

Mutation Accumulation in Growing Asexual Lineages

J. F. Fontanari,¹ A. Colato,¹ and R. S. Howard²

¹*Instituto de Física de São Carlos, Universidade de São Paulo, Caixa Postal 369, 13560-970 São Carlos São Paulo, Brazil*

²*Department of Biology, Middle Tennessee State University, Murfreesboro, Tennessee 37132, USA*

(Received 1 July 2003; published 18 November 2003)

The stochastic loss of entire classes of individuals bearing the fewest number of mutations — a process known as Muller’s ratchet — is studied in asexual populations growing unconstrained from a single founder. In the neutral regime, where mutations have zero effect on fitness, we derive a recursion equation for the probability distribution of the minimum number of mutations carried by individuals in the least-loaded class, and obtain an explicit condition for the halting of the ratchet. Next, we consider the case of deleterious mutations, and show that weak selection can actually accelerate the ratchet beyond that achieved for the neutral regime. This effect is transitory, however, as our results suggest that even weak purifying selection will eventually lead to the complete cessation of the ratchet. These results may have important implications for problems in biology and the medical sciences.

DOI: 10.1103/PhysRevLett.91.218101

PACS numbers: 87.10.+e, 02.50.Ey, 87.23.Kg

In the 1960s, Muller first described the potential for nonrecombining asexual lineages to deteriorate due to the irreversible accumulation of slightly deleterious mutations [1]. He reasoned that, in the absence of compensatory back mutation, the number of mutations (mutation load) carried by any individual can be no less than that of its parent. As a consequence, in finite populations the occasional loss of the entire class of individuals bearing the fewest number of mutations (the least-loaded class) is effectively guaranteed. This may occur either because all individuals in the class fail to reproduce, or because each offspring in the class receives at least one additional mutation. As this sequence of losses is reminiscent of the movement of a ratchet, the process became known as Muller’s ratchet [2].

In recent years, Muller’s ratchet has been shown to have important implications for a variety of problems in the biological and medical sciences. For example, it has been invoked as a possible explanation for the extinction of asexual lineages [3], the advantage of genetic recombination [4], infertility in human males [5], and the attenuation of the human immunodeficiency virus [6]. It may also contribute to the apparent advantage that sexual reproduction has over asexual reproduction in nature [7].

Following the seminal paper by Haigh, most theoretical studies of the ratchet have assumed a fixed population size [8]. This assumption warrants further scrutiny, however, as the efficiency of the ratchet is inextricably linked to population size. For example, a positive feedback between the speed of the ratchet and the overall rate of population decline, termed mutational meltdown, greatly reduces the chances of survival of finite asexual lineages [3]. Reductions in population size have also been shown to accelerate the ratchet during serial bottleneck passage experiments [9–11], and in parasite transmission bottlenecks where only a few pathogens are transmitted from one host to another [12].

Theoretical analyses of the ratchet in populations undergoing bottleneck events have been carried out recently, but all of them considered that the loss of the least-loaded individuals occurs in the transfer stage only [13–15]. In the present study, we take a different approach and consider the potential for Muller’s ratchet to operate in populations that are increasing in size. In Haigh’s framework, mutations are assumed to be slightly deleterious, and selection operates to remove from the population those individuals with the highest mutation loads; hence, the main effect of selection is the deceleration of the ratchet [8,16]. In this contribution we show that, for a population undergoing unconstrained growth, the effect of selection is more subtle. To study the influence of selection on the ratchet rate we resort to extensive agent-based simulations. However, in the neutral regime, where the effect of mutation is zero, we derive a recursion equation for the probability distribution of the minimum number of mutations n carried by individuals in generation t , denoted by $\mathcal{P}_n^{(t)}$. This distribution approaches a unique steady state for $t \rightarrow \infty$, provided a simple condition [see Eq. (12)] is satisfied.

Let us assume that for the case of the neutral regime, in the initial generation ($t = 0$) the population is composed of a single, mutation-free individual — the founder. This individual, as well as any of its descendants, is capable of producing r offspring, regardless of the number of mutations it carries, where $r = 0, 1, \dots$ is a random variable distributed according to the Poisson distribution $p(r; R)$. Here R is the mean number of offspring produced per individual (the per capita reproductive rate), and we have used the standard notation for the Poisson distribution

$$p(k; \lambda) = \exp(-\lambda) \frac{\lambda^k}{k!}, \quad k = 0, 1, \dots \quad (1)$$

In addition, the number of new mutations $l = 0, 1, \dots$ acquired by newborns is distributed by the Poisson

$f_l \equiv p(l; U)$ where U is the mean number of new mutations per individual per generation. After the individuals in the population had the chance to generate their progeny, they all died, being then replaced by their offspring. In this sense there is no overlap of generations in our setting.

Clearly, the population size N at generation t is a stochastic variable that can take on any integer value (if $N = 0$ the population is extinct). A useful result we derive now is the probability $S^{(t)}$ that a lineage survives at a given generation. Let $\mathcal{N}(k, t)$ denote the probability that the population is composed of k individuals in generation t . The number of offspring produced by these individuals (i.e., the population size at $t + 1$) is a sum of k independent Poisson distributed random variables of mean R , which is also a Poisson variable of mean kR . Hence,

$$\mathcal{N}(N, t + 1) = \sum_{k=0}^{\infty} p(N; kR) \mathcal{N}(k, t). \quad (2)$$

This recursion equation can be solved by introducing the generating function $g(z, t) = \sum_{N=0}^{\infty} z^N \mathcal{N}(N, t)$, which is easily shown to satisfy the recursion

$$g(z, t + 1) = g[e^{-R(1-z)}, t] \quad (3)$$

with the initial condition $g(z, 0) = z$. Solving Eq. (3) recursively and noting that $S^{(t)} = 1 - g(0, t)$ yields the recursion

$$S^{(t+1)} = 1 - \exp(-RS^{(t)}) \quad (4)$$

with $S^{(0)} = 1$. Hence a nonvanishing survival probability in the limit $t \rightarrow \infty$ is possible provided that $R > 1$, and in what follows we will consider this regime only.

Next we consider two statistical measures of the mutation accumulation process at the population level. The first, and most directly related to the ratchet, is the number of mutations carried by the least-loaded individuals in the population—the ratchet clicks whenever this quantity changes. The second is the average number of mutations in the population, which yields information on the rate of mutation accumulation. When population size is kept constant, it has been argued that both measures increase in time at the same rate, such that they are equally useful in estimating the rate at which the ratchet clicks [16]. We will show, however, that this is not so for an expanding population.

Let us consider a single individual, referred to as the parent, that carries k mutations. Assuming that the parent produces r progeny, the probability that the least-loaded offspring differ from the parent by m mutations is

$$P(m | r) = \left[f_m + \sum_{l=m+1}^{\infty} f_l \right]^r - \left[\sum_{l=m+1}^{\infty} f_l \right]^r. \quad (5)$$

Hence the probability that the least-loaded offspring carry m new mutations regardless of the progeny size is

given by

$$\begin{aligned} G_m(\{f_l\}) &= \sum_{r=0}^{\infty} P(m | r) p(r; R) \\ &= [\exp(Rf_m) - 1] \exp\left(-R \sum_{l=0}^m f_l\right). \end{aligned} \quad (6)$$

Note that if the parent is the original founder, in which case $k = 0$, one has simply

$$\mathcal{P}_n^{(1)} = G_n(\{f_l\}). \quad (7)$$

On the other hand, if the parent is a direct descendant of the founder (i.e., it belongs to the first generation) then the probability that it carries k mutations is simply f_k and so the probability that the least-loaded offspring in its progeny carry a total of n mutations is the convolution $\sum_{k=0}^n f_k G_{n-k}(\{f_l\}) = \sum_{k=0}^n f_k \mathcal{P}_{n-k}^{(1)}$. The idea is then to replace each one-level subtrees branching from the members of the first generation by a single virtual individual carrying a number of mutations distributed according to the convolution $\sum_{k=0}^n f_k \mathcal{P}_{n-k}^{(1)}$, and then proceed to find the probability distribution for the minimum n of those numbers, which is given by the function G_n calculated at this convolution. Similarly, in the third generation we replace each two-level subtree branching from the individuals of the first generation by virtual individuals carrying a number of mutations distributed by the convolution $\sum_{k=0}^n f_k \mathcal{P}_{n-k}^{(2)}$ and then determine the distribution of the minimum among those numbers. This procedure yields the recursion equation

$$\mathcal{P}_n^{(t+1)} = G_n\left(\left\{\sum_{k=0}^l f_k \mathcal{P}_{l-k}^{(t)}\right\}\right), \quad (8)$$

which in the cases $n = 0$ and $n = 1$ is written explicitly as

$$\mathcal{P}_0^{(t+1)} = [\exp(Rf_0 \mathcal{P}_0^{(t)}) - 1] \exp(-Rf_0 \mathcal{P}_0^{(t)}) \quad (9)$$

and

$$\begin{aligned} \mathcal{P}_1^{(t+1)} &= [\exp(Rf_0 \mathcal{P}_1^{(t)} + Rf_1 \mathcal{P}_0^{(t)}) - 1] \\ &\times \exp[-R(f_0 + f_1) \mathcal{P}_0^{(t)} - Rf_0 \mathcal{P}_1^{(t)}]. \end{aligned} \quad (10)$$

The question we address now is whether a stationary distribution $\mathcal{P}_n^{(t+1)} = \mathcal{P}_n^{(t)} = \mathcal{P}_n^{\infty} > 0$ exists or whether the only possibility is the usual situation in which the ratchet never halts, corresponding to the nonstationary distribution $\mathcal{P}_n^{\infty} = 0$ for any finite n . Expansion of the right-hand side of Eq. (9) in powers of \mathcal{P}_0^{∞} yields

$$\mathcal{P}_0^{\infty} \approx \frac{2}{R^2 f_0^2} (Rf_0 - 1), \quad (11)$$

which, when inserted into Eq. (8), results in $\mathcal{P}_n^{\infty} \propto (Rf_0 - 1)^{1/2^n}$. These results show that such a steady state distribution exists indeed, provided that $Rf_0 > 1$, i.e.,

$$U < \ln R. \quad (12)$$

More generally, this condition yields the threshold for the nonvanishing of the probability that the least-loaded individuals carry the same number of mutations as the founder, irrespective of that number. It is self-evident since it implies that the average number of offspring bearing no new mutations produced by an individual is one, so that on average the number of least-loaded individuals is unchanged [17].

Consider now the problem of measuring the average number of mutations carried by the least-loaded individuals in generation t , henceforth denoted by $k_m^{(t)}$, by simulating the evolution of many independent populations. The procedure is to record the mutation load of the least-loaded individuals for each population in generation t , add them all, and then divide the result by the number of populations that have survived at generation t . This is equivalent to writing

$$k_m^{(t)} = \sum_{n=0}^{\infty} n \frac{\mathcal{P}_n^{(t)}}{S^{(t)}}, \quad (13)$$

where the ratio between probabilities may be interpreted as the probability that the least-loaded individuals in generation t carry n mutations conditional to the survival of the population at this generation. The dependence of $k_m^{(t)}$ on the generation number t , depicted in Fig. 1, shows that the signature of the regime where the ratchet comes to a halt is the emergence of plateaus in these plots. For fixed U we find that the stationary value k_m^{∞} decreases exponentially with increasing R while for fixed R , as shown in the figure, k_m^{∞} increases with U and diverges at the threshold value $U = \ln R$. To corroborate the analytical predictions, this figure shows also the results of individual-based computer simulations of 10^6 independent lineages. Since on the average the population size increases exponentially, we can keep track of the lineage for only a few generations (typically $t = 15$). The inset of

Fig. 1 shows that when condition (12) is violated, and for sufficiently large t , $k_m^{(t)}$ increases linearly with t , the (constant) rate of increase being then the ratchet rate. As expected, this rate decreases with increasing R since the larger the progeny of an individual the larger its odds of producing offspring free of new mutations. For large U , it increases linearly with the mutation rate, but the rate of increasing is definitely lower than unity.

The calculation of the average number of mutations $\bar{k}^{(t)}$ in the population at generation t is much simpler. Since all individuals at generation t are equivalent, we need to know only the average number of mutations carried by a single individual at that generation, which is simply $\bar{k}^{(t)} = Ut$. This is the same result obtained for the case where the population size is kept constant [16]. Interestingly, while the mutations keep accumulating steadily with rate U in the population, regardless of the value of the per capita reproductive rate, in the regime $R > e^U$ the average mutation load of the least-loaded individuals remains fixed at k_m^{∞} . Moreover, in the regime where the ratchet does not stop, we find that the rate of increase of \bar{k} (i.e., U) is significantly larger than the actual rate of the ratchet. So, in contrast to results obtained for a fixed population size, in expanding populations the average mutation load yields no information on the rate at which the ratchet operates.

Figure 2 illustrates how the average mutation load of the least-loaded class, $k_m^{(t)}$, is modified by the introduction of selection against the mutations. As before, each symbol represents an average over 10^6 independent lineages. Specifically, we maintain that all individuals have the same average reproductive rate R , but assume that the probability of survival of the offspring depends on the number of mutations l they carry, being given by $w = (1 - s)^l$ where s is the deleterious effect of each mutation. A striking feature of this figure is the nontrivial

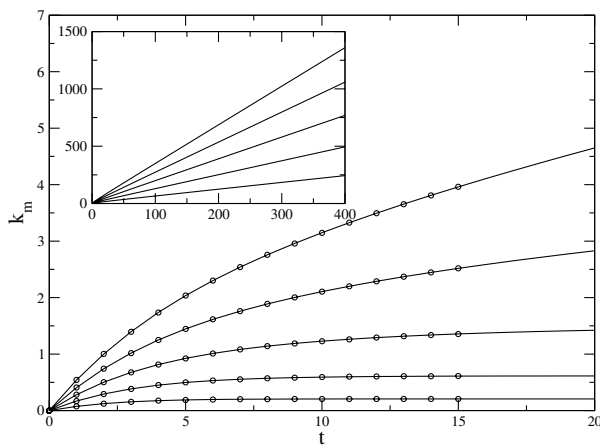


FIG. 1. Average mutation load of the least-loaded class against generation number for $R = 2$ and (bottom to top) $U = 0.2, 0.4, \dots, 1.0$ in the main graph, and $U = 2, 3, \dots, 6$ in the inset. The symbols are the simulation results.

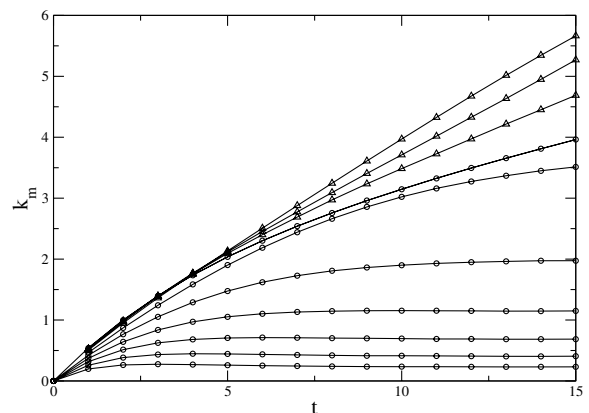


FIG. 2. Average mutation load of the least-loaded class against generation number for $R = 2$, $U = 1$ (Δ from top to bottom), $s = 0.1, 0.05, 0.03$, and (\circ from top to bottom) $s = 0, 0.2, 0.3, \dots, 0.7$. The lines connecting the symbols are guides to the eye.

dependence of $k_m^{(t)}$ on the selection parameter s : starting from the neutral value, $k_m^{(t)}$ first increases reaching a maximum value at $s \approx 0.1$ then decreases continuously towards zero as s increases further on. This suggests that selection against mutations of slight effect ($s < 0.1$) can actually accelerate the ratchet beyond that achieved for the neutral regime, although the effect is transitory. This counterintuitive result has a simple explanation.

There are three different mechanisms which can contribute to the loss of the least-loaded class: forward mutation pressure, stochastic loss, and elimination by selection. When selection is strong ($s > 0.1$), the variance in fitness among classes is large, and those with many mutations suffer a disproportionately higher rate of elimination from the population. As a consequence, the distribution of mutation number in the population is shifted toward the least-loaded classes, and the probability of their elimination by the stochastic mechanism is diminished. This is responsible for the expected slowing down of the ratchet. Weak selection, on the other hand, has a much smaller effect on the overall distribution of the mutation number in the population; hence, the risk of elimination of the least-loaded class by the stochastic mechanism approaches that achieved under the neutral regime. But at the same time, the loss of classes characterized by $k > 0$ mutations is slightly more likely under small s than for the neutral regime (since some offspring may die before reproduction). This mechanism operates only after the class $k = 0$ is eliminated, which accounts for a delay in the effect until the population has grown for a minimum number of generations, typically $t > 3$.

Finally, the average mutation load of the population can be calculated analytically and yields simply

$$\bar{k}^{(t)} = \frac{U}{s}(1-s)[1 - (1-s)^t], \quad (14)$$

regardless of the value of R . Since $k_m^\infty \leq \bar{k}^\infty = U(1-s)/s$, this result implies that $s > 0$ is a sufficient condition to halt the ratchet. This is consistent with the finding that the fraction of each fitness class reaches a stable value despite the lineage growth [15]. The ratchet acceleration due to the weak selection observed in Fig. 2 is thus limited to the initial generations only, and in the long run selection always brings the ratchet to a halt.

The analysis of asexual lineages growing unconstrained from a single founder has revealed at least two

remarkable effects. First, we find that for the neutral regime, the ratchet is halted completely provided condition (12) is met. Second, we show that for the case of deleterious mutations, selection eventually brings the ratchet to a complete halt. Interestingly, the ratchet undergoes an initial but transitory period of acceleration when the effect of mutation is small ($s < 0.1$). These results support the assumption that the loss of the least-loaded classes during serial bottleneck experiments occurs in the transfer stage only [13–15], and may have important implications for rates of mutation accumulation in organisms which undergo dynamic fluctuations in population size [7].

The research at São Carlos was supported by FAPESP, Project No. 99/09644-9. J. F. F. was supported in part by CNPq, and R. S. H. was supported by a summer grant from the MTSU College of Graduate Studies.

-
- [1] H. J. Muller, *Mutat. Res.* **1**, 1 (1964).
 - [2] J. Felsenstein, *Genetics* **78**, 737 (1974).
 - [3] M. Lynch, R. Bürger, D. Butcher, and W. Gabriel, *J. Hered.* **84**, 339 (1993).
 - [4] D. D. G. Gessler and S. Z. Xu, *Genet. Res.* **73**, 131 (1999).
 - [5] H. J. Cooke, *Rev. Reprod.* **4**, 5 (1999).
 - [6] E. Yuste, S. Sanchez-Palomino, C. Casado, E. Domingo, and C. Lopez-Galindez, *J. Virol.* **73**, 2745 (1999).
 - [7] R. S. Howard and C. M. Lively, *Nature (London)* **367**, 554 (1994); *J. Evol. Biol.* **15**, 648 (2002).
 - [8] J. Haigh, *Theor. Pop. Biol.* **14**, 251 (1978).
 - [9] L. Chao, *Nature (London)* **348**, 454 (1990).
 - [10] E. Duarte, D. Clarke, A. Moya, E. Domingo, and J. Holland, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6015 (1992).
 - [11] L. M. Wahl, P. J. Gerrish, and I. Saika-Voivod, *Genetics* **162**, 961 (2002).
 - [12] C. T. Bergstrom, P. McElhany, and L. A. Real, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5095 (1999).
 - [13] A. Colato and J. F. Fontanari, *Phys. Rev. Lett.* **87**, 238102 (2001).
 - [14] E. Lázaro, C. Escarmís, E. Domingo, and S. C. Manrubia, *J. Virol.* **76**, 8675 (2002).
 - [15] S. C. Manrubia, E. Lázaro, J. Pérez-Mercader, C. Escarmís, and E. Domingo, *Phys. Rev. Lett.* **90**, 188102 (2003).
 - [16] P. G. Higgs and G. Woodcock, *J. Math. Biol.* **33**, 677 (1995).
 - [17] A. L. Melzer and J. H. Koeslag, *Evolution* **45**, 649 (1991).