the observed value from computer simulation for all modes of reproduction examined. The accumulation of neutral mutations is, of course, independent of V_k, but Tables 4 and 6 show that for small values of Ns, say $Ns < 0.1$, V_k is close to the neutral expectation and thus could be used for computing $d\bar{k}/dt$. In general, however, this approach cannot be used in the presence of selection. When the selection coefficient is $s = \pm 0.001$, our simulations show no significant difference between selfing and hermaphroditic populations for $N = 100$, but with $s = \pm 0.01$, the variance V_k is clearly highest and selection is most effective in the hermaphroditic population (Table 4). Unlike the situation in selfing and asexual populations, the variance in a hermaphroditic population is greater for beneficial than for neutral mutations (Table 4).

We have shown how the multilocus variances can be approximated by using a single-locus model of slightly harmful mutations. The problem with this approach is that it is difficult to know how small the selection coefficient should be for the approximation to be justified. Our simulations with a hermaphroditic model with $s = 0.001$ and $= 0.01$ show that the total variance V_T is higher than that given by (3) and that the variance of the population mean, $V_{\bar{k}}$, is higher than that predicted by the single-locus stochastic variance from (8) (Table 2). The within-population variance V_k, however, can be close to the value computed from equation (4) (Table 2), although the latter value tends to be larger than the simulation value. Equation (1)

indicates that no accumulation of harmful mutations occurs when $V_k = 0.2(1 - \bar{X})/s$ and $v = 2 \times 10^{-5}$. In our simulations with $s = 0.01$, the observed variances are so close to this limit (Table 2) that small inaccuracies in predicting V_k can lead to erroneous predictions of the accumulation from (1).

Table 4 shows that the rate of increase of harmful mutations is higher in asexual organisms than in hermaphrodites, as expected, and that the ratio of the rate for the former to that of the latter increases as s increases. When $s = 0.01$, the rate of increase is about six times higher for asexual reproduction than for hermaphroditism. By contrast, advantageous mutations accumulate faster for hermaphrodites than for asexual organisms. In the case of $s = -0.01$, the rate of accumulation is more than three times higher for the former than for the latter. This clearly indicates that hermaphroditism is more advantageous than asexual reproduction. Selfing is obviously intermediate between the above two cases except for $s = -0.001$.

Felsenstein (1974) (see also Kimura & Ohta, 1971) presented an approximate formula for the rate of accumulation of mutations in a sexual population. In our hermaphroditic model with intermediate heterozygotes, it becomes

$$\frac{dX}{dt} = 2Nv(1 - e^{2s})/(1 - e^{4Ns}) = 4Nsv/(e^{4Ns} - 1). \quad (17)$$

In this equation, the rate of accumulation is a function of the product Ns, rather than N and s separately. As Felsenstein (1974) stated, this formula is expected to give overestimates. For $L = 5000$ and $v = 2 - 10^{-5}$, the expected rate of accumulation from (17) is 0.815 for $Ns = -1.0$ and 0.015 for $Ns = 1.0$, while the rate observed in our simulation with $N = 100$ is 0.712 ± 0.010 for $s = -0.01$ and 0.022 ± 0.006 for $s = 0.01$. For $N = 500$ the rate is 0.725 ± 0.032 for $s = -0.002$ (10 replications) and 0.061 ± 0.013 for $s = 0.002$ (20 replications). (The rate for $N = 500$ represents the average for generations 1250–1300 in

Table 5. *The effect of the recombination frequency on the accumulation of mutations*; 100 *replications* (50 *in automixis*) *with* $N = 100$, $L = 5000$, $v = 2 \times 10^{-5}$

	Rate of accumulation $d\bar{k}/dt$ (mean ± s.e.)[a]	
	Beneficial mutations ($s = -0.01$)	Harmful mutations ($s = 0.01$)
Automixis		
$R = 0.0$[b]	0.300 ± 0.007	0.130 ± 0.006
0.1	0.379 ± 0.008	0.076 ± 0.007
0.5	0.381 ± 0.009	0.072 ± 0.007
Partial selfing		
Outcrossing rate (%)		
0	0.367 ± 0.006	0.083 ± 0.005
5	0.466 ± 0.008	0.043 ± 0.006
10	0.518 ± 0.008	0.031 ± 0.006
100	0.712 ± 0.010	0.022 ± 0.006
Mixture of sexual and asexual reproduction		
All generations asexual	0.296 ± 0.005	0.125 ± 0.004
Every 10th generation sexual	0.540 ± 0.008	0.033 ± 0.006
Every 5th generation sexual	0.628 ± 0.010	0.025 ± 0.006
All generations sexual	0.712 ± 0.010	0.022 ± 0.006

[a] The entries are mean values for generations 450–500.
[b] R is the expected number of crossovers within a chromatid pair.

the case of beneficial and 950–1000 of harmful mutations.)

As mentioned earlier, we also studied automixis and two mixed models of reproduction, i.e. partial selfing and a mixture of sexual (hermaphroditic) and asexual (parthenogenetic) reproduction. Automixis can be regarded as a special type of mixed model, since the loci close to the centromere follow the asexual mode of reproduction and those at the end of the chromosome the selfing mode. Table 5 shows that when there is no recombination automixis is equivalent to asexual reproduction with respect to the rates of accumulation of mutant genes. However, when there is recombination, the rate of increase of harmful mutations is reduced substantially, whereas the rate of accumulation of beneficial mutations increases slightly. The rates of accumulation of mutant genes for partial selfing rapidly approach those for biparental reproduction, as the proportion of outcrossing increases. When the proportion of outcrossing is as high as 10%, there is not much difference between the two reproductive systems. Occasional sexual reproduction in largely asexual organisms also drastically changes the rates of accumulation of mutant genes. When sexual reproduction occurs in every fifth generation, the rates are similar to those of sexual reproduction.

The computer simulations discussed above require a large amount of computer time, so that we were forced to use small population sizes. In the case of asexual reproduction, however, it is possible to use a much larger population size, since there is no need to study individual loci separately. The observed total variance (V_T) for asexual reproduction tends to be smaller than that expected from (13), except for neutral mutations and for harmful mutations after the mutation–selection equilibrium has been attained (Table 6). A small population of size $N = 100$ maintains very little diversity. With a mutation rate $v = 2 \times 10^{-6}$, the population variance is $V_k = 1.7$ for neutral mutations, 1.3 for slightly beneficial mutations ($s = -0.01$) and 1.0 for slightly harmful mutations ($s = 0.01$). A population of size $N = 10^4$ can maintain a large amount of neutral variation ($V_k = 125.4$), but even weak selection reduces this variance substantially, V_k being 12.3 for $s = 0.001$ and 18.1 for $s = -0.001$ (Table 6). As a consequence, the rates of accumulation of mutations are affected more strongly than those in small populations.

5. Discussion

The accumulation of harmful mutations in a genome is often discussed in terms of Muller's ratchet. The problem at issue here is whether harmful mutations can accumulate beyond the proportion given by the deterministic mutation–selection equilibrium. In an equilibrium population of infinite size, the number of harmful genes per individual is expected to follow a Poisson distribution with mean $2Lv/s$ and is the same for both sexual and asexual populations (Kimura & Maruyama, 1966). If this parameter is small, the best genotype is present in the population with a high frequency and the multilocus genotypic distribu-

Table 6. *Simulated accumulation rates and variances of the number of mutant alleles per genome in an asexual population*; 100 *replicates with L* = 5000[a]

v	s	$N = 100$			$N = 10000$		
		$\dfrac{d\bar{k}}{dt}$	V_k	V_T-ratio[b]	$\dfrac{d\bar{k}}{dt}$	V_k	V_T-ratio[b]
2×10^{-6}	−0·010	0·037	1·3	0·71	0·072	5·6	0·27
	−0·001	0·022	2·0	0·74	0·037	18·1	0·31
	0·000	0·020	1·7	0·78	0·020	125·4	1·11
	0·001	0·018	1·4	0·80	0·008	12·3	0·46
	0·005	0·013	1·2	0·75	0·001	4·0	0·66
	0·010	0·009	1·0	0·84	0·000	2·0	1·00
2×10^{-5}	−0·010	0·291	10·5	0·55	0·416	22·3	0·27
	−0·001	0·208	16·5	0·73	0·281	79·8	0·61
	0·000	0·199	21·2	1·28	0·198	182·3	0·99
	0·001	0·181	18·5	0·84	0·134	67·2	0·63
	0·005	0·146	9·6	0·96	0·078	24·0	0·47
	0·010	0·120	7·3	0·68	0·053	14·4	0·54

[a] The entries of the table are calculated in generation 10000 for $v = 2 \times 10^{-6}$ and in generation 1000 for $v = 2 \times 10^{-5}$.
[b] V_T-ratio is the ratio of the V_T observed in the simulation to that calculated from equation (13).

tion remains stable. However, if $2Lv/s$ is large, the expected frequency of the genotype with the highest fitness is small, and this genotype is easily lost in a finite asexual population because of random drift (Haigh, 1978). Maynard Smith (1978) stated that harmful mutations would not accumulate in an asexual population with $2Lv/s \approx 1$ or less, but would accumulate in small populations with $2Lv/s \approx 10$, or in large populations with $2Lv/s \approx 20$ or greater. This conclusion is largely supported by our results in Table 6, if $N = 100$ is defined as small and $N = 10^4$ as large. We note that in the case of $v = 2 \times 10^{-6}$ and $s = 0·01$, $2Lv/s$ is only 2, but harmful mutations can accumulate in a population of $N = 100$, though at a slow rate. In the case of $v = 2 \times 10^{-5}$ and $s = 0·01$, $2Lv/s$ is 20 and harmful mutations can accumulate even for $N = 10000$. Harmful mutations accumulate in sexual organisms as well if the population size is small, but the rate of accumulation is lower, as shown in this paper.

Heller & Maynard Smith (1979) suggested that Muller's ratchet also works in selfing populations. Our results in Table 4 show that moderately or slightly harmful mutations can indeed accumulate in selfing populations, though the rate of accumulation is substantially lower than that for asexual populations. The results in Table 3 further show that the rate of accumulation of harmful mutations in an automictic population is similar to that for a selfing population. This contradicts Templeton's (1982, 1983) suggestion that harmful mutations might not accumulate in automictic populations.

One problem with the above argument is the assumption of no backward mutation. If we consider lethal or semilethal genes, the frequency of backward mutation seems to be about 1/100 the frequency of forward mutation (Atwood, Schneider & Ryan, 1951).

If this were the case for all harmful mutations, the effect of backward mutation would not be important. However, the ratchet effect works only for slightly harmful mutations with selection coefficients of about 0·01 or less. For such mutations the rate of backward mutation may be of the same order of magnitude as the rate of forward mutation. If this is the case, the difference in the rate of accumulation of harmful mutations between sexual and asexual reproduction would diminish considerably, particularly in large populations (Li, 1986). Of course, the advantage of sexual reproduction with respect to accumulation of beneficial mutations will remain.

Let us now examine the change in population fitness due to accumulation of beneficial and harmful mutations under the assumption of no backward mutation. Felsenstein (1974) and Takahata (1982) have earlier done similar comparisons for haploid organisms. Our results are largely consistent with those of Felsenstein, though the latter was based on only 10 replicates. Takahata simulated populations with either free or no recombination, but his results partly disagree with ours, apparently because of the small number of loci (10) he used. For instance, he detected the same rate of accumulation for harmful mutations with $s = 0·01$, regardless of whether the population is with free or with no recombination, whereas our simulations indicate a several-fold difference. We believe our results are more realistic, though the simulations were restricted to very small populations.

The evolutionary significance of the accumulation of mutations in a genome depends on how the accumulation affects fitness. With mutation rates $v = 2 \times 10^{-5}$ and 2×10^{-6} for harmful alleles with selection coefficients 0·001 −0·01, a biparental population of size $N = 10^4$ is expected to reach the

10-2

mutation–selection equilibrium relatively quickly, whereas an asexual population of the same size will accumulate harmful mutations for a considerable period of time. When the mutation rate is 2×10^{-6}, the average decrease of the population mean fitness (given here as a change per 1000 generations) is approximately $\Delta \bar{w} = -0.0075$ for $s = 0.001$, $\Delta \bar{w} = -0.005$ for $s = 0.005$, and $\Delta \bar{w} = 0$ for $s = 0.01$ (i.e. no accumulation takes place in the last case). This reduction in fitness is so small that it is unlikely to have a pronounced short-term evolutionary influence. If the mutation rate is as high as 2×10^{-5}, a severe reduction in the mean fitness can take place in a few hundred generations for an asexual population. Namely, the change per 1000 generations is $\Delta \bar{w} = -0.13$ for $s = 0.001$, $\Delta \bar{w} = -0.39$ for $s = 0.005$, and $\Delta \bar{w} = -0.53$ for $s = 0.01$.

Crow & Simmons (1983) have argued that beneficial mutations are rare and have smaller effects on fitness than harmful mutations. Letting $s = -0.001$, we can compare the changes in the mean fitness for $N = 10^4$ by using our simulation results for an asexual population and equation (17) for a sexual population. If the mutation rate is $v = 2 \times 10^{-7}$, the change of the mean fitness per 1000 generations is 0.006 in an asexual population and 0.083 in a sexual population. The difference is relatively small but will become large as t increases, say $t > 10^5$ generations. Thus, unless the v value is much lower than 2×10^{-7}, even slightly beneficial mutations may cause marked differences in fitness between sexual and asexual populations in long-term evolution.

It should be noted that while the rate of accumulation of mutations in a genome varies with the mode of reproduction, this may not be the major factor for a change of reproductive system in an evolutionary lineage. To understand the actual cause of this change, we must study the transitional stage of reproductive systems within a population (Bernstein *et al.* 1985). It should also be noted that there are several other factors that may affect the evolutionary success of a population with a given mode of reproduction (Maynard-Smith, 1978; Bell, 1982; Templeton, 1982; Lynch, 1984). However, our results can explain why sexual reproduction has not been replaced by parthenogenesis. Unless the population size is very large, parthenogenetic populations are likely to suffer from an elevated accumulation of harmful and a retarded accumulation of beneficial mutations, compared to a population with biparental reproduction. If there is competition between such populations, the parthenogenetic population has a disadvantage which in the long run can contribute to the likelihood of its extinction. These factors can partly explain why parthenogenetic lineages are generally short-lived (Lynch, 1984).

References

Asher, J. H. Jr. (1970). Parthenogenesis and genetic variability. II. One-locus models for various diploid populations. *Genetics* **66**, 369–391.

Atwood, K. C., Schneider, L. K. & Ryan, F. J. (1951). Selective mechanism in bacteria. *Cold Spring Harbor Symposia on Quantitative Biology* **16**, 345–355.

Bell, G. (1982). *The Masterpiece of Nature.* Los Angeles: University of California Press.

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

Crow, J. F. & Kimura, M. (1965). Evolution in sexual and asexual populations. *American Naturalist* **99**, 439–450.

Crow, J. F. & Kimura, M. (1970). *An Introduction to Population Genetics Theory.* New York: Harper and Row.

Crow, J. F. & Simmons, M. J. (1983). The mutation load in Drosophila. In *The Genetics and Biology of Drosophila*, vol. 3c (ed. M. Ashburner, H. L. Carson and N. J. Thompson Jr), pp. 1–35. New York: Academic Press.

Ewens, W. J. (1979). *Mathematical Population Genetics.* Heidelberg: Springer-Verlag.

Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics* **78**, 737–756.

Felsenstein, J. & Yokoyama, S. (1976). The evolutionary advantage of recombination. II. Individual selection for recombination. *Genetics* **83**, 845–859.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection.* Oxford: Clarendon Press.

Haigh, J. (1978). The accumulation of deleterious genes in a population – Muller's ratchet. *Theoretical Population Biology* **14**, 251–267.

Heller, R. & Maynard Smith, J. (1979). Does Muller's ratchet work with selfing? *Genetical Research* **32**, 289–293.

Hill, W. G. & Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetical Research* **8**, 269–294.

Kimura, M. & Maruyama, T. (1966). The mutational load with epistatic gene interactions in fitness. *Genetics* **54**, 1337–1351.

Kimura, M. & Ohta, T. (1971). *Theoretical Aspects of Population Genetics.* Princeton: Princeton University Press.

Li, W.-H. (1980). Rate of silencing at duplicate loci: a theoretical study and interpretation of data from tetraploid fishes. *Genetics* **95**, 237–258.

Li, W.-H. (1986). Models of nearly neutral mutations with particular implications for nonrandom usage of synonymous codons. *Journal of Molecular Evolution.* (In the Press.)

Lynch, M. (1984). Destabilizing hybridization, general-purpose genotypes and geographic parthenogenesis. *Quarterly Review of Biology* **59**, 257–290.

Lynch, M. & Gabriel, W. (1983). Phenotypic evolution and parthenogenesis. *American Naturalist* **122**, 745–764.

Maynard Smith, J. (1978). *The Evolution of Sex.* Cambridge: Cambridge University Press.

Muller, H. J. (1932). Some genetic aspects of sex. *American Naturalist* **8**, 118–138.

Muller, H. J. (1964). The relation of recombination to mutational advance. *Mutation Research* **1**, 2–9.

Nei, M. (1963). Effect of selection on the components of

genetic variance. In *Statistical Genetics and Plant Breeding*, ed. W. D. Hanson and H. F. Robinson, pp. 501–515. Washington, D.C.: National Academy of Science–National Research Council.

Suomalainen, E. (1950). Parthenogenesis in animals. *Advances in Genetics* 3, 193–253.

Takahata, N. (1982). Sexual recombination under joint effects of mutation, selection and random sampling drift. *Theoretical Population Biology* 22, 258–277.

Takahata, N. & Slatkin, M. (1983). Evolutionary dynamics of extranuclear genes. *Genetical Research* 42, 257–265.

Templeton, A. (1982). The prophecies of parthenogenesis. In *Evolution and Genetics of Life Histories*, ed. H. Dingle and J. P. Hegmann, pp. 75–102. Heidelberg: Springer-Verlag.

Templeton, A. (1983). Natural and experimental parthenogenesis. In *The Genetics and Biology of Drosophila*, vol. 3c, ed. M. Ashburner, H. L. Carson and N. J. Thompson, Jr, pp. 343–398. New York: Academic Press.

Appendix

According to equation (1), the rate of accumulation of new mutations depends on the within-population variance of mutant genes per individual, V_k. We here derive this variance, assuming that all loci are independent of each other. This assumption is discussed in the text.

A. *Biparental reproduction*

The following derivation follows the suggestion given by P. Avery. Let k_{mij} be a random variable that takes the value of 1 if the jth gene of locus i in population m is a mutant and 0 otherwise ($j = 1, 2, \ldots, 2N$). The probability that $k_{mij} = 1$ is equal to $\overline{X}$, the mean proportion of mutant genes among $2L$ genes in an individual. Therefore, the mean and variance of k_{mij} are $E(k_{mij}) = \overline{X}$ and $V(k_{mij}) = \overline{X}(1-\overline{X})$, respectively. k_{mij} may be written as

$$k_{mij} = \overline{X} + a_{mi} + b_{mij}, \qquad (A\ 1)$$

where a_{mi} and b_{mij} are the population effect and the allelic effect, respectively, with $E(a_{mi}) = E(b_{mij}) = 0$. Therefore, the variance of k_{mij} is given by

$$\mathrm{var}(k_{mij}) = \overline{X}(1-\overline{X}) = V_a + V_b, \qquad (A\ 2)$$

where V_a and V_b are the variances of a_{mi} and b_{mij}, respectively.

The frequency of mutant genes at locus i in population m is

$$x_{mi} = \frac{1}{2N} \sum_{j=1}^{2N} k_{mij}. \qquad (A\ 3)$$

The variance of x_{mi} is V_S by definition (see text). It can also be written as

$$\mathrm{var}(x_{mi}) = \frac{1}{4N^2} \mathrm{var}\left(\sum_{j=1}^{2N} k_{mij} \right). \qquad (A\ 4)$$

On the other hand, the mean number of mutant genes per individual in population m is

$$\overline{k}_m = \frac{1}{N} \sum_{i=1}^{L} \sum_{j=1}^{2N} k_{mij}. \qquad (A\ 5)$$

Thus, the variance ($V_{\overline{k}}$) of $\overline{k}_m$ is

$$V_{\overline{k}} = \frac{L}{N^2} \mathrm{var}\left(\sum_{j=1}^{2N} k_{mij} \right) = 4LV_S. \qquad (A\ 6)$$

We also note that

$$\sum_{j=1}^{2N} k_{mij} = 2N\overline{X} + 2Na_{mi} + \sum_{j=1}^{2N} b_{mij}. \qquad (A\ 7)$$

Hence,

$$V_{\overline{k}} = \frac{L}{N^2}(4N^2 V_a + 2N V_b) = 4LV_a + \frac{2L}{N} V_b. \qquad (A\ 8)$$

Therefore, from (A 2), (A 6) and (A 8), we have $V_a \approx V_S$ and $V_b \approx \overline{X}(1-\overline{X}) - V_S$ for large N.

Let us now derive the total variance (V_T), i.e. the variance of the total number of mutant genes in an individual ($k_{m.j}$). $k_{m.j}$ is given by

$$k_{m.j} = \sum_{i=1}^{L} \{k_{mij} + k_{mi(j+1)}\}, \quad j = 1, 3, \ldots, 2N-1.$$

Then,

$$\begin{aligned}
V_T &= \mathrm{var}\left[\sum_{i=1}^{L} \{k_{mij} + k_{mi(j+1)}\} \right] \\
&= \mathrm{var}\left\{ \sum_{i=1}^{L} 2\overline{X} + 2a_{mi} + b_{mij} + b_{mi(j+1)} \right\} \\
&= L(4V_a + 2V_b) \\
&\approx L(4V_S + 2\overline{X}(1-\overline{X}) - 2V_S) \\
&= 2L\overline{X}(1-\overline{X}) + 2LV_S. \qquad (A\ 9)
\end{aligned}$$

Hence, the within-population variance in the text is given by

$$V_k = V_T - V_{\overline{k}} = 2L\{\overline{X}(1-\overline{X}) - V_S\}. \qquad (A\ 10)$$

B. *Selfing*

In selfing organisms, heterozygous genotypes do not last long, but segregate into the two homozygous types. In the following we assume that no heterozygous loci exist. Let $\overline{X}$ be the proportion of loci homozygous for the mutant allele (note that $\overline{X}$ is also the proportion of mutant genes per individual). Because each homozygous locus carries two copies of the allele, the random variable k_{mij} takes the value of 2 when the locus has the mutant allele and 0 otherwise ($j = 1, 2, \ldots, N$). Thus

$$k_{mij} = 2(\overline{X} + a_{mi} + b_{mij}),$$

where b_{mij} refers to the effect of an individual. The variance of k_{mij} is given by

$$\mathrm{var}(k_{mij}) = 4\overline{X}(1-\overline{X}) = 4(V_a + V_b). \qquad (A\ 11)$$

The single-locus variance of the gene frequency (x_{mi}) becomes

$$\begin{aligned}
V_S &= \mathrm{var}\left(\frac{1}{2N} \sum_{j=1}^{N} k_{mij} \right) \\
&= \frac{1}{4N^2} \mathrm{var}\left(\sum_{j=1}^{N} k_{mij} \right)
\end{aligned}$$

Therefore we obtain

$$V_{\bar{k}} = 4LV_S, \tag{A 12}$$

as in the case of biparental reproduction. The total variance becomes

$$V_T = \mathrm{var}\left(\sum_{i=1}^{L} k_{mij}\right) = \mathrm{var}\left[\left(\sum_{i=1}^{L} 2(\overline{X} + a_{mi} + b_{mij})\right)\right]$$
$$= 4L(V_a + V_b) = 4L\overline{X}(1-\overline{X}) \tag{A 13}$$

Therefore we have

$$V_k = 4L[\overline{X}(1-\overline{X}) - V_S]. \tag{A 14}$$

C. *Asexual reproduction*

Decomposition of the total variance for the case of asexual reproduction can be studied by assuming that the mutant allele always exists in heterozygous condition. Let Q be the proportion of heterozygous loci per individual. We then have $Q = 2\overline{X}$. The variable k_{mij} takes the value of 1 when a locus is heterozygous and 0 otherwise. k_{mij} can be written as

$$k_{mij} = Q + a_{mi} + b_{mij},$$

where b_{mij} refers to the effect of an allele. The single-locus stochastic variance of the frequency of heterozygotes is

$$V_{S'} = \mathrm{var}\left(\frac{1}{N}\sum_{j=1}^{N} k_{mij}\right) = \frac{1}{N^2}\mathrm{var}\left(\sum_{j=1}^{N} k_{mij}\right).$$

$V_{\bar{k}}$ can be derived in the same way as in the derivation of (A 6), and becomes

$$V_{\bar{k}} = LV_{S'}.$$

Because the number of heterozygous loci per individual for this model equals the number of mutant alleles per individual, we obtain

$$V_T = \mathrm{var}\left(\sum_{i=1}^{L} k_{mij}\right) = \mathrm{var}\left(\sum_{i=1}^{L} Q + a_{mi} + b_{mij}\right)$$
$$= L(V_a + V_b) = LQ(1-Q). \tag{A 15}$$

Therefore

$$V_k = V_T - V_{\bar{k}} = L[Q(1-Q) - V_{S'}]. \tag{A 16}$$

Replacement of Q by $2\overline{X}$ gives equation (14) in the text.