Germline Bottlenecks and the Evolutionary Maintenance of Mitochondrial Genomes

Carl T. Bergstrom and Jonathan Pritchard[1]

Department of Biological Sciences, Stanford University, Stanford, California 94305

Manuscript received August 15, 1997
Accepted for publication May 6, 1998

ABSTRACT

Several features of the biology of mitochondria suggest that mitochondria might be susceptible to Muller's ratchet and other forms of evolutionary degradation: Mitochondria have predominantly uniparental inheritance, appear to be nonrecombining, and have high mutation rates producing significant deleterious variation. We demonstrate that the persistence of mitochondria may be explained by recent data that point to a severe "bottleneck" in the number of mitochondria passing through the germline in humans and other mammals. We present a population-genetic model in which deleterious mutations arise within individual mitochondria, while selection operates on *assemblages* of mitochondria at the level of their eukaryotic hosts. We show that a bottleneck increases the efficacy of selection against deleterious mutations by increasing the variance in fitness among eukaryotic hosts. We investigate both the equilibrium distribution of deleterious variation in large populations and the dynamics of Muller's ratchet in small populations. We find that in the absence of the ratchet, a bottleneck leads to improved mitochondrial performance and that, over a longer time scale, a bottleneck acts to slow the progression of the ratchet.

THE mitochondrial genome features a mode of reproduction and transmission markedly different from that of the nuclear genome. In mammals, it appears that mitochondrial genomes have no recombination and predominantly uniparental inheritance (Wolstenholme 1992). Moreover, because of their mode of transmission, mitochondrial populations are subject to rapid genetic drift: Numerous studies have observed that the genotype frequencies in heteroplasmic mitochondria can shift greatly in a single transmission from mother to offspring. While the precise mechanics of this process are unknown, the rapid shift of mitochondrial genotype frequencies suggests that somewhere in the host germ line there is a mitochondrial "bottleneck." The effective size of this bottleneck in transmission number from mother to offspring is quite tight—on the order of 1 to 30 in humans (Howell *et al.* 1992; Bendall *et al.* 1996, 1997; Ivanov *et al.* 1996; Marchington *et al.* 1997; Parsons *et al.* 1997), and slightly higher in cattle and mice (reviewed in Bendall *et al.* 1997).

Mitochondrial DNA generally has a mutation rate higher than that of nuclear DNA (Avise 1991; Wolstenholme 1992); these mutations produce significant deleterious variation ranging in effect from mild (Nachmann *et al.* 1996; Lynch 1996) to severe or even lethal (Poulton and Marchington 1996; Brown 1997; Sherratt *et al.* 1997). Moreover, DNA repair in mitochondria is limited (Avise 1991). It is not surprising that mitochondrial genes have been shown to feature higher rates of evolution and higher ratios of nonsynonymous to synonymous substitution than do nuclear genes (Lynch 1996).

By analogy with the population genetics of diploid organisms, one might expect that these characteristics of mitochondrial transmission should make mtDNA highly susceptible to genetic degradation. In particular, several authors (Gabriel *et al.* 1993; Lynch *et al.* 1993; Hurst and McVean 1996; Lynch 1996) have suggested that a bottleneck should reduce the effective population size of mitochondrial or endosymbiont populations and hence hasten genetic degradation. Hurst and McVean, in particular, challenge evolutionists to explain how endosymbionts (such as mitochondria) have managed to stave off the ruinous effects of Muller's ratchet and mutational meltdown (Gabriel *et al.* 1993; Lynch *et al.* 1993) since the time that they first became incorporated into eukaryotic hosts.

The analogy to nuclear genome evolution can be misleading, however. Because of their high copy number and the differences in their reproductive biology, mitochondrial and nuclear genes may respond differently to natural selection. Here, we construct a model of mitochondrial evolution with which to explore these differences. In this article, we focus in particular on the role of the mitochondrial bottleneck.

We demonstrate that, rather than hastening genetic degradation, a bottleneck may be essential in maintaining mitochondrial genetic quality over evolutionary time. We show that while a bottleneck indeed increases

Corresponding author: Jonathan Pritchard, Department of Biological Sciences, Stanford University, Stanford, CA 94305.
E-mail: jkp@charles.stanford.edu

[1] The authors contributed equally to this paper.

the rate of genetic degradation *within* a particular lineage, it also serves to strengthen selection among lineages, and hence has a net effect of resisting genetic decay. We suggest that uniparental inheritance has a similar effect. The bottleneck process in our model is closely related to a process of within-generation drift in extranuclear genomes that was studied in a simulation model by Takahata and Slatkin (1983). In their model, they found that within-generation drift slowed the accumulation of deleterious extranuclear mutations and facilitated the fixation of advantageous extranuclear mutations.

We proceed as follows: In Section 1, we present a basic model of mitochondrial reproduction, including selection, mutation, and the bottleneck phase. In Section 2, we use this model to examine the consequences of a bottleneck on the distribution of variation and fitness in a large population. We also consider the impact of paternal leakage. In Section 3, we argue that in the absence of paternal leakage or recombination, the genetic decay of mitochondria proceeds as a *double ratchet*. The first (host-level) ratchet turns at the level of host individuals, when a parent fails to transmit its best mitochondria to its offspring. The second (population-level) ratchet turns when the best host individuals either fail to reproduce or fail to pass their best mitochondria to the next generation. In this context, we examine the process of genetic decay over evolutionary time.

Two distinct processes drive the accumulation of deleterious mutations and hence pose a threat to the genetic integrity of these endosymbionts. Deleterious alleles may be fixed within the population, or, alternatively, Muller's ratchet may operate (Charlesworth *et al.* 1993). In appendix a we provide an argument that for many fitness models the rates of the two processes converge in the long run. [A similar result was previously suggested by Higgs and Woodcock (1995) on the basis of a simulation study.] This result is useful because to estimate the long-term rates of genetic degradation, it is not necessary to track both the ratchet rate and the fixation rate. Therefore in the present analysis, we restrict ourselves to an analysis of Muller's ratchet. In appendix b, we present detailed derivations of the recursions presented in Section 2.

1. THE MODEL

In this section we introduce a simple model of mitochondrial evolution with which to explore the consequences of varying bottleneck size on the short-term and long-term evolution of mitochondria and their eukaryotic hosts. In this model, the fitness of eukaryotic host-individuals is a decreasing function of the total number of mutations carried by their mitochondria. A schematic diagram of the model is given in Figure 1. Populations are composed of N host individuals, each containing M mitochondria. All hosts transmit mitochondria. Each generation is composed of three steps, as follows:

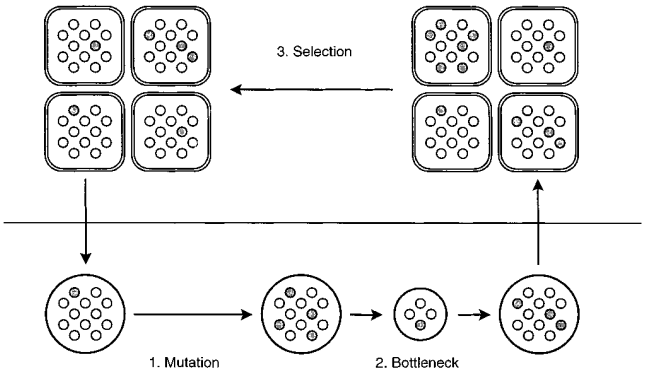

Figure 1.—A schematic representation of the model. The large circles below the horizontal line represent the germline of a single individual; the boxes above the line represent a population of eukaryotic host individuals. The little circles are mitochondria; these are shaded for some mitochondria to indicate the accumulation of deleterious mutations. The model is composed of three stages: (1) Mutation occurs within the mitochondria of each host; (2) each host passes its germline mitochondria through a bottleneck; (3) selection favors hosts with fewer deleterious mutations in their mitochondria.

1. Mutation. Deleterious mutations occur independently in each mitochondrion. The number of new mutations per mitochondrion is Poisson with mean μ per host generation. There is no back mutation.
2. Bottleneck. The mitochondrial compositions of the hosts in the next generation are formed by two rounds of sampling with replacement: the first from the original number M mitochondria per host down to the bottleneck number B, and the second from B back up to M. In this study we examine the consequences of a bottleneck by varying B, while holding the other parameters in the model constant.
3. Selection. Hosts are chosen to reproduce by sampling with replacement with probability proportional to their fitnesses. The fitness of a host is determined by a function of the total number of mutations n in its mitochondria. We assume a linear fitness function: $W = 1 - \alpha n$ with $W \equiv 0$ when $n > \alpha^{-1}$.

We follow previous analyses (Takahata and Slatkin 1983; Clark 1988; Gabriel *et al.* 1993) in assuming that these mildly deleterious mutations have no effect on the within-host replication of individual mitochondria. Experimental data on this point are equivocal (Clark and Lyckegaard 1988; Shoubridge *et al.* 1990; Jenuth *et al.* 1996), but as discussed below, this assumption is not crucial to our conclusions. For simplicity, we assume that the fitness of the host depends only on the number of mutations carried by its mitochondria and not on the arrangement of the mutations among the mitochondria.

Suppose that the mitochondria in a germline cell before the bottleneck phase (*i.e.*, after mutation) of our model contain n_m mutations and that following the bottleneck the daughter cell contains n_b mutations. Then it can be shown that regardless of B, the expectation $E[n_b] = n_m$. This means that the bottleneck procedure itself does not change the *expected* fitness under a linear fitness model. However, nonlinear fitness models will generally cause mean fitness to change through the bottleneck. We base our analysis on a linear model to make the different bottleneck sizes precisely comparable. In the discussion we consider the implications of other kinds of fitness models.

In the absence of knowledge of the precise mechanism of the bottleneck process (see, *e.g.*, Poulton and Marchington 1996), we follow Bendall *et al.* (1996, 1997) in treating this as a two-cell-generation procedure, as described below. When the effective bottleneck is severe, the bottleneck procedure might actually require numerous cell generations, giving rise to intragenerational drift. Our model does not explicitly consider the empirical observation that a given mitochondrion often contains multiple mtDNA copies. There is currently some dispute as to whether these packages of mtDNAs are stable across host generations. If they *are* stable (Hayashi *et al.* 1994; Takai *et al.* 1997), then heteroplasmy within individual mitochondria will be short lived (Preiss *et al.* 1995; Lightowlers *et al.* 1997) because of the small number of mtDNAs per mitochondrion. In that case, Bendall's estimates of effective bottleneck size—and similarly, our model parameters—refer to the number of mitochondrial copies. Alternatively, if mtDNA packaging is ephemeral (Yoneda *et al.* 1994; Attardi *et al.* 1995), then these estimates can be taken to refer to the number of distinct mtDNA molecules. One should also note that most estimates of bottleneck size derive from the intergenerational changes in heteroplasmy ratios in somatic rather than germ cells; somatic evolution may therefore bias these bottleneck estimates.

When no members of the population's best mitochondrial class are transmitted from one generation to the next, the population-level ratchet is said to have turned. For parameters that cause the population-level ratchet to turn slowly, the distribution of mitochondrial mutation number may approach a steady state between turns of the ratchet [analogous to that specified by Stephan *et al.* (1993) for the nuclear genome of asexual haploids]. We call this steady state a *quasi-equilibrium* distribution.

2. THE DISTRIBUTION OF MUTATIONS AT QUASI-EQUILIBRIUM

We begin by examining the quasi-equilibrium distribution of deleterious mitochondrial variation in a large population, between turns of the population-level ratchet. To understand the role that bottleneck size plays in determining the magnitude and distribution of deleterious variation, we follow three variables. The mean fitness of individuals in the population, $\overline{w}$, is of obvious interest from the perspective of host-level selection, and also reflects the total number of deleterious mutations in the mitochondria of each host. The variance among hosts in the total number of mitochondrial mutations, σ^2, indicates the potential for natural selection. The average (sample) variance within hosts in the number of mutations per mitochondrion, s^2, provides a measure of the effect of the bottleneck in generating between-host variation. We will be interested in the values of these quantities at different stages of the life cycle, denoted using the subscripts m (after mutation), b (after the bottleneck), and s (after selection), paying particular attention to how these variables depend on the bottleneck size B. We can relate these variables to each other using a series of deterministic recursions (derived in appendix b).

1. Mutation: As described in Section 1, the number of mutations is Poisson with mean and variance μ per mitochondrion (and hence $M\mu$ per host). Mutations therefore lower mean fitness and increase the variance among and within hosts as follows:

$$\overline{w}_m = \overline{w}_s - \alpha M\mu \tag{1}$$

$$\sigma_m^2 = \sigma_s^2 + M\mu \tag{2}$$

$$s_m^2 = s_s^2 + \mu. \tag{3}$$

2. Bottleneck: The bottleneck process does not alter mean fitness under the linear fitness model. (If we assumed a concave-up fitness function, such as a multiplicative model, we would find an increase in mean fitness; a concave-down function would cause a decrease in mean fitness through the bottleneck.) The variance between hosts increases monotonically as B decreases, due to the greater sampling; the magnitude of the increase depends on the amount of within-host variation. Meanwhile, within-host variance decreases monotonically with tighter bottleneck size. These predictions are consistent with the empirical data showing that mammals, which have tight bottlenecks, generally have low heteroplasmy (low s^2), but considerable variation among hosts (high σ^2):

$$\overline{w}_b = \overline{w}_m \tag{4}$$

$$\sigma_b^2 = \sigma_m^2 + s_m^2\left(\frac{M^2}{B} + M - 1\right)\left(1 - \frac{1}{M}\right) \quad \text{for } B > 1$$

$$= \sigma_m^2 + s_m^2\, M(M-1) \quad \text{for } B = 1 \tag{5}$$

$$s_b^2 = s_m^2\left(1 - \frac{1}{B}\right)\left(1 - \frac{1}{M}\right) \quad \text{for } M > 1 \text{ and } B > 1$$

$$= 0 \quad \text{for } B = 1. \tag{6}$$

3. Selection: The increase in mean fitness due to selection is proportional to the variance among hosts.

The variance within and between hosts is not a simple function of the other variables and introduces two additional variables: $\overline{w_b^3}$ (the mean value of $w_{b,k}^3$, where $w_{b,k}$ is the fitness of the kth individual in the population following the bottleneck), and $\mathrm{Cov}(s_{b,k}^2, w_{b,k})$ (where $s_{b,k}^2$ is the variance in mutation number among the mitochondria in the kth individual in the population following the bottleneck). Unfortunately, the presence of these additional variables prevents us from solving analytically for the equilibrium values. The selection recursions are

$$\overline{w}_s = \overline{w}_b + \frac{\alpha^2\sigma_b^2}{\overline{w}_b} \tag{7}$$

$$\sigma_s^2 = \left(\frac{\overline{w_b^3}}{\overline{w}_b} - (\overline{w}_s)^2\right)/\alpha^2 \tag{8}$$

$$s_s^2 = s_b^2 + \frac{\mathrm{Cov}(s_{b,k}^2, w_{b,k})}{\overline{w}_b}. \tag{9}$$

Using this system, we can compute the shift in mean fitness over the course of a single generation, starting from an arbitrary distribution (*i.e.*, not necessarily at equilibrium). We census a population before mutation and find $\overline{w}_s$, σ_s^2, and s_s^2. Then the mean fitness one complete generation later, following the next round of selection (call this $\overline{w}_s^*$), is

$$\overline{w}_s^* = \overline{w}_s - \alpha M\mu + \frac{\alpha^2}{\overline{w}_s - \alpha M\mu}(\sigma_s^2 + M\mu + (s_s^2 + \mu)\phi(B,M)), \tag{10}$$

where

$$\phi(B,M) = \left(\frac{M^2}{B} + M - 1\right)\left(1 - \frac{1}{M}\right) \quad \text{for } B > 1,$$
$$= M(M-1) \quad \text{for } B = 1. \tag{11}$$

Note that for all $M > 1$, $\phi(B,M)$ decreases monotonically with increasing B: Hence, starting from an arbitrary mutation distribution, mean population fitness after one generation increases with tighter bottleneck size. This occurs because a tight bottleneck increases the variance in fitness among individuals (Equation 5). The fitness improvement due to selection is proportional to σ_b^2 (Equation 7); hence the increased variance leads to higher fitness.

These recursions tell us how $\overline{w}$, σ^2, and s^2 change over a single host generation as a function of bottleneck size B. Using simulations, we have asked a different question: How does bottleneck size affect the values of these variables at the quasi-equilibrium? Details of the parameter values used are given in the figure legends. The population-level ratchet did not turn a single time during the runs displayed in Figures 2 and 3, and thus the plotted values approximate the quasi-equilibrium values. The recursions (1–9) were used to check the accuracy of each phase of the simulations.

In Figure 2a, we show the variance among hosts σ_b^2 and average variance within hosts (s_m^2) Within-host variance is censused before the bottleneck, because this variance generates the between-host variance σ_b^2 (Equation 5); among-host variance is censused after the bottleneck because that is the variance that produces the selective response (Equation 7). Note that the among-host variance decreases as B increases, while the within-host variance increases with B. These results are in accordance with Equations 5 and 6.

In Figure 2b we plot the values of $\overline{w_b^3}$ and $\mathrm{Cov}(s_{b,k}^2,$

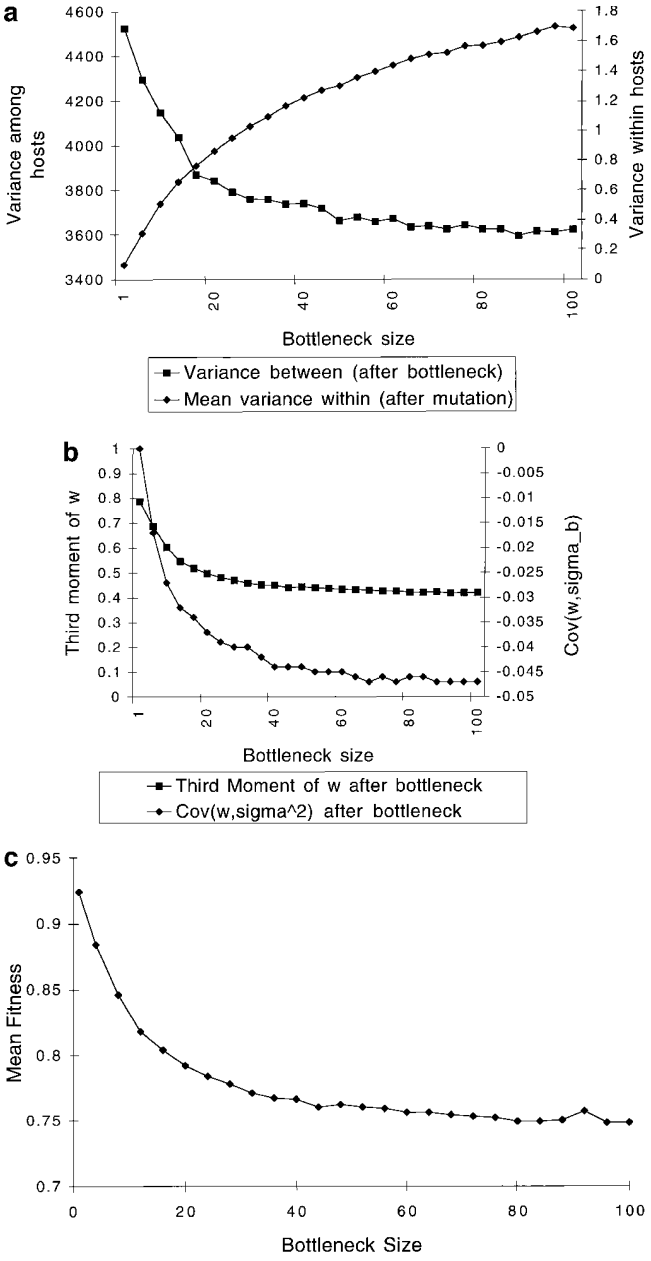

Figure 2.—Simulation results for parameter values $N = 500$, $M = 100$, $\mu = 0.1$, and a linear fitness function $W = 1 - \alpha n$ with $\alpha^{-1} = 500$, $W \equiv 0$ when $n > \alpha^{-1}$. The plotted values are averages over 2000 generations, following an initial 100-generation period allowing the population to settle into quasi-equilibrium. (a) Variance among hosts (σ_b^2) and variance within hosts (s_m^2) are shown. (b) The values of $\overline{w_b^3}$ and $\mathrm{Cov}(s_{b,k}^2, w_{b,k})$ are shown. (c) Mean fitness after selection ($\overline{w}_s$).

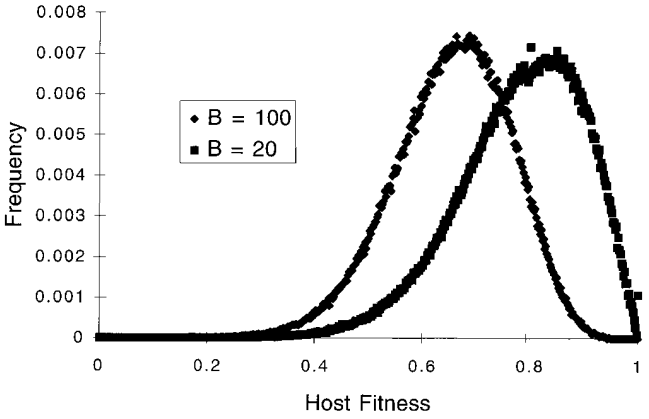

Figure 3.—Distribution of host fitnesses after selection (w_s). The model parameters are the same as those in Figure 2.

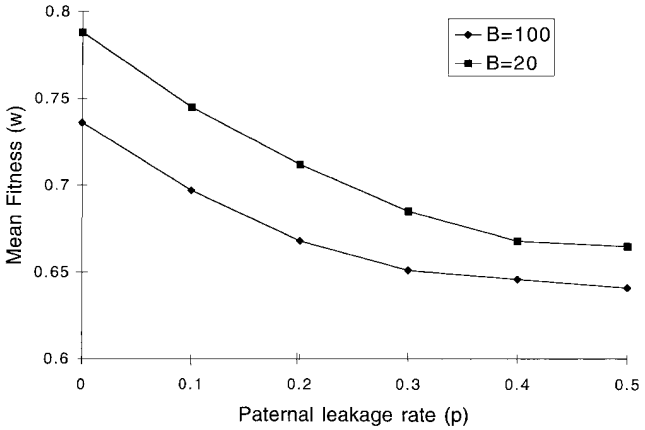

Figure 4.—Mean fitness as a function of paternal leakage rate. The model parameters are as those in Figure 2, but with $N = 1000$ to account for the presence of two sexes.

$w_{b,k}$), to complete our description of the parameters governing the system. Note that $\mathrm{Cov}(s^2_{b,k}, w_{b,k}) < 0$ for $B > 1$, showing that the within-host variance increases with the total number of mutations in a host.

In Figure 2c we plot the mean fitness after selection ($\bar{w}_s$) as a function of the bottleneck size. As expected from the single-generation response to selection and from the monotonicity of σ^2_b at quasi-equilibrium (Figure 2a), the highest equilibrium mean fitness is achieved at bottleneck size 1, where all mitochondria in a zygote are identical. Mean fitness decreases steadily as B increases.

Besides looking at mean fitness values at quasi-equilibrium, we have examined the *distribution* of fitnesses in a quasi-equilibrium population. Two such distributions are shown in Figure 3. Note that the tighter bottleneck size ($B = 20$ instead of $B = 100$) is shifted toward higher fitness. Fitness distributions with even tighter bottlenecks are also shifted, but are harder to interpret graphically, because the extreme sampling imposes sharp discontinuities on the distribution.

In summary, we have found that a tight bottleneck improves mean fitness in a population and shifts the fitness distribution toward higher fitness. This result can be explained by the fact that a mitochondrial bottleneck—like the intragenerational drift process modeled by Takahata and Slatkin (1983)—increases the variance among host offspring. Another way to look at this is that when a new mutation arises in a host, it is initially at low frequency, so that selection against the host is relatively weak. A tight bottleneck either eliminates the mutation through random sampling or exposes the mutation to much stronger selection (at the host level). Our analysis has not addressed the question of whether a bottleneck is advantageous to the individual; we will return to this point in the discussion section. Our results in this section are related to results obtained by Kondrashov (1994a,b,c), who also found a reduction in genetic load with increased sampling in a model of vegetative reproduction.

Paternal leakage: In this context, it is also interesting to consider the effect of paternal leakage on fitness at the quasi-equilibrium. To model paternal leakage, we now designate half the population as male and half as female. We then select $N/2$ mating pairs and from each pair generate two offspring. Mitochondrial transmission is as follows: A fraction p of the mitochondria in each zygote are drawn from the father's mitochondria, sampled without replacement; a fraction $1 - p$ are drawn from the mother, again sampled without replacement. (We require that Mp be an integer.) Leakage occurs immediately following the bottleneck in the life cycle. We assume sampling without replacement so that when $p = 0$ this reduces to the model studied above.

We can anticipate the results as follows. Note that paternal leakage typically reduces the variance (σ^2) among hosts, because it produces offspring whose mitochondria are a mixture of the parental mitochondria. (The variance is reduced provided that on average mitochondria within hosts are more similar than mitochondria from different hosts—as expected given the shared ancestry of mitochondria within a host.) This reduction in variance will reduce the efficacy of selection. Acting to mitigate this effect is the fact that leakage will typically increase the within-host variance (s^2).

In Figure 4 we show population mean fitness at quasi-equilibrium as a function of paternal leakage rate. Note that mean fitness decreases as the paternal contribution tends toward 0.5. We found that in the simulations, the population-level ratchet turned occasionally when the assumed rate of paternal leakage was high. The mean fitness was adjusted as described in the legend to Figure 6.

3. RATCHET DYNAMICS

So far we have been examining the effect that the mode of mitochondrial transmission has on fitness in the quasi-equilibrium state. That is most relevant to understanding the distribution of fitnesses in a large

population, at a single time point. We now turn our attention to the accumulation of mutations in a population over evolutionary time. We explore the dynamics of Muller's ratchet in mitochondria, with a particular focus on how these dynamics are impacted by a germline bottleneck.

In conventional models of Muller's ratchet in asexual haploids (Felsenstein 1974; Haigh 1978), the ratchet turns when the individuals with the fewest mutations—those in the best *mutation class*—fail to reproduce. This is an irreversible step, or *ratchet*, because in the absence of back-mutation, this best mutation class can never be recovered. Similarly, the ratchet turns in the present model when a population of mitochondria fails to pass on its best mitochondrion. However, in the absence of paternal leakage and recombination, we can consider two levels of population structure in this model: the population of mitochondria in a single host, or the population of all mitochondria in the entire host population. An irreversible ratchet process operates at each level, and therefore in our model there is a *double ratchet*. We now consider how the bottleneck affects the rate of each ratchet.

Examining the population of mitochondria within a single host, we have a *host-level ratchet*. The host-level ratchet turns when an offspring individual fails to inherit its parent's best mitochondrion or mitochondria; this establishes a lineage that will never have a mitochondrion as good as the best mitochondrion in the parent.

Alternatively, if we ignore the partitioning of mitochondria into specific hosts and consider the population of mitochondria across all hosts in the host population, the *population-level ratchet* turns when the best mitochondrion or mitochondria in the entire host population at time t_i are not transmitted to any member of the host population at time t_{i+1}.

The population-level ratchet need not turn with every turn of the host-level ratchet. A particular lineage may fail to transmit its best mitochondrion, but as long as this lineage does not contain the only copy of the best mitochondrion in the entire host population, the best mitochondrion can still be passed into the next generation in some parallel lineage.

When considering questions of long-term evolutionary persistence of populations in the face of genetic degradation, one is primarily interested in the progression of the population-level ratchet. However, it is necessary to first understand the behavior of the host-level ratchet. In Figure 5a, we show the average rates at which the host-level ratchet turned during mutation and bottleneck sampling in simulations. The host-level ratchet turned more slowly for larger bottleneck sizes. This result was expected, because a tight bottleneck increases the probability of failing to sample the best mitochondrial class. The maximum rates occurred at bottleneck size 1, where the host-level ratchet rate equals the probability of at least one mutation arising in a given mito-

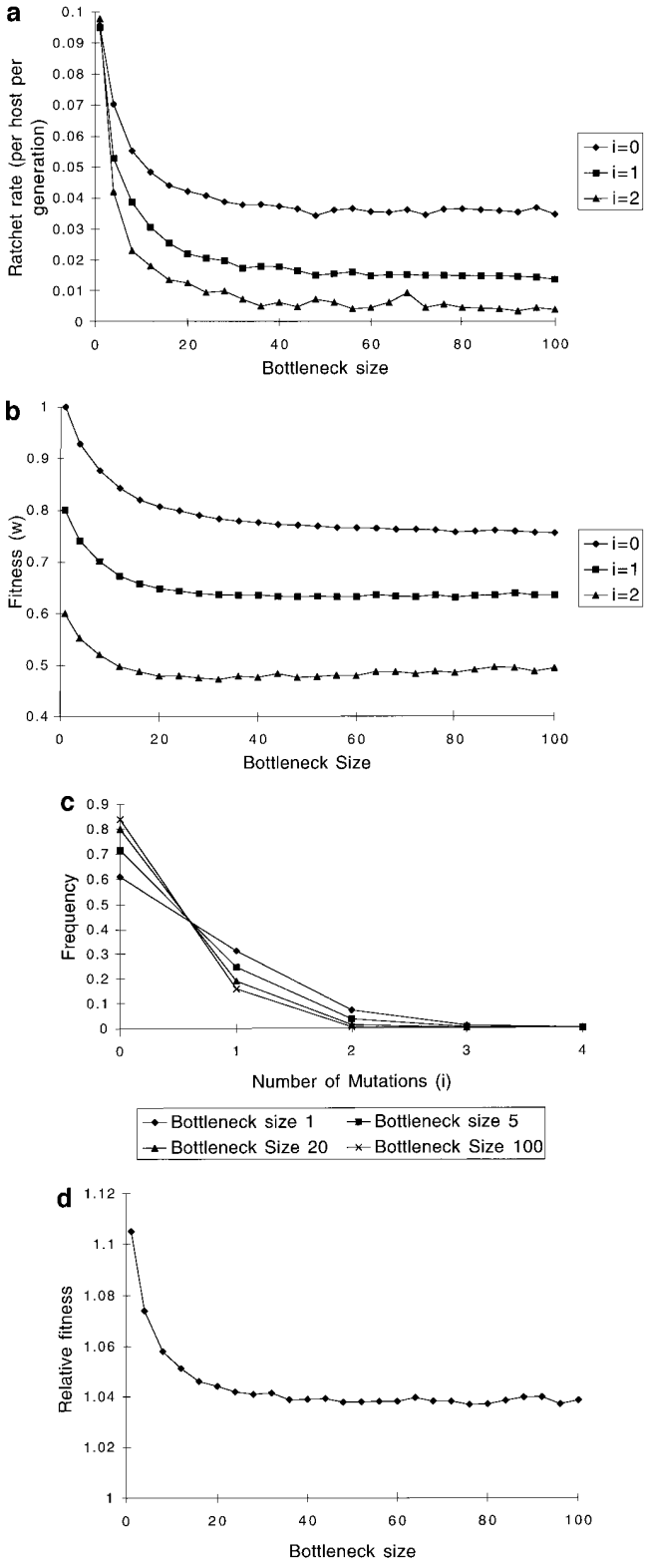

Figure 5.—Simulation results using the same parameter values as in Figures 2 and 3. (a) Host-level ratchet rates are shown for $i = 0, 1, 2$ mutations in the host's best mitochondrion. (b) Mean fitnesses are shown conditional on host-level ratchet state for $i = 0, 1, 2$. (c) Distributions of $p_b(i)$ values for bottleneck sizes of 1, 5, 20, and 100 are shown. (d) The ratio $\hat{w}_b(0)/\overline{w}_b$, as a function of bottleneck size, provides a measure of the relative fitness advantage of the best class of individuals.

chondrion. For a Poisson mutation rate of 0.1 (as in the simulations), this corresponds to an expected ratchet rate of 0.095. As the host-level ratchet state progresses from $i = 0$ mutations in an individual's best mitochondrion to $i = 1$, $i = 2$, etc., the distribution of within-host mitochondrial distributions shifts toward increasing representation of the best mitochondria present. As a consequence, the ratchet rate slows with increasing i.

Figure 5b plots the mean fitnesses conditional on host-level ratchet state for $i = 0, 1, 2$. As expected, the mean fitnesses conditional on i increase as i decreases. Also, the mean fitnesses conditional on each i decrease with increasing bottleneck size, as expected from the results of Section 2.

These ratchet rates and mean conditional fitnesses determine the quasi-equilibrium fraction of individuals in each host-level ratchet state. Figure 5c shows the distribution of $p_b(i)$ values from the same simulations, where $p_b(i)$ is defined as the frequency before selection (*i.e.*, following the bottleneck) of individuals whose best mitochondrial class carries i mutations. It is apparent from Figure 5c that a tight bottleneck leads to a broader range of host-level ratchet states, with fewer individuals containing perfect mitochondria. This results from the more rapid rate of host-level ratchet turning associated with tighter bottlenecks.

Thus, while tight bottlenecks are associated with higher mean fitness averaged across the population, and a better overall distribution of mitochondrial quality in the population as a whole, they also increase the rate at which the host-level ratchet turns within individual lineages (note that within each lineage, the rate of the host-level ratchet is independent of host population size). For this reason, they lead to a poorer quasi-equilibrium distribution of host-level ratchet states in the population. We now examine the consequences of this on the rate of the population-level ratchet.

Recall the link between the host-level and population-level ratchets. Suppose that the best mitochondria in the population contain i mutations. Then the best host-level ratchet state is i. If by chance this best class is lost, then the distribution of host-level ratchet states slides to the right. This corresponds to a turn of the population-level ratchet.

In a haploid asexual model the size of the best class may be useful in predicting the rate of Muller's ratchet in the nuclear genome (Haigh 1978). However, it is also crucial to consider the fitness differences between the classes. Mutations of small effect may result in rapid rates of fitness loss due to Muller's ratchet, because the small fitness advantage of the best class is overwhelmed by the stochastic effects (Kondrashov 1994a; Butcher 1995).

In our model, the fitness differences between the classes are crucial in understanding the rate of Muller's ratchet. Let $\hat{w}_b(0)$ be the mean fitness of individuals with no deleterious mutations in their best mitochondrion before selection and $\overline{w}_b$ be the mean fitness of the entire population, also before selection. In Figure 5d we plot $\hat{w}_b(0)/\overline{w}_b$, as a function of bottleneck size. This corresponds to the relative fitness advantage of the best class (*i.e.*, the class with host-level ratchet state of 0). Note that the advantage of the best class is highest for small bottleneck sizes. So while the frequency of hosts carrying mitochondria with no mutations is smallest in the presence of a tight bottleneck, this is precisely when those individuals enjoy their greatest fitness advantage.

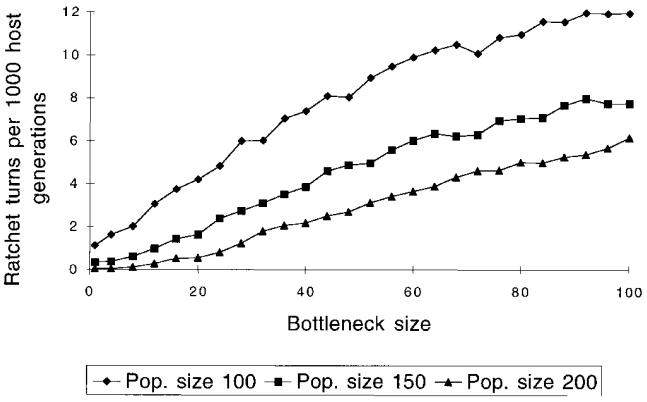

Figure 6.—Turns of the ratchet at the population level per thousand host generations, as a function of bottleneck size. Here the population-level ratchet turned repeatedly; to keep the rate of the ratchet's progression constant, it was necessary to keep the strength of selection independent of the distance the ratchet had progressed (Kondrashov 1994a). To do this, we used the linear fitness function $W = 1 - \alpha(n - n_{best})$, where $\alpha^{-1} = 1000$ and n_{best} is the total number of mutations in the best host present in the population. As in the previous figures, $M = 100$ and $\mu = 0.1$. As in Figures 2–5, the simulations were run for 2000 generations following an initial 100-generation period. The data shown are averages over 60 replicates for population sizes $N = 100$, $N = 150$, and $N = 200$, respectively.

It turns out that the high relative fitness of the best class more than compensates for the smaller size of the best class. We have conducted simulations with small populations of $N = 100$, 150, and 200 individuals, to study the rate of turning of the population-level ratchet (Figure 6). We used a corrected fitness function to keep the rate of ratchet progression constant over time (see Figure 6 legend). We found that the rate of turning of the population-level ratchet increased monotonically with bottleneck size. This implies that a tight germline bottleneck reduces the long-term damage due to Muller's ratchet.

DISCUSSION

Empirical studies in several species have found evidence for a germline bottleneck in the number of mitochondria passed from mother to daughter. In this article we ask what role these bottlenecks play in maintaining the genetic integrity of mitochondria. Our results dem-

onstrate that a bottleneck acts to improve the distribution of mitochondrial qualities and host fitnesses at quasi-equilibrium, between turns of the population-level ratchet. Moreover, while a bottleneck hastens the progress of the host-level ratchet within individual lineages, intensified selection among these lineages more than compensates for this acceleration, reducing the turning rate of the population-level ratchet and slowing the rate of genetic degradation within the population.

Our goal in this article was to study the evolutionary fate of nearly neutral deleterious mutations at quasi-equilibrium, where the mutations are held at a stochastic balance by selection and drift, and where the ratchet progresses at a slow rate. The presence of such mutations has been inferred empirically by Nachmann *et al.* (1996) and Lynch (1996). In our model, we assumed that all mutations have the same selective effect. In practice, the selective effects of deleterious mutations will vary, with some being effectively neutral and others strongly deleterious; our results apply to those that are intermediate in effect and whose fate depends both on selection and drift. These are the mutations that make the greatest contribution to Muller's ratchet (Butcher 1995).

We have illustrated the features of our model using a particular set of parameter values (for N, μ, and α). To make the simulations computationally tractable we used rather small population sizes (at most 10^3) and high mutation rates (10^{-1} per mitochondrial genome). Because we were interested in mutations at selection-drift balance, we adjusted the strength of selection (by increasing α) accordingly. To check the general validity of our qualitative results, we ran additional simulations over a range of parameter values. Those simulations did not suggest that the impact of the bottleneck depends strongly on the parameters, provided that the parameters chosen lead to a state of selection-drift quasi-equilibrium. As might be intuitively obvious, the bottleneck has little effect on mutations that are either virtually neutral or strongly deleterious.

In addition, we have found similar results using multiplicative (concave-up) and quadratic (concave-down) fitness functions (simulation results not shown). Concave-down functions are of particular interest; if there is dominance masking or mutation complementarity (Takai *et al.* 1997) in the mitochondrial genomes, we might expect fitness functions to take this general form.

We have also shown (appendix a) that over extended periods of evolutionary time the rate of fixation of deleterious mutations converges to the rate of turning of Muller's ratchet. Thus, our conclusions on the effect of bottleneck size on Muller's ratchet are immediately applicable to studies on fixation rates of deleterious mutations in mammalian mitochondria (Lynch 1996), and other endosymbionts such as Buchnera (Moran 1996).

When bottleneck size is reduced to one, the dynamics of our model become identical to those in a conventional model of Muller's ratchet, where the genome size is that of a single mitochondrion, and selection is scaled appropriately ($W = 1 - \alpha Mn$). This allows a comparison of the rates of progression of the ratchet in mitochondria and asexuals. From our results, it follows that for bottleneck sizes greater than one, Muller's ratchet turns faster, and deleterious mutations are fixed more often in the mitochondrial genome than on an asexual chromosome of equal size.

Throughout this discussion we have assumed deleterious mutation to be neutral with respect to the intrahost replication of mitochondria. However, it is conceivable that there might be selfish genotypes with rapid replication (suggested, for example, in data by Shoubridge *et al.* 1990). In that case, a process analogous to interdemic selection (Wright 1931, 1935; Wilson 1975, 1977) will operate, as mitochondria compete not only at the between-host level but also within individual hosts.

While we have not attempted a detailed analysis of this situation, it is interesting to note that the bottleneck not only slows the ratchet, but often discourages within-host competition as well. A bottleneck serves to increase mitochondrial genetic variance among hosts and reduce variance within hosts (Equations 5 and 6, and Figure 2a). These conditions favor greater "cooperation" within hosts (Frank 1996a,b), and hence a bottleneck should once again improve the distribution of host fitnesses.

Another simplification that we have made is to allow the bottleneck to occur in only two cell generations. More gradual bottleneck processes will produce additional intragenerational drift with effects similar to those of the bottleneck itself. Takahata and Slatkin (1983) have shown that intragenerational drift serves to increase the variance in host fitness and hence strengthens selection against deleterious mutations.

While we have demonstrated that a bottleneck improves the distribution of mitochondrial qualities and host fitnesses, and impedes the progress of Muller's ratchet in mitochondrial genomes, we have not demonstrated that the host trait of imposing a bottleneck is itself directly favored by natural selection. Indeed, for certain selection functions (*e.g.*, concave down) imposing a bottleneck will be at an immediate selective disadvantage. Even in this case, however, host lineages in which bottlenecks arise will eventually enjoy better fitness distributions (simulation data not shown) and will be less susceptible to the operation of Muller's ratchet on their mitochondria. Despite the initial selective disadvantages, these lineages may be stochastically favored in the long run. It follows from the results of Eshel (1973) that even in this case a host gene causing a bottleneck could spread in an asexual population; however, the problem is further complicated in sexual populations because genes controlling bottleneck size may be separated from the associated mitochondria.

Our results also generate an interesting prediction

for mitochondrial transmission in birds and butterflies with WZ sex chromosomal systems. If fitness is a concave-down function of mitochondrial mutation number, a bottleneck will be directly selected against because it increases the variance in offspring mutation number, but will be selectively favored in the long run for the reasons discussed above. Since mitochondria are maternally transmitted, there will be no advantage to a bottleneck when producing sons; the long-term advantage will apply only to the production of daughters. If the fitness function was indeed concave down, there would be a long-term selective advantage to employing a bottleneck when producing daughters, but not when producing sons. This effect could be exploited only in WZ systems, where the female is the heterogametic sex and sex determination is controlled by the female gamete. There we would predict an unequal segregation of mitochondria into W (male-producing) and Z (female-producing) gametes, with W gametes receiving a disproportionately large fraction of the mitochondria.

Another interesting application of our results involves evolution in bacterial endosymbionts. Moran (1996) has recently demonstrated that bacterial endosymbionts—which share a number of reproductive traits with mitochondria—also suffer from accelerated rates of genetic deterioration. Moran found that, compared with their free-living relatives, endosymbiotic members of the *Buchnera aphidicola* complex feature accelerated rates of ribosomal DNA evolution and higher ratios of nonsynonymous to synonymous substitutions. She argues that these data indicate accelerated rates of accumulation of mildly deleterious mutations in the endosymbiotic species.

On the basis of the models treated here, we suggest that a bottleneck in population size between host generations may act in a similar fashion to help these species avoid genetic degradation, as long as they are beneficial to their hosts. When the endosymbionts' genetic integrity affects host fitness, the bottleneck acts to strengthen selection and oppose Muller's ratchet and the fixation of deleterious mutations. In contrast, when the endosymbionts' genetic integrity has no effect on host fitness, a bottleneck serves only to reduce the effective population size of the endosymbionts, accelerate the within-host ratchet, and hasten genetic decay. *B. aphidicola* species are indeed beneficial or essential to their hosts (Dadd 1985; Douglas 1989), and there appears to be a bottleneck in these species (Buchner 1965; Moran 1996). Other cytoplasmically inherited organisms, such as Wolbachia species, do not appear to be beneficial to their hosts (Moran and Baumann 1994); a bottleneck, if present for these species, would accelerate genetic deterioration.

The authors thank M. Feldman, M. Lachmann, B. Levin, and M. Tanaka for numerous helpful comments and discussions and J. Coyle for help in preparing the manuscript. C. Bergstrom and J. Pritchard are supported by Howard Hughes Predoctoral Fellowships and by National Institutes of Health grant GM 28016 to M. Feldman.

LITERATURE CITED

Attardi, G., M. Yoneda and A. Chomyn, 1995 Complementation and segregation behavior of disease-causing mitochondrial-DNA mutations in cellular-model systems. Biochim. Biphys. Acta. **1271**(1): 241–248.

Avise, J. C., 1991 Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. Annu. Rev. Genet. **25:** 45–69.

Bendall, K. E., V. A. Macaulay, J. R. Baker and B. C. Sykes, 1996 Heteroplasmic point mutations in the human mtDNA control region. Am. J. Hum. Genet. **59:** 1276–1287.

Bendall, K. E., V. A. Macaulay and B. C. Sykes, 1997 Variable levels of a heteroplasmic point mutation in individual hair roots. Am. J. Hum. Genet. **61:** 1303–1308.

Brown, G. K., 1997 Bottlenecks and beyond: mitochondrial DNA segregation in health and disease. J. Inherited Metab. Dis. **20:** 2–8.

Buchner, P., 1965 *Endosymbiosis of Animals with Plant Microorganisms* (revised English ed.). Interscience Publishers, New York.

Butcher, D., 1995 Muller's ratchet, epistasis and mutation effects. Genetics **141:** 431–437.

Charlesworth, D., M. T. Morgan and B. Charlesworth, 1993 Mutation accumulation in finite outbreeding and inbreeding populations. Genet. Res. **61:** 39–56.

Clark, A. G., 1988 Deterministic theory of heteroplasmy. Evolution **42**(3): 621–626.

Clark, A. G., and E. M. S. Lyckegaard, 1988 Natural selection with nuclear and cytoplasmic transmission. III. Joint analysis of segregation and mtDNA in *Drosophila melanogaster*. Genetics **118:** 471–481.

Dadd, R. H., 1985 Nutrition: organisms, pp. 313–390 in *Comprehensive Insect Physiology, Biochemistry, and Pharmacology*, Vol. 4, Chapter 8, edited by G. A. KerKut and L. I. Gilbert. Pergamon Press, Oxford.

Douglas, A. E., 1989 Mycetocyte symbiosis in insects. Biol. Rev. Camb. Philos. Soc. **64:** 409–434.

Eshel, I., 1973 Clone selection and the evolution of modifying features. Theor. Popul. Biol. **4:** 196–208.

Felsenstein, J., 1974 The evolutionary advantages of recombination. Genetics **78:** 737–756.

Frank, S. A., 1996a Host control of symbiont transmission: the separation of symbionts into germ and soma. Am. Nat. **1338:** 1113–1124.

Frank, S. A., 1996b Host-symbiont conflict over the mixing of symbiotic lineages. Proc. R. Soc. Lond. Ser. B **263:** 339–344.

Gabriel, W., M. Lynch and R. Burger, 1993 Muller's ratchet and mutational meltdowns. Evolution **47:** 1744–1757.

Haigh, J., 1978 The accumulation of deleterious genes in a population. Theor. Popul. Biol. **14:** 251–267.

Hayashi, J.-I., M. Takemitsu, Y. Goto and I. Nonaka, 1994 Human mitochondria and mitochondrial genome function as a single dynamic cellular-unit. J. Cell Biol. **125**(1): 53–50.

Higgs, P. G., 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.

Howell, N., S. Halvorson, I. Kubacka, D. A. McCullough, L. A. Bindoff and D. M. Turnbull, 1992 Mitochondrial gene segregation in mammals: is the bottleneck always narrow? Hum. Genet. **90:** 117–120.

Hudson, R. R., and N. L. Kaplan, 1995 Deleterious background selection with recombination. Genetics **141:** 1605–1617.

Hurst, L. D., and G. T. McVean, 1996 ... and scandalous symbionts. Nature **381:** 650–651.

Ivanov, P., M. Wadhams, R. Roby, M. M. Holland, V. Weedn *et al.*, 1996 Mitochondrial DNA sequence heteroplasmy in the Grand Romanov establishes the authenticity of the remains of Tsar Nicholas II. Nat. Genet. **12**(4): 417–420.

Jenuth, J., A. C. Peterson, K. Fu and E. Shoubridge, 1996 Random genetic drift in the female germline explains the rapid segrega-

tion of mammalian mitochondrial DNA. Nat. Genet. **14**(2): 146–151.
Kondrashov, A. S., 1994a Muller's ratchet under epistatic selection. Genetics **136:** 1469–1473.
Kondrashov, A. S., 1994b Mutation load under vegetative reproduction and cytoplasmic inheritance. Genetics **137:** 311–318.
Kondrashov, A. S., 1994c Sex and deleterious mutation. Nature **369:** 99–100.
Lightowlers, R. N., P. F. Chinnery, D. M. Turnbull and N. Howell, 1997 Mammalian mitochondrial genetics: heredity, heteroplasmy and disease. Trends Genet. **13**(11): 450–455.
Lynch, M., R. Berger, D. Butcher and W. Gabriel, 1993 The mutational meltdown in asexual populations. J. Hered. **84:** 339–344.
Lynch, M., 1996 Mutation accumulation in transfer RNAs: Molecular evidence for Muller's ratchet in mitochondrial genomes. Mol. Biol. Evol. **13**(1): 209–220.
Marchington, D. R., G. M. Hartshorne, D. Barlow and J. Poulton, 1997 Homopolymeric tract heteroplasmy in mtDNA from tissues and single oocytes: support for a genetic bottleneck. Am. J. Hum. Genet. **60:** 408–416.
Moran, N., and P. Baumann, 1994 Phylogenetics of cytoplasmically inherited microorganisms of arthropods. Trends Ecol. Evol. **9**(1): 15–20.
Moran, N. A., 1996 Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. Proc. Natl. Acad. Sci. USA **93:** 2873–2878.
Nachmann, M. W., W. M. Brown, M. Stoneking and C. F. Aquadro, 1996 Nonneutral mitochondrial DNA variation in humans and chimpanzees. Genetics **142:** 953–963.
Nordborg, M., 1997 Structured coalescent processes on different time scales. Genetics **146:** 1501–1514.
Parsons, T. J., D. S. Muniec, K. Sullivan, N. Woodyatt, R. Alliston-Greiner *et al.*, 1997 A high observed substitution rate in the human mitochondrial DNA control region. Nat. Genet. **15:** 363–368.
Poulton, J., and D. R. Marchington, 1996 Prospects for DNA-based prenatal diagnosis of mitochondrial disorders. Prenatal Diagn. **16**(13): 247–256.
Preiss, T. A., S. A. Lowerson, K. Weber and R. N. Lighttowlers, 1995 Human mitochondria: distinct organelles or dynamic network? Trends Genet. **11**(6): 211–212.
Sherratt, E. J., A. W. Thomas and J. C. Alcolado, 1997 Mitochondrial-DNA defects: a widening clinical spectrum of disorders. Clin. Sci. **92**(3): 225–235.
Shoubridge, E. A., G. Carpati and K. M. Hastings, 1990 Deletion mutants are functionally dominant over wild-type mitochondrial genomes in skeletal muscle fiber segments in mitochondrial disease. Cell **62:** 43–49.
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**(3): 225–231.
Takahata, N., and M. Slatkin, 1983 Evolutionary dynamics of extranuclear genes. Genet. Res. **42:** 257–265.
Takai, D., K. Inoue, Y. Goto, I. Nonaka and J. I. Hayashi, 1997 The interorganellar interaction between distinct human mitochondria with deletion mutant mtDNA from a patient with mitochondrial disease and with hela mtDNA. J. Biol. Chem. **272(9):** 6028–6033.
Wilson, D. S., 1975 A theory of group selection. Proc. Natl. Acad. Sci. USA **72**(1): 143–146.
Wilson, D. S., 1977 Structured demes and the evolution of group-advantageous traits. Am. Nat. **111:** 157–185.
Wolstenholme, D. R., 1992 Animal mitochondrial-DNA: structure and evolution. Int. Rev. Cytol. **141:** 173–216.
Wright, S., 1931 Evolution in mendelian populations. Genetics **16**(2): 97–159.
Wright, S., 1935 Evolution in populations in approximate equilibrium. J. Genet. **30:** 257–266.
Yoneda, M., T., Miyatake and G. Attardi, 1994 Complementation of mutant and wild-type human mitochondrial DNAs coexisting since the mutation event and lack of complementation of DNAs introduced separately into a cell within distinct organelles. Mol. Cell. Biol. **14:** 2699–2712.

Communicating editor: A. G. Clark

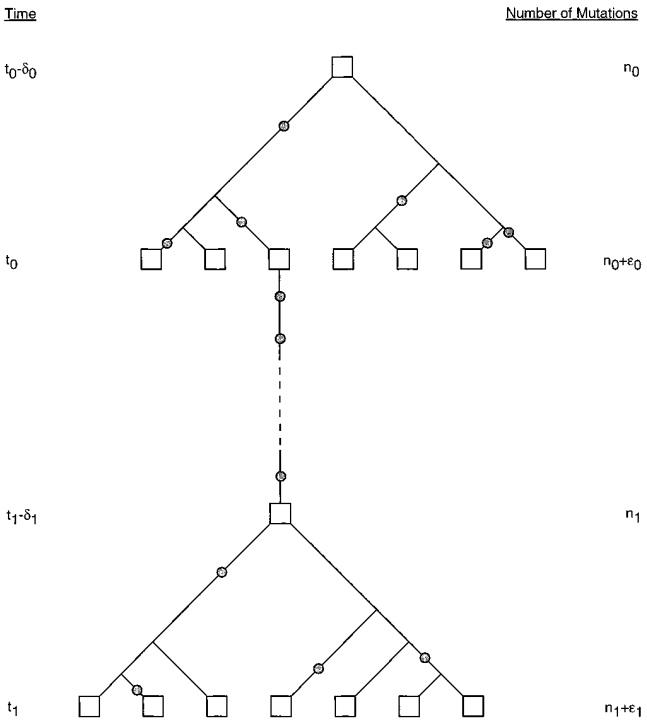

Figure 7.—The relationship between the rate of fixation and the rate of Muller's ratchet. Individuals are represented by squares; these are connected by lines to indicate paths of descent. Shaded circles are mutations. The population is censused twice (at times t_0 and t_1). The most recent common ancestors for each of those two time points occur at $t_0 - \delta_0$ and $t_1 - \delta_1$. The common ancestors contain n_0 and n_1 mutations, respectively; the kth individuals at times t_0 and t_1 carry $n_0 + \varepsilon_{0,k}$ and $n_1 + \varepsilon_{1,k}$ mutations, respectively.

APPENDIX A

Here we provide an argument that over long periods of evolutionary time, the rate of fixation of deleterious alleles converges to the rate of Muller's ratchet. This result was previously suggested by Higgs and Woodcock (1995), who performed simulations of the multiplicative fitness case.

Consider a single Wright-Fisher population containing a constant number N haploid asexual individuals. Deleterious mutations arise at a rate μ per individual per generation, with no back-mutation. There is no recombination. To begin with, assume multiplicative fitness (that is, an individual carrying n mutations has fitness s_n, where $0 < s < 1$). In this case, there are no synergistic fitness interactions (Kondrashov 1994a), and so the expected rate of progression of the ratchet is independent of the mutation load in the population.

Suppose that we sample the population at two times, t_0 and t_1 (as shown in Figure 7), where t_0 is earlier than t_1. The population at time t_0 has its most recent common ancestor at time $t_0 - \delta_0$. Similarly, the population at time t_1 has a common ancestor at time $t_1 - \delta_1$. Here, δ_0 and δ_1 are random variables arising from the coalescent process.

Suppose that the common ancestor at time $t_0 - \delta_0$ carries n_0 mutations and that the kth individual in the population at time t_0 carries $n_0 + \varepsilon_{0,k}$ mutations. Here, $\varepsilon_{0,k}$ is the number of deleterious mutations that occurred on the lineage connecting individual k to the common ancestor at time $t_0 - \delta_0$. Likewise, we say that the common ancestor at time $t_1 - \delta_1$ carries n_1 mutations and that the kth individual in the population at time t_1 carries $n_1 + \varepsilon_{1,k}$ mutations. Since there is no back-mutation, $\varepsilon_{0,k} \geq 0$, $\varepsilon_{1,k} \geq 0$ for all k, and $n_1 \geq n_0$.

At time t_0, the state of the ratchet is $n_0 + \min(\varepsilon_{0,k})$, and the state of the ratchet at time t_1 is $n_1 + \min(\varepsilon_{1,k})$. It also follows that exactly n_0 deleterious mutations are fixed at time t_0, and n_1 mutations are fixed at time t_1.

So the mean fixation rate ($\bar{F}$) can be estimated as

$$\bar{F} - \frac{n_1 - n_0}{t_1 - t_0}. \tag{A1}$$

The mean rate at which Muller's ratchet turns ($\bar{R}$) is estimated by

$$\bar{R} = \frac{(n_1 - n_0)}{t_1 - t_0} + \frac{(\min \varepsilon_{1,k}) - \min(\varepsilon_{0,k})}{t_1 - t_0}. \tag{A2}$$

It follows from Haigh (1978), Hudson and Kaplan (1995), and Nordborg (1997) that the distribution of $\min(\varepsilon_{x,k})$ (where $x = 0, 1$), is constant in time for the multiplicative fitness function and that its expectation is finite. [In a population with deleterious mutations, the coalescent times scale roughly as a factor $e^{-\mu/s}$ of the neutral coalescent times (Hudson and Kaplan 1995).] Thus,

$$\frac{\min (\varepsilon_{1,k}) - \min(\varepsilon_{0,k})}{t_1 - t_0} \to 0, \tag{A3}$$

as $t_1 - t_0 \to \infty$. However, the expectation of $n_1 - n_0$ is asymptotically linear in time (Kondrashov 1994a). Hence, as $(t_1 - t_0) \to \infty$,

$$\bar{R} \to \bar{F}. \tag{A4}$$

That is, the rate of Muller's ratchet is asymptotically equal to the rate of fixation of deleterious mutations.

Note that this argument holds not only for the multiplicative fitness function, but in general for fitness functions for which (1) $n_1 - n_0$ increases linearly in time (if it does not, the rate is not well defined anyway), and (2) Equation (A3) holds.

The first criterion is met for fitness functions in which there is a constant distribution of relative fitnesses, such as for the multiplicative model, or the *corrected* linear fitness model used for the ratchet simulations in this article. This does not hold when there are synergistic fitness interactions (Kondrashov 1994a). We anticipate that the second criterion will fail only in those rather special models that do not have a finite expected population coalescent time. Such a situation could arise if at one locus there is a protected polymorphism with no mutation.

From this argument, we conclude that when considering short-term evolution, it is crucial to distinguish between the rate of fixation of deleterious mutations and the rate of Muller's ratchet, as pointed out by Charlesworth *et al.* (1993). However, when considering sufficiently long-term evolution (including the question treated in this paper, that of long-term evolutionary maintenance of genetic integrity), the rates of these two processes converge and therefore the rate of either process suffices to specify the rate of genetic decay.

APPENDIX B

Here we provide derivations for the recursions (1–9).

Equation 1: Since we are using a linear fitness function, the loss in mean fitness due to mutation is proportional to the mean number of new mutations per host: $M\mu$.

Equations 2 and 3: The two variances increase by the variance in the number of mutations (per host, and per mitochondrion). Since mutation number is Poisson in our model, these two variances are $M\mu$ and μ, respectively.

Equation 4: The sampling procedure does not change the *expected* number of mutations in an individual. Since we use a linear fitness function, the bottleneck does not change the mean fitness.

Equation 5: We begin by computing the variance in the number of mutations in a single zygote after completion of the bottleneck process, conditional on the number and distribution of mutations in the parent cell before the bottleneck. Recall that the bottleneck process is composed of two stages. Suppose that the mitochondria in the parent cell before the first stage (total of M mitochondria) contain $\mathcal{X} = \{x_0, x_1, \ldots, x_M\}$ mutations. After the first stage (bottleneck down to B mitochondria), the mitochondria contain $\mathcal{Y} = \{y^0, y^1, \ldots, y_B\}$ mutations. Let $\bar{x}$ be the mean number of mutations per mitochondrion before the first stage, $\bar{y}$ be the mean number after the first stage, and $\bar{z}$ be the mean number after the second (final) stage. We start by finding $\mathrm{Var}(\bar{z}|\mathcal{X})$.

Since the two stages of sampling are independent, we can write

$$\mathrm{Var}(\bar{z}|\mathcal{X}) = \mathrm{Var}(\bar{z}|\mathcal{Y}) + \mathrm{Var}(\bar{y}|\mathcal{X}), \tag{B1}$$

where $\mathrm{Var}(\bar{z}|\mathcal{Y}) \equiv 0$ if $B = 1$. Since each stage of sampling is performed by random sampling with replacement, we have

$$\mathrm{Var}(\bar{z}|\mathcal{Y}) = \mathrm{Var}(y_i)/M, \tag{B2}$$

$$\mathrm{Var}(\bar{y}|\mathcal{X}) = \mathrm{Var}(x_i)/B. \tag{B3}$$

In order to compute $\mathrm{Var}(y_i)$ in terms of $\mathcal{X}$, define the mean square difference D: $D(x) \equiv E[(x_i - x_j)^2]$, and

$D(y) \equiv E[(y_i - y_j)^2]$, where $i \neq j$. It is easy to relate $D(x)$ and $D(y)$ using a coalescent argument. In comparing two mitochondria from $\mathcal{Y}$, there are two cases: With probability M^{-1} they have the same parent mitochondrion in $\mathcal{X}$, in which case they are identical; with probability $1 - M^{-1}$ they have different parents, in which case the mean square difference is $D(x)$. Hence, $D(y) = (1 - M^{-1})D(x)$. Using the property that the mean square difference is twice the variance, we obtain the result that

$$\mathrm{Var}(y_i|\mathcal{X}) = (1 - M^{-1})\mathrm{Var}(x_i). \tag{B4}$$

Let Σz_i be the total number of mutations in a particular individual following the bottleneck. Then $M^2 \mathrm{Var}(\bar{z}) = \mathrm{Var}(\Sigma z_i)$. From this we find that

$$\begin{aligned}\mathrm{Var}(\Sigma z_i|\mathcal{X}) &= \mathrm{Var}(x_i)\left(\frac{M^2}{B} + M - 1\right) \quad \text{for } B > 1\\ &= \mathrm{Var}(x_i)M^2 \qquad\qquad\qquad \text{for } B = 1.\end{aligned} \tag{B5}$$

The expected variance among hosts following the bottleneck (σ_b^2) equals the variance among hosts before the bottleneck (σ_m^2) plus a component due to the bottleneck sampling

$$\sigma_m^2 = \sigma_b^2 + \frac{1}{N}\sum_k \mathrm{Var}(\Sigma z_i|\mathcal{X}_k), \tag{B6}$$

where N is the population size, and $\mathcal{X}_k$ gives the distribution of mutations in the kth host before the bottleneck phase. If we let $\mathrm{Var}(x_{i,k})$ be the within-host variance in the kth host, then this becomes

$$\begin{aligned}\alpha_m^2 &= \alpha_b^2 + \frac{1}{N}\left[\sum_k \mathrm{Var}(x_{i,k})\right]\left(\frac{M^2}{B} + M - 1\right) \quad \text{for } B > 1\\ &= \sigma_b^2 + \frac{1}{N}\left[\sum_k \mathrm{Var}(x_{i,k})\right]M^2 \qquad\qquad \text{for } B = 1.\end{aligned} \tag{B7}$$

Note that the average sample variance, s_m^2 equals $(1 - M^{-1})\,\Sigma_k \mathrm{Var}(x_{i,k})/N$, which completes the derivation of Equation 5.

Equation 6: Let z_i and z_j be the numbers of mutations in two mitochondria drawn from a single host following the second (final) stage of the bottleneck process ($i \neq j$). Define the mean square difference D between these as $D(z) \equiv E[(z_i - z_j)^2]$. By the coalescent argument used above (Equation 5 proof), the probability of the two mitochondria having the same parent mitochondrion among the y_i is B^{-1}; the probability of them having different parent mitochondria is $1 - B^{-1}$, in which case the mean square difference is $D(y)$. So $D(z) = (1 - B^{-1})D(y)$, and substituting the value of $D(y)$ obtained above,

$$D(z) = (1 - B^{-1})(1 - M^{-1})D(x) \tag{B8}$$

for $(B > 1)$, $(M > 1)$. Recalling that the expected squared difference is twice the variance, and taking the average across individuals, we obtain Equation 6.

Equation 7: Let $q_b(n)$ be the frequency of individuals carrying n mutations after the bottleneck phase (*i.e.*, before selection) and $q_s(n)$ be the frequency of that class after selection. Let $w(n)$ be the fitness of an individual carrying n mutations [hence $w(n)$ is given by the selection function]. Then

$$q_s(n) = \frac{w(n)q_b(n)}{\bar{w}_b}, \tag{B9}$$

and so

$$\begin{aligned}\bar{w}_s &= \bar{w}_b + (\bar{w}_s - \bar{w}_b)\\ &= \bar{w}_b + \sum_n w(n)\left(\frac{w(n)q_b(n)}{\bar{w}_b}\right) - \sum_n w(n)q_b(n)\\ &= \bar{w}_b + \frac{\mathrm{Var}(w_b)}{\bar{w}_b},\end{aligned} \tag{B10}$$

where $\mathrm{Var}(w_b)$ is the variance in fitness after the bottleneck, and equals $\alpha^2\,\sigma_b^2$.

Equation 8: The variance in fitness following selection ($\mathrm{Var}[w_s]$) is given by

$$\begin{aligned}\mathrm{Var}[w_s] &= \sum_n w(n)^2 q_s(n) - (\bar{w}_s)^2\\ &= \sum_n w(n)^3 q_b(n)/\bar{w}_b - (\bar{w}_s)^2\\ &= \overline{w_b^3}/\bar{w}_b - (\bar{w}_s)^2,\end{aligned} \tag{B11}$$

from which we can obtain (8), using the fact that $\mathrm{Var}[w_s] = \alpha^2\sigma_s^2$.

Equation 9: Let the within-host variance of the kth individual (before selection) be $s_{b,k}^2$ and let its fitness be $w_{b,k}$. Then

$$\begin{aligned}\mathrm{E}[s_s^2] &= \frac{1}{N}\sum_k s_{b,k}^2\, w_{b,k}/\bar{w}_b\\ &= \mathrm{E}[s_{b,k}^2,\, w_{b,k}]/\bar{w}_b\\ &= \mathrm{Cov}[s_{b,k}^2,\, w_{b,k}]/\bar{w}_b + s_b^2.\end{aligned} \tag{B12}$$