Mutation bias can shape adaptation in large asexual populations experiencing clonal interference

Kevin Gomez*, Jason Bertram†, and Joanna Masel [1,‡]

*Graduate Interdisciplinary Program in Applied Mathematics, University of Arizona
†Environmental Resilience Institute, Indiana University
†Department of Biology, Indiana University
‡Department of Ecology & Evolutionary Biology, University of Arizona

February 18, 2020

Abstract

The Extended Evolutionary Synthesis invokes a role for development in shaping adaptive evolution, which in population genetics terms corresponds to mutation-biased adaptation. Critics have claimed that clonal interference makes mutation-biased adaptation rare. We consider the behavior of two simultaneously adapting traits, one with larger mutation rate U, the other with larger selection coefficient s, using asexual traveling wave models. We find that adaptation is dominated by whichever trait has the faster rate of adaptation v in isolation, with the other trait subject to evolutionary stalling. Reviewing empirical claims for mutation-biased adaptation, we find that not all occur in the "origin-fixation" regime of population genetics where v is only twice as sensitive to s as to U. In some cases, differences in U are ten times larger than differences in s, causing mutation-biased adaptation to occur even in the "multiple mutations" regime. Surprisingly, when $U > s$ in the "diffusive-mutation" regime, the required sensitivity ratio is also only two, despite pervasive clonal interference. Given two traits with identical v, the benefit of having higher s is surprisingly small, occurring largely when one trait is at the boundary between the origin-fixation and multiple mutations regimes.

1 Introduction

What shapes the course of adaptive evolution? The neo-Darwinian position was that natural selection is preeminent; variation is plentiful and appears in gradual increments, providing raw material that natural selection can shape in any way. The idea that biases in the introduction of variation had a significant influence on the course of evolution was mocked as "the revolt of the clay against the power of the potter" [1]. This selectionist view persisted through the Modern Synthesis [2], following the metaphor of populations as vast "gene pools" with ample amounts of genetic variation already available for natural selection to act on. With abundant mutations each of tiny effect already present and available to recombine to produce any possible phenotype, bias in variation was thought irrelevant.

While the influence of mutation bias on neutral evolution was later acknowledged as part of Neutral Theory, selectionism with respect to the course of adaptation is still upheld by advocates for "Standard Evolutionary Theory", on the grounds that other evolutionary processes like mutation, drift, and migration, are random with respect to the adaptive direction favored by selection [3]. In contrast, advocates for an "Extended Evolutionary Synthesis" invoke a significant role for developmental processes [4, 5, 6], which shape the phenotypic effects that mutations are more or less likely to have. Exchanges between the two camps have been heated [5, 7].

A role for differences in beneficial mutation rates in shaping the nature of adaptations that evolve has been called "survival of the likeliest" [8], or "first-come-first-served" [9]. Here we refer to it as mutation-biased adaptation. This phenomenon now has substantial empirical support [10]. Evidence for adaptation aligned with mutation bias has been found during the experimental evolution of microvirid bacteriophage [11, 12], *Escherichia coli* [13], and *Pseudomonas aeruginosa* [14, 15]. Studies of parallel adaptation in natural populations also reveal patterns suggestive of mutation bias, e.g. in sodium pump ATPα1 adaptations enabling the consumption of glycosine toxins by insects [15], antibiotic resistance in *Mycobacterium tuberculosis* [9], and hemoglobin adaptations for high altitude in birds [16].

A formal population genetic description of mutation-biased adaptation was given by Yampolsky and Stoltzfus [17] for the case where the product of the beneficial mutation rate U and the population size N is small. In this "origin-fixation" or "strong-selection weak-mutation" parameter regime, each adaptive mutation is either lost or fixed before the next appears [17, 8]. Mutations of each kind appear at their own characteristic mutation rate U, and fix with probability proportional to the selection coefficient s [18]. Differences in mutation rate and differences in selection thus have exactly equal quantitative influence on the flux of adaptive mutations, which is proportional to UNs.

Yampolsky and Stoltzfus [17] originally considered a simple sign epistatic landscape of two loci each with two

[1]masel@email.arizona.edu, Dpt. Ecology & Evolutionary Biology, University of Arizona, 1041 E Lowell St Tucson AZ 85721 USA.

alleles, beginning in a valley, such that mutating to reach one peak precluded reaching the other, and found mutation bias even when simulating scenarios with $UN > 1$. However, sign epistasis is not required to argue on the basis of the sensitivities of the flux of adaptive mutants to U and to s [19]. The theoretical argument for mutation-biased adaptation in this case is, however, strongly dependent on the assumption that the influx of adaptive mutations is small, such that $UN\log(Ns) \ll 1$. When two adaptive mutations escape drift to reach appreciable frequencies at the same time, clonal interference will favor the fixation of the one with higher s [20]. For this reason, previous theoretical arguments in favor of mutation-biased adaptation have been restricted to the parameter regime $UN\log(Ns) \ll 1$, and the phenomenon has been criticized on this basis [3]. However, clonal interference may occur in at least some of empirically documented cases of mutation-biased adaptation [14, 21, 12]. There is therefore a need for theory to understand how mutation-biased adaptation might be possible in the presence of clonal interference.

Models that have previously been used to dismiss mutation-biased adaptation when $UN\log(Ns) \geq 1$ [3] are simple, and do not capture the multi-locus complexities included in more recent traveling wave models of population genetics [22, 23, 24, 25]. Here we build on a recent two-dimensional travelling wave model [26] to formally investigate the circumstances under which mutation-biased adaptation can occur.

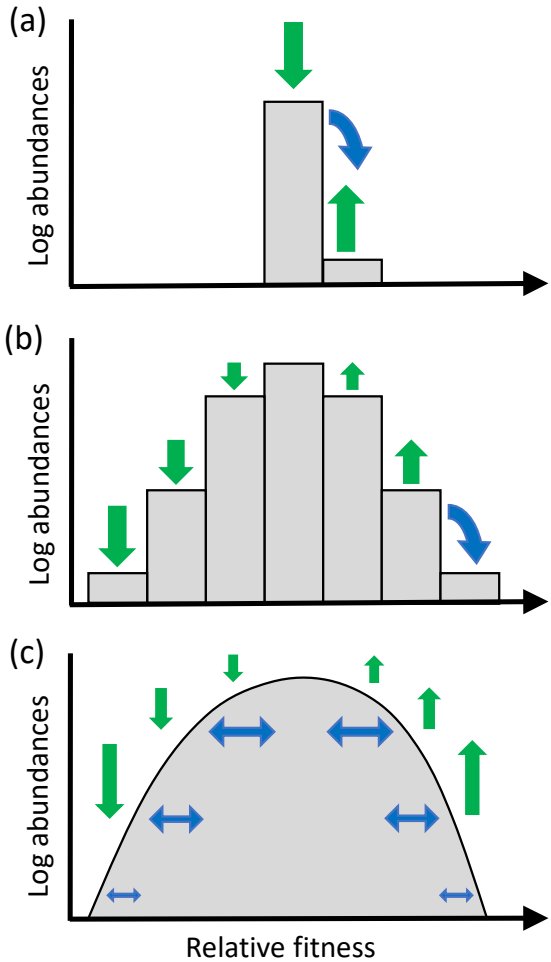

Figure 1: Contribution of mutation fluxes to adaptive evolution in three regimes of adaptation: (a) origin-fixation regime, (b) multiple mutations regime, and (c) diffusive-mutation regime. In all three regimes, the mean fitness moves to the right (adaptation) due to selection. In (a) and (b), changes in fitness class abundances are dominated by selection (green arrows). Mutation fluxes are negligibly small by comparison, except at the front (blue arrow) where they have the crucial role of producing new fitter lineages, allowing the travelling fitness wave to continue. In (c), mutation fluxes when $U \gg s$ increase fitness variance throughout the population (blue double-sided arrow).

2 Materials and Methods

(a) Elasticities

We quantify the sensitivity of a population's rate of adaptation v to changes in U and s by computing elasticities, a metric in common usage in economics (e.g. price elasticity) and applied mathematics, but not in evolutionary biology. Elasticity is the percent change in a function's output due to a percent change in its input. Given the function $v(U,s)$, the U-elasticity E_U and s-elasticity E_s of v are

$$E_U = \frac{\Delta_u v/v}{\Delta U/U} \qquad E_s = \frac{\Delta_s v/v}{\Delta s/s}, \tag{1}$$

where $\Delta_u v = v(U+\Delta U, s) - v(U,s)$ and $\Delta_s v = v(U, s+\Delta s) - v(U,s)$. The ratio E_s/E_U measures the relative sensitivity of adaptation to selection versus mutation.

Consider the origin-fixation regime (figure 1a), where the rate of adaptation is

$$v(U,s) = UN \times \left(\frac{1-(1+s)^{-2}}{1-(1+s)^{-N}}\right) \times s \tag{2}$$

for a population of size N. Here UN is the total influx of beneficial mutations per generation and $[1-(1+s)^{-2}]/[1-(1+s)^{-N}] \approx 2s$ is the fixation probability of each mutation [27, Equation 2]. Thus, approximately $UN \times 2s$ mutations are expected to fix in the population each generation, with each increasing relative fitness by s, yielding Eq. (2). The U- and s-elasticities of v in the origin-fixation regime are $E_U \approx 1$ and $E_s \approx 2$ since $\Delta_u v/v \approx \Delta U/U$ and $\Delta_s v/v \approx 2\Delta s/s$. The latter expressions follow from applying the first order Taylor approximations

$$\begin{aligned} \Delta_u v &\approx \frac{\partial v}{\partial U}\Delta U \approx N2s^2\Delta U \\ \Delta_s v &\approx \frac{\partial v}{\partial s}\Delta s \approx UN4s\Delta s, \end{aligned}$$

and dividing through by $v \approx UN2s^2$.

In the origin-fixation regime, in which there is no interference between the adaptation of different traits, mutation-biased adaptation will occur when relative differences in U ($\Delta U/U$) exceed relative differences in s ($\Delta s/s$) by at least a factor of $E_s/E_U \approx 2$. In the two sections below, we present analytical expressions for $v(U,s)$ in two more parameter value regimes, allowing us to apply an elasticity-based analysis to them to identify where the influence of mutation and selection on adaptation rates is comparable.

(b) Multiple mutations regime analytics

When beneficial mutations are more common ($UN\log(Ns) \geq 1$), there is competition among lineages, each of which may have accumulated multiple beneficial mutations. The probability of fixation is then no longer a simple function of the selection coefficient of the focal mutation alone, but instead depends on the presence of other beneficial mutations. The resulting complexities change the relationship between v and the parameters N, U and s.

In the staircase model (Desai et al. 28; figure 1b), beneficial mutations appear at rate U per birth and each confer a relative fitness advantage of s (one "step" on the staircase). The population is divided into fitness classes that are composed of individuals with the same number of (interchangeable) beneficial mutations, and hence the same fitness.

Populations in the multiple mutations regime ($UN\log(Ns) \geq 1$) will consist of many fitness classes. A traveling fitness wave results from beneficial mutations adding new fitness classes and selection changing genotype frequencies and eventually culling the least fit classes [23]. A new fitness class appears by mutation at the "nose" of the travelling wave, initially populated by few individuals relative to their selective advantage Q over mean population fitness, subject to branching process dynamics dominated by random genetic drift. However, if a new class reaches a size of $n > 1/Q$, growth due to selection becomes approximately deterministic. Deterministic classes can be modeled using exponential growth equations, simplifying the analysis of the traveling wave and allowing for efficient numerical simulation as described below in the Subsection Simulations.

Under the assumption that $U \ll s$, Desai and Fisher [23, Eq. 41] obtained

$$v(U,s;N) \approx s^2 \frac{\log(N^2 sU)}{\log(s/U)}, \tag{3}$$

from which the ratio E_s/E_U can be calculated.

(c) Diffusive-mutation regime analytics

When $U \gg s$, the mutational fitness flux between neighbouring fitness classes becomes comparable to the change in fitness class abundance due to selection, breaking down the distinction between deterministic established fitness classes and the stochastic nose. This causes Eq. (3) to break down.

A useful simplifying assumption in this "diffusive-mutation" regime (figure 1c) is to set the beneficial and deleterious mutation rates equal, so that mutation bias or "pressure" [29, 30] alone in the absence of selection does not drive the fitness distribution forward. Mutational steps then follow unbiased random walks with equal chance of reducing or increasing fitness by s. The resulting symmetric mutational diffusion process creates fitness variance which selection acts on to drive the wave forward. The diffusion coefficient for this mutational diffusion process is $D = Us^2/2$, because the variance in the position of the mutational walk is s^2 per step, and mutations occur at rate U. The dependence of v on U and s is then approximately given by [Appendix F in ref 25]

$$v(U,s;N) \sim D^{2/3} \log\left(ND^{1/3}\right)^{1/3}. \tag{4}$$

Note that due to the logarithmic second term, the rate of adaptation (variance of the fitness distribution) is approximately proportional to $D^{2/3}$. As we will see, this substantially alters the relative roles of mutation and selection, compared to the multiple mutations regime.

(d) Simulations

The analytical, elasticity-based analysis described above considers evolution in one trait only. But when there is clonal interference, i.e. outside the origin-fixation regime, traits with different U and s do not evolve independently.

We therefore also simulate the simultaneous evolution of two traits ($k = 1, 2$), with beneficial mutations of fitness effect s_k appearing at rate U_k, using an extension of the method of Gomez et al. [26]. We consider strong selection relative to population size, such that $1/N \ll s_k \leq 1$. With no pleiotropy or epistasis in this model, Malthusian fitness is $r_{i,j} = is_1 + js_2$ for an individual with i beneficial mutations in trait $k = 1$ and j beneficial mutations in trait $k = 2$. We group individuals with the same numbers of beneficial mutations into classes denoted by subscripts (i, j), with abundances and frequencies at time t given by $n_{i,j}(t)$ and $p_{i,j}(t) = n_{i,j}(t)/\sum_{i',j'} n_{i',j'}(t)$, respectively. Thus, the selective advantage of an individual in class (i, j) is $Q_{i,j}(t) = r_{i,j} - (\sum_{i,j} n_{i,j}/N)\bar{r}(t)$ where $\bar{r}(t) = \sum_{i,j} p_{i,j}(t) r_{i,j}$ is population mean fitness. Selection according to $Q_{i,j}$ ensures logistic regulation of population size [31, page 27]. The set of abundances forms a two-dimensional distribution in (i, j) space.

Given abundances $n_{i,j}(t)$ at generation t, we calculate the abundances at generation $t+1$ in three steps. First, we calculate the expected change due to selection as

$$\tilde{n}_{i,j}(t) = n_{i,j}(t) \frac{Ne^{Q_{i,j}(t)}}{\sum_{i',j'} n_{i',j'}(t) e^{Q_{i',j'}(t)}}.$$

Second, we determine the expected net flux of mutations $\Delta_u \tilde{n}_{i,j}$ into class $\tilde{n}_{i,j}$. We also include deleterious mutations in our simulation; these occur at the same mutation rate U_k and decrease fitness by s_k, yielding

$$\begin{aligned}\Delta_u \tilde{n}_{i,j}(t) = {} & U_1 \left[\tilde{n}_{i-1,j}(t) + \tilde{n}_{i+1,j}(t)\right] \\ & + U_2 \left[\tilde{n}_{i,j-1}(t) + \tilde{n}_{i,j+1}(t)\right] \\ & - 2(U_1 + U_2)\tilde{n}_{i,j}(t)\end{aligned}$$

Expected abundances after selection and mutation are $n^*_{i,j}(t) = \tilde{n}_{i,j}(t) + \Delta_u \tilde{n}_{i,j}(t)$. In the last step we set abundances of deterministic classes ($n^*_{i,j} > 1/Q_{i,j}$) to $n_{i,j}(t+1) = \text{Round}\left[n^*_{i,j}(t)\right]$, while classes that grow stochastically ($n^*_{i,j} \leq 10/Q_{i,j}$) are instead sampled from a Poisson distribution with mean $n^*_{i,j}(t)$.

We implemented the simulation method described above using Matlab code originally developed by Pearce and Fisher [32], modified by Gomez et al. [26] to deal with two traits, and here modified to allow those traits to have distinct mutation rates and selection coefficients.

When simulating the evolution of one trait for comparison with theoretical predictions for $v(U,s)$ given by Eq. (2), Eq. (3) and Eq. (4), we simply set $s_2 = 0$ and $U_2 = 0$. We numerically determined $v(U,s)$ values by applying iterative root-finding methods that allow for noise in $v(U,s)$ [33]. Time-averages of rates of adaptation were taken over runs lasting 1-2 million generations to ensure convergence, with burn-in periods of 5,000 generations.

Our code ran on Matlab 2016R installed on a desktop computer running Linux. All scripts used to generate our results are available at `https://github.com/MaselLab/Gomez_et_al_2020`.

3 Results

(a) Elasticities suggest mutation-biased adaptation in the diffusive-mutation regime

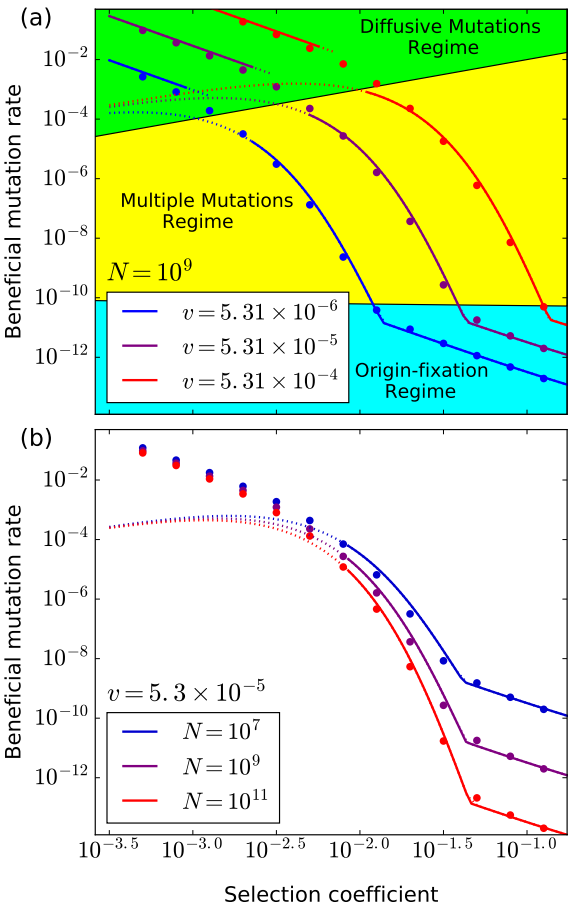

Figure 2: Contour lines of $v(U,s)$ in U-s parameter space reveal the possibility of mutation-biased adaptation being just as easy in the diffusive-mutation regime as in the origin-fixation regime. (a) shows distinct v contours and (b) shows contours of the same v given different values of N. Regions of U-s parameter space in (a) are labeled to indicate where $UN\log(Ns) < 1$ (origin-fixation; aqua blue), $s > U \geq 1/N\log(Ns)$ (multiple mutations; yellow), and $U \geq s$ (diffusive-mutation; neon green). In (b), these regime boundaries are not shown since they depend on N.

To visualize the relative influence of mutation and selection on the rate of adaptation v, we plot contour lines of constant $v(U,s)$ in U-s parameter space (figure 2), as obtained from analytical theory (Section 2.2; colored piecewise curves) and numerical simulations (colored points). The slopes of these contours as shown in log-log space are closely related to the ratio of elasticities E_s/E_U, discussed in the Methods as a metric of the relative sensitivities of the adaptation rate to the selection coefficients and appearance rate of beneficial mutations. As shown in the Appendix, we have

$$\lim_{\Delta v \to 0} \frac{E_s}{E_U} = -\frac{s}{U}\frac{dU}{ds}\bigg|_{v(U,s)=\text{constant}}. \tag{5}$$

Thus, steeper contours in figure 2 imply that selection has a greater effect on v.

The slopes of the contours depend on which of the three parameter regimes of U and s described in the Methods applies, as shown as background color in figure 2a for different values of v (figure 2a) and N (figure 2b). In the origin-fixation regime (aqua), simulation confirm that $E_s/E_U \approx 2$, as previously obtained analytically. As expected from clonal interference, selection becomes more important (v contours become much steeper, as high as $E_s/E_U \approx 10$) in the multiple mutations regime (yellow region, figure 2).

However, surprisingly — and in contrast to arguments suggesting that mutation-biased adaptation is impossible with clonal interference [3] — for sufficiently high mutation rates, the steepness of the contour lines declines and we return to $E_s/E_U \approx 2$ (green diffusive-mutation region, figure 2a). This regime could be thought of as a "gene flood" of new mutations, in contrast to the "gene pool" analogy of standing genetic variation in the Modern Synthesis. The modest slope is a result of the fact that $v \propto (Us^2)^{2/3}$ (Eq. (4)). While this differs from origin-fixation formula $v \propto Us^2$, due to the presence of the 2/3 exponent, the ratio of elasticities for $v \propto (Us^2)^{2/3}$ is nevertheless identical to that of $v \propto Us^2$. We see this by solving for U in $v \propto (Us^2)^{2/3}$, which provides $U \propto v^{3/2}s^{-2}$. Applying Eq. (1) to $U \propto v^{3/2}s^{-2}$ then provides

$$\frac{E_s}{E_U} \approx -\frac{s}{v^{3/2}s^{-2}}\left(-2v^{3/2}s^{-3}\right) = 2,$$

as in the origin-fixation regime.

The origin of the $v \propto D^{2/3}$ dependence in Eq. (4) can be seen directly from the diffusive nature of mutations in the $U \gg s$ regime. In the bulk of the population, the steady-state abundance $\overline{n}$ of a fitness class with fitness x relative to the population mean changes according the the diffusion equation $\partial\overline{n}(x)/\partial t = D\partial^2\overline{n}(x)/\partial x^2 + x\overline{n}$. Taking the first moment gives $v = \sigma^2$, where σ is the standard deviation in fitness (Fisher's Fundamental Theorem). Similarly, taking the second moment gives $D \propto \int x^3\overline{n}(x)dx = \mu_3\sigma^3$ where μ_3 is Pearson's moment coefficient of skewness of the steady-state fitness distribution $\overline{n}(x)$. Therefore, $v \propto (D/\mu_3)^{2/3}$, with a skew factor that depends more weakly on the parameters (Eq. (4)).

(b) Clonal interference creates little bias between two traits of equal v

Our single-trait analysis described above shows the existence of a parameter regime in which the rate of adap-

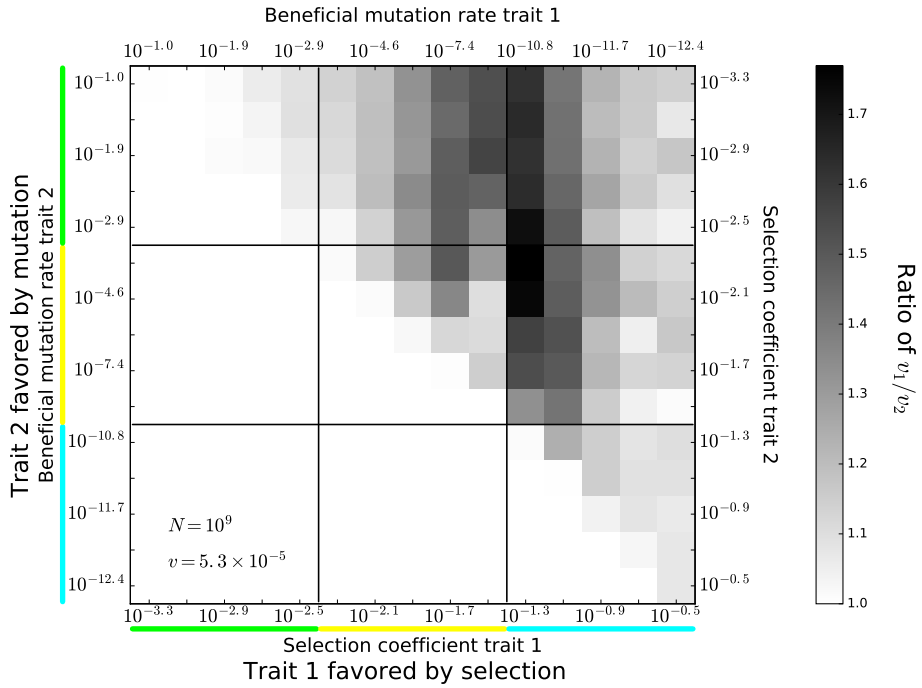

Figure 3: Heat map showing relative rates of adaptation in two simultaneously adapting traits, both of whose values of U and s would in isolation produce the same rate of adaptation $v = 5.3 \times 10^{-5}$. Axes are labeled by color to show the regime (from figure 2a) that a trait would be in were it evolving on its own. Darker cell color indicates how much faster the selection-favored trait (v_1) evolves relative to the mutation-favored trait (v_2). Cell values in the lower left are redundant and not shown, as these would amount to switching trait labels. The fact that ratios are near 1 makes stochasticity difficult to eliminate.

tation is sensitive to the mutation rate even in the presence of clonal interference, suggestive of mutation-biased adaptation. But as previously discussed (see Materials and Methods), clonal interference may alter how differences in s and U shape the course of adaptation, and this must be fully accounted for.

We do this by simulating the simultaneous evolution of two fitness-associated traits, each of which would on its own evolve at the same rate as the other ($\{s,U\}$ values corresponding to the purple dots in figure 2). We then measured the reduction of a trait's rate of adaptation due to clonal interference with the second trait. Specifically, we calculated the rate of adaptation for the higher-s trait v_1 and divided it by the rate of adaptation of the higher-U trait v_2, and examined how this ratio changed across parameter space (figure 3). A ratio higher than one indicates the degree to which clonal interference makes selection more important than indicated by the one-trait elasticity analysis. When both traits are in the origin-fixation regime, we see $v_1/v_2 \approx 1$, as expected in the absence of clonal interference (figure 3; pale color in the aqua blue vs. aqua blue lower right).

Surprisingly, we see even smaller deviations from $v_1/v_2 = 1$ when both traits fall within the diffusive-mutation regime (pale color in the neon green vs. neon green upper left of figure 3 never exceeds $v_1/v_2 = 1.09$). Our finding of $E_s/E_U \approx 2$ already suggested susceptibility to mutation-biased adaptation; here we see that this persists even when clonal interference is fully accounted for.

Indeed, deviations from $v_1/v_2 = 1$ are mild throughout, with the highest value of 1.77 observed when the high-s trait is near the boundary between origin-fixation and multiple-mutation regimes. At this maximum value, adaptation in the mutation-favored trait still accounts for 36% of the total rate of adaptation. Clonal interference thus has a limited ability to block mutation-driven adaptation, beyond that captured by the sensitivity of v to s.

(c) Evolutionary stalling is driven by differences in v not s

The previous section shows that two traits with the same independent rate of evolution v, but drastically different values of s and U, will continue to adapt at similar rates even in the presence of clonal interference. Next we ask how clonal interference operates between two traits that differ not only in their values of s and U, but also in their values of v.

Venkataram et al. [34] has claimed that "evolutionary stalling" will occur in whichever trait has smaller s. In other words, only the trait with the largest s will continue to adapt significantly, and traits with only small-s beneficial mutations available will effectively stop adapting. This argument implicitly assumes either that traits have comparable U, or that differences in U are not important for evolutionary stalling because it is driven by clonal interference.

To test whether evolutionary stalling occurs, and if so whether the identity of the dominant trait depends on s alone, we quantify how a focal trait's adaptation suffers from clonal interference from a second trait, as a function of that second trait's value of s and U (figure 4). The top, middle, and bottom panels of figure 4 show focal traits with the same v taken from the origin-fixation, multiple-mutations, and diffusive-mutation regimes, re-

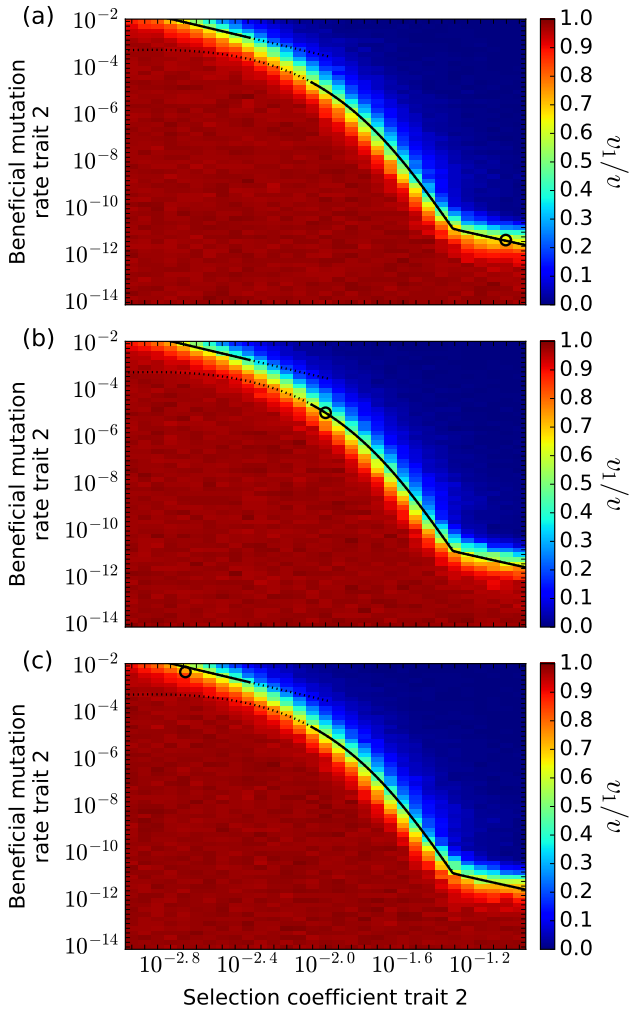

Figure 4: The trait with larger v is little affected by clonal interference with another trait, while that with smaller v stalls. The focal trait, shown as a dot, would have the same value of $v_1 = 5.31 \times 10^{-5}$ in each of the three panels, if it were evolving on its own. The focal trait falls into the origin-fixation regime, multiple mutations regime, and diffusive-mutation regime in the top, middle, and bottom panel, respectively. Colors show the extent to which adaptation in the focal trait slows down as a function of the selection coefficient s_2 and mutation rate U_2 of a second clonally interfering trait. The theoretical v contour for $v = 5.31 \times 10^{-5}$ from figure 2a is shown as a black piecewise curve in each panel. v_1 transitions from a negligible impact of clonal interference from trait two (dark red region) to having its evolution nearly halted by evolution in trait two (dark blue region). The decline in v_1 follows patterns set by v-contours, not s_2 or U_2, indicating that the rate of adaptation of trait two, when evolving independently, determines the decline in v_1 from clonal interference between traits. In all panels $N = 10^9$.

spectively (open circles). These three panels are striking in their similarity. Adaptation in the focal trait is essentially unaffected by clonal interference from a trait with smaller v, but severely affected by a trait with larger v. In other words, we see marked evolutionary stalling that depends on v.

Under a distribution of beneficial effect sizes, adaptation depends on the probability of fixation as a function of s, which is dominated by an exponential term $\exp(sT_c)$ involving the expected time to coalescence T_c [35, 36]. T_c is also the sweep time, and hence closely related to rate of adaptation. In our model, the rate of adaptation is driven by the trait with the largest v. Thus, v may simply be an accessible heuristic for a deeper mathematical dependence on T_c. The fact that the functional form of the probability of fixation is exponential in T_c explains why the drop in trait one's adaptation shown in figure 4 is so dramatic.

4 Discussion

When two adaptive traits evolve together, adaptation is dominated by whichever has the higher rate of adaptation v in isolation, not by whichever has the higher selection coefficient s. If differences in v among fitness-associated traits were dominated by differences in s, then this would be a distinction without a difference. But if differences in U are sufficiently large relative to differences in s, then mutation-biased adaptation will occur. How much larger differences in U need to be is well summarized using simple equations for the adaptation rate v, and depends on which parameter value regime the population is in. For both the origin-fixation regime with $UN\log(Ns) \ll 1$, and the diffusive-mutation regime with $U > s$, differences in U need to be twice as large as differences in s. For values of U in the multiple mutations regime between the two, ratios of as much as 10 might be required. Our theoretical calculations in the multiple mutations and diffusive-mutation regime are for the strong linkage disequilibrium produced in asexual microbes — in the absence of strong linkage disequilibrium in more sexual populations, the origin-fixation regime is likely a good approximation.

(a) Empirical examples

Table 1 summarizes a variety of proposed empirical examples of mutation-biased adaptation. Here we examine each example in the light of our theoretical results, attempting to quantitatively estimate N, s and U where possible, and if not to qualitatively assess which parameter value regime applies, as well as to infer what we can about missing parameter values from the evidence for mutation-biased adaptation.

The first two entries in table 1 are expected to lie in the origin-fixation regime, albeit for different reasons. N is likely reasonably small for the natural bird populations studied by Storz et al. [16], where there were more adaptive transitions than expected if each one-step mutation were equally likely to be adaptive. While N is much larger for the natural viral populations studied by Willis and Masel [37], the rate U of the beneficial *de novo* birth of new overlapping genes is likely so low that $UN \ll 1$. There were 5 times more birth events in the +1 reading frame, which has more and longer ORFs, than there were in the +2 reading frame, on which the genetic code bestows a tendency toward the favored protein property of high intrinsic structural disorder. Neither study quantified differences in U or s — instead, differences in occurrence, i.e. in v/s, were counted.

Stoltzfus and McCandlish [15] also found that adaptive transitions were over-represented relative to the expected ratio of transversions predicted by the null model,

System/Organism	Natural / Lab	Population size (N)	Selection coeff. (s)	Beneficial mut. rate (U)	Regime	Estimated mutation bias	Comment
70 avian taxa; adaptation to elevation via Hb-O2 affinity Storz et al. [16]	natural	small to moderate	moderate to large	small because few genes	OF	unavailable	
80 viral species: 6 ssDNA, 5 retroviruses, 81 RNA viruses; gene birth Willis & Masel [37]	natural	large	unknown	tiny	OF because tiny U	$> 465:92 = 5$	Mutation bias greater when longer ORFs are more likely to be adaptive
Meta-analysis of adaptive mutations Stoltzfus & McCandlish [15]	natural & lab	mixed	mixed	mixed	Mixed, OF + MM ($U \ll s$)	$2.2 - 5.7$ for which mutation, $3.6 - 5.8$ for its frequency	If mutation bias didn't matter, expect transition bias of 0.5.
M. tuberculosis; antibiotic resistance Payne et el. [9]	natural, sequenced genomes, clinical isolates	large	large 100-fold increase in MIC	small $10^{-9} - 10^{-7}$	OF or MM ($U \ll s$)	$0.10 - 7.45$ Basal data; $0.10 - 41.6$ Manson data.	
P. aeruginosa; antibiotic resistance MacLean et al. [14]	lab	$10^4 - 10^9$	$0.12 - 0.96$	$8.3 \times 10^{-10} - 2.5 \times 10^{-8}$	OF or MM ($U \ll s$)	50 fold	Mutation rates provided for specific alleles.
ssDNA bacteriophages ID8, ID11, NC13, WA13; phage growth Sackman et al. [12]	lab	$10^5 - 10^9$	$0.11 - 0.64$	$1.1 \times 10^{-3} - 5.3 \times 10^{-3}$	OF or MM ($U \ll s$)	$29:5 = 5.8$ $74:6 = 12.33$	
ssDNA bacteriophage ID11; phage growth Rokyta etl al. [11]	lab	$10^4 - 10^8$	$0.11 - 0.39$	$1.1 \times 10^{-3} - 5.3 \times 10^{-3}$	OF or MM ($U \ll s$)	8-9 fold	U is the total mutation rate. Bias is calculated from relative mutation rates.
E. coli; antibiotic resistance Schober et al. [38]	lab	$6 \times 10^6 - 7.2 \times 10^8$	large	high	MM ($U \ll s$)	unavailable	
E. coli; adaptation following inactivation of DHFR Rodrigues et al. [39]	lab	$1.6 \times 10^7 - 4.8 \times 10^7$	large	small	MM ($U \ll s$)	unavailable	
E. coli; antibiotic resistance Couce et al. [13]	lab	antibiotic concentration titrated to allow proliferation. $7.2 \times 10^7 - 3.6 \times 10^9$	large $0.00158 - 1$	high $3 \times 10^{-4} - 1 \times 10^{-2}$	MM ($U \ll s$) or DM	$100 - 300$ fold $500 - 10{,}000$ fold	Different mutational profiles of $\Delta mutT$ (elevates transitions, excl. usual transversion adaptation) v. $\Delta mutH$ (elevates most).

Table 1: Empirical examples of mutation-biased adaptation. The beneficial mutation rate for [13] was assumed to be 3% of the total point mutation rate in the mutator strain studied. The population sizes reported for Schober et al. [38] and Rodrigues and Shakhnovich [39] are CFU values calculated from the optical densities (OD_{600}) reported in their papers. We used the formula $\mathrm{CFU} = \mathrm{OD}_{600} \cdot (8 \times 10^8) \cdot \mathrm{volume}$, where volume refers to culture volume (ml) in a turbidostat [38], or fresh media volume in 96-well flat-bottom plates [39]. The factor 8×10^8 is the scaling constant yielding $\mathrm{CFU/ml} = \mathrm{OD}_{600} \cdot 8 \times 10^8$.

in a meta-analysis of 15 studies (10 natural and 5 experimental) of different taxonomic groups, potentially representing a mixture of origin-fixation and multiple mutations regimes. Stoltzfus and McCandlish detected a 2-fold excess of transitions paths (3-fold for events) in the natural cases, and a 4-fold excess (7-fold for events) in the experimental cases.

The next five entries in table 1 involve microbial experimental evolution, which generally takes place in the multiple mutations regime [40, 41, 42]. Among evolved replicate lineages of the phage ID11, the specific adaptive mutation with the higher mutation rate and with the most fixation events appeared 6 times, while the adaptive mutation with largest selection coefficient appeared only once Rokyta et al. [11], Sackman et al. [12]. [12] examined 3 other phages along with ID11 and found a 24-fold over-representation of adaptive transitions over transversions, in a study in which the measured value of s of a particular mutation did not predict the number of lineages in which it became fixed. MacLean et al. [14] measured 50-fold differences in beneficial mutation rates [10], while selection coefficients only differed by 2- to 7-fold, during the experimental evolution of antibiotic resistance in *P. aeruginosa*. Given these vast differences between mutation rates (i.e. E_s/E_U is in the range between 2 and 10), mutation-biased adaptation can take place even in the multiple mutations regime. The *E. coli* populations studied by Rodrigues and Shakhnovich [39] and Schober et al. [38] also show evidence for mutation-biased adaptation, and can be presumed to be in the multiple mutations regime.

These last two studies examined experimental evolution to recover from a genetic loss. Both Rodrigues and Shakhnovich [39] and Schober et al. [38] found it to be mutationally easier to inactivate more loci, rewiring metabolic pathways in the process, than to recover the efficiency of the original pathway. In contrast, Venkataram et al. [34] partly disabled the translational machinery and then observed the active evolution of this seemingly "stalled" complex as it evolved to recover. But as adaptive improvements accumulated, diminishing returns epistasis led to a drop in s to the point that evolution again stalled. This was interpreted in terms of a dominant role for clonal interference and s in determining which modules would undergo adaptation. Our theoretical results support the concept of the stalling of a less-important module, but on the basis of differences in v instead of differences in s.

The only candidate for the diffusive-mutation regime in table 1 is Couce et al.'s [13] experiments with two mutator strains of *E. coli*, i.e. with artificially high U. One mutator strain (ΔmutH) has transition rates 100-300 times larger than wild-type [43, 44], the other (ΔmutT) has transversion rates 500-10,000 times larger than wild-type [43, 45]. These high mutation rates make $U > s$ and hence the diffusive-mutation regime a possibility, although not certain.

These empirical examples of mutation-biased adaptation collectively span two or even three regimes of adaptation, and in particular, include evolution experiments known to include rampant clonal interference in the multiple mutations regime. Mutation-biased adaptation in this regime requires large mutation bias relative to selective differences. Unsurprisingly, where relative mutation rates have been quantified, biases of appropriate magnitude were found. We can therefore infer in cases where relative mutation rates are not known, but mutation-biased adaptation has been documented in the multiple mutations regime, that similarly large mutation bias must exist. Thus, one reason that mutation-biased adaptation may be more widespread than previously claimed [3] is that differences in U can be much larger than differences in s.

(b) Variation in s

We have assumed throughout that all beneficial mutations affecting the same trait have the same selection coefficient. Good et al. [35] and Good and Desai [36] have shown that when there is a continuous distribution of fitness effects (DFE), one obtains effective parameters U^* and s^* provided that the beneficial DFE decays faster than exponentially. These effective parameters correspond to the typical beneficial mutations that drive adaptation, and provide expressions $v(U^*, s^*)$ similar in form to Equations (2)-(4). This suggests that the results of our analysis are also likely to apply using U^* and s^* instead of U and s.

(c) Deleterious mutations

For mathematical convenience in the diffusive-mutation regime, we assumed throughout that the deleterious mutation rate is equal to the beneficial mutation rate. In fact, deleterious mutations are much more common. Representing this asymmetry would have no effect on the dynamics of adaptation in the origin-fixation regime. In the other two regimes, assuming symmetry in effect sizes, larger deleterious mutation rates would primarily alter the shape of the traveling wave [23, 46, 36], rather than alter v. Deleterious mutations do contribute to v when their effect sizes are smaller, but do so in a linear manner, resulting in $v = v_b - v_d$ [36]. Our analysis then has its basis in (v_b), such that the elasticities we calculate should remain valid.

(d) Conclusion

Both selective advantages s and beneficial mutation rates U determine adaptation rates, which in turn determine which trait will dominate the adaptive process. Differences in U need to be 2 to 10 times as large, depending on the parameter value regime, in order to swamp differences in s. Both molecular and developmental biases can create such large differences in U, leading to mutation-biased adaptation. Diverse case studies suggest that mutation bias significantly shapes which adaptations occur, even in populations with strong clonal interference. While adaptation does not occur without natural selection, which adaptation occurs among the many possibilities has more complex causes.

Acknowledgments

We thank Benjamin Good for helpful discussions and Arlin Stolzfus for comments on the manuscript. KG, JB, and JM were funded by the National Science Foundation (DEB-1348262), KG by the National Institutes of Health (T32 GM084905) and JB by the Environmental Resilience Institute at Indiana University.

References

[1] Edward Bagnall Poulton. *Essays on evolution 1889-1907*. Oxford: Claredon Press, 1908.

[2] Julian Huxley. *Evolution. The modern synthesis*. New York; London: Harper & brothers, 1943.

[3] Erik I Svensson and David Berger. The role of mutation bias in adaptive evolution. *Trends in ecology & evolution*, 34(5): 422–434, 2019. doi: 10.1016/j.tree.2019.01.015.

[4] Massimo Pigliucci and Leonard Finkelman. The extended (evolutionary) synthesis debate: where science meets philosophy. *BioScience*, 64(6):511–516, may 2014.

[5] Kevin Laland, Tobias Uller, Marc Feldman, Kim Sterelny, Gerd B. Müller, Armin Moczek, Eva Jablonka, and John Odling-Smee. Does evolutionary theory need a rethink? - POINT Yes, urgently. *Nature*, 514(7521):161–164, 2014. ISSN 14764687. doi: 10.1038/514161a.

[6] Kevin Laland, Marcus W Feldman, Kevin N Laland, Tobias Uller, Marcus W Feldman, Kim Sterelny, Gerd B Mu, Armin Moczek, Eva Jablonka, and John Odling-smee. The extended evolutionary synthesis: Its structure, assumptions and predictions. *Proceedings of the Royal Society B: Biological Sciences*, 282(1813):20151019, aug 2015. ISSN 0962-8452. doi: 10.1098/rspb.2015.1019.

[7] Gregory A. Wray, Hopi E. Hoekstra, Douglas J. Futuyma, Richard E. Lenski, Trudy F.C. Mackay, Dolph Schluter, and Joan E. Strassmann. Does evolutionary theory need a rethink? - COUNTERPOINT No, all is well. *Nature*, 514 (7521):161–164, 2014. ISSN 14764687. doi: 10.1038/514161a.

[8] David M McCandlish and Arlin Stoltzfus. Modeling Evolution Using the Probability of Fixation: History and Implications. *The Quarterly Review of Biology*, 89(3):225–252, sep 2014. ISSN 0033-5770. doi: 10.1086/677571. URL `https://doi.org/10.1086/677571`.

[9] Joshua L Payne, Fabrizio Menardo, Andrej Trauner, Sonia Borrell, Sebastian M Gygli, Chloe Loiseau, Sebastien Gagneux, and Alex R Hall. Transition bias influences the evolution of antibiotic resistance in Mycobacterium tuberculosis. *PLOS Biology*, 17(5):e3000265, may 2019. URL `https://doi.org/10.1371/journal.pbio.3000265`.

[10] Arlin Stoltzfus. Reactionary fringe meets mutation-biased adaptation. 1. The empirical case, 2019. URL `https://sandwalk.blogspot.com/2019/06/reactionary-fringe-meets-mutation\_26.html`.

[11] Darin R Rokyta, Paul Joyce, S Brian Caudle, and Holly A Wichman. An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus. *Nature Genetics*, 37(4):441–444, 2005. ISSN 1546-1718. doi: 10.1038/ng1535. URL https://doi.org/10.1038/ng1535.

[12] Andrew M Sackman, Lindsey W McGee, Anneliese J Morrison, Jessica Pierce, Jeremy Anisman, Hunter Hamilton, Stephanie Sanderbeck, Cayla Newman, and Darin R Rokyta. Mutation-Driven Parallel Evolution during Viral Adaptation. *Molecular Biology and Evolution*, 34(12):3243–3253, 2017. ISSN 0737-4038. doi: 10.1093/molbev/msx257. URL https://doi.org/10.1093/molbev/msx257.

[13] Alejandro Couce, Alexandro Rodríguez-Rojas, and Jesús Blázquez. Bypass of genetic constraints during mutator evolution to antibiotic resistance. *Proceedings. Biological sciences*, 282(1804):20142698, apr 2015. ISSN 1471-2954. doi: 10.1098/rspb.2014.2698. URL https://www.ncbi.nlm.nih.gov/pubmed/25716795.

[14] R C MacLean, G G Perron, and A Gardner. Diminishing returns from beneficial mutations and pervasive epistasis shape the fitness landscape for rifampicin resistance in Pseudomonas aeruginosa. *Genetics*, 186(4):1345–1354, dec 2010. ISSN 1943-2631. doi: 10.1534/genetics.110.123083. URL https://www.ncbi.nlm.nih.gov/pubmed/20876562.

[15] Arlin Stoltzfus and David M McCandlish. Mutational biases influence parallel adaptation. *Molecular Biology and Evolution*, 34(9):2163–2172, sep 2017.

[16] Jay F. Storz, Chandrasekhar Natarajan, Anthony V. Signore, Christopher C. Witt, David M. McCandlish, and Arlin Stoltzfus. The role of mutation bias in adaptive molecular evolution: insights from convergent changes in protein function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374(1777):20180238, jul 2019. doi: 10.1098/rstb.2018.0238.

[17] Lev Y Yampolsky and Arlin Stoltzfus. Bias in the introduction of variation as an orienting factor in evolution. *Evolution & Development*, 3(2):73–83, 2001.

[18] J B S Haldane. A Mathematical Theory of Natural and Artificial Selection, Part V: Selection and Mutation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 23(7): 838–844, 1927. doi: 10.1017/S0305004100015644.

[19] Arlin Stoltzfus. Reactionary fringe meets mutation-biased adaptation. 6. What "limits" adaptation?, 2019. URL https://sandwalk.blogspot.com/2019/08/reactionary-fringe-meets-mutation.html.

[20] Philip J Gerrish and Richard E Lenski. The fate of competing beneficial mutations in an asexual population. *Genetica*, 102/103:127–144, 1998. ISSN 1573-6857. doi: 10.1023/A:1017067816551.

[21] Craig R Miller, Paul Joyce, and Holly A Wichman. Mutational effects and population dynamics during viral adaptation challenge current models. *Genetics*, 187(1):185–202, 2011.

[22] Igor M Rouzine, John Wakeley, and John M Coffin. The solitary wave of asexual evolution. *Proceedings of the National Academy of Sciences*, 100(2):587–592, 2003.

[23] Michael M Desai and Daniel S Fisher. Beneficial mutation–selection balance and the effect of linkage on positive selection. *Genetics*, 176(3):1759–1798, 2007.

[24] Su-Chan Park, Damien Simon, and Joachim Krug. The speed of evolution in large asexual populations. *Journal of Statistical Physics*, 138(1):381–410, 2010.

[25] Oskar Hallatschek. The noisy edge of traveling waves. *Proceedings of the National Academy of Sciences*, 108(5):1783–1787, 2011.

[26] Kevin Gomez, Jason Bertram, and Joanna Masel. Directional selection rather than functional constraints can shape the G matrix in rapidly adapting asexuals. *Genetics*, 211(2):715–729, 2019. URL https://doi.org/10.1534/genetics.118.301685.

[27] Guy Sella and Aaron E Hirsh. The application of statistical physics to evolutionary biology. *Proceedings of the National Academy of Sciences*, 102(27):9541–9546, 2005.

[28] Michael M Desai, Aleksandra M Walczak, and Daniel S Fisher. Genetic Diversity and the Structure of Genealogies in Rapidly Adapting Populations. *Genetics*, 193(2):565–585, 2013. ISSN 0016-6731. doi: 10.1534/genetics.112.147157. URL http://www.genetics.org/content/193/2/565.

[29] Ronald Aylmer Fisher. *The Genetical Theory of Natural Selection: A Complete Variorum Edition*. Oxford University Press, Oxford, 1930.

[30] J. B. S. Haldane. *The causes of evolution*. Longmans, Green and Co., New York, 1932.

[31] James F Crow and Motoo Kimura. *An introduction to population genetics theory*. Harper & Row, New York, 1970.

[32] Michael T Pearce and Daniel S Fisher. Rapid adaptation in large populations with very rare sex: Scalings and spontaneous oscillations. *Theoretical Population Biology*, 129:18–40, October 2019. ISSN 0040-5809. doi: https://doi.org/10.1016/j.tpb.2017.11.005. URL http://www.sciencedirect.com/science/article/pii/S0040580917301442.

[33] Harold Kushner and G George Yin. *Stochastic approximation and recursive algorithms and applications*, volume 35. Springer Science & Business Media, 2003.

[34] Sandeep Venkataram, Ross Monasky, Shohreh H Sikaroodi, Sergey Kryazhimskiy, and Betul Kacar. Evolutionary stalling in the optimization of the translation machinery. *bioRxiv*, page 850644, 2019.

[35] Benjamin H Good, Igor M Rouzine, Daniel J Balick, Oskar Hallatschek, and Michael M Desai. Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. *Proceedings of the National Academy of Sciences*, 109(13):4950–4955, 2012.

[36] Benjamin H Good and Michael M Desai. Deleterious passengers in adapting populations. *Genetics*, 198(3):1183–1208, 2014.

[37] Sara Willis and Joanna Masel. Gene Birth Contributes to Structural Disorder Encoded by Overlapping Genes. *Genetics*, 210(1):303 LP – 313, sep 2018. doi: 10.1534/genetics.118.301249. URL http://www.genetics.org/content/210/1/303.abstract.

[38] Andrew F Schober, Andrew D Mathis, Christine Ingle, Junyoung O Park, Li Chen, Joshua D Rabinowitz, Ivan Junier, Olivier Rivoire, and Kimberly A Reynolds. A two-enzyme adaptive unit within bacterial folate metabolism. *Cell reports*, 27(11):3359–3370, 2019.

[39] João V Rodrigues and Eugene I Shakhnovich. Adaptation to mutational inactivation of an essential gene converges to an accessible suboptimal fitness peak. *Elife*, 8, 2019.

[40] Sasha F Levy, Jamie R Blundell, Sandeep Venkataram, Dmitri A Petrov, Daniel S Fisher, and Gavin Sherlock. Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature*, 519(7542):181–186, mar 2015. ISSN 0028-0836. URL http://dx.doi.org/10.1038/nature14279.

[41] Benjamin H. Good and Michael M. Desai. The impact of macroscopic epistasis on long-term evolutionary dynamics. *Genetics*, 199(1):177–190, 2015. ISSN 19432631. doi: 10.1534/genetics.114.172460.

[42] Benjamin H Good, Michael J McDonald, Jeffrey E Barrick, Richard E Lenski, and Michael M Desai. The dynamics of molecular evolution over 60,000 generations. *Nature*, 551 (7678):45–50, 2017.

[43] C. G. Cupples and J. H. Miller. A set of lacz mutations in escherichia coli that allow rapid detection of each of the six base substitutions. *Proc Natl Acad Sci USA*, 86(14): 5345, July 1989. URL `http://www.pnas.org/content/86/14/5345.abstract`.

[44] Tracey Seier, Dana R. Padgett, Gal Zilberberg, Vincent A. Sutera, Noor Toha, and Susan T. Lovett. Insights into mutagenesis using escherichia coli chromosomal lacz strains that enable detection of a wide spectrum of mutational events. *Genetics*, 188(2):247, June 2011. URL `http://www.genetics.org/content/188/2/247.abstract`.

[45] J. John Vidmar and Claire G. Cupples. Muty repair is mutagenic in mutt - strains of escherichia coli. *Can. J. Microbiol.*, 39(9):892–894, January 2020. ISSN 0008-4166. doi: 10.1139/m93-133. URL `https://doi.org/10.1139/m93-133`.

[46] Daniel S Fisher. Asexual evolution waves: fluctuations and universality. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(01):P01011, 2013.

Appendix

Here we derive Eq. (5), which states that the derivative of the contour of v gives the ratio of elasticities. Consider the solution $U(s)$ to the equation $v(U(s), s) = v_c$ (v_c constant), where Eq. (2), Eq. (3), and Eq. (4) give $v(U, s)$ in their respective regimes. Implicit differentiation of both sides of $v(U(s), s) = v_c$ with respect to s, and application of the chain rule provides

$$\frac{\partial v}{\partial U}\frac{dU}{ds} + \frac{\partial v}{\partial s} = 0.$$

Multiplying this expression by $\frac{(\Delta s)^2}{Uv}\frac{dU}{ds}$ gives

$$0 = \left(\frac{\Delta s}{v}\frac{\partial v}{\partial U}\right)\left(\frac{\Delta s}{U}\frac{dU}{ds}\right) + \left(\frac{\Delta s}{v}\frac{\partial v}{\partial s}\right)\left(\frac{\Delta s}{s}\right)\left(\frac{s}{U}\frac{dU}{ds}\right). \quad (6)$$

For small changes in the selection coefficient Δs, $\frac{\Delta U}{U} \approx \frac{\Delta s}{U}\frac{dU}{ds}$, and consequently

$$\frac{\Delta_u v}{v} \approx \frac{\Delta s}{v}\frac{\partial v}{\partial U} \qquad \text{and} \qquad \frac{\Delta_s v}{v} \approx \frac{\Delta s}{v}\frac{\partial v}{\partial s}.$$

In addition,

$$\frac{d(\log U)}{d(\log s)} = \frac{s}{U}\frac{dU}{ds} \quad (7)$$

also holds, where $\frac{d(\log U)}{d(\log s)}$ is the slope of the v contour in the $\log(U) - \log(s)$ plane. Substituting for these terms in Eq. (6) and rearranging provides

$$-\frac{d(\log U)}{d(\log s)} = \lim_{\Delta v \to 0} \frac{\left(\frac{\Delta_s v}{v}\right) / \left(\frac{\Delta s}{s}\right)}{\left(\frac{\Delta_u v}{v}\right) / \left(\frac{\Delta U}{U}\right)}.$$

The left side is the slope of a v contour line in $\log U$ vs. $\log s$ space, scaled by -1 and evaluated at s, while the right side is the ratio E_s/E_U in Eq. (5). This establishes Eq. (5). It is worth noting that relative changes are approximately equal to changes in the logarithm of a quantity for small changes, eg. $\Delta s/s \approx \Delta \log s$.