Natural selection can favor the evolution of ratchet robustness over evolution of mutational robustness

Yinghong Lan¹*, Aaron Trout², Daniel M Weinreich¹, C Scott Wylie¹

1 Department of Ecology and Evolutionary Biology, and Center for Computational Molecular Biology, Brown University, Providence, RI, USA
2 Department of Mathematics, Chatham University, Pittsburgh, PA, USA

* yinghong_lan@brown.edu

Abstract

The vast majority of mutations are deleterious. How natural populations evolve to cope is a question of fundamental interest. Previous studies have reported the evolution of mutational robustness, that is, natural selection favoring mutations that reduce effects of deleterious mutations elsewhere in the genome. Here, we demonstrate that finite, asexual populations’ ability to purge recurrent deleterious mutations declines with increased mutational robustness. Consequently, higher mutational robustness leads to higher risk of extinction by Muller’s ratchet. We therefore hypothesize that in the long run, natural populations may instead evolve robustness against Muller’s ratchet by increasing sensitivity to deleterious mutations, despite the short-term fitness consequences. We call this phenomenon “ratchet robustness”. Using individual-based simulations, we first confirm that ratchet robustness is inversely correlated with mutational robustness on fitness landscapes without epistasis. Next, we demonstrate that negative epistasis increases ratchet robustness, precisely because on fitness landscapes with negative epistasis sensitivity to deleterious mutations increases as mutations accumulate. We also show that on a fitness landscape with adjacent fitness peaks exhibiting exclusively positive and negative epistasis, the populations will converge on the latter. On the other hand, introducing even a vanishingly small region
of positive epistasis on a single-peaked fitness landscape that otherwise exhibits negative epistasis is enough to entirely ablate ratchet robustness, exposing the population to Muller’s ratchet. We conclude that while regions of the fitness landscape with higher mutational robustness could be favored in the short term, purely because of temporary fitness advantage, in the long run mutational robustness may render a population vulnerable to extinction. Some empirical support exists for this prediction.

Introduction

The fitness landscape maps genotype to fitness, and describes fitness effects of all possible mutations. In doing so, it determines the extent to which populations can sustain their fitness when perturbed by mutations, namely their mutational robustness. Mutational robustness has been viewed as equivalent to neutrality, that is, mutational robustness means that mutations have little to no deleterious effects \( [1,2] \). In other words, mutational robustness is manifest as comparatively flat fitness landscapes. Previous studies have concluded that under the pressure of high mutation rate, natural selection will cause populations to evolve mutational robustness. Equivalently, it has been proposed that populations that experience high mutation rates will evolve to occupy flatter regions of the fitness landscape \( [3,4] \). However, finite asexual populations on flatter landscapes are also more vulnerable to extinction by Muller’s ratchet \( [5] \). This suggests the possibility that natural selection might also favor lineages that reside in regions of the fitness landscape that protect them from such extinction, a property we here designate “ratchet robustness”. To our knowledge, this tension between the evolution of mutational robustness and ratchet robustness has never been explored \( [3,4,6–9] \).

In infinite populations, where drift can be neglected, mean fitness depends only on mutation rate and is independent of fitness landscape \( [10,11] \). Consequently, equilibrium fitness of infinite populations is insensitive to mutations that affect mutational robustness. On the other hand, finite asexual populations risk Muller’s ratchet \( [12,13] \), because even the fittest genotype can be lost to genetic drift. Whenever that happens, the ratchet clicks. Even in the presence of beneficial mutations, Muller’s ratchet can still happen if deleterious mutations are sufficiently common.
Indeed, only if the proportion of beneficial mutations is high enough, can finite asexual populations halt the ratchet and reach Mutation-Selection-Drift Equilibrium (MSDE) ([5,14–16]).

Here, we confirm first that other things being equal, populations on locally steeper regions of the fitness landscape exhibit more ratchet robustness. Next, we show that negative pairwise epistasis increases ratchet robustness (as previously seen in a model lacking beneficial mutations [17]), while positive epistasis reduces it and introduces a novel, accelerating form of Muller’s ratchet. As a result, populations converge to regions of the landscape with pure negative epistasis. Finally, we demonstrate that even the smallest amount of positive epistasis on a fitness landscape otherwise dominated by negative epistasis is enough to ablate a population’s ratchet robustness. We conclude that while in the short term mutational robustness can be selected for, in the long run natural selection sacrifices mutational robustness in favor of ratchet robustness.

Methods

Evolutionary model

For all simulations, we implemented discrete-time Wright-Fisher evolutionary model with custom Python code. Within populations of size $N$, each individual’s genotype is solely identified by the number of deleterious mutations it has, which is denoted by $i$, $i = 0, 1, ..., L$. $L$ represents the genome length, which is also the maximum number of deleterious mutations possible. Populations “live in” vectors $V$ of length $L + 1$, where each bucket $V_i$ records the number of individuals with $i$ deleterious mutations. Certainly, $\sum_i V_i = N$. Individuals with $i$ mutations have fitness denoted by

$$W_i = e^{-si^{1-\epsilon}}$$

with selection coefficient $s$ and epistasis parameter $\epsilon$. $\epsilon = 0$ in the absence of epistasis (Fig 1), $\epsilon < 0$ with negative epistasis (Fig 2A&B: $\epsilon = -0.25$), $\epsilon > 0$ with positive epistasis (Fig 2C&D: $\epsilon = 0.25$).

Each generation of evolution starts with reproduction with selection. Number of offspring with $i$ mutations is a random variable proportional to $V_i$, i.e., number of
parents carrying \(i\) mutations, times \(W_i\), i.e., fitness of these parents. Then, mutations are imposed by sampling from Poisson distribution with parameter \(U_{del}\) to generate number of deleterious mutations, and from Poisson distribution with parameter \(U_{ben} = 0.01U_{del}\) to generate number of beneficial mutations. Eventually, a new vector \(V'\) is “born”, where \(V'_i\) records number of individuals with \(i\) mutations at the new generation. Simulations are performed for a prespecified number of generations, usually 10,000, to ensure the result is unaffected by transient effects. Simulations in Fig 1 and Fig 2 are performed on simple fitness landscapes with \(L = 300\), landscapes that have only one peak (as opposed to two in Fig 3) and the only peak is isotropic (as opposed to hybrid peak in Fig 4).

In Fig 1B, we presented variance of fitness across time for simulations conducted in Fig 1A. Notably, critical \(U_{del}\), i.e., the highest \(U_{del}\) under which populations could resist Muller’s ratchet, or the lowest under which populations start to succumb to Muller’s ratchet, seems to overlap with the \(U_{del}\) under which time variance of fitness reaches maximum. This is expected, as when \(U_{del}\) is below or above critical \(U_{del}\), populations are either at MSDE or extinction, both showing minimal time variance of fitness. At critical \(U_{del}\), the “tug-of-war” between the two possible state shows maximum stochasticity.

We utilized the above observation to locate critical \(U_{del}\) in Fig 1C, across different population size \(N\) and selection coefficient \(s\). For each \(N\) and \(s\) combination, we recorded time variance of fitness under different \(U_{del}\), and reported the \(U_{del}\) associated with highest time variance as critical \(U_{del}\). There are three sources of uncertainty during our analysis. First, it’s impossible to conduct simulations under every possible \(U_{del}\) value. In practice, we sampled \(U_{del}\) intevalled by 0.025, meaning that the “true” critical \(U_{del}\) could fall in between two sampled \(U_{del}\) values. Second, because of the intrinsic granularity, detected maximum time variance may not be the true maximum. We measure such uncertainty as the range of \(U_{del}\) values that show time variance above half of the detected maximum. Third, there is inevitable variation in recorded time variance across replicated simulations, due to the stochastic nature of the simulation. We found that the second source of variation dominates the other two by at least one order of magnitude. Therefore, we portrayed only the second type of uncertainty as error bars in Fig 1C.
Multi-peak fitness landscape and hybrid peaks

For simulations in Fig 3 as shown in Fig 3A, we fused two isotropic peaks together: on the positive epistasis side, $\epsilon = 0.25$ and $L = 270$; on the negative epistasis side, $\epsilon = -0.25$ and $L = 29$. Mutation operations are applied with beneficial mutations always going towards the peak at the same side of the valley, and deleterious mutations always going towards the valley.

For simulations in Fig 4 as shown in Fig 4A, fraction $p$ of the first mutations puts individuals in the positive epistasis domain, where $\epsilon = 0.25$ and $L = 300$, while the rest $1 - p$ puts individuals at the negative epistasis domain where $\epsilon = -0.25$ and $L = 300$. If an individual is brought back to the peak by beneficial mutations, the same $p$ and $1 - p$ apply, i.e., individuals can “travel” between the two domains through the peak.

Results

In this study, we used individual-based simulations to evolve asexual finite populations on four kinds of fitness landscapes with increasing complexity, in order to study how fitness landscape affects mutational and ratchet robustness. On all fitness landscapes, we implement a constant proportion of beneficial mutation rate relative to deleterious mutation rate ($U_{\text{ben}}/U_{\text{del}} = 0.01$, where mutation rate represents expected number of mutations per genome per duplication, see Methods). This is equivalent to assuming that the number of loci carrying deleterious alleles remains small compared to the total genome size. (Of course, changing $U_{\text{ben}}/U_{\text{del}}$ also influences ratchet robustness, [5], and see Discussion.) We define a fitness peak as isotropic if the fitness of any genotype only depends on its Hamming distance from the peak, i.e., on number of deleterious mutations, but not their identity (Eq 1). We first examined evolutionary behavior on fitness landscapes with a single isotropic peak in absence of epistasis (Eq 1 with $\epsilon = 0$, Fig 1A inset). Next, we introduced a pairwise epistatic term to the isotropic peak (Eq 1, $\epsilon \neq 0$), distinguishing between negative ($\epsilon < 0$, Fig 2A) and positive ($\epsilon > 0$, Fig 2C) epistasis. We then examined evolutionary behavior on fitness landscapes consisting of mutationally adjacent isotropic peaks with negative and positive epistasis (Fig 3A). Finally, we relaxed our assumption of isotropic peaks and modeled populations evolving on hybrid fitness peaks where a fraction ($p$) of all first deleterious mutations leaving the
peak lead to domains of positive epistasis, while the rest \((1 - p)\) lead to domains of negative epistasis (Fig 4A).

Populations residing on flatter fitness landscapes exhibit reduced ratchet robustness
Fig 1. In the absence of epistasis, populations residing on flatter fitness landscapes exhibit reduced ratchet robustness. A. Equilibrium fitness of finite populations under different mutation rates ($U_{del}$) on fitness landscapes with different selection coefficients ($s$). Dashed black line: theoretical prediction for equilibrium fitness in infinite populations ($e^{-U_{del}}$). Steeper fitness landscapes (blue), i.e., ones with larger $s$, maintain MSDE under higher $U_{del}$. Beneficial mutation rate $U_{ben} = 0.01U_{del}$. Population size $N = 1000$. Average fitness is recorded after evolving for 10,000 generations (see Methods). Error bars: standard deviation over 50 replicates. Inset: comparison of the two fitness landscapes, each of which is composed of one isotropic peak without epistasis. B. Variance of fitness over time for the same populations as in A. Variance peaks around critical $U_{del}$, i.e., the largest $U_{del}$ under which populations maintain MSDE. Error bars: standard deviation over 50 replicates. All parameters as in Panel A. C. Critical $U_{del}$ values from simulations (points) and analysis (lines, derived numerically from Eq 7 in [5]), for landscapes with different selection coefficients in the absence of epistasis. Error bars: uncertainty associated with locating critical $U_{del}$ from simulation results (see Methods). Regardless of population size, steeper landscapes always protect populations from Muller’s ratchet under higher $U_{del}$.

We start by studying whether populations on flatter fitness landscapes, i.e., those with greater mutational robustness ([4]), are more resilient to Muller’s ratchet than ones on steeper landscapes. We constructed two fitness landscapes, each composed of one isotropic peak but with different steepness (selection coefficients $s$) in the absence of epistasis (Fig 1A inset). We simulated the evolution of finite asexual populations on such landscapes under different deleterious mutation rates (recall $U_{del} = 0.99U$), and recorded their fitness at equilibrium (Fig 1A).

Under low $U_{del}$, selection is more than sufficient to purify deleterious mutations, and drift (and thus, the ratchet) is negligible on both flat and steep landscapes. This is because loss of fittest class is extremely rare: purifying selection is strong relative to mutation. Moreover, the fittest class is quickly restored by beneficial mutations if lost. Consequently, equilibrium fitness lies close to the well-known infinite population expectation $w_{max}e^{-U_{del}}$, where $w_{max}$ is fitness at the peak and is set to 1 throughout the paper ([10][11], black dashed line in Fig 1A), which depends only on $U_{del}$. However, after $U_{del}$ exceeds some critical value (Fig 1A: $U_{del} \approx 0.4$ for $s = 0.1$ and $U_{del} \approx 1.9$ for $s = 0.4$), rapidly accumulating deleterious mutations overwhelm selection, leading to extinction via Muller’s ratchet. We name the highest $U_{del}$ under which populations are able to resist Muller’s ratchet “critical $U_{del}$”. Fixing population size ($N$) and the proportion of beneficial mutations, populations on steeper landscapes demonstrate higher critical $U_{del}$ and thus are able to resist Muller’s ratchet under higher $U_{del}$.
(Fig 1A&C). In other words, although populations on flatter landscapes have higher mutational robustness ([1,2]), they fail to maintain MSDE under higher $U_{del}$ and thus exhibit lower ratchet robustness.

**Populations on fitness landscapes with negative epistasis have higher ratchet robustness than ones with positive epistasis**
Fig 2. Fitness landscapes with negative epistasis have higher ratchet robustness than ones with positive epistasis. A. Cartoon of evolutionary dynamics of populations on fitness landscape composed of one isotropic peak with negative epistasis. Curved dashed line: fitness landscape. Straight tangent dashed line: fitness landscape without epistasis that shares the same selection coefficient at the peak. Note that selection coefficient (the slope of the fitness landscape) increases with Hamming distance from the peak. Golden vertical line and horizontal dashed line: $e^{-U_{del}}$. B. Time course in average Hamming distance in 35 populations initialized at random points on isotropic peak with negative epistasis during 10,000 generations ($N = 1000, s = 0.2, U_{del} = 1.0, U_{ben} = 0.01U_{del}, \epsilon = -0.25$). Golden vertical line: same as in A. All populations converge to the attractor wherever they are initiated (green traces). C. Cartoon of evolutionary dynamics of populations on fitness landscape composed of one isotropic peak with positive epistasis. Curved dashed line: fitness landscape. Straight tangent dashed line: fitness landscape without epistasis that shares the same selection coefficient at the peak. Note that selection coefficient (equal to the slope of the fitness landscape) decreases with Hamming distance from the peak. Cyan vertical line and horizontal dashed line: $e^{-U_{del}}$. Violet vertical line: predicted point of no return (derived numerically from Eq 7 in [5]). D. Time course in average Hamming distance in 100 populations initialized at random points on isotropic peak with positive epistasis during 10,000 generations ($N = 1000, s = 0.5, U_{del} = 0.5, U_{ben} = 0.01U_{del}, \epsilon = 0.25$). Cyan and violet vertical lines: same as in C. Populations initiated above the point of no return tend to evolve to the peak (blue traces), whereas ones initialized below it tend to succumb to Muller’s ratchet (pink traces). Realizations that fluctuate across the point of no return are colored in brighter blue and pink.

To understand how epistasis affects ratchet robustness, we added pairwise epistasis to a single isotropic peak and studied two simple cases: peaks with only negative epistasis (Fig 2A), and ones with only positive epistasis (Fig 2C). Note that when epistasis is negative (positive), the local strength of purifying selection increases (decreases) with Hamming distance to the peak, and consequently deleterious mutations are less (more) likely to accumulate. Since Hamming distance itself increases with mutation rate, this suggests that a population evolving on an isotropic fitness peak with negative epistasis might enjoy heightened ratchet robustness as a consequence of negative feedback between mutation rate and the tendency to accumulate deleterious mutations (Fig 2A). Conversely, a population evolving on an isotropic fitness peak with positive epistasis might be particularly susceptible to Muller’s ratchet as a consequence of positive feedback between mutation rate and the tendency to accumulate deleterious mutations (Fig 2C).

Consistent with this intuition, we find that on isotropic peaks with negative epistasis, selection will drive the population back to the peak (Fig 2B) regardless of where on the landscape it is initialized, if $U_{del}$ is below critical $U_{del}$ at the peak (i.e., the highest $U_{del}$...
at which the peak could be sustained by the population). On the other hand, even if
$U_{del}$ is above the critical $U_{del}$ at the peak, there exists a point on the landscape where
selection exactly offsets such $U_{del}$, because purifying selection increases monotonically
from the peak. If a population is initialized below this point, selection locally will be
strong enough to push populations upward until this point is reached. If, instead, a
population is initialized above this point, mutation will be strong enough to push the
population downward to this point. This point thus represents an attractor (5,18), a
stable equilibrium that must be achieved regardless of starting point on the landscape.

Conversely, on isotropic peaks with positive epistasis, if $U_{del}$ is below critical $U_{del}$ at
the peak, populations initiated at the peak can maintain MSDE. However, because
purifying selection decreases monotonically from the peak, there exists a point on the
landscape where selection exactly offsets such $U_{del}$. Populations initiated above this
point would benefit from selection stronger than required and adapt to the peak, while
ones initiated below this point suffer from selection weaker than needed, and succumb
to Muller’s ratchet (Fig 2D). Therefore, we refer to this point as “point of no return”.

Importantly, even if populations have equilibrated around the peak, stochastic
fluctuations will eventually take them across this point. Moreover, as $U_{del}$ increases,
selection required to oppose mutation naturally increases as well. Correspondingly, as
$U_{del}$ increases, the point of no return migrates towards the peak. This imposes a greater
danger of succumbing to Muller’s ratchet for populations in the vicinity of the peak via
stochastic fluctuations. Finally, once $U_{del}$ is above the critical $U_{del}$ at the peak, the
point of no return overlaps with the peak and mutation overwhelms selection
everywhere on the fitness landscape.

**Populations converge to peak with negative epistasis on
multi-peak fitness landscape**
A

![Graph showing fitness landscape and point of no return](image)

- **Logarithm of Fitness**
- **Hamming Distance From Corresponding Peak**

**Peak of Positive Epistasis Side**

**Peak of Negative Epistasis Side**

**Point of No Return**

B

![Generations vs. Hamming Distance](image)

- **Generations**
- **Hamming Distance From Corresponding Peak**
Fig 3. Populations converge to peak with negative epistasis on multi-peak fitness landscape. A. Cartoon of evolutionary dynamics of populations on multi-peak fitness landscape. Both peaks are isotropic, one with only positive epistasis (left) and the other with only negative epistasis (right). Curved dashed lines on both sides: fitness landscape. Straight tangent dashed line on both sides: slope corresponding to selection coefficient at peaks on both sides. Horizontal dashed line: $e^{-U_{del}}$ on both sides. Cyan vertical line: $e^{-U_{del}}$ on the positive epistasis side. Violet vertical line: numerically derived point of no return on the positive epistasis side (Eq 7 in [5]). Black vertical line: valley of the landscape. Golden vertical line: $e^{-U_{del}}$ on the negative epistasis side. B. Time course in average Hamming distance in 100 populations initiated at random points on multi-peak landscape during 10,000 generations ($N = 10,000$, $s = 0.3$, $U_{del} = 0.5$, $U_{ben} = 0.01U_{del}$, positive epistasis side $\epsilon = 0.25$, negative epistasis side $\epsilon = -0.25$). Cyan, violet, black and golden vertical lines: same as in B. Populations are color coded based on their starting point (blue traces: above the point of no return on the positive epistasis side, pink traces: below the point of no return on the positive epistasis side, bright pink and blue traces: realizations that fluctuate across the point of no return on the positive epistasis side, green traces: populations initialized on the negative epistasis side). The vast majority of populations below the point of no return rapidly cross the valley and converge to the attractor on the negative epistasis side. Due to uniformly stronger selection on the negative epistasis side of the valley, populations there exhibit less stochasticity in their trajectories.

To validate the observation that regions of the landscape with negative epistasis have intrinsically higher ratchet robustness, and to demonstrate that populations could evolve ratchet robustness via occupying such regions, we constructed a fitness landscape composed of two mutationally adjacent isotropic peaks featuring opposite signs of epistasis (Fig 3A). Populations finding themselves below the point of no return on the positive epistasis side will initially decline in fitness and come to the valley, similar to populations declining to the bottom of the landscape in (Fig 2D). However, at this point, strongly beneficial mutations become available, drawing populations onto the negative epistasis side of the valley, after which they quickly climb to the attractor (Fig 3B). (Note that $U_{ben} = 0.01U_{del}$ on both sides of the valley, and consequently this behavior is driven entirely by natural selection.) Moreover, populations above the point of no return on the positive epistasis side will nevertheless experience stochastic fluctuations (Fig 2D). Eventually, they will be carried over the point of no return, at which point they will experience selection lower than required to offset current mutation rate and decline to the bottom of the landscape due to Muller’s ratchet, followed by convergence to the negative epistasis side. Note that the two peaks share identical selection at the peak, meaning that the positive epistasis side has uniformly higher mutational robustness but uniformly lower ratchet robustness.
Mutational robustness and ratchet robustness on hybrid peaks

A

![Graph A - Hamming Distance vs Logarithm of Fitness](image)

- Hamming Distance
- Logarithm of Fitness

B

![Graph B - Equilibrium Percentages](image)

- Equilibrium Percentages
- Deleterious Mutation Rate
- negative epistasis
- peak
- positive epistasis

C

![Graph C - Equilibrium Percentages](image)

- Equilibrium Percentages
- Deleterious Mutation Rate
- negative epistasis
- peak
- positive epistasis
Fig 4. Mutational robustness and ratchet robustness on hybrid peaks. A. Hybrid peaks: while all first mutations leaving the peak share the same fitness effect, a fraction $p$ of them cause subsequent mutations to exhibit positive pairwise epistasis, and the remaining $1 - p$ cause subsequent mutations to exhibit negative epistasis. Populations are always initiated at the peak. Here $p = 0.5$. B. Equilibrium proportions of individuals at the peak (black), the negative epistasis region (blue), and the positive epistasis region (pink), under different $U_{del}$ ($U_{ben} = 0.01U_{del}$, $p = 0.5$, $N = 1000$, $s = 0.35$, negative epistasis region: $\epsilon = -0.25$, positive epistasis region: $\epsilon = 0.25$, evolved for 10,000 generations, error bars: standard deviation across 50 replicates). Under $U_{del}$ less than critical $U_{del}$ at the peak (here $\sim 1.1$; see main text), the subpopulation on the negative epistasis region exists in mutation-selection balance and relies on continual mutational input from subpopulation on the peak. After $U_{del}$ exceeds critical $U_{del}$ at the peak, the subpopulation on the peak goes extinct, and with it, the subpopulation on the negative epistasis region. At this point, the remaining population finds itself beyond the point of no return on the positive epistasis region (see main text), and it succumbs to Muller’s ratchet. C. Equilibrium proportions of individuals at the peak (black), the negative epistasis region (blue), and the positive epistasis region (pink), under different $U_{del}$ ($U_{ben} = 0.01U_{del}$, $p = 0.01$, $N = 1000$, $s = 0.35$, negative epistasis region: $\epsilon = -0.25$, positive epistasis region: $\epsilon = 0.25$, evolved for 10,000 generations, error bars: standard deviation across 50 replicates). Population dynamics resemble $p = 0.5$ (see main text).

Biologically realistic landscape peaks are unlikely to be isotropic: fitness effects of mutations at any Hamming distance usually depend on current genome background and follow complex distributions. As a first attempt to capture part of the reality on our model landscape with one peak, we now allow the sign of epistatic effects of mutations to be dependent on the first mutation away from the peak. Specifically, a certain fraction ($p$) of the first mutations now place the evolving population on a region of the landscape exhibiting positive epistasis, while the rest ($1 - p$) place the evolving population on a region of the landscape exhibiting negative epistasis (Fig 4A).

Concretely, among all the paths leaving from the peak, $p$ of them show positive epistasis, while the rest $1 - p$ show negative epistasis. We assume that all first mutations share identical fitness effects, so that there is no immediate fitness advantage in choosing the region with negative or positive epistasis. However, for all Hamming distances greater than one, fitness is necessarily higher in regions of the landscape with positive epistasis than in regions with negative epistasis (Eq 1). In other words, regions of the landscape with positive epistasis have uniformly higher fitness and mutational robustness, but lower ratchet robustness compared with ones with negative epistasis. Note that the only mutational path between regions of this landscape exhibiting positive and negative
epistasis is through the peak. We evolved populations on such hybrid peak with $p = 0.5$ and report the proportions of populations residing exactly at the peak, in the negative epistasis region, and in the positive epistasis region as a function of $U_{del}$ (Fig 4B).

When $U_{del}$ is so low that majority of the individuals carry zero or one mutation (here $U_{del} \approx 0.2$, Fig 4B, S2 FigA), the proportion of the population on the negative epistasis region is very close to that on the positive epistasis region. This merely reflects the fact that $p$ equals 0.5, since the fitness cost of the first mutation is the same. However, as $U_{del}$ rises to moderate level (here $\sim 0.2 < U_{del} < \sim 1.1$, Fig 4B, S2 FigB), the proportion of the population on the negative epistasis region begins to drop below $p = 0.5$. This reflects selective enrichment for the subpopulation experiencing positive epistasis: all mutations after the first are always less deleterious on the positive epistasis region than on the negative epistasis region. Nevertheless, in this intermediate range of values of $U_{del}$, a subpopulation on the negative epistasis region is still sustained despite lower fitness, thanks to net mutational inflow from the subpopulation at the peak (S1 Table). In essence, the subpopulation on the negative epistasis region is at mutation-selection balance within the population: constantly being purified by selection but being regenerated by mutation from the peak.

However, after $U_{del}$ increases to so high that the peak can no longer be sustained (here, $U_{del} \approx 1.1$), the proportion of the population at the peak becomes negligible (Fig 4B, S2 FigC). As a result, the negative epistasis region is disconnected from mutation input from the peak and is quickly wiped out by selection. The remaining population now occupies only the positive epistasis region. However, since $U_{del}$ has overwhelmed selection at the peak, it necessarily does so also at every other point on the positive epistasis region and the population quickly succumbs to Muller’s ratchet. This threshold recapitulates results seen when the point of no return overlaps with the peak for the positive epistasis region (Fig 2C&D, although the numeric value of the critical mutation rate differs here, reflecting its dependence on the contours of the fitness landscape).

More importantly, the observed pattern does not depend on the particular value $p = 0.5$: even when there is only a very small fraction of paths leaving the peak with positive epistasis, subpopulations on the positive epistasis region of the landscape will always be favored due to short term fitness advantage. Such advantage is amplified by
higher $U_{del}$, so long as it remains less than the critical $U_{del}$ at the peak, i.e., the $U_{del}$ under which the peak is lost. In this regime, the two subpopulations accumulate more mutations, reach larger Hamming distances from the peak, and thus experience increased fitness differences. Eventually, populations on the positive epistasis region dominate. Indeed, this phenomenon apparently occurs with much smaller $p$ ($p = 0.01$, Fig 4C). At low $U_{del}$, a fraction approximately equal to $p$ of the population not on the peak resides in the positive epistasis region. But as $U_{del}$ increases, this fraction increases, again because fitness is higher there. And as we observe when $p = 0.5$ (Fig 3), once $U_{del}$ exceeds the critical mutation rate at the peak, the mutational connection between subpopulations on the fitness landscape is extinguished. At this point both subpopulations are doomed. The (lower-fitness) subpopulation on the region of the landscape with negative epistasis will lose its mutational input and go extinct. At the same time, because $U_{del}$ exceeds the critical mutation rate at the peak, it’s certainly higher than critical mutation rate everywhere else on the positive epistasis region, meaning that the point of no return has reached the peak and the remaining (higher-fitness) subpopulation on the region with positive epistasis will necessarily succumb to Muller’s ratchet.

**Discussion**

Our findings suggest that although mutational robustness may be favored in the short term, resilience against Muller’s ratchet, i.e., ratchet robustness, can evolve in the long term. We showed first that in the absence of epistasis, fitness landscapes with higher mutational robustness are more susceptible to Muller’s ratchet, meaning that they offer lower ratchet robustness, and vice versa (Fig 1). We next demonstrated that landscapes with negative epistasis provide higher ratchet robustness, while landscapes with positive epistasis are intrinsically unstable (Fig 2 and 3). Finally, while mutational robustness may be selected for in the short term, this can lead to population extinction in the long term (Fig 4).
Negative feedback in mutation or selection can confer ratchet robustness

In the presence of negative epistasis, decreasing fitness leads to increasing selection strength, which halts fitness decline (Fig 2A). Such negative feedback therefore protects the population from Muller’s ratchet, thereby providing superior ratchet robustness relative to a landscape with positive epistasis (see Results). As previously noted (5), similar negative feedback could also be achieved through increasing the ratio of beneficial mutation rate to deleterious mutation rate ($U_{ben}/U_{del}$) with decreasing fitness. While negative epistasis enables more effective purifying selection, increasing $U_{ben}/U_{del}$ reduces occurrence of deleterious mutations. In other words, while negative epistasis operates at the selection level, increasing $U_{ben}/U_{del}$ functions at the level of mutational input.

Natural populations will equilibrate on fitness landscapes with sufficient ratchet robustness

Results in Fig 1 strongly suggest that natural populations cannot survive on isotropic fitness peaks with arbitrarily high mutational robustness, or equivalently, insufficient ratchet robustness. To our knowledge, this cost of mutational robustness has not previously been articulated. Furthermore, results in Fig 2 and Fig 3 indicate that natural populations cannot permanently survive on isotropic peaks with positive epistasis, since here ratchet robustness decreases with mutation rate. However, results in Fig 4 show that mutational robustness provides short term fitness advantage, since, by definition, increased mutational robustness means that mutations have less deleterious effect (1). This effect results in populations becoming more susceptible to Muller’s ratchet. Importantly, even when mutational robustness is only available on small fraction of the fitness landscape, populations are still blind to the long term perils of mutational robustness (Fig 4C).

For simplicity, in Fig 4 we constructed fitness landscape with two domains, one of which only has negative epistasis, the other positive epistasis. However, fitness landscapes in reality are highly unlikely to be composed of a few domains of distinct mutational robustness (19). Nevertheless, our conclusion that populations will favor
the short-term advantage of mutational robustness in spite of the long-term hazard of Muller’s ratchet may apply to the local mutational neighborhood in which a biological population finds itself on the landscape. Specifically, we predict that extant populations should find themselves on regions of the fitness landscape on which the fitness (and hence, mutational robustness) is maximized, subject to the constraint of sufficient ratchet robustness as determined by the organism’s mutation rate.

**Widespread empirical observations of mutational robustness do not necessarily demonstrate selection for mutational robustness**

We find that mutational robustness is unlikely be selected for in the long term at the expense of ratchet robustness. However, mutational robustness is seen at many levels of biological systems ([20]). For example, it has been observed that many proteins are tolerant of single mutations, a finding taken as evidence for selection for mutational robustness (e.g., [21]). We note however that the existence of mutational robustness need not imply selection for mutational robustness ([22],[23]). Following others ([7]), we suggest instead that mutational robustness may often evolve as a correlated consequence of selection for environmental robustness, i.e. an organism’s ability to sustain fitness against environmental perturbations. These perturbations can be external, such as temperature or rainfall variation, or internal, such as thermal noises of microenvironments inside an organism. In the presence of environmental noise, expected reproductive success should be measured as average fitness under different environments weighted by the probabilities of each environments appearing, while also considering relative timescale between generation and environmental change ([24]). In most cases of fluctuating environments, lineages with higher environmental robustness have higher overall reproductive success, and environmental robustness will be selected for ([25]). In fact, diverse mechanisms of achieving environmental robustness have evolved in response to various forms of environmental perturbations. While an exhaustive survey of this work is outside the scope of this study, the interested reader is directed to [20].

Instead, our focus here is on the relationship between environmental and mutational robustness. Theoretically and empirically, environmental robustness has been shown to give rise to mutational robustness ([7],[26],[27]). For example, RNA molecules that can
sustain their secondary structure despite thermal noises also show mutational robustness \((28–30)\). And proteins that evolved to be robust against transcription errors can also tolerate deleterious mutations \((31)\). Consequently, selection for environmental robustness can give rise to mutational robustness in nature, even if selection is unlikely to favor mutational robustness \textit{per se}.

Why can natural selection favor the evolution of environmental but not mutational robustness? The key distinction is that mutational robustness requires tolerating heritable perturbations, which inevitably alters the “starting point” of future generations. Such heritable decay is intrinsic to Muller’s ratchet. By contrast, selection for environmental robustness entails non-heritable environmental perturbation. Consequently, the short term advantage of environmental robustness is not offset by any long term cost, accounting for the absence of an “environmental ratchet”. In summary, while mutational robustness may be widespread in nature, we suggest an alternative interpretation for its evolution: namely as a correlated consequence of selection for environmental robustness \((7)\).

**Previous theoretical studies on the evolution of mutational robustness**

Interestingly, a few previous theoretical studies have uncovered the long-term cost of mutational robustness \((1,32,33)\), using different methods from ours. However, while we interpret empirical evidence of mutational robustness as reflecting selection for environmental robustness, some theoretical studies have suggested the possibility of selection for mutational robustness. Among these, the most well-known are work by van Nimwegen et al. using neutral networks \((3)\) and work by Wilke et al. using the computational model of evolution by natural selection called Avida \((4)\). We reconcile our conclusions with these previous studies next.

In the neutral network study of van Nimwegen et al. \((3)\), the fitness landscape consists of a subset of genotypes sharing identical non-zero fitness, while all other genotypes are inviable. The authors represent the viable subset of genotypes by a connected graph \(G\), where each vertex corresponds to a genotype and two vertices are connected by an edge when they can be reached from each other via a single mutation.
In other words, all edges within $G$ represent neutral mutations; hence $G$ is a neutral network. Mutational robustness of each genotype is proportional to the number of mutational neighbors on $G$ (the degree of corresponding vertex), since higher degree means higher proportion of neutral mutations. van Nimwegen et al. ([3]) prove that populations where $N\mu \ll 1$ (where $N$ is population size, $\mu$ is total mutation rate) effectively take a random walk on $G$, experiencing the average mutational robustness over all genotypes in $G$. On the other hand, if $N\mu \gg 1$, populations evolving on $G$ equilibrate on highly connected regions of $G$, i.e. in regions of high mutational robustness.

As previously noted ([34]), these findings are actually in agreement with our conclusions here. Specifically, by construction Muller’s ratchet is impossible in van Nimwegen et al.’s model, since any mutation off of $G$ is lethal and so is instantaneously eliminated. Put another way, under this model, evolving populations enjoy the benefits of increased mutational robustness observed in Fig 4 without the otherwise concomitant risk of succumbing to Muller’s ratchet. Relaxing van Nimwegen et al.’s assumption of strict lethality for all genotypes off of $G$ recovers exactly our predicted behavior: at equilibrium mutational robustness declines with mutation rate (see Fig S5 in [34]). The above argument also applies to other studies where any deleterious mutation is lethal (e.g., [35]).

Wilke et al. ([4]) examined the evolution of mutational robustness using Avida, a platform for conducting in silico evolutionary experiments. Digital organisms in Avida are computer programs capable of self-replicating, and they compete for the limiting resource – CPU cycles – to reproduce. Their genotypes are the instructions making up the program, and their phenotypes, evaluated via executing the program, determine their fitness, i.e., how many CPU cycles will be allocated to them and thus how fast they can replicate. Mutations influence their fitness by changing instructions inside the program, i.e., their genotypes, mimicking biological mutations.

Wilke et al. evolved 40 paired populations (each initialized from a common genotype) under low and high mutation rates for 1000 generations. Next, the most abundant genotype was extracted from the low- and high-mutation rate populations; these were designated as A and B in ([4]). In each pair, A almost always had higher fitness than B, reflecting the lower mutational load, and Wilke et al. focused on 12
cases in which the fitness of A was at least 1.5-fold higher than that of B. Competitions were conducted for each such pair under different mutation rates.

Not surprisingly, at low mutation rate A eliminated B, since A always had higher fitness. However, the study’s key result was that in each case, B outcompeted A when mutation rate is high. Representative results are shown in Fig 1 of [4], and these were taken to suggest that B has higher mutational robustness. To confirm this interpretation, Fig 2 in [4] presented fitness distributions of genotype in the mutational neighborhood of a representative A and B. The authors write that “the competitive reversal [at high mutation rate] reflects a shift toward less fit genotypes, which is more pronounced for A than B,” implying that “A occupied a higher but narrower fitness peak, whereas B was on a lower but broader peak” (p. 322 in [4]).

However, to our mind Fig 2 is actually more consistent with the evolution of ratchet robustness than of mutational robustness. Specifically, the fitness distribution of genotypes surrounding B appears to contain more lethal mutations (Fig 2 in [4]: lower panels show larger fractions of black, which correspond to more mutants with zero fitness). Furthermore, A seems more prone to mutation accumulation than B (Fig 2 in [4]: right upper panel skews towards more mutations compared with right lower panel, which is only plausible if A has more weakly deleterious mutations available). Thus we interpret Fig 2 in [4] to suggest that B actually has more lethal and deleterious mutations, i.e., lower mutational robustness, and consequently higher ratchet robustness, which is consistent with our results. Unfortunately all reagents used in that previous study are no longer available (C. Wilke, pers comm) and we have been unable to replicate those earlier findings. Thus we are unable to rigorously test our alternative interpretation.

There also exist studies attempting to experimentally validate conclusions in Wilke et al.. Here, two viroid ([36]) or RNA virus ([37]) populations were evolved in regimes analogous to those for populations A and B in Wilke et al., and then competed against each other under different mutation rates. Unfortunately, these studies fail to accurately characterize A or B’s fitness landscapes, thus cannot support the claim that population B has higher mutational robustness than A. Specifically, only nonlethal genotypes are considered, while lethal mutations are ignored.
Two different kinds of MSDE in the presence of negative epistasis

Attentive readers may have noticed that two distinct kinds of MSDE are observed on fitness landscapes with negative epistasis. When $U_{del}$ is lower than critical $U_{del}$ at the peak, the first kind of MSDE, which we call MSDE$_{d@p}$ is sustained at the peak (i.e., at MSDE$_{d@p}$ fitness equals $e^{-U_{del}}$). MSDE$_{d@p}$ is characterized by selection greater than that required for offsetting current $U_{del}$. Populations reside at the MSDE$_{d@p}$ merely because genotypes at the peak have exhausted their supply of beneficial mutations (5).

In contrast, when $U_{del}$ grows larger than critical $U_{del}$ at the peak, populations move downwards away from the peak. But importantly, because of the negative feedback between mutation and the strength of selection induced by negative epistasis (see Results), the population will again equilibrate, but now at the second kind of MSDE, which we call MSDE$_{n@p}$. Unlike MSDE$_{d@p}$, MSDE$_{n@p}$ exists where purifying selection exactly offsets deleterious mutation and genetic drift. The position of this second equilibrium was recently solved quantitatively (5), although in that study the essential negative feedback was realized by increasing $U_{ben}/U_{del}$ with decreasing fitness, rather than increasing $s$ locally, as on our landscape.

Importantly, MSDE$_{n@p}$ does not exist on fitness landscape with positive epistasis, because positive epistasis means positive feedback between mutation rate and the strength of purifying selection. Consequently, populations below the point of no return experience accumulation of deleterious mutations accelerated by weakening of selection, quickly succumbing to Muller’s ratchet. Populations above the point of no return encounter stronger and stronger selection while driven upward by selection, until halted by the peak, and establish MSDE$_{d@p}$ around the peak, where selection is the strongest.

Finally, the log-linear fitness landscape, or landscape with no epistasis, presents a still subtler scenario, since log-linear fitness landscapes have no feedback between $U_{del}$ and $s$. Thus at first glance, it seems that since selection is identical everywhere on the landscape, critical $U_{del}$ at the peak should be the same as critical $U_{del}$ anywhere else. This suggests that no MSDE$_{n@p}$ should exist in the absence of epistasis, since once MSDE$_{d@p}$ is no longer sustainable, mutation should overwhelm selection anywhere else as well. However, as noted above, the genotype at the fitness peak uniquely lacks
beneficial mutations. Consequently, critical $U_{\text{del}}$ at the peak is actually slightly lower than critical $U_{\text{del}}$ anywhere else, allowing for existence of MSDE$_{\text{nup}}$ at $U_{\text{del}}$ higher than critical $U_{\text{del}}$ at the peak but lower than critical $U_{\text{del}}$ elsewhere. This effect can be seen in Fig 1A, where at intermediate values of $U_{\text{del}}$ population equilibrates at fitness values between $e^{-U_{\text{del}}}$ and 0 (See also S1 Fig).

Fitness landscapes with negative epistasis are common in nature

We note that many surveys of biological fitness landscapes find extensive evidence for negative epistasis, precisely the property that our results predict facilitate long-term survival. For example, a genome wide study of pairwise epistasis in Saccharomyces cerevisiae found twice as many cases of negative epistasis as positive epistasis ([38]), and another study on evolving Escherichia coli populations reported negative epistasis between beneficial mutations ([39]). Relatedly, biophysical principles predict negative epistasis among mutations in protein-coding genes, mediated by their effect on protein folding stability ([34]). On the other hand, metabolic control theory ([40]) implies that whether or not negative epistasis is favored depends on metabolic pathway topology ([41]). It’s worth noting that, although we predict selection for ratchet robustness favors landscapes with negative epistasis, selection for other evolutionary traits may have opposite effects.

Conclusion

Our simulations demonstrate that ubiquitous pairwise negative epistasis profoundly influences ratchet robustness (Fig 2, Fig 3), with the negative feedback between mutation and selection strength. However, epistasis present in real fitness landscapes are often more complex than simply pairwise interactions, and it has been previously proposed that higher order interactions could play an important role in evolutionary dynamics ([42,43]). At this point, we are unable to even speculate on the characteristics of complex landscapes that promote high ratchet robustness.

Supporting information
S1 Table. Average number of individuals leaving or entering peak from either negative epistasis region or positive epistasis region. $U_{del} = 0.5$, $U_{ben} = 0.01U_{del}$, $p = 0.5$, $N = 1000$, $s = 0.35$, negative epistasis region: $\epsilon = -0.25$, positive epistasis region: $\epsilon = 0.25$, average over 10,000 generations. The mutational net effect results in increase in individuals at the negative epistasis region, which offsets decrease in individuals at the negative epistasis region due to lower fitness (see Results).

<table>
<thead>
<tr>
<th>Peak → Negative</th>
<th>Negative → Peak</th>
<th>Peak → Positive</th>
<th>Positive → Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>57.9168</td>
<td>0.5154</td>
<td>57.8357</td>
<td>0.5256</td>
</tr>
</tbody>
</table>

S1 Fig. Average fitness over time for populations evolving on fitness landscape in the absence of epistasis. Lines of different colors represent five independent replicates ($N = 1000$, $s = 0.1$, $U_{del} = 0.4$, $U_{ben} = 0.01U_{del}$). Upper dashed line denotes $w_{max}e^{-U_{del}}$ ($w_{max}$ is fitness at the peak), whereas lower dashed line denotes average equilibrium fitness at generation 10,000 over five replicates, which equals $w_{s}e^{-U_{del}}$ ($w_{s}$ is fitness of genotype one mutation away from the peak). This shows that current $U_{del}$ exceeds critical $U_{del}$ at the peak, but is still lower than critical $U_{del}$ everywhere else (see Discussion).
S2 Fig. Equilibrium distribution of individuals at negative epistasis region, peak and positive epistasis region on hybrid peaks under different mutation rates. Counts of individuals at different parts of the landscape for A: $U_{\text{del}} = 0.3$, B: $U_{\text{del}} = 0.7$, C: $U_{\text{del}} = 1.0$ ($U_{\text{ben}} = 0.01U_{\text{del}}$, $p = 0.5$, $N = 1000$, $s = 0.3$, negative epistasis region: $\epsilon = -0.25$, positive epistasis region: $\epsilon = 0.25$, evolved for 10,000 generations). There are no individuals present at omitted parts. Note that for panel C, every individual is at the “bottom” of the positive epistasis region after succumbing to Muller’s ratchet.

References


