Rare beneficial mutations can halt Muller’s ratchet

Sidhartha Goyal‡, Daniel J. Balick‡, Elizabeth R. Jerison‡, Richard A. Neher§, Boris I. Shraiman∗†, and Michael M. Desai‡

‡Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA, †Department of Physics, University of California, Santa Barbara, CA, Departments of Organismic and Evolutionary Biology and of Physics, and FAS Center for Systems Biology, Harvard University, Cambridge MA, and ﬁMax Planck Institute for Developmental Biology, Tübingen, Germany

Submitted to Proceedings of the National Academy of Sciences of the United States of America

The vast majority of mutations are deleterious, and are eliminated by purifying selection. Yet in finite asexual populations, purifying selection cannot completely prevent the accumulation of deleterious mutations due to Muller’s ratchet: once lost by stochastic drift, the most-fit class of genotypes is lost forever. If deleterious mutations are weakly selected, Muller’s ratchet turns into a mutational “meltdown” leading to a rapid degradation of population fitness. Evidently, the long term stability of an asexual population requires an influx of beneficial mutations that continuously compensate for the accumulation of the weakly deleterious ones. Here we propose that the stable evolutionary state of a population in a static environment is a dynamic mutation-selection balance, where accumulation of deleterious mutations is on average offset by the influx of beneficial mutations. We argue that this state exists for any population size $N$ and mutation rate $U$. Assuming that beneficial and deleterious mutations have the same fitness effect $\epsilon$, we calculate the fraction of beneficial mutations, $\epsilon$, that maintains the balanced state. We find that a surprisingly low $\epsilon$ suffices to maintain stability, even in small populations in the face of high mutation rates and weak selection. This may explain the maintenance of mitochondria and other asexual genomes, and has implications for the expected statistics of genetic diversity in these populations.

Purifying selection maintains well-adapted genotypes in the face of deleterious mutations [1]. Yet in asexual populations, random genetic drift in the most-fit class of individuals will occasionally lead to its irreversible extinction, a process known as Muller’s ratchet [2, 3]. The repetitive action of the ratchet leads to the accumulation of deleterious mutations, despite the action of purifying selection. This ratchet effect has been extensively analyzed [4, 5, 6, 7, 8] and has been observed in experiments [9, 10, 11, 12] and in nature [13, 14, 15]. In small populations when deleterious mutation rates are high or selection pressures are weak, the ratchet can proceed quickly, causing rapid degradation of asexual genomes and a collapse in population size: the so-called “mutational meltdown” [16, 17, 18]. Hence, Muller’s ratchet has been described as a central problem for the maintenance of asexual populations such as mitochondria [19]. Avoiding this mutational catastrophe is thought to be a major benefit of sex and recombination (see [20, 21] for reviews).

In the absence of recombination [22] or epistasis [23], the only forces that can check the accumulation of deleterious mutations are back and compensatory mutations. These mutations are typically assumed to be rare and are either neglected in models of the ratchet, or assumed to only slightly slow its rate [1, 24, 25]. However, extensive experimental work [26, 27, 28, 29, 30, 31, 32] and theoretical expectations [33, 34, 35, 36] suggest that the number of available back and compensatory mutations increases as a population accumulates deleterious mutations and declines in fitness. The increase in beneficial mutation rates with declining fitness will eventually stop the progress of Muller’s ratchet, as illustrated in Fig. 1. Furthermore, since back and compensatory mutations are selectively favored, they can balance deleterious mutations even while they are relatively rare, thereby maintaining a well-adapted population.

Despite the potential relevance of compensatory mutations, only a few studies have quantitatively analyzed the simultaneous accumulation of both beneficial and deleterious mutations. These include simulations by Antezana and Hudson [37] and Wagner and Gabriel [38], experimental work by Silander et al. [28], and both simulations and analytical work by Rouzine et al. [39, 40]. In each of these analyses, simulations or experiments showed that the mean population fitness settles at a steady-state value $\omega_t$ — the point where beneficial mutations arise often enough to compensate for newly-fixed deleterious mutations.

While the absolute fitness, $\omega_t$, of the population in this dynamic equilibrium state is model dependent, the existence of the stable steady state is generic. Intuitively, suppose a fraction

Reserved for Publication Footnotes
e of mutations are beneficial, and the critical fraction needed to maintain stability is $\epsilon_c$. In a poorly adapted population, a higher fraction of mutations will be beneficial, $\epsilon > \epsilon_c$. This excess of beneficial mutations will push the population towards higher fitness. At the same time, adaptation will deplete the available pool of beneficial mutations, until $\epsilon_c$ is reached, as shown in Fig. 1. Conversely, in an “over-adapted” population, we expect $\epsilon < \epsilon_c$, so that deleterious mutations dominate, reducing population fitness until $\epsilon_c$ is recovered. Thus we expect the dynamic mutation-selection balance point to be a stable evolutionary “attractor.” In contrast to $\omega_c$, the critical $\epsilon_c$ provides a general measure of how well adapted a population is, independent of the model of compensatory mutations.

Below, we show analytically how $\epsilon_c$ depends on population size, mutation rate, and strength of selection and compare our predictions to numerical simulations. We show that even for high mutation rate and small population size, the stable state occurs at high fitness, where most mutations are deleterious. Muller’s ratchet notwithstanding, selection enables rare beneficial mutations to compensate for more frequent deleterious mutations, and to maintain a well adapted population.

Model

Our analysis is based on the standard discrete generation Wright-Fisher model with population size fixed at $N$. We assume that beneficial and deleterious mutations increase or decrease fitness by $s$ (where $s > 0$ by convention). Each individual has some number, $k$, deleterious mutations relative to the perfectly adapted state. We define $\lambda$ to be the mean number of deleterious mutations per individual, so that an individual with $k$ deleterious mutations has fitness $s(\lambda - k)$ relative to the mean, as shown in Fig. 2. The population can be characterized by the number of individuals $n_k$ in each of these fitness classes. We define $U$ to be the total mutation rate, and $\epsilon$ to be the fraction of mutations that are beneficial, so that the beneficial mutation rate is $U_b = U\epsilon$, and the deleterious mutation rate is $U_d = U(1 - \epsilon)$.

Although mutations and genetic drift are stochastic processes, in the bulk of the distribution — where fitness classes contain many individuals — the population dynamics are well-captured by a deterministic approximation. On average, the number of individuals in fitness class $k$ evolves as:

$$\frac{dn_k}{dt} = s(\lambda - k)n_k - U_b n_k + U_d n_{k+1} + U_b n_{k-1}.$$  \[1\]

The terms on the right hand side of this equation describe the different processes acting on fitness class $k$: (I) the effect of selection relative to the mean fitness, (II) mutational load, (III) deleterious mutations from more-fit individuals, and (IV) beneficial mutations from less-fit individuals. We assume that the selective effect of a single mutation is small, $s \ll 1$.

In general, we expect $\epsilon$ to increase with $k$, since less-fit genotypes have a higher frequency of beneficial back and compensatory mutations. However, at any given time there will only be a relatively narrow range of $k$ present in the population. Thus in Eq. (1) we neglect the $k$-dependence of $\epsilon$; we have tested this approximation against a simulation incorporating the $k$-dependence of $\epsilon$ (see SI). We use this approximation to compute the $\epsilon$ and corresponding distribution $n_k$ at which the population stays on average at a constant fitness. As we have argued above, the stability of this equilibrium state is ensured by the $k$-dependence of $\epsilon$ on larger scales.

Results

In steady state, the fitness distribution $n_k$ stays constant on average: $\frac{dn_k}{dt} = 0$ for all $k$. Solving Eq. (1) for the steady state in the case of $\epsilon = 0$ leads to the familiar mutation-selection balance [1], with a Poisson distribution of fitness: $n_k = Ne^{-\lambda} / \lambda^k$, where $\lambda = U/s$ and $k = 0$ corresponds to the most-fit class.

This deterministic approximation corresponds to an infinite population where Muller’s ratchet (i.e. the stochastic extinction of the fittest genotype) does not operate. In the absence of such stochastic extinction, any beneficial mutation from the fittest class would move the distribution towards higher fitness, in conflict with the steady state assumption. Hence, in infinite populations, steady state can only be achieved with $\epsilon_c = 0$. Conversely, in a finite population where the fittest class can be lost due to genetic drift, Muller’s ratchet will eventually lead to a decrease in fitness if no beneficial mutations are available. Thus we must have $\epsilon_c > 0$ to be in the dynamic steady state.

A correct description of the dynamic equilibrium state therefore requires a suitable treatment of genetic drift. These stochastic effects are particularly important in the most-fit edge of the distribution, which we call the “nose,” where the number of individuals with a particular fitness is small. Our analysis is based on matching a stochastic treatment of the nose with a deterministic description of the bulk of the fitness distribution (as governed by Eq. (1) with $\epsilon > 0$). The stochastic dynamics of the nose are determined by two competing processes: the random extinction of the most-fit class, versus the establishment of a new more-fit class due to a beneficial mutation. At stationarity, the rates of these two processes have to be equal; this condition determines the number of individuals in the nose class. We match this stochastic condition with the number of individuals in the nose class calculated based on the deterministic distribution. This determines the critical fraction of beneficial mutations $\epsilon_c$ as a function of the population parameters.
The shape of the fitness distribution. We begin with a deterministic analysis of the shape of the fitness distribution at steady state. Since only relative fitnesses matter, it is convenient to set the origin of $k$ such that the mean of the fitness distribution is $k = \lambda$. We solve for the steady state of Eq. (1) using Fourier analysis (see SI). We find

$$n_k = Ne^{-\lambda(1-2\epsilon)k \log \sqrt{\pi} J_1(2\epsilon k)},$$  \hspace{1cm} [2]

where $J_1$ denotes the Bessel function of order $k$ and we have defined $\alpha = \lambda \sqrt{\pi(1 - \epsilon)}$. A comparison of this solution to simulation results is shown in Fig. 3. For $k$ in the vicinity of $k^*$ this distribution is approximately Gaussian with variance $\sigma^2_k = \lambda(1 - 2\epsilon)$. While the the low fitness tail of the distribution decays less rapidly than a Gaussian, the high fitness side decays more rapidly than a Gaussian.

It is important to note that in the deterministic limit there is no true steady state for arbitrary $\epsilon > 0$; this manifests itself in the lost of positivity of the solution given by Eq. (2) for high-fitness classes above the position of the nose class. While the position of the nose, which we will call $k^*$, requires a more careful treatment, Eq. (2) gives a very accurate description of the bulk of the fitness distribution (i.e. $n_k$ for $k > k^*$).

To determine the position of the nose, $k^*$, we observe that if classes with $k < k^*$ carry no individuals, fitness class $k^*$ does not receive an influx of mutations from the more-fit class $k^* - 1$. That is, term III in Eq. (1) is absent, yielding $(k - k^* - \lambda)n_{k^*} + \lambda e n_{k^*+1} = 0$. This along with Eq. (2), gives the following equation for $k^*$

$$k^* J_k(2\epsilon) = \sqrt{\pi} J_{k+1}(2\epsilon)$$  \hspace{1cm} [3]

This must in general be determined numerically. However, in SI we show that

$$k^* \approx \begin{cases} \frac{\lambda^2 \epsilon}{2\lambda \sqrt{\epsilon(1 - \epsilon)}} & \text{for } \epsilon \lambda^2 \ll 1 \\ \frac{\lambda^2 \epsilon}{2\lambda \sqrt{\epsilon(1 - \epsilon)}} & \text{for } \epsilon \lambda^2 \gg 1. \end{cases}$$  \hspace{1cm} [4]

To gain some intuition into the significance of $k^*$, note that $k^* = 0$ in an infinite population with $\epsilon = 0$. In this case, the fitness distribution is the familiar Poisson distribution and the fittest class contains a fraction $e^{-\lambda}$ of all individuals. In a finite population, $Ne^{-\lambda}$ will be smaller than one for sufficiently large $\lambda$, which implies that at steady state the mutation free genotype is typically absent and therefore $k^*$ is larger than 0. This adjustment of $k^*$ for finite population size is reflected in Eq. (4), which states that the most-fit class gets closer to the population mean as $\epsilon$ increases.

The stochastic matching condition. We seek to determine the $\epsilon = \epsilon_*$ at which the distribution is stationary for a given finite $N$. Since the most-fit class is populated by a comparatively small number of individuals, fluctuations due to genetic drift may change its occupancy significantly. In particular, the population at the nose can go extinct, in which case the next fitness class becomes the new nose. Alternatively, a lucky beneficial mutation can cause the nose to advance by one class. Stationarity can therefore only be achieved in an average sense; the rate of advancing the nose has to equal the rate of extinction of the nose.

Since these two processes depend sensitively on the number of individuals in the nose, requiring the equality of extinction and establishment rates determines the average size of the population of the nose, $n_{k^*}$ as a function of $U$, $s$, and $\epsilon$. We match this $n_{k^*}$ to the deterministic solution determined above to find $\epsilon_*(N, s, U)$.

The nature of the dynamics at the nose depends qualitatively on $n_{k^*}$. If $n_{k^*}$ is large enough that its dynamics are dominated by selection, it is rarely lost and Muller’s ratchet is slow. Conversely, if $n_{k^*}$ is small, the class turns over neutrally and is easily lost and reseeded. We begin by considering the slow ratchet regime where $n_{k^*}$ is relatively large, and then turn to the opposite fast ratchet regime.

Slow ratchet regime. In the regime where $n_{k^*} > 1$, the fittest class is only rarely lost due to drift. Note that this regime corresponds to $\epsilon \lambda^2 < 1$, which implies $n_{k^*} \approx Ne^{-\lambda}$ and $k^* \approx 0$. In this regime the mean extinction rate for the nose, $r_\epsilon$, is due to rare large fluctuations which overcome the “restoring force” due to selection trying to preserve mutation selection balance. Estimates of $r_\epsilon$ have been obtained via diffusion theory [1, 6, 4] and have the form (see SI)

$$r_\epsilon = e^{\gamma Ne^{-\lambda}} \sqrt{sNe^\lambda}/\sqrt{\pi},$$

where $\gamma$ is a parameter characterizing the effective strength of selection on the fittest class, introduced by Haigh [1].

To impose our stochastic condition, we must also calculate the rate at which a new more-fit class establishes, $r_\epsilon$. This is given by the product of the rate at which new beneficial mutants are generated and the probability they establish, $r_\epsilon = n_{k^*} U \gamma_\epsilon$. Since $k^* \approx 0$ in this regime, we have $P_{\text{ext}} \approx 2$. Equating the two rates yields the stochastic condition

$$\epsilon_* = \frac{\gamma^2}{2\lambda} e^{\gamma Ne^{-\lambda}} \sqrt{sNe^{-\lambda}}.$$

[5]

This condition suggests that $\epsilon_*$ in the slow ratchet regime is a function of a combination of population parameters: $\Lambda = (U, N, s)/\sqrt{sNe^{-\lambda}}$. We find that contrary to earlier studies [1, 4] constant $\gamma$ is not consistent with simulations. Instead, as shown in Fig. 4A the simulation data is well explained by $\gamma = 1/\sqrt{\lambda}$. Our empirical $\lambda$-dependence is however consistent with the approximation $\gamma \approx 0.6$ used by Gordo et al. [4], within the range of parameters $(0.1 \leq 1/\sqrt{\lambda} \leq 0.8)$ addressed by their study. Since we consider a much broader range of parameters $(0.1 \leq 1/\sqrt{\lambda} \leq 1)$, accounting for the dependence of $\gamma$ on $\lambda = U/\lambda$ becomes important. Our analytical solution Eq. (5), shown as dotted line in Fig. 4A, is in good agreement with the simulation data across this wide range of parameters.

In this regime, $\epsilon_*$ depends exponentially on $N e^{-\lambda}$, and hence rapidly approaches zero as $\lambda$ becomes small, approaching the infinite population limit where Muller’s ratchet does not operate.

Fast ratchet regime. For larger mutation rate, smaller population size, or weaker selection, the occupancy of the most
fit class decreases, thereby increasing its rate of extinction. Consequently, a higher rate of beneficial mutations (larger $\epsilon$) is required to match the extinction rate. The resulting rapid turnover of the population at the nose leads to the failure of the quasi-static approximation we used in the slow ratchet regime.

As the occupancy of the nose decreases, in particular when $n_{hs}s \leq 1$, the dynamics of the fittest class is governed by drift. The rate of extinction, $r_-$, can therefore be estimated from neutral diffusion: $r_- \approx 1/(2n_{hs})$. The rate at which a new more-fit class is established is given by the same formula as before, $r_+ = eU\bar{n}_h P_{oa}$. However, $P_{oa}$ now refers to the probability that a new mutant lineage reaches $\frac{1}{2\Delta s}n_{hs}$ individuals. At this point, the lineage crosses over from stochastic to deterministic dynamics, entering the domain described by the deterministic solution Eq. (2). Thus $P_{oa} = 2(k_+ + 1)s$. Equating $r_- = r_+$ yields the stochastic condition:

$$n_{hs} \approx \frac{1}{\sqrt{4ekU\bar{n}_h}}. \quad [6]$$

This condition, together with the solution Eq. (2) for the distribution, allows us to determine $\epsilon_c(N, s, U)$. As shown in Fig. 4, our solution Eq. (6) is in excellent agreement with simulations for most of the data. However, our solution overestimates $\epsilon_c$ for populations with small population size and small $\lambda$ that still satisfy $n_{hs}s \leq 1$. A better description in this regime will require a more careful analysis of fluctuations beyond the nose.

**Large population size limit.** Large populations can maintain a well adapted genome with $\epsilon_c < \frac{1}{2}$ even for moderately large $\lambda$. In the limit of large populations and large $\lambda$, with the limit taken so that $\epsilon_c\lambda^2 > 1$ while $\epsilon_c < \frac{1}{2}$, the matching condition reduces to $\epsilon_c^* = \lambda^{-1}\log(N\bar{s}e^{-c(1-2c)})$. This equation has the approximate solution

$$\epsilon_c^* \approx \frac{z^2}{\log^2(z/\log(z^{-1})))} \quad 2z \equiv 1 - \frac{\log N\bar{s}}{\lambda} \quad [7]$$

This result can be made more precise through iteration (see SI for derivations and comparisons to simulation data). Note here that $\epsilon_c$ depends only very weakly on $N\bar{s}$ and on $\lambda$, in contrast to the slow ratchet regime where $\epsilon_c$ declines exponentially with $N\bar{s}e^{-c}$.

**High mutation rate or weak purifying selection.** In this limit, the dynamic equilibrium tends towards the state where beneficial and deleterious mutations are equally frequent, $\epsilon = \frac{1}{2}$. This corresponds to a population that can no longer maintain a well-adapted state. In the SI, we derive an approximate expression for $\epsilon_c$ in the fast-ratchet regime in the limit of large $\lambda$:

$$\epsilon_c \approx \frac{1}{2} - \left(\frac{3}{4\lambda} \log N\bar{s}\right)^{1/3} \quad [8]$$

From this expression we can immediately determine the point at which the population can no longer maintain a well-adapted state, as defined by a maximum fraction of beneficial mutations, $\epsilon_{max}$. In order to have $\epsilon_c < \epsilon_{max}$, we require

$$\frac{\lambda}{\log N\bar{s}} \leq 3\left(4\left[\frac{1}{0.5 - \epsilon_{max}}\right]^3\right) \quad [9]$$

valid for $\epsilon_{max} \geq 0.2$. Note that for $\epsilon_{max} = 1/4$, the RHS of this expression is 48; for $\epsilon_{max} = 1/3$, the RHS is 162. Thus even small populations with $N\bar{s} \geq 1$ can maintain relatively well-adapted genomes in the face of high mutation rates and weak selection (i.e. when $U \gg \bar{s}$).

**Discussion**

In infinite populations it has long been recognized that the balance between mutational pressure and purifying selection leads to a fitness equilibrium [1, 41]. Our analysis demonstrates that such an equilibrium also exists in a finite population, despite the action of genetic drift and Muller’s ratchet. Previous work on these finite populations has focused primarily on the rate of Muller’s ratchet in the absence of beneficial mutations, and has generally assumed that back and compensatory mutations are rare enough to be neglected [42, 43, 6, 4, 5, 25]. However, we expect that beneficial mutations will become more common as a population accumulates deleterious mutations. Our analysis has illustrated how these beneficial mutations balance Muller’s ratchet in asexual populations, to create a dynamic equilibrium state.

We have shown that for any population size, mutation rate, and selection pressure, this dynamic equilibrium state is characterized by the proportion of beneficial mutations, $\epsilon_c$, necessary to balance the ratchet. This $\epsilon_c$ is a measure of how optimal
a genome a finite population can maintain in the presence of mutations, selection, and genetic drift. Larger $\epsilon$ indicates that the population needs more beneficial mutations to be in the stable state, and hence must be less fit, and vice versa. In the limit $\epsilon \rightarrow 0$, the population is perfectly adapted, and no beneficial mutations are possible. The opposite case of mutational meltdown, where a population cannot maintain an adapted genome, corresponds to $\epsilon \rightarrow 1/2$, where a random mutation is as likely to be beneficial as deleterious. There are two qualitatively different regimes for $\epsilon$, as shown in Fig. 5. In the slow ratchet regime, selection stabilizes the nose, so $\epsilon$ is small and depends exponentially on population parameters $Ns$ and $\lambda$. On the other hand, in the fast ratchet regime, $\epsilon$ must be larger to balance the faster ratchet, and depends more weakly on the population parameters $Ns$ and $\lambda$. The boundary of the two regimes is roughly given by $Ns^{-1} = 1$.

We note that Rouzine et al. [39, 40] studied population dynamics under the combined action of purifying and positive selection in a model similar to ours, finding both simulation and analytical evidence for the existence of the dynamic balance state. However, their analysis focuses on the rate of the ratchet (with $\epsilon \ll \epsilon_c$) or the rate of adaptation (with $\epsilon \gg \epsilon_c$), and does not apply as $\epsilon \rightarrow \epsilon_c$. Our work, by contrast, focuses exclusively on the dynamic equilibrium state and the critical $\epsilon_c$ required to maintain it.

Our analytic results for $\epsilon_c$ in the dynamic equilibrium were made possible by the explicit calculation of the shape of this fitness distribution, given in Eq. (2) and illustrated in Fig. 3. With increasing $N\epsilon$s and decreasing $N$ this distribution deviates from the Poisson distribution for the classic mutation-selection balance [1], most notably reducing the fitness of the top class, $k_\ast$, relative to the mean. Examining the properties of the dynamic mutation-selection equilibrium over the full parameter range, shown in Fig. 5, has revealed a strong asymmetry between beneficial and deleterious mutations. For example, for a small population $N\epsilon \sim 10^4$ with high mutation rate $U/s \sim 8$, just 2% beneficial mutations are enough to counteract the effect of deleterious mutations. This indicates that purifying selection is remarkably effective even for conditions where Muller’s ratchet would proceed extremely quickly in the absence of back and compensatory mutations.

Our analysis implies that populations with $\epsilon > \epsilon_c$ should adapt, while populations with $\epsilon < \epsilon_c$ should decline in fitness. Experimental evolution of model organisms in controlled laboratory environments appears to be consistent with this expectation. In particular, Silander et al. [28] showed that bacteriophage $\phi X 174$ converged to a population-size dependent fitness plateau, as our model would predict. For their population parameters ($s \sim 0.08$, $U \sim 0.13$) our theory predicts $\epsilon_c$ ranging from 2% to 0.2% for their experiments with $N$ ranging from 100 to 200, consistent with the beneficial mutation rates they infer. Their experiments with lower population sizes ($N \leq 30$) are in the Fast ratchet regime and estimating $\epsilon_c$ requires further analysis of fluctuations beyond the nose.

Other experiments with vesicular stomatitis virus found that large populations adapt while small populations melt, also consistent with our analysis [44]. Similar results have been observed in yeast [45, 46]. In the long-term evolution experiments of Lenski and collaborators in E. coli, parameters are such that we expect the critical $\epsilon_c$ to be very small, $\epsilon_c \ll 10^{-10}$, consistent with their observation of continuous adaptation even after tens of thousands of generations. On the other hand, mutator strains of E. coli shows decrease in fitness [47], which could be used to further test the model. In Fig. 5, we show rough estimates for the parameter regimes in which each of these experimental systems lie.

Population parameters for HIV place it close to the boundary between the slow and fast ratchet regimes (Fig. 5), so that its evolutionary dynamics depend sensitively on the strength of selection. This intriguing case highlights two important issues. First, our analysis has assumed that all mutations have the same selective effect $s$, illustrating how well a particular strength of purifying selection can maintain well-adapted states. In general, however, mutations have a range of selective effects. Mutations with a particular characteristic strength may tend to dominate the dynamics, if more strongly selected mutations are quickly eliminated by selection while more weakly selected mutations are effectively neutral given interference from selection on the other mutations. However, the effect is non-trivial and generalizing our results to an arbitrary distribution of mutational effects remains an important avenue for future work. Second, our analysis has focused on asexual populations, and neglected recombination. Since even weak recombination has the potential to significantly slow Muller’s ratchet, it would be interesting to generalize our dynamic balance to include its effects. Increasing recombination rates would presumably allow the population to maintain better-adapted genotypes in this steady state.

We have argued that the dynamic mutation selection balance is a stable evolutionary attractor. If the population size is independent of $\epsilon$, this is immediately obvious from Fig. 1. Under some conditions, however, the population size might be a function of absolute fitness and hence of $\epsilon$. Muller’s ratchet might then lead to a decrease in fitness, resulting in a smaller population size, which in turn leads to a faster ratchet. This phenomenon of mutational meltdown might cause rapid population extinction [16]. Eq. (8), however, shows that the degree of adaptation depends only weakly on the population size and a stable equilibrium can be maintained even if $N$ depends on $\epsilon$ rather strongly.

The dynamic balance state has important implications for patterns of molecular evolution. Recent analysis of the effects of purifying selection on the structure of genealogies has suggested that Muller’s ratchet plays a crucial role in determining the structure of genetic variation in asexual populations or on short distance scales in the genomes of sexual organisms [48, 49]. However, these expectations must be revised if populations exist instead in the dynamic state in which both beneficial and deleterious mutations fix, without any continuous net accumulation of deleterious mutations. This means...
that even though no change in fitness occurs, signatures of both positive and negative selection are likely to be found in patterns of molecular evolution, as has been suggested by earlier studies [50, 37]. We argue that this state is the natural null expectation for the effects of mutations and purifying selection on patterns of genetic variation; efforts to look for positive selection which represent “true” adaptation should look for deviations from this situation.

Finally, we note that although the dynamic balance we have analyzed is stable, a population will typically fluctuate around this steady state. A few beneficial mutations may become established at the nose by chance, leading to a temporary increase in mean fitness which is later balanced by a reduction in $e$ at this higher fitness, restoring the population to the equilibrium state. Conversely, a few clicks of Muller’s ratchet will occasionally lead to temporary reductions in fitness behavior before the corresponding increase in $e$ restores the steady state. These fluctuations around the dynamic balance may be important to the clonal structure of the population, and hence are likely to play a key role in patterns of molecular evolution and in understanding the effects of recombination.

Acknowledgements

We thank Pierre Neveu, Adel Dayarian, Aleksandra Walczak and Paul Sniegowski for many useful discussions. SG, DB and BIS were supported by NIGMS and HSFP. ERJ acknowledges support from an NSF graduate research fellowship. RAN is supported by the ERC through grant StG-2010-260666. MMD acknowledges support from the James S. McDonnell Foundation, the Alfred P. Sloan Foundation, and the Harvard Milton Fund.

1. Haigh, J. (1978) The accumulation of deleterious genes in a population. Theor. Popul. Biol. 14, 251–264.
2. Muller, H. (1964) The relation of recombination to mutational advance. Mutat. Res. 1, 2–9.
3. Felsenstein, J. (1974) The evolutionary advantage of recombination. Genetics 78, 157–159.
4. Gordo, I & Charlesworth, B. (2000) The degeneration of asexual haploid populations and the speed of muller’s ratchet. Genetics 154, 1379–1387.
5. Gordo, I & Charlesworth, B. (2000) On the speed of muller’s ratchet. Genetics 156, 2137–2140.
6. Stepney, W & Kim, Y. (2002) Recent Applications of Diffusion Theory to Population Genetics, eds. Slatkin, M & Veilleux, M. (Oxford University Press, Oxford, UK), pp. 72–93.
7. Charlesworth, B & Charlesworth, D. (1997) Rapid fixation of deleterious alleles can be caused by muller’s ratchet. Genetical Research 70, 63–72.
8. Gessler, D. (1995) The constraints of finite size in asexual populations and the rate of the ratchet. Genetical Research 49, 135–146.
9. Chao., L. (1990) Fitness of RNA virus decreased by muller’s ratchet. Nature 348, 454–455.
10. Duarte, E, Clarke, D, Moya, A, Domingo, E, & Holland, J. (1992) Rapid fitness loss in mammalian RNA virus clones due to muller’s ratchet. Proc. Natl. Acad. Sci. 89, 6015–6019.
11. Andersson, D & Hughes, D. (1996) Muller’s ratchet decreases fitness of a dna-based microbe. Proc. Natl. Acad. Sci. 93, 906–907.
12. Zeyl, C, Mizosko, M, & de Visser, J. (2001) Mutational meltdown in laboratory yeast populations. Evolution 55.
13. Rice, W. (1994) Degeneration of a nonrecombining chromosome. Science 263, 230–232.
14. Lynch, M. (1996) Mutation accumulation in transfer msa: molecular evidence for muller’s ratchet in mitochondrial genomes. Molecular Biology and Evolution 13, 209–220.
15. Howe, D & Denver, D. (2008) Muller’s ratchet and compensatory mutation in caenorhabditis briggsae mitochondrial genome evolution. BMC Evolutionary Biology 8, 62.
16. Lynch, M, Burger, R, Butcher, D, & Gabriel, W. (1993) The mutational meltdown in asexual populations. J. Hered. 84, 339–344.
17. Gabriel, W & Burger, R. (1993) Muller’s ratchet and mitochondrial meltdowns. Evolution 47, 1744–1757.
18. Lynch, M, Conery, J, & Burger, R. (1995) Mutation accumulation and the extinction of small populations. The American Naturalist 146, 489–518.
19. Loewe, L. (2006) Quantifying the genomic decay paradox due to muller’s ratchet in human mitochondrial dna. Genet. Res., Camb. 87, 133–159.
20. Barton, N & Charlesworth, B. (1998) Why sex and recombination? Science 281, 1968–1990.
21. de Visser, J & Elena, S. (2007) The evolutionof sex: empirical insights into the roles of epistasis and drift. Nature Reviews Genetics 8, 139–149.
22. Bell, G. (1998) Recombination and the immortality of the germ line. J. Evol. Biol. 1, 67–82.
23. Kondrashov, A. (1994) Muller’s ratchet under epistatic selection. Proc. Natl. Acad. Sci. 91, 1469–1473.
24. Kondrashov, A. (1995) Contamination of the genome by very slightly deleterious mutations: why have we not died 100 times over? Journal of Theoretical Biology 175, 583–594.
25. Bachtgro, D & Gordo, I. (2004) Adaptive evolution of asexual populations under muller’s ratchet. Evolution Int. J. Org. Evolution 58, 1403–1413.
26. Barrick, J, Kauth, M, Strollof, C, & Lenski, R. (2010) Escherichia coli rpoB mutants have increased evolvability in proportion to their fitness defects. Molecular Biology and Evolution 27, 1338–1347.
27. Schoustra, S, Hallmark, T, Gifford, D, & Kassen, R. (2009) The properties of adaptive walks in evolving populations of fungus. PLoS Biology 7, e1000250.
**Supplemental information**

**Simulation methods.** Our simulations are done using a custom written Python code available on request. We implement a discrete time Wright-Fisher model where the population is represented by a vector \( n_k \) with elements corresponding to the number of individuals in fitness class \( k \). Each generation consists of separate selection and mutation steps. To implement selection, the vector \( n_k \) is multiplied by \( e^{\epsilon(k-k)-a} \) to obtain a vector \( \tilde{n}_k \) containing the expected number of offspring in class \( k \). Here \( (k - \bar{k}) \) is the fitness of class \( k \) relative to the population mean, and \( \alpha = \sum_{k>N} \) maintains an approximately constant population size around \( N \).

Our model is simplified by using beneficial and deleterious mutation with effect \( \pm s \) and mutation rate \( U_e \) and \( U(1 - e) \), respectively. To implement mutation, we calculate the probability \( P(i, j) \) of a genome being hit by \( i \) beneficial and \( j \) deleterious mutations, which are Poisson distributed. This mutation matrix is then applied to \( \tilde{n}_k \), i.e. the parts of \( \tilde{n}_k \) are moved up or down according to the net number of mutations they accrued in this time step. Having constructed the expected number of individuals in fitness class \( k \) after selection and mutation, we draw a population sample from each class from a Poisson distribution. If necessary, the population is re-centered in the discrete vector \( n_k \). This prevents occupied classes from running off the grid due to accumulated increase or decrease in their absolute fitness.

The above process is repeated for a specified number of generations. The speed of adaptation and other features of the dynamics are measured after an equilibration time to remove transient effects from the initial conditions. In the parameter regimes studied, we found that \( 10^4 \) generations was generally sufficient.

To solve for \( \epsilon, (U, N, s) \), we re-run the simulation while iteratively adjusting \( \epsilon \) to get \( v \) as close to zero as possible. Since \( v \) is a stochastic quantity, \( \epsilon \) can only be determined with limited accuracy, but this can be improved by increasing the number of generations for each run. We have also run simulations in which \( \epsilon \) varies with fitness (i.e. with \( k \)) to demonstrate that the population indeed evolves towards the values of \( k \) at which they have the appropriate \( \epsilon \).

**Stationary distribution in the presence of deleterious and beneficial mutations.** The steady state distribution satisfies

\[
0 = (k - \bar{k})n_k - \lambda n_k + \lambda(1 - \epsilon)n_{k+1} + \lambda e n_{k-1}
\]

which we solve by introducing a Fourier transform

\[
\hat{h}(q) = \sum_k e^{-iqk} n_k
\]

which yields an ODE in \( \hat{h}(q) \)

\[
\frac{d}{dq} \hat{h}(q) = -i\lambda [(1 - \epsilon)e^{-iq} + e^{iq}]\hat{h}(q).
\]

The ODE is readily solved to give

\[
\ln \hat{h}(q) = \lambda \left[ (1 - \epsilon)e^{-iq} - e^{iq} \right] - \lambda(1 - 2\epsilon)
\]

where the last term is a constant added to enforce normalization \( \sum_k n_k = 1 \).

Fourier transforming back to solve for \( n_k \) gives

\[
n_k = e^{-\lambda(1-\epsilon)} \int_{-\pi}^{\pi} \frac{dq}{2\pi} \exp \left[ i\bar{q}k + \lambda(1 - \epsilon)e^{-iq} - \lambda e^{iq} \right].
\]

This can be re-written by shifting the integration contour upward into the complex plane by \( Q = \ln (1 - e^{-\epsilon})/\epsilon \), i.e. change integration variable \( q = \bar{q} - iQ \), which yields

\[
n_k = e^{-\lambda(1-2\epsilon)} \int_{-\pi}^{\pi} \frac{dq}{2\pi} \exp \left[ i\bar{q}k + \frac{\lambda(1 - \epsilon)}{2\epsilon} \bar{q}^2 - i\lambda Q \right] \]

\[
= e^{-\lambda(1-2\epsilon)} \int_{2\pi}^{\infty} \frac{dz}{2\pi i} \left( 1 - \frac{\lambda(1 - \epsilon)}{2\epsilon} z^2 + \ldots \right)
\]

where \( \alpha = \lambda \sqrt{1/e} \) and the Bessel function is identified from its integral representation [51].

While this solution holds without any restriction on parameters, it is instructive to examine the \( \epsilon \to 0 \) limit, for which we have a regular perturbative expansion in \( \epsilon \ll 1 \):

\[
n_k = e^{-\lambda(1-2\epsilon)} e^{\epsilon(1-\epsilon)/(2\pi)} \left( 1 - \frac{\lambda(1 - \epsilon)}{4\pi^2} + \ldots \right)
\]

which in the \( \epsilon \to 0 \) recovers the familiar Poisson mutation/selection balance.

**Rate of extinction of the nose in slow ratchet regime.** The rate of extinction for the nose can be calculated using the stochastic dynamics of the fittest class, which can approximately be modeled by a diffusion equation:

\[
\partial_t p(x, t) = -\partial_x \left[ D_1(x)p(x, t) \right] + \partial_x^2 \left[ D_2(x)p(x, t) \right]
\]

where \( p(x) = p(n_x/N = x) \) is the probability distribution of the nose occupancy and \( x_{\text{fit}} = e^{-\epsilon(1-e)} \) is the mean fraction of the population, \( D_1(x) = 3x(1 - x/k_n) \), and \( D_2(x) = x(1 - x)/2N \). Both \( x = 1 \) (fixation) and \( x = 0 \) (extinction) are absorbing boundary conditions [6]. The density function for fixation/extinction time \( \phi(t; x) \) satisfies the backward Kolmogorov equation

\[
\partial_t \phi(t; x) = D_1(x) \partial_t \phi(t; x) + D_2(x) \partial_x^2 \phi(t; x)
\]

where \( t \) is the time interval between initial state \( x \) and fixation or extinction. The mean time to fixation/extinction starting at \( x = y \) at \( t = 0 \) is given by \( \tilde{f}(y) = \int_0^y t\phi(t; y)dt \), which satisfies

\[
-1 = D_1(y) \partial_y \tilde{f}(y) + D_2(y) \partial_y^2 \tilde{f}(y).
\]

Using the integrating factor \( \phi(x) = e^{2\int_0^x D_2(z)/D_1(z)} \), yields for the mean time to extinction of the fittest class:

\[
\tilde{f}(x_{\text{fit}}) = \int_0^{x_{\text{fit}}} d\zeta \frac{\phi(\zeta)}{D_2(\zeta)}.
\]

The second integral in the limit \( Ns \gg 1 \) can be approximated as

\[
2N \int_y^{x_{\text{fit}}} d\zeta \left( 1 - \frac{x_{\text{fit}}}{\zeta} \right) e^{\phi(\zeta)} \approx 2N \int_y^{x_{\text{fit}}} d\zeta e^{\phi(\zeta)}\left( 1 - \frac{x_{\text{fit}}}{\zeta} \right).
\]

Using \( \alpha = Ns\bar{x}_x, n_{x_1} = N\bar{x}_x \) and changing the variables, \( \phi = y/x_{\text{fit}} - 1, z = \zeta/x_{\text{fit}} - 1 \) yields:

\[
\tilde{f}(x_{\text{fit}}) = 2n_{x_1} \int_0^{x_{\text{fit}}} d\phi e^{\phi^2} \int_0^{\sqrt{2\pi\alpha}} d\zeta e^{-\zeta^2/2}.
\]

For \( \alpha \gg 1 \) the integral the second integral can be readily approximated yielding

\[
\tilde{f}(x_{\text{fit}}) \approx 2n_{x_1} \sqrt{\frac{\pi}{4\alpha}} e^{-\alpha/4}\left[ \text{erf}(\sqrt{\alpha}\phi) - \text{erf}(\sqrt{\alpha}\phi) \right]
\]
where \( \beta = \frac{1}{\Gamma} - 1 \). Using \( \phi^2 = 1 - \theta^2 / \alpha \) allows the evaluation of the integral in the \( \alpha \gg 1 \) limit and yields

\[
\tilde{f}(x) \approx n_{\alpha} \sqrt{\pi} \alpha^{-3/2} e^x. \tag{25}
\]

Assuming an exponential distribution for time of extinction of the fittest class, the rate of extinction is given by \( r = 1/\tilde{f}(x) \), which is used in the Main Text.

**Asymptotic expressions for \( \epsilon_c \)** The matching condition used to determine \( \epsilon_c \) in the fast-ratchet regime can be rewritten as follows.

\[
\frac{1}{\sigma \sqrt{\epsilon_k}} = Ne^{-\lambda(1-2\epsilon)} \frac{1}{2} \log \frac{1-\epsilon}{\epsilon} J_0(2\epsilon). \tag{26}
\]

In the limit of \( \alpha \gg 1 \), the zero of the Bessel function is approximately at \( k_{\ast} + 1 \approx \alpha \) and the Bessel function at \( k_{\ast} \) evaluates to roughly \( \sim k^{2/3} \) (Abramowitz 9.3.33 and 9.1.27). Hence we have

\[
\lambda(1-2\epsilon) - \lambda \sqrt{\epsilon(1-\epsilon)} \log \frac{1-\epsilon}{\epsilon} = \log N s
\]

where we have neglected powers to the \( 1/6 \) and \( O(1) \) factors inside the logarithm. In the limit \( \epsilon \ll 1 \) (but \( \lambda^2 \epsilon \gg 1 \)), the matching simplifies to

\[
1 - 2\epsilon + \sqrt{\epsilon} \log \epsilon = \lambda^{-1} \log N s
\]

which simplifies further to

\[
\sqrt{\epsilon} \log \epsilon \approx \lambda^{-1} \log(N s) - 1 = -2z
\]

where we defined \( z \) for convenience. This can be solved for \( \epsilon_c \):

\[
\epsilon_c = \frac{z^2}{W(-z)^2} \approx \frac{z^2}{\log(z) - \log(-\log(z))^2} \tag{30}
\]

where \( W(z) \) is the \( -1 \) branch of Lambert’s \( W \)-function, i.e., the solutions of \( W(z)e^{W(z)} = z \). The linear correction in Eq. 28 can be incorporated iteratively.

\[
\epsilon_{i+1} = \frac{(z + \epsilon_i)^2}{W(-z + \epsilon_i)^2}. \tag{31}
\]

This iteration converges for small \( C \) and \( \epsilon \). At larger \( C \), the branch of the \( W(z) \) function is lost. The result \( \epsilon_2 \) obtained after the first iteration, starting with Eq. 30, is also shown in Fig. 6 as dashed red line.

The other limit that is amenable to analytic calculations is the limit \( \epsilon \to 1/2 \). To this end, we define \( \delta = \frac{1}{2} - \epsilon \) and expand the right hand side of Eq. 28

\[
1 - 2\epsilon - \sqrt{\epsilon(1-\epsilon)} \log \frac{1-\epsilon}{\epsilon} = 4 \delta^3 + \frac{44}{15} \delta^5 + O(\delta^7) \equiv \lambda^{-1} \log N s. \tag{32}
\]

From this, we find

\[
\epsilon \approx \frac{1}{2} - \left( \frac{3}{4\lambda} \log(N s) \right)^{1/3}. \tag{33}
\]

This expression is compared to simulation results in Fig. 6. The term proportional to \( \delta^3 \) can again be included by iteration and the result is shown as the dashed green line in Fig. 6.

---

**Fig. 6.** Dynamic mutation-selection balance. The figure shows \( \epsilon_c \) for many combinations of \( U, s \) and \( N \) as a function of \( \frac{1}{\lambda} \log N s \), while the color codes for \( \epsilon_c U^2 s^{-2} \). If \( \epsilon_c U^2 s^{-2} \gg 1 \), \( \epsilon_c \) is solely a function of \( \frac{1}{\lambda} \log N s \) and is well described by the numerical solution of Eq. 6 shown as a black line. The asymptotic approximations for large \( \epsilon \) (Eq. 33) and small \( \epsilon \) (Eq. 30) are shown as green and red lines, respectively. The dashed lines correspond to the more accurate version mentioned in the main text.