Muller's Ratchet and the Pattern of Variation at a Neutral Locus

Isabel Gordo,*,[1] Arcadio Navarro[†] and Brian Charlesworth[†]

**Instituto Gulbenkian da Ciência, P-2781-901 Oeiras, Portugal and [†]Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom*

Manuscript received August 27, 2001
Accepted for publication February 22, 2002

ABSTRACT

The levels and patterns of variation at a neutral locus are analyzed in a haploid asexual population undergoing accumulation of deleterious mutations due to Muller's ratchet. We find that the movement of Muller's ratchet can be associated with a considerable reduction in genetic diversity below classical neutral expectation. The extent to which variability is reduced is a function of the deleterious mutation rate, the fitness effects of the mutations, and the population size. Approximate analytical expressions for the expected genetic diversity are compared with simulation results under two different models of deleterious mutations: a model where all deleterious mutations have equal effects and a model where there are two classes of deleterious mutations. We also find that Muller's ratchet can produce a considerable distortion in the neutral frequency spectrum toward an excess of rare variants.

EVERY population is continuously exposed to newly occurring mutations, the majority of which are probably deleterious. Muller (1964) pointed out that, in the absence of recombination and back mutation, a finite population would suffer an accumulation of deleterious mutations due to the irreversible loss of the classes of individuals with the least numbers of mutations—the least-loaded classes. Consider the simplest case of a nonrecombining haploid population of N breeding individuals, subject to mutations with a deleterious effect s, occurring at a rate U per individual per generation. With multiplicative fitness interactions between loci and a Poisson distribution of the number of mutations occurring per individual per generation, the population is divided into classes of individuals according to the number of mutations they carry. In an effectively infinite population at deterministic equilibrium, the frequencies of the classes, after mutation but prior to selection, are given by a Poisson distribution with mean U/s (Haigh 1978). In particular, the size of the class free of deleterious mutations is $n_0 = Nf_0 = N\exp(-U/s)$. If the size of this class is not very large, genetic drift will cause it to be lost. After such an event, it will be replaced by the next least-loaded class. Then a new deterministic equilibrium will be approached and, after some time, the least-loaded class will be lost again.

This is the repetitive irreversible process known as Muller's ratchet. Muller's ratchet has been thought to play a role in providing an evolutionary advantage to sex and recombination and to cause the degeneration of large nonrecombining portions of the genome of sexual organisms, such as *Y* chromosomes, and the extinction of small asexual populations (Muller 1964; Felsenstein 1974; Charlesworth 1978; Maynard Smith 1978; Lynch *et al.* 1993; Gessler and Xu 1999). A great deal of theoretical work has been done to understand the circumstances under which Muller's ratchet can produce evolutionary effects. In particular, one can ask the question: How many generations on average will it take for an asexual population to lose its present best class of individuals (Haigh 1978; Pamilo *et al.* 1987; Stephan *et al.* 1993; Gessler 1995; Higgs and Woodcock 1995; Charlesworth and Charlesworth 1997; Gordo and Charlesworth 2000a,b)?

Although the biological importance of the ratchet can be assessed by the quantification of this time and the associated decline in mean fitness, its extremely high sensitivity to small changes in the parameters (Gordo and Charlesworth 2000a,b), together with our lack of knowledge of the exact values of the relevant parameters of mutation and selection (Keightley and Eyre-Walker 1999), make it hard to draw definitive conclusions about the ratchet's role. For example, if the majority of deleterious mutations are mildly deleterious (with $s \ll 1\%$), then, with our present knowledge of the deleterious mutation rate, Muller's ratchet could potentially be a major process driving the degeneration of *Y* chromosomes, even in very large populations such as those of Drosophila (Gordo and Charlesworth 2000b). But if this is not the case, its operation may be biologically negligible in that context.

One signature of the operation of Muller's ratchet is the fixation of deleterious alleles as a consequence of the recurrent loss of the best class (Higgs and Woodcock 1995; Charlesworth and Charlesworth 1997; Bergstrom and Pritchard 1998). In this article, we try to evaluate another signature of its operation. It is

[1]*Corresponding author:* Isabel Gordo, Instituto Gulbenkian da Ciência, Rua da Quinta Grande 6, Apartado 14, P-2781-901 Oeiras, Portugal.

well known that the elimination of strongly deleterious mutations can substantially reduce variation levels at linked neutral sites—an effect known as "background selection" (Charlesworth *et al.* 1993). Background selection, as classically stated, assumes that no irreversible accumulation of deleterious mutations occurs, which is not the case when the ratchet is turning. We therefore asked the following question: What is the level and pattern of neutral variation in a population where Muller's ratchet is operating? Observations on neutral variability in asexual populations or on *Y* chromosomes may detect the signature of processes such as the ratchet (Charlesworth and Charlesworth 2000), so that it is important to have theoretical predictions of what to expect.

This question is examined using Monte Carlo stochastic simulations of a neutral locus embedded in a set of selected loci that accumulate deleterious mutations by the ratchet mechanism. Variability at the neutral locus is measured and compared with both analytical and simulated results based on the structured coalescent. Tajima's *D* statistic (Tajima 1989) and Fu and Li's *D** statistic (Fu and Li 1993), commonly used to test deviations from the standard neutral model, are calculated, and their power to reject neutrality in a population undergoing a ratchet mechanism is assessed. Recent studies on the rate of occurrence and selection coefficients of deleterious mutations have suggested that a simple discrete distribution with two classes of mutation effects seems to fit the data better than a continuous distribution (Keightley 1996; Davies *et al.* 1999). We therefore also studied a model in which the population is subject to deleterious mutations with two major types of effects and compared the level of neutral variability under such a model to an analytical approximation.

SIMULATION METHODS

Multilocus Monte Carlo simulations: Following the previous work of Gordo and Charlesworth (2000a,b), a haploid, nonrecombining population of N chromosomes was simulated with the following life cycle: mutation, reproduction, and selection. In each generation, mutations to deleterious alleles occur according to a Poisson distribution with mean U. Multiplicative fitness effects of the deleterious mutations are assumed. Two kinds of model of the effects of deleterious mutations are considered: a model in which all mutations have the same selection coefficient (s) and a model where mutations can have two types of deleterious effect (s_s and s_b, with $s_s < s_b$; Gordo and Charlesworth 2001). At a neutral locus, mutations are generated according to the infinite-sites model; *i.e.*, each mutation occurs at a new site, at a total rate of μ at the locus, which has 250 neutral sites. Every N generations, a sample of size $n = 25$ chromosomes is taken from the population, and variability measures at the neutral locus are computed. For each set of parameter values, 10 simulation runs were done, with 5 samples taken per run (no significant correlation between samples was observed). A total of 50 samples was used for the computations.

Coalescent simulations: A model in which all mutations have identical selection coefficients (s) was simulated using the coalescent process. Our method is based on the structured coalescent described in Charlesworth *et al.* (1995), with a slight modification. The program follows the genealogies of a sample of neutral alleles backward in time. These can move between the different classes of individuals defined by the number of deleterious mutations harbored by each individual. The frequencies of these classes, after selection, are given by their deterministic expectation, *i.e.*, according to a Poisson distribution with parameter $\lambda = U(1 - s)/s$ when $n_0 > 1$. When $n_0 = Ne^{-\lambda} < 1$, the distribution is replaced by a shifted Poisson distribution with parameter $\lambda = U(1 - s)/s - k$, where $k = \min\{i, Nf_i^* > 1\}$ (Gessler 1995; see below for details). The first step in the program is to generate a transition matrix, Q_{ij}, of the probabilities that an individual with i mutations in a given generation has an ancestor with j mutations in the previous generation, according to Equation 3 of Charlesworth *et al.* (1995). Then, the haplotypes in which the neutral alleles in the sample are embedded are generated randomly from the equilibrium distribution after selection (this is different from Charlesworth *et al.* 1995, who considered the equilibrium distribution after mutation and prior to selection—although in practice, this makes no difference for the values of s considered here).

After these two preliminary steps, the backward process starts. Every generation, the number of mutations in the ancestor of each individual is obtained randomly by using the probabilities in the matrix Q_{ij} as expected values. When all the ancestors have been assigned to a class, coalescence is allowed to occur between individuals belonging to the same class, with probability $k_i(k_i - 1)/2/Nf_i^*$, where k_i is the number of lineages with i deleterious mutations present in the sample at a given generation and Nf_i^* is the deterministic equilibrium size of class i, after selection. The possibility of more than one coalescent event within a class is neglected. As commonly implemented in previous algorithms for the structured coalescent (Hudson 1990; Charlesworth *et al.* 1995), simultaneous coalescent events are possible in the same generation only if they occur in different classes. Once the most recent common ancestor of the whole sample is reached, neutral mutations are distributed over the gene tree generated by the simulation according to the infinite-sites model. For each set of parameter values, 10^4 trees were generated.

Measures of genetic diversity at a neutral locus: Two measures of genetic variation in a sample of alleles at the neutral locus are considered: the mean number of

pairwise differences between randomly sampled sequences, k, and the number of segregating sites, S. Under the infinite-sites model in the absence of deleterious mutations, the expectations of these quantities for a haploid population are

$$k_0 = \theta$$

$$S_0 = \theta \sum_{i=1}^{n-1} \frac{1}{i}$$

(Ewens 1979), where $\theta = 2N\mu$, where the subscript 0 refers to the strictly neutral model. In the absence of recombination, these expectations are reduced by a factor of $\sim f_0 = \exp(-U/s)$ in a large population that is at equilibrium between recurrent mutation to strongly deleterious alleles and their elimination by purifying selection (background selection). This approximation was shown previously to be accurate in a population where Muller's ratchet does not operate (Charlesworth *et al.* 1993).

The mutational frequency spectrum: Selection against deleterious mutations is expected to affect k more than S, since k is weighted toward variants at intermediate frequencies (Tajima 1989; Charlesworth *et al.* 1993). This is observed when computing statistics, such as Tajima's D, designed to test deviations from the frequency spectrum predicted under strict neutrality. Tajima's D is defined as

$$D = \frac{k - \theta_w}{\sqrt{\mathrm{Var}(k - \theta_w)}},$$

where $\theta_w = S/\Sigma_{i=1}^{n-1} 1/i$ and $\mathrm{Var}(k - \theta_w)$ is calculated assuming no recombination (Tajima 1989). Negative values of D are associated with a skew in the distribution of frequencies of neutral mutations toward an excess of rare variants. Because we expect a negative D in the presence of purifying selection (Charlesworth *et al.* 1993), we asked how often, in the presence of a ratchet, we can reject the neutral model because of very negative D's.

For a given θ value, we ran standard coalescent simulations (Hudson 1990) of the neutral infinite-sites mutational model and calculated the critical values [at the 95% confidence interval (C.I.)] of the statistic D. We then used forward simulations to compute the rejection power, given by the proportion of forward simulations whose observed D was lower than the critical value expected under neutrality: Pow 1. Although in some species, such as *Drosophila melanogaster*, one has information about genome-wide levels of variability, from which one can estimate θ, in others such information is not available. Because θ is generally not known, we also performed a power analysis assuming a fixed number of segregating sites, S (Hudson 1993). We generated standard neutral genealogies, distributed S mutations onto them, and obtained the 95% critical values of D. Afterward, we ran structured coalescent simulations, also distributing S mutations onto the genealogical trees, and calculated Pow 2, the proportion of times the value of D obtained in the structured coalescent simulations was lower than the critical D in the neutral simulations. A total of 10^4 genealogical trees were run for every set of parameters.

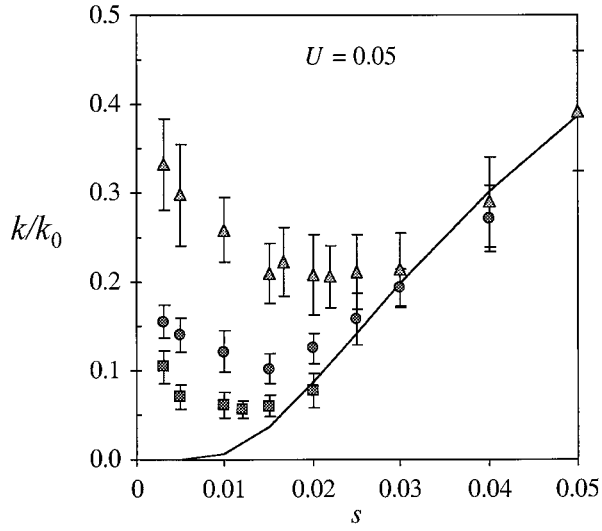

Figure 1.—Simulations of the effect of Muller's ratchet on mean number of pairwise differences at a neutral locus k (relative to that under strict neutrality, k_0) as a function of the selection coefficient against deleterious mutations (s). The mutation rate is 0.05 and N is 1000 (triangles), 3000 (circles), and 8000 (squares). We start observing clicks of the ratchet when $s < 0.03$ for $N = 1000$, $s < 0.02$ for $N = 3000$, and $s < 0.015$ for $N = 8000$. The continuous line is the frequency of the least-loaded class at the mutation-selection deterministic equilibrium, f_0. The error bars correspond to two standard errors.

RESULTS

Muller's ratchet and genetic diversity: Suppose that the accumulation of deleterious alleles is occurring due to the repetitive "clicks" of Muller's ratchet. What is the expected level of variability at a locus evolving neutrally? In Figure 1 we show the reduction in the mean number of pairwise differences caused by deleterious mutations, *i.e.*, the ratio of the observed k to that expected in the absence of purifying selection, k_0, as a function of s. We also plot the deterministic equilibrium frequency of the least-loaded class, f_0. For any value of N, with a sufficiently large value of s the reduction in k is independent of N and is very well approximated by f_0. With recurrent mutations with very large deleterious effects, the rate at which the ratchet operates is extremely low (if it operates at all), and the level of variation at a neutral locus reflects the size of the class of individuals with the highest fitness. This is because any neutral variant arising in less fit classes is quickly driven to extinction (Fisher 1930, p. 122), or, putting it in another way, any neutral variant sampled from mutated classes is very recently derived from gametes belonging to the fittest class.

Thus, for these cases we recover the classical background selection approximation $E(k) \approx 2Nf_0\mu$ (Charlesworth *et al.* 1993).

For intermediate selection coefficients, Muller's ratchet starts to operate at a reasonable rate. Two phenomena start to occur: The size of the best class fluctuates around its deterministic equilibrium value and is driven to 0 with a time lag that varies stochastically (Haigh 1978), and fixations of deleterious alleles in the population start to occur (Higgs and Woodcock 1995; Charlesworth and Charlesworth 1997). Therefore, neither is the size of the least-loaded class constant nor is the frequency of every deleterious allele predicted by the mutation-selection deterministic equilibrium. For example, for $s = 0.01$ and $U = 0.05$, as in Figure 1, the ratchet clicks on average every 110 generations for $N = 1000$, every 170 generations for $N = 3000$, and every 331 generations for $N = 8000$. We find that, under conditions favorable to the operation of Muller's ratchet, variability is always higher than the value predicted by f_0 (although always lower than the expectation under the strictly neutral model) and that the reduction in variability is dependent on N.

The conditions for the operation of the ratchet require that s is not very large and/or Nf_0 is relatively small. This implies that the mean time that a gamete with a deleterious mutation persists in the population can be larger than the mean coalescent time within the least-loaded class ($1/s \gg Nf_0$), which means that more loaded classes can contribute significantly to variability. Hence, in these circumstances the relative genetic diversity (k/k_0) is higher than the value predicted simply by f_0, as seen in Figure 1. For a given N and U, there is a value of s, $s_{\min}$, that produces a minimum in diversity. If we plot the results of Figure 1 as a function of Nf_0s we observe that the minimum occurs for $Nf_0s \sim 1$. When $Nf_0s \gg 1$, increasing s increases diversity through the increase in f_0; when $Nf_0s \ll 1$, decreasing s increases diversity due to the contribution of classes other than the least-loaded one. With very weak selection, the reduction in variability becomes very small and is negligible in the limiting case $s \ll 1/N$, as deleterious alleles then become effectively neutral and do not interfere with the dynamics of the linked neutral locus at which variation is being measured (Crow and Kimura 1970, p. 322).

We try to approximate the reduction in genetic diversity as follows. Because $E(k) = 2\mu T_2$, where T_2 is the expected time to the most recent common ancestor of two randomly sampled gametes, we approximate T_2 using the coalescent approach of Hudson and Kaplan (1994) by assuming that a population subject to recurrent deleterious mutations can be thought of as a subdivided population in which mutation plays the role of migration. Under conditions where an approximate mutation-selection balance can be attained, *i.e.*, when $n_0 \gg 1$, (Stephan *et al.* 1993; Gessler 1995), the sizes of the mutational classes are close to their deterministic expectation most of the time, and the assumption of mutation-selection balance to calculate the coalescent time should produce reasonable results. The expressions for the coalescent time are given in appendix a, which are equivalent to Equation 12 of Hudson and Kaplan (1994).

When $Nf_0 < 1$, the distribution of mutations can deviate considerably from a Poisson with mean U/s. For these cases, Gessler (1995) has suggested that a shifted Poisson distribution of mean λ, where $\lambda = U/s - k$ with $k = \min\{i,\ Nf_i > 1\}$, is a better approximation. For these cases we replace Equation A2 of appendix a by

$$Q_{i,i-1} = \frac{U}{U + ((1 - s)/i)\lambda}, \qquad (1)$$

where $Q_{i,i-1}$ is the probability that a gamete with i mutations derives from a gamete with $i - 1$ mutations. T_2 can be calculated in the same way as before, but using Equation 1.

In Figure 2, we compare the results of these analytical approximations (leading to Equation A5 in appendix a) with those from the exact Monte Carlo forward simulations. The deleterious mutation rate is 0.05, and two values of s are considered: $s = 0.005$ and $s = 0.015$. Figure 2 shows that variability is more reduced for larger values of N and slowly approaches the value f_0 as $N \to \infty$ (absence of the ratchet). The analytical expressions provide reasonably good approximations to the simulation results. Note that, for the case $s = 0.005$ in the range of values of N considered, the deterministic value of n_0 is <1, so that Equation 1 was used to calculate the mean coalescent time. Simulations of the structured coalescent were run and compared with both the results of the exact Monte Carlo forward simulations and Equation A5. As expected, no difference is observed between the mean number of pairwise differences predicted by Equation A5 and the one obtained in the coalescent simulations, since they are based on the same assumptions (results not shown).

In Figure 3, we consider the effects of different values of U with a constant population size of 2000 individuals. As expected, the larger the value of U, the bigger the reduction in variability, for any given value of s. The reduction in expected variability predicted by the coalescent approach is a reasonably good approximation to the means obtained in the forward simulations, even for cases where $n_0 < 1$. However, for the cases where $n_0 \ll 1$, with the smaller values of s and large values of U in Figure 3, coalescent predictions (and, therefore, coalescent simulations) underestimate the mean pairwise differences in the forward simulations. For example, in Figure 3 with $U = 0.1$ and $s = 0.003$, the reduction in the mean pairwise differences observed in forward simulations is 0.217 (with 95% C.I. 0.027) while the prediction from Equation A5 is 0.164. A similar behavior is detected upon close examination of Figure 2, although the difference there is much smaller.

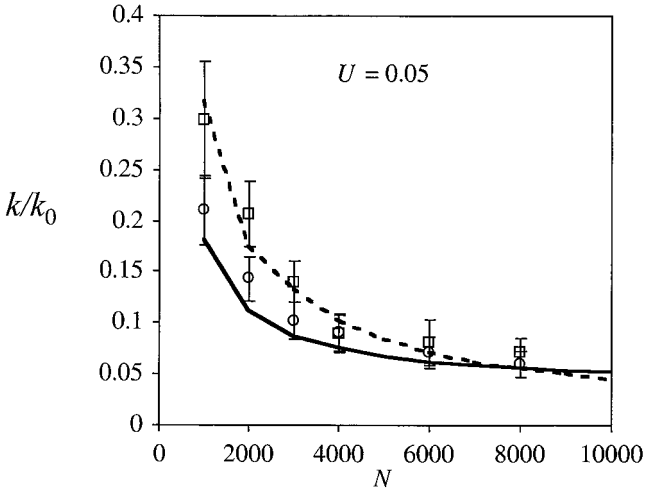

FIGURE 2.—Relation between the mean number of pairwise differences, relative to that under strict neutrality, and population size (N). The deleterious mutation rate is 0.05. The simulation results for $s = 0.005$ (squares) and 0.015 (circles) are shown. The dashed and solid lines are the corresponding theoretical values calculated using Equations A2–A5. For sufficiently large values of N, k/k_0 would be $\sim 5 \times 10^{-5}$ for $s = 0.005$ and ~ 0.04 for $s = 0.015$. The error bars correspond to two standard errors.

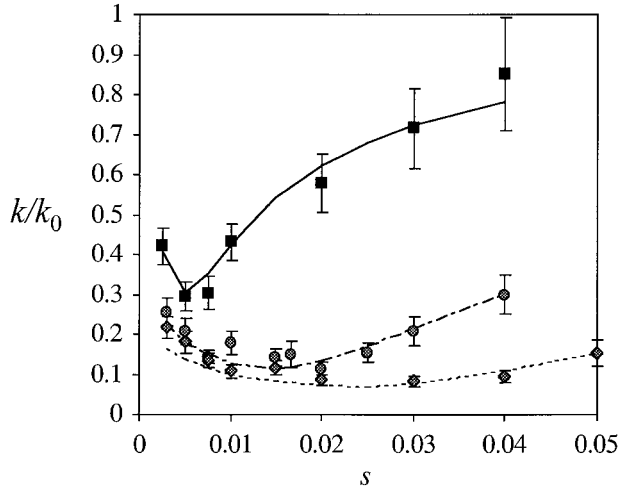

FIGURE 3.—The dependence of the relative reduction in mean number of pairwise differences, k/k_0, on the deleterious mutation rate and selection coefficient, with $N = 2000$. Squares are the simulation results for $U = 0.01$, and the solid line is the analytical prediction; circles and the dashed-dot line are the simulation and analytical results for $U = 0.05$; diamonds and the dashed line are the results for $U = 0.1$. Error bars represent two standard errors. Clicks of the ratchet were observed in the simulations when $s < 0.01$ for $U = 0.01$, $s < 0.025$ for $U = 0.05$, and $s < 0.04$ for $U = 0.1$.

There are at least two reasons to expect a discrepancy between the coalescence approximations and the forward simulations in these cases. The first is that, when $n_0 \ll 1$, the time between clicks of the ratchet is so small that it is very difficult to maintain the stability assumed in the approximations over reasonable periods of time. The second is that, due to this fact, the frequency of the least-loaded class experiences large fluctuations and spends a considerable amount of time above the expected value assumed in the coalescent approximations. This implies that the level of genetic diversity is likely to be underestimated by the coalescent approach. We observe such underestimation whenever selection is very weak and the mutation rate is very high, so that the ratchet clicks >100 times over N generations.

From the results presented here, we conclude that Muller's ratchet can considerably reduce genetic diversity at a neutral locus. The extent to which this variation is reduced depends strongly on s (Figure 1), N (Figure 2), and U (Figure 3). For large values of U, the reduction is essentially unaffected by changes in s over a wide range of intermediate selection coefficients (Figure 3), which is important since the exact value of s is poorly known.

Muller's ratchet and the frequency spectrum: We now consider the effect of Muller's ratchet on the frequency spectrum of mutations at the neutral locus. As explained above, we examined Tajima's D, which is widely used for this purpose (Fu 1997). In Table 1, we show, for different values of s, the time between clicks of the ratchet and average values of D obtained from forward and coalescent simulations, assuming a fixed value of θ. Coalescent simulations were run to compare with the results obtained from forward simulations. As can be seen in Table 1, they agree quite well with each other. The power to detect deviations from neutrality in samples of size 25 is also shown in Table 1. Table 2 shows the results of coalescent simulations where a fixed number of segregating mutations, S, was distributed over the trees (see SIMULATION METHODS).

We find that the operation of Muller's ratchet produces negative values of Tajima's D in samples of realistic size. The mean value of D for different values of N and intermediate values of s is ~ -1. For the θ and S values considered, with intermediate values of Ns there is considerable power to detect deviations from neutrality in samples of size 25. For sample sizes of 10, however, we generally found no power to reject neutrality (results not shown). For a given N, the maximum negative average value of Tajima's D is observed for intermediate values of s. We observe that, as the time between turns of the ratchet becomes very large, by increasing s and N (or decreasing U), the average value of Tajima's D becomes less negative and the frequency spectrum becomes closer to that expected under neutrality ($D \rightarrow 0$), as expected from previous results on background selection (CHARLESWORTH *et al.* 1995). We also computed values of Fu and Li's D^* (not shown) and found that this statistic is less powerful than Tajima's D to detect deviations from neutrality, for the sample size considered and this range of parameter values.

TABLE 1

Mean Tajima's *D* and power with fixed θ

	$N = 2000$				$N = 3000$				$N = 8000$			
s	T (gen)	D coal (2 SE)	D forw (2 SE)	Pow 1 (%)	T (gen)	D coal (2 SE)	D forw (2 SE)	Pow 1 (%)	T (gen)	D coal (2 SE)	D forw (2 SE)	Pow 1 (%)
0.003	50	−0.97 (0.04)	−1.01 (0.18)	25	56	−0.96 (0.04)	−1.05 (0.17)	24	73	−1.02 (0.04)	−1.03 (0.23)	26
0.005	72	−1.08 (0.04)	−1.26 (0.15)	30	80	−1.11 (0.04)	−1.02 (0.19)	31	102	−1.10 (0.04)	−1.05 (0.20)	23
0.01	144	−1.20 (0.04)	−1.30 (0.16)	35	165	−1.21 (0.04)	−1.14 (0.13)	35	293	−1.14 (0.04)	−1.03 (0.17)	21
0.015	325	−1.22 (0.04)	−0.96 (0.22)	33	611	−1.15 (0.04)	−1.11 (0.18)	30	2239	−0.90 (0.06)	−0.80 (0.18)	17
0.02	1292	−1.03 (0.04)	−1.15 (0.20)	27	3478	−0.91 (0.06)	−0.95 (0.28)	23	—	−0.55 (0.06)	−0.65 (0.23)	10
0.03	—	−0.60 (0.06)	−0.74 (0.20)	13	—	−0.53 (0.06)	−0.80 (0.22)	12				
0.04	—	−0.33 (0.06)	−0.29 (0.24)	9								

The mean values of Tajima's *D* for different *s* values (two standard errors are shown below the means), based on forward Monte Carlo simulations (*D* forw) and coalescent simulations (*D* coal). Pow 1 is the percentage of coalescent simulations that yielded values of *D* lower than the critical value (at the 95% probability level) obtained as explained in SIMULATION METHODS. *T* is the time between turns of Muller's ratchet calculated from the simulations. —, no clicks were observed in the simulations. Sample size is 25. $\theta = 2N\mu = 8$, for the neutral locus as a whole. The deleterious mutation rate is 0.05.

Muller's ratchet with two types of deleterious mutations: Assume now that there are two major types of deleterious mutations: one class of mutations causing very strongly deleterious effects (s_b) and another class with weak deleterious effects (s_s), occurring at rates U_b and U_s, respectively. Although this mutational model is probably too simplistic biologically, it has been suggested that it provides a reasonably good fit to data from experiments on the fitness effects of induced mutations, at least in *Caenorhabditis elegans* (DAVIES *et al.* 1999). In addition, it allows us to explore the combined operation of two processes: Muller's ratchet and background selection (see below; CHARLESWORTH 1996b; GORDO and CHARLESWORTH 2001). The deterministic equilibrium frequency of the class with *i* mutations of effect s_s and *j* mutations of effect s_b, after selection, is the product of the relevant Poisson distributions (JOHNSON 1999). In particular the size of the least-loaded class, after selection, is

$$n_{00}^{*} = f_{0_s}^{*} f_{0_b}^{*} = Ne^{-U_s(1-s_s)/s_s}\, e^{-U_b(1-s_b)/s_b}.$$

We can easily extend the coalescent approach used above to this two-type mutation model. The expression for the mean number of pairwise differences relative to the neutral case is given in APPENDIX B.

If $n_{00} > 1$, the population will be close to the deterministic equilibrium most of the time and the sizes of the classes can be well approximated by Equation B3. When s_s is small and/or U_s is large, such that $n_{00} < 1$, we approximate the distribution of the classes with respect to these mutations by a shifted Poisson with parameter λ_s (see APPENDIX B; GESSLER 1995).

In Table 3, we show the mean number of pairwise differences relative to the neutral case, in populations of size 3000 and 6000 subject to both types of deleterious mutations. We also show the case when the deleterious mutations with selection coefficient s_b are absent, for comparison, and the results from Equation B3, which are referred to as "theoretical." The distortion of the neutral frequency spectrum, as measured by the mean Tajima's *D*, is given for every set of parameters.

There are several distinct cases that can occur in a two-type mutational model. The first is the accumulation of mutations of effect s_s in the presence of much more strongly deleterious mutations, for which there is no ratchet—*i.e.*, the combined operation of Muller's ratchet and background selection (GORDO and CHARLESWORTH 2001). The large effect mutations are expected to reduce variability by a fraction f_{0_b}, and the additional presence of the other mutations, which are accumulating due to Muller's ratchet, is expected to reduce variability even more. For the cases where this occurs (a, b, g, h, i, and k in Table 3), we see that Equation B3 gives good predictions of the relative diversity observed in the simulations. Strongly deleterious mutations reduce diversity at neutral sites by a fraction f_{0_b}, but they also reduce the effective population size experienced by the small effect mutations by approximately the same amount (GORDO and CHARLESWORTH 2001). The small effect mutations will then cause a reduction in genetic diversity according to this new effective size (Nf_{0_b}). It follows that, in this

TABLE 2

Mean Tajima's *D* and power with fixed *S*

	$N = 2000$		$N = 3000$		$N = 8000$	
s	*D* coal (2 SE)	Pow 2 (%)	*D* coal (2 SE)	Pow 2 (%)	*D* coal (2 SE)	Pow 2 (%)
			$S = 30$			
0.003	−1.17	33	−1.31	43	−1.54	58
	(0.01)		(0.01)		(0.01)	
0.005	−1.33	41	−1.43	50	−1.68	70
	(0.01)		(0.01)		(0.01)	
0.01	−1.53	57	−1.66	69	−1.88	86
	(0.01)		(0.01)		(0.01)	
0.015	−1.60	65	−1.61	65	−1.42	50
	(0.01)		(0.01)		(0.01)	
0.02	−1.31	42	−1.20	36	−0.83	19
	(0.01)		(0.01)		(0.01)	
0.03	−0.71	15	−0.59	12	−0.34	8
	(0.02)		(0.02)		(0.02)	
0.04	−0.43	9	−0.36	8	−0.20	5
	(0.02)		(0.02)		(0.02)	
			$S = 10$			
0.003	−1.03	27	−1.17	35	−1.37	47
	(0.01)		(0.01)		(0.01)	
0.005	−1.19	35	−1.26	40	−1.49	56
	(0.01)		(0.01)		(0.01)	
0.01	−1.36	46	−1.47	53	−1.65	68
	(0.01)		(0.01)		(0.01)	
0.015	−1.40	48	−1.43	50	−1.26	41
	(0.01)		(0.01)		(0.04)	
0.02	−1.17	34	−1.06	30	−0.72	17
	(0.01)		(0.02)		(0.02)	
0.03	−0.65	15	−0.53	13	−0.31	8
	(0.02)		(0.02)		(0.02)	
0.04	−0.37	9	−0.31	8	−0.18	6
	(0.02)		(0.02)		(0.02)	

The mean values of Tajima's *D* for different *s* values (two standard errors are shown below the means), based on simulations of the structured coalescent are shown. Pow 2 is the percentage of simulations that yielded values of *D* lower than the critical value (at the 95% probability level) obtained by coalescent simulations of the neutral model, as explained in SIMULATION METHODS. The number of segregating sites, *S*, in the sample is fixed. $U = 0.05$.

case, the resulting reduction in the mean number of pairwise differences caused by both types of mutations is given by

$$f_{0_b} \frac{k}{k_0}(Nf_{0_b}, U_s, s_s)$$

with k/k_0 calculated with Equation A5.

The average values of Tajima's *D* are ~ -0.9 and there is some power to reject neutrality in samples of reasonable size (25 chromosomes and $\theta = 6$, in the cases in Table 3).

The second case occurs when both types of mutations are accumulating due to the ratchet. In Table 3, we show some examples of this (c, d, e, f, l, m, o, and p). We see that Equation B3 predicts the expected mean number of pairwise differences relative to that under strict neutrality reasonably well. Average values of Tajima's *D* are between -0.8 and -1, for the θ value considered, and there is some power to detect a distortion in the frequency spectrum, for a sample size of 25.

The third case occurs if the effects of both types of mutations are very large and/or the mutation rates are very small, such that none will accumulate. This corresponds to the classical background selection model, with no recombination and two mutational classes. In Table 3, we see that, when we did not observe any clicks of the ratchet (cases j and n) and when $n_{00} \gg 1/s_s$ and $n_{00} \gg 1/s_b$, the reduction in genetic diversity is well approximated by f_{00} (as expected from the expressions in APPENDIX B). Note that this is the result expected

TABLE 3

The reduction in the mean number of pairwise differences (k/k_0) due to Muller's ratchet with two classes of mutations

s_s	$N = 3000$: k/k_0 (2 SE)	$U_b = 0.03$: k/k_0 theoretical	$s_b = 0.0275$: Ts_s	Ts_b	$U_s = 0.05$: D (2 SE)	Pow 1 (%)	k/k_0 (2 SE): $U_b = 0$
a. 0.003	0.124 (0.023)	0.136	44	—	−0.91 (0.21)	16	0.156 (0.019)
b. 0.005	0.116 (0.019)	0.110	57	—	−0.95 (0.17)	22	0.140 (0.020)
c. 0.01	0.082 (0.018)	0.080	115	6753	−0.93 (0.21)	25	0.104 (0.024)
d. 0.015	0.086 (0.016)	0.064	231	3858	−0.98 (0.24)	36	0.102 (0.017)
e. 0.02	0.083 (0.018)	0.067	494	4260	−0.77 (0.24)	21	0.125 (0.018)
f. 0.03	0.085 (0.022)	0.093	4560	4913	−0.93 (0.19)	25	0.193 (0.022)
s_s	$N = 3000$: k/k_0 (2 SE)	$U_b = 0.03$: k/k_0 theoretical	$s_b = 0.0275$: Ts_s	Ts_b	$U_s = 0.01$: D (2 SE)	Pow 1 (%)	k/k_0 (2 SE): $U_b = 0$
g. 0.003	0.195 (0.035)	0.211	402	—	−0.65 (0.23)	16	0.273 (0.036)
h. 0.005	0.164 (0.024)	0.158	656	—	−0.86 (0.23)	24	0.291 (0.045)
i. 0.01	0.15 (0.032)	0.172	4313	—	−0.86 (0.24)	30	0.401 (0.056)
j. 0.02	0.246 (0.039)	0.233	—	—	−0.36 (0.28)	14	0.565 (0.089)
s_s	$N = 6000$: k/k_0 (2 SE)	$U_b = 0.03$: k/k_0 theoretical	$s_b = 0.0275$: Ts_s	Ts_b	$U_s = 0.05$: D (2 SE)	Pow 1 (%)	k/k_0 (2 SE): $U_b = 0$
k. 0.005	0.079 (0.018)	0.059	69	—	−0.86 (0.23)	26	0.081 (0.011)
l. 0.01	0.060 (0.017)	0.044	154	11657	−0.79 (0.21)	16	0.051 (0.013)
m. 0.02	0.062 (0.014)	0.049	1508	15188	−0.99 (0.190)	26	0.090 (0.019)
n. 0.03	0.056 (0.013)	0.081	—	—	−0.65 (0.25)	19	0.171 (0.030)
s_s	$N = 3000$: k/k_0 (2 SE)	$U_b = 0.03$: k/k_0 theoretical	$s_b = 0.01$: Ts_s	Ts_b	$U_s = 0.03$: D (2 SE)	Pow 1 (%)	k/k_0 (2 SE): $U_b = 0$
o. 0.003	0.135 (0.021)	0.156	68	480	−1.18 (0.16)	32	0.190 (0.030)
p. 0.005	0.111 (0.024)	0.123	101	325	−1.14 (0.20)	30	0.186 (0.030)

Ts_b and Ts_s are the average time (in generations) between clicks of the ratchet with respect to each type of deleterious mutation. —, no clicks were observed during the runs. Other symbols are as in Table 1. Pow 1 is based on forward simulations with fixed $\theta = 6$. The last column contains the reduction of neutral diversity in the absence of mutations with selection coefficient s_b.

from a one-class deleterious mutational model in which the relevant selection coefficient is the harmonic mean of the selection coefficients in the two-class mutational model (Charlesworth 1996a).

The fourth case occurs when the presence of strongly deleterious mutations reduces the effective population size by such a large amount that the smaller mutations become effectively neutral, *i.e.*, $Nf_{0_b} s_s < 1$ (Charlesworth 1996b). Under these conditions, genetic drift is the major force determining the dynamics of the small effect mutations and driving them to fixation. Some examples of this case are considered in Table 4, with two different mutational models for the small effect mutations: one considering irreversible mutation and another, more realistic model, allowing for back mutation (McVean and Charlesworth 2000). In these cases, the reduction in the mean number of pairwise differences is very close to the one caused by the strong mutations, since the weak ones are effectively neutral (Kimura 1983) and do not have any significant effect on the neutral locus at which variation is being measured. Therefore, k/k_0 can essentially be approximated by f_{0_b}. In this case, the average Tajima's D is much less negative than in some of the previous cases, and it is very difficult to detect distortions in the frequency spectrum (*cf.* Charlesworth *et al.* 1995), especially when allowing for back mutation.

As in the previous model, in this two-type mutational model we also find that, when Muller's ratchet starts to operate, the level of k/k_0 is roughly the same over intermediate values of the selection coefficient, for a fixed population size and mutation rate.

DISCUSSION

Muller's ratchet and neutral variation: Genetic diversity at a neutral locus results from the balance between the rate at which variation is generated (mutation pressure) and the rate with which it is lost (genetic drift). In

TABLE 4

The effects of background selection on weakly selected mutations with and without back mutation

		Without back mutation				$U_s = 0$		
U_{back}	s_s	k/k_0 (2 SE)	f_{0_b}	D (2 SE)	Pow 1 (%)	k/k_0 (2 SE)	D (2 SE)	Pow 1 (%)
0	0.001	0.127 (0.032)	0.115	−0.76 (0.24)	22	0.130 (0.029)	−0.46 (0.23)	8
0	0.0005	0.102 (0.025)	0.115	−0.61 (0.26)	21			
		With back mutation				$U_s = 0$		
U_{back}	s_s	k/k_0 (2 SE)	f_{0_b}	D (2 SE)	Pow 1 (%)	k/k_0 (2 SE)	D (2 SE)	Pow 1 (%)
0.001	0.0005	0.115 (0.027)	0.115	−0.46 (0.29)	14	0.130 (0.029)	−0.46 (0.23)	8
0.01	0.0005	0.146 (0.029)	0.115	−0.39 (0.22)	8			

U_{back} is the mutation rate to back mutations, with respect to the small-effect mutational type. Power is based on forward simulations with $\theta = 6$. Other symbols are as in Table 1. Parameter values are $N = 3000$, $U_b = 0.09$, $s_b = 0.04$, $U_s = 0.01$. The simulation results due simply to the presence of mutations with effect s_b at rate U_b (*i.e.*, without weak mutations) are presented in the last column. The value of U_b was chosen to make the effect of background selection sufficiently strong that the weak mutations become effectively neutral in the presence of the strong mutations.

a population that is permanently at equilibrium under recurrent mutation to deleterious alleles, in the absence of recombination neutral genetic diversity is expected to be smaller than the strict neutral expectation (Charlesworth *et al.* 1993). This results from the fact that a large fraction of individuals in such a population are destined to be eliminated relatively quickly, so that its effective size is reduced to the class of individuals that do not carry deleterious mutations, Nf_0 (Charlesworth *et al.* 1993).

In this article, we have quantified the expected genetic diversity when a population is not permanently at equilibrium, but is losing its least-loaded class at a given rate. We have shown that the operation of Muller's ratchet is consistent with a considerable reduction in genetic diversity. The extent to which such variation is reduced is a function, not only of the relevant mutation and selection parameters, but also of population size. In particular, in a population where Muller's ratchet does not operate, or does so at an exceedingly slow rate, which is expected when $Nf_0s \gg 10$ (Gordo and Charlesworth 2000a,b), the effective size is well approximated by Nf_0. But when the ratchet starts to operate, the effective size is higher than Nf_0. Although it has been suggested that the operation of Muller's ratchet causes a different value of genetic diversity from that given by the simple $2Nf_0\mu$ approximation (Charlesworth *et al.* 1993), this study is the first attempt to formally demonstrate that it does so and to estimate by how much.

We have shown that the mean coalescent time of two randomly sampled alleles derived from a structured coalescent model with fixed class size (Hudson and Kaplan 1994) is a good predictor of expected genetic diversity when the ratchet is operating. Just as is observed in the full Monte Carlo simulations, the analytical approximation predicts a minimum genetic diversity for an intermediate value of the selection coefficient. Our results are closely related to those of Higgs and Woodcock (1995), who studied the effect of deleterious mutations on genealogies in very small populations and showed that the probability of common parentage was maximal for some intermediate value of the selection coefficient against deleterious alleles. The results from Figure 8 of Higgs and Woodcock (1995) can be obtained by our approximation, provided that we correct them for sampling after selection.

Our results are also related to those of Tachida (2000), who found that diversity at neutral sites was minimal for an intermediate strength of selection, although his model is different from the one we consider here. In Tachida's independent multicodon (IMC) model, a gene is composed of a set of completely linked sites. One-third of the sites are neutral and two-thirds are selected, with selection coefficients drawn from a normal distribution. In our model, the selection coefficient is constant, but the qualitative effect on neutral diversity is the same. The simulation results regarding genetic diversity at neutral sites in Table 1 of Tachida (2000) can be obtained by our approximation, if we substitute s in our approximation by the value corresponding to the mean strength of selection (α) considered in his Table 1 ($\alpha \approx 2Ns$). With our formula, we obtain good estimates of the average genetic diversity at neutral sites observed in his simulations, except when $\alpha < 5$. As an example, with $u = 1 \times 10^{-5}$ per site (implying $U = 0.002$ for the whole nonrecombining region) and $\alpha = 5$ (implying $s = 0.005$), the value of neutral variability observed in Tachida's simulations is 0.00857 and the value predicted by our approximation is 0.00836.

In contrast to the classical background selection model with strong selection (Hudson and Kaplan 1994; Charlesworth *et al.* 1995), if Muller's ratchet is op-

erating under weak selection, a considerable distortion of the frequency spectrum at the neutral locus, toward an excess of rare variants, is expected in samples of realistic size (as seen in Table 1). But such an effect may be difficult to detect when the ratchet causes a very large reduction in variation, as may be the case in large populations (see below). This signature of the ratchet is quite close to that of selective sweeps, but not as extreme (see below). For the parameter values tested we found more power for Tajima's D test than for Fu and Li's D^* test to reject neutrality under the operation of the ratchet than in its absence.

Because a model that considers that all deleterious mutations have the same effect on fitness is a simplification, we also studied the pattern of neutral variation under a two-type deleterious mutational model (Gordo and Charlesworth 2000b). We considered several distinct cases. In the case where none of the deleterious mutations accumulate, the classical background selection scenario, we recover the expected prediction: The reduction in genetic diversity is well approximated by considering the harmonic mean of the selection coefficients of the two-type model. Tajima's D is negative on average, but distortions of the frequency spectrum are hard to detect (see Table 3). In the case where one of the mutational types is sufficiently strongly selected against that it does not accumulate, but the other type of deleterious mutations does accumulate, we essentially observe the effects of the ratchet in a population of reduced size Nf_0. In the case where both mutational types accumulate, we again observe negative values of Tajima's D and obtain reasonably good predictions of the genetic diversity by the extended coalescent approach.

It would be of interest to study a potentially more realistic case that considers a continuous distribution of selection coefficients, but we have not pursued this here. In qualitative terms, one would expect that a continuous distribution with a high frequency of weakly deleterious mutations, such as the case of an exponential distribution, would produce similar results to those in the two-type model considered here: *i.e.*, negative Tajima's D values and a reduction in genetic diversity. But the quantitative effects will depend on the shape of the distribution of selection coefficients and the deleterious mutation rate, which are presently a matter for debate (Davies *et al.* 1999; Keightley and Eyre-Walker 1999).

Muller's ratchet and the *Y* chromosome: It has been suggested that Muller's ratchet has been a potentially major process in shaping the evolution of Y chromosomes (Charlesworth 1978). It is therefore interesting to try to quantify the levels of neutral variation expected under its operation. It is of special interest to ask about the diversity levels expected under the ratchet

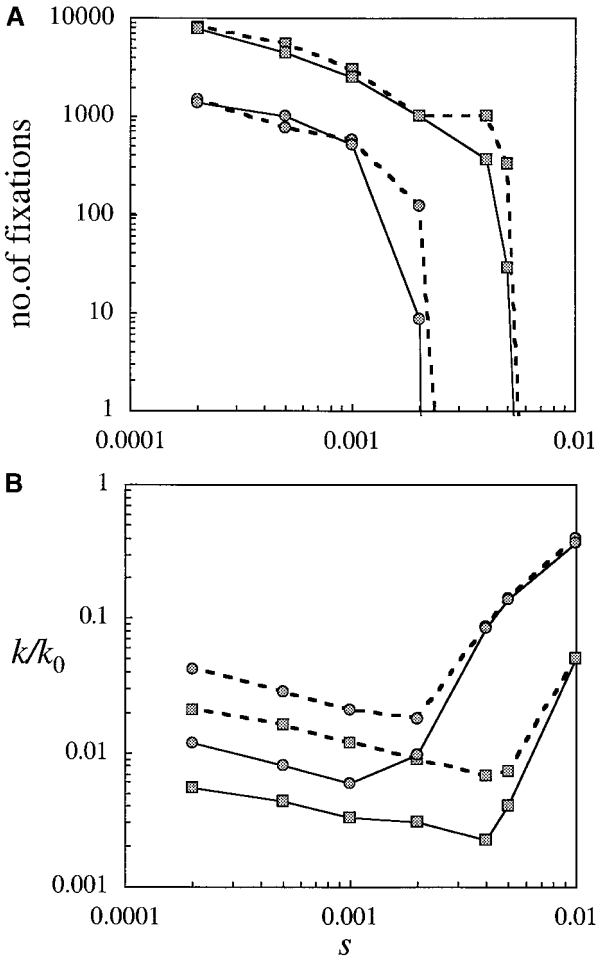

Figure 4.—The signatures of Muller's ratchet in large hypothetical populations of nonrecombining Y or neo-Y chromosomes. The parameters are $N = 125{,}000$ for dashed lines and $N = 500{,}000$ for solid lines, with $U = 0.01$ for circles and $U = 0.03$ for squares. (A) The expected number of fixations over a period of 500,000 generations. These are based on the expressions for the time between clicks of Muller's ratchet in Gordo and Charlesworth (2000a,b) for the cases when $Nf_0 > 1$ and the results of Gessler (1995) for the cases when $Nf_0 < 1$. (B) Expected mean number of pairwise differences relative to that in the absence of deleterious mutations, calculated using the analytical prediction.

in systems with relatively young Y chromosomes (Rice 1996; Charlesworth and Charlesworth 2000). Some examples of these systems are the Y chromosomes of the plant species *Silene latifolia* and *S. dioica* (Filatov *et al.* 2000) and the neo-Y chromosomes (resulting from fusions between an autosome and the old Y chromosome) of some Drosophila species such as *D. miranda* (Charlesworth and Charlesworth 2000).

In Figures 4 and 5 we show some expectations for the signatures of Muller's ratchet. Figure 4A shows the expected number of fixations of deleterious alleles over a period of 500,000 generations (in the case of *D. miranda*, this corresponds to ~0.1 million years). Figure

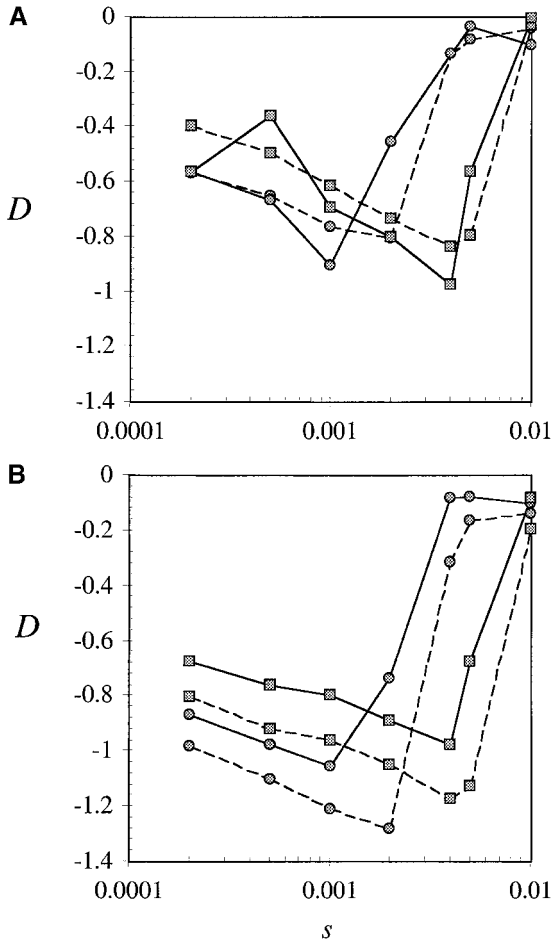

FIGURE 5.—Average value of Tajima's *D* for a sample size of 12 for $\theta = 5$ (A) and sample size 12 for $\theta = 100$ (B). Lines and symbols are as in Figure 4.

4B shows the reduction in the mean number of pairwise differences. Figure 5, A and B, shows the average values of Tajima's *D* caused by the ratchet. The number of fixations is estimated from the number of clicks of Muller's ratchet over the time period, since there is a one-to-one correspondence between clicks and fixation events (CHARLESWORTH and CHARLESWORTH 1997). The parameter values *N* and *U* were assigned in the light of the data presently available (DRAKE *et al.* 1998; KEIGHTLEY and EYRE-WALKER 1999; FILATOV *et al.* 2000; YI 2000). As previously discussed (GORDO and CHARLESWORTH 2000b) substantial declines in fitness of the nonrecombining chromosome can be produced, especially for intermediate *s* values.

Associated with the fixation of deleterious mutations by the ratchet, a reduction in genetic diversity of ~10- to 100-fold is expected (as calculated by Equation B3 and the simulations of the structured coalescent). This is expected across a wide range of values of selection coefficients for which the ratchet can operate, since the mutation rate is the major determinant of the level of variation expected (Figure 4B). In samples of moderate size ($n = 12$ in Figure 5), average Tajima's *D* values of ~ -1 are expected when the selection coefficient is intermediate. For larger samples, the average values of Tajima's *D* become more negative. As an example, with a sample size of 40, with $N = 125{,}000$, $U = 0.01$, and $s = 0.1$–0.2%, we obtained an average value of Tajima's *D* of -1.7 for $\theta = 50$ and -1.9 for $\theta = 100$. The power to reject neutrality for these two examples was >80% (assuming a fixed θ). A large amount of sequence information and large samples are, however, needed to detect this effect. For example, in Drosophila, where normal levels of variability are ~1–3% per nucleotide site (MORIYAMA and POWELL 1996), the above example implies sequencing ~5000–10,000 neutral sites. One can ask if increasing sample size (n) will produce higher power than increasing θ by increasing the number of sites sequenced. From simulations of the structured coalescent with $n\theta$ held constant, we found that increasing n seems to give more power than increasing θ.

The additional presence of much more strongly deleterious mutations, causing background selection, will result both in an increase in the number of fixations (GORDO and CHARLESWORTH 2001) and in a bigger reduction in genetic diversity, as expected from the results presented before. There is evidence for considerably reduced levels of variability in some *Y* chromosome systems (CHARLESWORTH and CHARLESWORTH 2000). For the *Y* chromosome of the dioecious plant *S. latifolia* and the neo-*Y* chromosome of *D. miranda*, nucleotide variability is 20- to 30-fold lower than for the *X* chromosome (FILATOV *et al.* 2000; BACHTROG and CHARLESWORTH 2002). If the simple process we have studied was the sole cause of the observed reduction, the results in Figure 4B imply that the deleterious mutation rate for such nonrecombining chromosomes is unlikely to be >0.01.

Selective sweeps *vs.* the ratchet: A large reduction in variability could, of course, be caused by another process, such as a recent selective sweep. When an advantageous mutation arises and goes to fixation in a nonrecombining population, it wipes out linked neutral variation—the hitchhiking effect (MAYNARD SMITH and HAIGH 1974). After such a sweep, variation is slowly restored by mutation, with most of the new neutral variants being at low frequency. Selective sweeps therefore cause distortions of the neutral frequency spectrum (BRAVERMAN *et al.* 1995; SIMONSEN *et al.* 1995), just as with repetitive clicks of the ratchet. In Table 5, we compare the pattern of variability under the ratchet and under a recent sweep. We assume knowledge of the neutral equilibrium value of θ, in the absence of any of these processes, and study conditions under which genetic diversity is reduced by ~20- to 30-fold. As is clear from Table 5, a recent sweep generally produces more negative average Tajima's *D* than the ratchet, for

TABLE 5

Comparison of the ratchet with the hitchhiking model

	Ratchet: $N = 125{,}000$, $U = 0.01$			Hitchhiking: $N = 125{,}000$		
n	S	D	Pow 1 (%)	S	D	Pow 1 (%)
	25-fold reduction					
	$s = 0.0002$, $k/k_0 = 0.04$, $\theta = 5$			$T_{hh} = 0.04N$ gen, $\theta = 5$		
n	S	D	Pow 1 (%)	S	D	Pow 1 (%)
12	0.79	−0.56	2	1.12	−1.22	12
40	1.66	−0.84	10	2.93	−1.51	48
100	3.00	−1.11	32	5.39	−1.71	84
	$\theta = 10$			$\theta = 10$		
n	S	D	Pow 1 (%)	S	D	Pow 1 (%)
12	1.63	−0.63	7	2.16	−1.34	31
40	3.38	−1.06	37	6.01	−1.85	91
100	5.92	−1.36	57	11.1	−2.10	99
	$\theta = 20$			$\theta = 20$		
n	S	D	Pow 1 (%)	S	D	Pow 1 (%)
12	3.21	−0.73	26	4.26	−1.55	59
40	6.78	−1.29	45	11.93	−2.19	99
100	11.98	−1.67	73	22.08	−2.38	100
	33-fold reduction					
	$s = 0.0005$, $k/k_0 = 0.03$, $\theta = 5$			$T_{hh} = 0.03N$ gen, $\theta = 5$		
n	S	D	Pow 1 (%)	S	D	Pow 1 (%)
12	0.58	−0.67	2	0.82	−1.18	6
40	1.29	−0.91	11	2.27	−1.43	37
100	2.35	−1.12	29	4.66	−1.69	80
	50-fold reduction					
	$s = 0.001$, $k/k_0 = 0.02$, $\theta = 5$			$T_{hh} = 0.02N$ gen, $\theta = 5$		
n	S	D	Pow 1 (%)	S	D	Pow 1 (%)
2	0.46	−0.69	2	0.53	−1.19	4
40	1.09	−1.04	9	1.63	−1.36	25
100	2.01	−1.15	24	3.53	−1.56	65

The reduction in genetic diversity was chosen to be of the order of that observed in the *D. miranda* neo-*Y* chromosome data (Bachtrog and Charlesworth 2002). The *S* and *D* values are means over 1000 coalescent trees for each model. For the ratchet, results are based on the structured coalescent; for hitchhiking, results are based on coalescent simulations (Hudson 1990) from a population that expanded from size 1 to *N* instantaneously at time T_{hh} in the past. T_{hh} was chosen to produce the same mean level of variability as with the ratchet.

a given reduction in diversity. In small samples, a sweep is more likely to be detected than the operation of the ratchet, but it is clear that there is a wide range of parameter space in which no unambiguous conclusion can be drawn. Other statistics, such as patterns of linkage disequilibrium, could also be helpful in trying to distinguish between these and other models (Charlesworth and Charlesworth 2000). Further work is needed to assess the power of such statistics to distinguish between background selection, Muller's ratchet, and selective sweeps.

We thank Peter Andolfatto, Carolina Bartolome, Xulio Maside, and Gil McVean for helpful discussions throughout the course of this study. I. Gordo specially thanks Frantz Depaulis for numerous comments that greatly improved the manuscript. I. Gordo is supported by the Gulbenkian Foundation and program PRAXIS XXI of Portugal, A. Navarro by the BBSRC/EPSRC, and B. Charlesworth by the Royal Society.

LITERATURE CITED

Bachtrog, D., and B. Charlesworth, 2002 Reduced adaptation of a non-recombining neo-*Y* chromosome. Nature **416:** 323–326.

Bergstrom, C. T., and J. Pritchard, 1998 Germline bottlenecks

and the evolutionary maintenance of mitochondrial genomes. Genetics **149:** 2135–2146.
Braverman, J. M., R. R. Hudson, N. L. Kaplan, C. H. Langley and W. Stephan, 1995 The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. Genetics **140:** 783–796.
Charlesworth, B., 1978 Model for evolution of *Y* chromosomes and dosage compensation. Proc. Natl. Acad. Sci. USA **75:** 5618–5622.
Charlesworth, B., 1996a Background selection and patterns of genetic diversity in *Drosophila melanogaster*. Genet. Res. **68:** 131–149.
Charlesworth, B., 1996b The evolution of chromosomal sex determination and dosage compensation. Curr. Biol. **6:** 149–162.
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., and D. Charlesworth, 2000 The degeneration of *Y* chromosomes. Philos. Trans. R. Soc. Lond. B Biol. Sci. **355:** 1563–1572.
Charlesworth, B., M. T. Morgan and D. Charlesworth, 1993 The effect of deleterious mutations on neutral molecular variation. Genetics **134:** 1289–1303.
Charlesworth, D., B. Charlesworth and M. T. Morgan, 1995 The pattern of neutral molecular variation under the background selection model. Genetics **141:** 1619–1632.
Crow, J. F., and M. Kimura, 1970 *An Introduction to Population Genetics Theory*. Burgess Publishing/Alpha Editions, Minneapolis.
Davies, E. K., A. D. Peters and P. D. Keightley, 1999 High frequency of cryptic deleterious mutations in *Caenorhabditis elegans*. Science **285:** 1748–1751.
Drake, J. W., B. Charlesworth, D. Charlesworth and J. F. Crow, 1998 Rates of spontaneous mutation. Genetics **148:** 1667–1686.
Ewens, W. J., 1979 *Mathematical Population Genetics*. Springer-Verlag, Berlin.
Felsenstein, J., 1974 The evolutionary advantage of recombination. Genetics **78:** 737–756.
Filatov, D. A., F. Moneger, I. Negrutiu and D. Charlesworth, 2000 Low variability in a *Y*-linked plant gene and its implications for *Y*-chromosome evolution. Nature **404:** 388–390.
Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford.
Fu, Y., 1997 Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics **147:** 915–925.
Fu, Y., and W. Li, 1993 Statistical tests of neutrality of mutations. Genetics **133:** 693–709.
Gessler, D. D., 1995 The constraints of finite size in an asexual population and the rate of the ratchet. Genet. Res. **66:** 241–253.
Gessler, D. D., and S. Xu, 1999 On the evolution of recombination and meiosis. Genet. Res. **73:** 119–131.
Gordo, I., and B. Charlesworth, 2000a The degeneration of asexual haploid populations and the speed of Muller's ratchet. Genetics **154:** 1379–1387.
Gordo, I., and B. Charlesworth, 2000b On the speed of Muller's ratchet. Genetics **156:** 2137–2140.
Gordo, I., and B. Charlesworth, 2001 The speed of Muller's ratchet, with background selection and the degeneration of *Y* chromosomes. Genet. Res. **78:** 149–162.
Haigh, J., 1978 The accumulation of deleterious genes in a population—Muller's Ratchet. Theor. Popul. Biol. **14:** 251–267.
Higgs, P., and G. Woodcock, 1995 The accumulation of mutations in asexual populations and the structure of genealogical trees in the presence of selection. J. Math. Biol. **33:** 677–702.
Hudson, R. R., 1990 Gene genealogies and the coalescent process. Oxf. Surv. Evol. Biol. **7:** 1–45.
Hudson, R. R., 1993 The how and why of generating neutral genealogies, pp. 23–36 in *Mechanisms of Molecular Evolution*, edited by N. Takahata and A. G. Clark. Japan Scientific Society, Tokyo.
Hudson, R. R., and N. K. Kaplan, 1994 Gene trees with background selection, pp. 140–153 in *Non-Neutral Evolution: Theories and Molecular Data*, edited by B. Golding. Chapman & Hall, New York.
Johnson, T., 1999 The approach to mutation-selection balance in an infinite asexual population, and the evolution of mutation rates. Proc. R. Soc. Lond. Ser. B **266:** 2389–2397.
Keightley, P. D., 1996 Nature of deleterious mutation load in Drosophila. Genetics **144:** 1993–1999.
Keightley, P. D., and A. Eyre-Walker, 1999 Terumi Mukai and the riddle of deleterious mutation rates. Genetics **153:** 515–523.
Kimura, M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, UK.
Lynch, M., R. Burger, D. Butcher and W. Gabriel, 1993 The mutational meltdown in asexual populations. J. Hered. **84:** 339–344.
Maynard Smith, J., 1978 *The Evolution of Sex*. Cambridge University Press, Cambridge, UK.
Maynard Smith, J., and J. Haigh, 1974 The hitch-hiking effect of a favourable gene. Genet. Res. **23:** 23–35.
McVean, G. A., and B. Charlesworth, 2000 The effects of Hill-Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. Genetics **155:** 929–944.
Moriyama, E. N., and J. R. Powell, 1996 Intraspecific nuclear DNA variation in Drosophila. Mol. Biol. Evol. **13:** 261–277.
Muller, H. J., 1964 The relation of recombination to mutational advance. Mutat. Res. **1:** 2–9.
Pamilo, P., M. Nei and W. H. Li, 1987 Accumulation of mutations in sexual and asexual populations. Genet. Res. **49:** 135–146.
Rice, W. R., 1996 Evolution of the Y sex chromosome in animals. Bioscience **46:** 331–343.
Simonsen, K. L., G. A. Churchill and C. F. Aquadro, 1995 Properties of statistical tests of neutrality for DNA polymorphism data. Genetics **141:** 413–429.
Stephan, W., L. Chao and J. G. Smale, 1993 The advance of Muller's ratchet in a haploid asexual population: approximate solutions based on diffusion theory. Genet. Res. **61:** 225–231.
Tachida, H., 2000 Molecular evolution in a multisite nearly neutral model. J. Mol. Evol. **50:** 69–81.
Tajima, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123:** 585–595.
Yi, S., 2000 Molecular variation and the evolution of the neo-sex chromosomes of *Drosophila miranda*. Ph.D. Dissertation, University of Chicago, Chicago.

Communicating editor: N. Takahata

APPENDIX A

Assume that there are m mutational classes in the population, so that a sample may contain gametes with 0, 1, 2, up to m mutations. The probability that a gamete with i mutations derives from a gamete with $i-1$ mutations is

$$Q_{i,i-1} = \frac{f^*_{i-1}Ue^{-U}}{f^*_{i-1}Ue^{-U} + f^*_i e^{-U}} = \frac{(f_{i-1}(1-s)^{i-1}/e^{-U})\,Ue^{-U}}{(f_{i-1}(1-s)^{i-1}/e^{-U})\,Ue^{-U} + (f_i(1-s)^i/e^{-U})\,e^{-U}},$$

which is a special case of Equation 3 in Charlesworth *et al.* 1995, and f^*_i is the frequency of the i class, after selection. (We assume here that the mutation rate is sufficiently low that we can neglect mutations from classes other than the adjacent one.)

If the distribution of the frequencies of the classes is close to the deterministic expectation most of the time, then

$$f^*_0 = e^{-(U/s)(1-s)} \quad \text{and} \quad f^*_i = f^*_{i-1}\frac{U(1-s)/s}{i} \tag{A1}$$

so that

$$Q_{i,i-1} \approx \frac{is}{1 + (i-1)s}. \tag{A2}$$

Suppose that we sample randomly two individuals and that these belong to classes i and j. If i and $j \neq 0$, there are two possible mutational events in the previous generation: Either gamete i came from the $i-1$ class (with probability $Q_{i,i-1}$), or gamete j came from the $j-1$ class (with probability $Q_{j,j-1}$); if $i = j$ they can also coalesce, with probability $1/Nf_i^*$, since the size of class i, after selection (which is when we are sampling) is Nf_i^*. Let $T_{i,j}$ be the mean time (in generations) back to the common ancestor of a sample of two gametes with i and j ($i, j \geq 0$) mutations. We then have

$$T_{i,j} = \left(1 - Q_{i,i-1} - Q_{j,j-1} - \frac{\delta_{ij}}{Nf_i^*}\right)(T_{i,j} + 1) + Q_{i,i-1}(T_{i-1,j} + 1) + Q_{j,j-1}(T_{i,j-1} + 1) + \frac{\delta_{ij}}{Nf_i^*},$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. Rearranging, we have

$$T_{i,j} = \left(\begin{array}{l} \dfrac{1}{Q_{i,i-1} + Q_{j,j-1} + \delta_{ij}/Nf_i^*} \\ \quad + \dfrac{Q_{i,i-1}}{Q_{i,i-1} + Q_{j,j-1} + \delta_{ij}/Nf_i^*}\, T_{i-1,j} \\ \quad + \dfrac{Q_{j,j-1}}{Q_{i,i-1} + Q_{j,j-1} + \delta_{ij}/Nf_i^*}\, T_{i,j-1}, \end{array}\right), \tag{A3}$$

which is equivalent to Equation 12 of Hudson and Kaplan (1994) for a sample size of two, with the difference that we are counting individuals as postselection adults. The mean time for the most recent common ancestor of two randomly sampled sequences is then

$$T_2 = \sum_{i,j=0}^{m} f_i^* f_j^* T_{i,j} \tag{A4}$$

and the resulting mean number of pairwise differences relative to the neutral expectation will be

$$\frac{k}{k_0}(N, U, s) = \frac{T_2}{N}. \tag{A5}$$

APPENDIX B

Suppose that we take a random sample of two individuals from a population subject to recurrent mutations with two types of effect, s_s and s_b, occurring at two different rates, U_s and U_b, respectively. Suppose one individual carries i mutations of type s_s and k mutations of type s_b and the other carries j mutations of type s_s and l mutations of type s_b. Let $T_{i,k,j,l}$ be the time to the most recent common ancestor of these individuals. If the population is close to the deterministic equilibrium this time will be given by

$$T_{i,k,j,l} = \left(\begin{array}{l} \dfrac{1}{Q_{s_{i,i-1}} + Q_{b_{k,k-1}} + Q_{s_{j,j-1}} + Q_{b_{l,l-1}} + \delta_{ij}\delta_{kl}/Nf_i^* f_k^*} \\ \quad + \dfrac{Q_{s_{i,i-1}}}{Q_{s_{i,i-1}} + Q_{b_{k,k-1}} + Q_{s_{j,j-1}} + Q_{b_{l,l-1}} + \delta_{ij}\delta_{kl}/Nf_i^* f_k^*}\, T_{i-1,k,j,l} \\ \quad + \dfrac{Q_{b_{k,k-1}}}{Q_{s_{i,i-1}} + Q_{b_{k,k-1}} + Q_{s_{j,j-1}} + Q_{b_{l,l-1}} + \delta_{ij}\delta_{kl}/Nf_i^* f_k^*}\, T_{i,k-1,j,l} \\ \quad + \dfrac{Q_{s_{j,j-1}}}{Q_{s_{i,i-1}} + Q_{b_{k,k-1}} + Q_{s_{j,j-1}} + Q_{b_{l,l-1}} + \delta_{ij}\delta_{kl}/Nf_i^* f_k^*}\, T_{i,k,j-1,l} \\ \quad + \dfrac{Q_{b_{l,l-1}}}{Q_{s_{i,i-1}} + Q_{b_{k,k-1}} Q_{s_{j,j-1}} + Q_{b_{l,l-1}} + \delta_{ij}\delta_{kl}/Nf_i^* f_k^*}\, T_{i,k,j,l-1} \end{array}\right), \tag{B1}$$

where

$$Q_{s_{i,i-1}} = \frac{is_s}{1 + (i-1)s_s} \quad \text{and} \quad Q_{b_{i,i-1}} = \frac{is_b}{1 + (i-1)s_b},$$

which is the extension of the previous approximation for mutations of equal effects. When $n_{00} < 1$, because $Nf_{0_s} < 1$, we use, as previously, the shifted Poisson distribution with parameter $\lambda_s = U_s/s_s - K_s$, where $K_s = \min\{k: Nf_{0_b} f_{k_s} \geq 1\}$, so that

$$Q_{s_{i,i-1}} = \frac{U_s}{U_s + ((1 - s_s)/i)\lambda_s}. \tag{B2}$$

Using these approximations, the mean time to the most recent common ancestor of two random gametes is

$$T_2 = \sum_{i,j,k,l} f_{ik}^* f_{jl}^* T_{i,j,k,l} \tag{B3a}$$

and

$$\frac{k}{k_0}(N, U_b, s_b, U_s, s_s) = \frac{T_2}{N}, \tag{B3b}$$

with

$$f_{ij}^* = f_{i_s}^* f_{j_b}^* = \left(\frac{U_s(1 - s_s)}{s_s}\right)^i \frac{e^{-U_s(1-s_s)/s_s}}{i!} \times \left(\frac{U_b(1 - s_b)}{s_b}\right)^j \frac{e^{-U_b(1-s_b)/s_b}}{j!}.$$