

Evolution of microbial growth traits under serial dilution

Jie Lin,¹ Michael Manhart,^{2,3} and Ariel Amir¹

¹*John A. Paulson School of Engineering and Applied Sciences,
Harvard University, Cambridge, MA 02138, USA*

²*Department of Chemistry and Chemical Biology,
Harvard University, Cambridge, MA 02138, USA*

³*Institute of Integrative Biology, ETH Zurich, 8092 Zurich, Switzerland*

(Dated: August 4, 2020)

Selection of mutants in a microbial population depends on multiple cellular traits. In serial-dilution evolution experiments, three key traits are the lag time when transitioning from starvation to growth, the exponential growth rate, and the yield (number of cells per unit resource). Here we investigate how these traits evolve in laboratory evolution experiments using a minimal model of population dynamics, where the only interaction between cells is competition for a single limiting resource. We find that the fixation probability of a beneficial mutation depends on a linear combination of its growth rate and lag time relative to its immediate ancestor, even under clonal interference. The relative selective pressure on growth rate and lag time is set by the dilution factor; a larger dilution factor favors the adaptation of growth rate over the adaptation of lag time. The model shows that yield, however, is under no direct selection. We also show how the adaptation speeds of growth and lag depend on experimental parameters and the underlying supply of mutations. Finally, we investigate the evolution of covariation between these traits across populations, which reveals that the population growth rate and lag time can evolve a nonzero correlation even if mutations have uncorrelated effects on the two traits. Altogether these results provide useful guidance to future experiments on microbial evolution.

Laboratory evolution experiments in microbes have provided insight into many aspects of evolution [1–3], such as the speed of adaptation [4], the nature of epistasis [5], the distribution of selection coefficients from spontaneous mutations [6], mutation rates [7], the spectrum of adaptive genomic variants [8], and the preponderance of clonal interference [9]. Despite this progress, links between the selection of mutations and their effects on specific cellular traits have remained poorly characterized. Growth traits — such as the lag time when transitioning from starvation to growth, the exponential growth rate, and the yield (resource efficiency) — are ideal candidates for investigating this question. Their association with growth means they have relatively direct connections to selection and population dynamics. Furthermore, high-throughput techniques can measure these traits for hundreds of genotypes and environments [10–13]. Numerous experiments have shown that single mutations can be pleiotropic, affecting multiple growth traits simultaneously [14, 15]. More recent experiments have even measured these traits at the single-cell level, revealing substantial non-genetic heterogeneity [10, 13, 16]. Several evolution experiments have found widespread evidence of adaptation in these traits [17–20]. This data altogether indicates that covariation in these traits is pervasive in microbial populations.

There have been a few previous attempts to develop quantitative models to describe evolution of these traits. For example, Vasi *et al.* [17] considered data after 2000 generations of evolution in *Escherichia coli* to estimate how much adaptation was attributable to different growth traits. Smith [21] developed a mathematical model to study how different traits would allow strains to

either fix, go extinct, or coexist. Wahl and Zhu [22] studied the fixation probability of mutations affecting different growth traits separately (non-pleiotropic), especially to identify which traits were most likely to acquire fixed mutations and the importance of mutation occurrence time and dilution factor. However, simple quantitative results that can be used to interpret experimental data have remained lacking. More recent work [23, 24] derived a quantitative relation between growth traits and selection, showing that selection consists of additive components on the lag and growth phases. However, this did not address the consequences of this selection for evolution, especially the adaptation of trait covariation.

In this work we investigate a minimal model of evolutionary dynamics in which cells interact only by competition for a single limiting resource. We find that the fixation probability of a mutation is accurately determined by a linear combination of its change in growth rate and change in lag time relative to its immediate ancestor, rather than depending on the precise combination of traits; the relative weight of these two components is determined by the dilution factor. Yield, on the other hand, is under no direct selection. This is true even in the presence of substantial clonal interference, where the mutant’s immediate ancestor may have large a fitness difference with the population mean. We provide quantitative predictions for the speed of adaptation of growth rate and lag time as well as their evolved covariation. Specifically, we find that even in the absence of an intrinsic correlation between growth and lag due to mutations, these traits can evolve a nonzero correlation due to selection and variation in number of fixed mutations.

I. MATERIALS AND METHODS

A. Model of population dynamics

We consider a model of asexual microbial cells in a well-mixed batch culture, where the only interaction between different strains is competition for a single limiting resource [23, 24]. Each strain k is characterized by a lag time L_k , growth rate r_k , and yield Y_k (see Fig. 1a for a two-strain example). Here the yield is the number of cells per unit resource [17], so that $N_k(t)/Y_k$ is the amount of resources consumed by time t by strain k , where $N_k(t)$ is the number of cells of strain k at time t . We define R to be the initial amount of the limiting resource and assume different strains interact only by competing for the limiting resource; their growth traits are the same as when they grow independently. When the population has consumed all of the initial resource, the population reaches stationary phase with constant size. The saturation time t_c at which this occurs is determined by $\sum_{\text{strain } k} N_k(t_c)/Y_k = R$, which we can write in terms of the growth traits as

$$\sum_{\text{strain } k} \frac{N_0 x_k e^{r_k(t_c - L_k)}}{Y_k} = R, \quad (1)$$

where N_0 is the total population size and x_k is the frequency of each strain k at the beginning of the growth cycle. In Eq. 1 we assume the time t_c is longer than each strain's lag time L_k . Note that some of our notation differs from related models in previous work, some of which used g for growth rate and λ for lag time [23], while others used λ for growth rate [25]. Although it is possible to extend the model to account for additional growth traits such as a death rate or lag and growth on secondary resources, here we focus on the minimal set of traits most often measured in microbial phenotyping experiments [10–12, 14–16, 18, 26].

We define the selection coefficient between each pair of strains as the change in their log-ratio over the complete growth cycle [27, 28]:

$$\begin{aligned} s_{ij} &= \ln \left(\frac{N_i^{\text{final}}}{N_j^{\text{final}}} \right) - \ln \left(\frac{N_i^{\text{initial}}}{N_j^{\text{initial}}} \right) \\ &= r_i(t_c - L_i) - r_j(t_c - L_j), \end{aligned} \quad (2)$$

where N_i^{initial} is the population size of strain i at the beginning of the growth cycle and N_i^{final} is the population size of strain i at the end. After the population reaches stationary phase, it is diluted by a factor of D into a fresh medium with amount R of the resource, and the cycle repeats (Fig. 1a). We assume the population remains in the stationary phase for a sufficiently short time such that we can ignore death and other dynamics during this phase [29, 30].

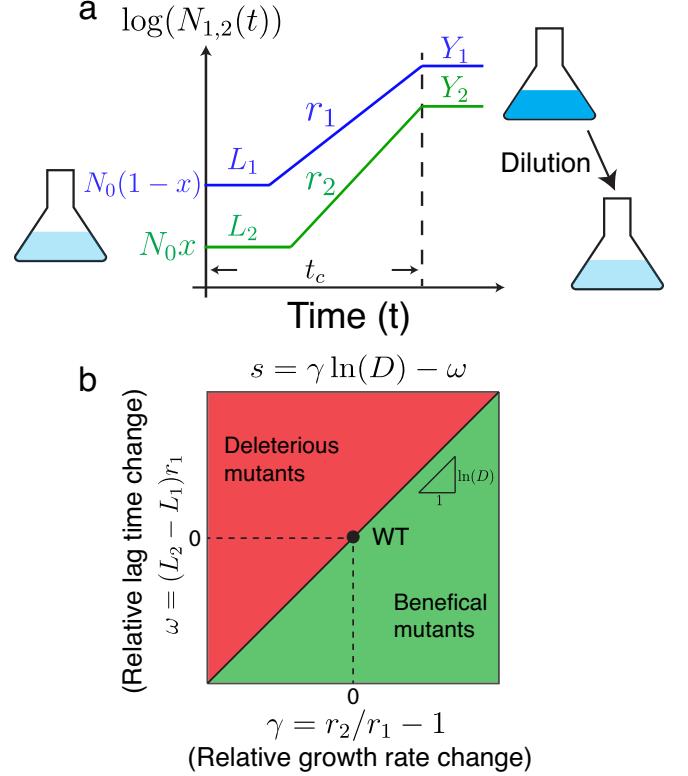


FIG. 1. Model of selection on multiple microbial growth traits. (a) Simplified model of microbial population growth characterized by three traits: lag time L , growth rate r , and yield Y . The total initial population size is N_0 and the initial frequency of the mutant (strain 2) is x . After the whole population reaches stationary phase (time t_c), the population is diluted by a factor D into fresh media, and the cycle starts again. (b) Phase diagram of selection on mutants in the space of their growth rate $\gamma = r_2/r_1 - 1$ and lag time $\omega = (L_2 - L_1)r_1$ relative to a wild-type. The slope of the diagonal line is $\ln D$.

Over many cycles of growth, as would occur in a laboratory evolution experiment [1, 28, 31], the population dynamics of this system are characterized by the set of frequencies x_k for all strains as well as the matrix of selection coefficients s_{ij} and the total population size N_0 at the beginning of each cycle. In Supplementary Methods (Secs. I, II, III) we derive explicit equations for the deterministic dynamics of these quantities over multiple cycles of growth for an arbitrary number of strains. In the case of two strains, such as a mutant and a wild-type, the selection coefficient is approximately

$$s \approx \gamma \ln D - \omega, \quad (3)$$

where $\gamma = (r_2 - r_1)/r_1$ is the growth rate of the mutant relative to the wild-type and $\omega = (L_2 - L_1)r_1$ is the relative lag time. The approximation is valid as long as the growth rate difference between the mutant and the wild-type is small (Supplementary Methods Sec. IV), which is true for most single mutations [6, 32]. This equation

shows that the growth phase and the lag phase make distinct additive contributions to the total selection coefficient, with the dilution factor D controlling their relative magnitudes (Fig. 1b). This is because a larger dilution factor will increase the amount of time the population grows exponentially, hence increasing selection on growth rate. Neutral coexistence between multiple strains is therefore possible if these two selection components balance ($s = 0$), although it requires an exact tuning of the growth traits with the dilution factor (Supplementary Methods Sec. III) [23, 24]. With a fixed dilution factor D , the population size N_0 at the beginning of each growth cycle changes according to (Supplementary Methods Sec. I):

$$N_0 = \frac{R\bar{Y}}{D}, \quad (4)$$

where $\bar{Y} = (\sum_{\text{strain } k} x_k/Y_k)^{-1}$ is the effective yield of the whole population in the current growth cycle. In this manner the ratio R/D sets the bottleneck size of the population, which for serial dilution is approximately the effective population size [31], and therefore determines the strength of genetic drift.

B. Model of evolutionary dynamics

We now consider the evolution of a population as new mutations arise that alter growth traits. We start with a wild-type population having lag time $L_0 = 100$ and growth rate $r_0 = (\ln 2)/60 \approx 0.012$, which are roughly consistent with *E. coli* parameters where time is measured in minutes [17, 31]; we set the wild-type yield to be $Y_0 = 1$ without loss of generality. As in experiments, we vary the dilution factor D and the amount of resources R , which control the relative selection on growth versus lag (set by D , Eq. 3) and the effective population size (set by R/D , Eq. 4). We also set the initial population size of the first cycle to $N_0 = RY_0/D$.

The population grows according to the dynamics in Fig. 1a. Each cell division can generate a new mutation with probability $\mu = 10^{-6}$; note this rate is only for mutations altering growth traits, and therefore it is lower than the rate of mutations anywhere in the genome. We generate a random waiting time τ_k for each strain k until the next mutation with instantaneous rate $\mu r_k N_k(t)$. When a mutation occurs, the growth traits for the mutant are drawn from a distribution $p_{\text{mut}}(r_2, L_2, Y_2 | r_1, L_1, Y_1)$, where r_1, L_1, Y_1 are the growth traits for the background strain on which the new mutation occurs and r_2, L_2, Y_2 are the traits for the new mutant. Note that since mutations only arise during the exponential growth phase, beneficial or deleterious effects on lag time are not realized until the next growth cycle [20]. After the growth cycle ceases (once the resource is exhausted according to Eq. 1), we randomly choose cells, each with probability $1/D$, to form the population for the next growth cycle.

We will assume mutational effects are not epistatic and scale with the trait values of the background strain, so that $p_{\text{mut}}(r_2, L_2, Y_2 | r_1, L_1, Y_1) = p_{\text{mut}}(\gamma, \omega, \delta)$, where $\gamma = (r_2 - r_1)/r_1$, $\omega = (L_2 - L_1)/L_1$, and $\delta = (Y_2 - Y_1)/Y_1$ (Supplementary Methods Sec. V). Since our primary goal is to scan the space of possible mutations, we focus on uniform distributions of mutational effects where $-0.02 < \gamma < 0.02$, $-0.05 < \omega < 0.05$, and $-0.02 < \delta < 0.02$. In the Supplementary Methods we extend our main results to the case of Gaussian distributions (Sec. V) as well as an empirical distribution of mutational effects based on single-gene deletions in *E. coli* (Sec. VI) [33].

C. Data Availability

Data and codes are available upon request. File S1 contains the Supplementary Methods. File S2 contains data of growth traits presented in Figure S3.

II. RESULTS

Fixation of mutations

We first consider the fixation statistics of new mutations in our model. In Fig. 2a we show the relative growth rates γ and the relative lag times ω of fixed mutations against their background strains, along with contours of constant selection coefficient s from Eq. 3. As expected, fixed mutations either increase growth rate ($\gamma > 0$), decrease lag time ($\omega < 0$), or both. In contrast, the yield of fixed mutations is the same as the ancestor on average (Fig. 2b); indeed, the selection coefficient in Eq. 3 does not depend on the yields. If a mutation arises with significantly higher or lower yield than the rest of the population, the bottleneck population size N_0 immediately adjusts to keep the overall fold-change of the population during the growth cycle fixed to the dilution factor D (Eq. 4). Therefore mutations that significantly change yield have no effect on the overall population dynamics.

Figure 2a also suggests that the density of fixed mutations in the growth-lag trait space depends solely on their selection coefficients, rather than the precise combination of traits, as long as other parameters such as the dilution factor D , the total amount of resource R , and the distribution of mutational effects are held fixed. Mathematically, this means that the fixation probability $\phi(\gamma, \omega)$ of a mutation with growth effect γ and lag effect ω can be expressed as $\phi(\gamma, \omega) = \phi(\gamma \ln D - \omega) \equiv \phi(s)$. To test this, we discretize the scatter plot of Fig. 2a and compute the fixation probabilities of mutations as functions of γ and ω (Supplementary Methods Sec. VII). We then plot the resulting fixation probabilities of mutations as functions of their selection coefficients calculated by Eq. 3 (Fig. 2c,d,e,f). We test the dependence of the fixation probability on the selection coefficient over a range

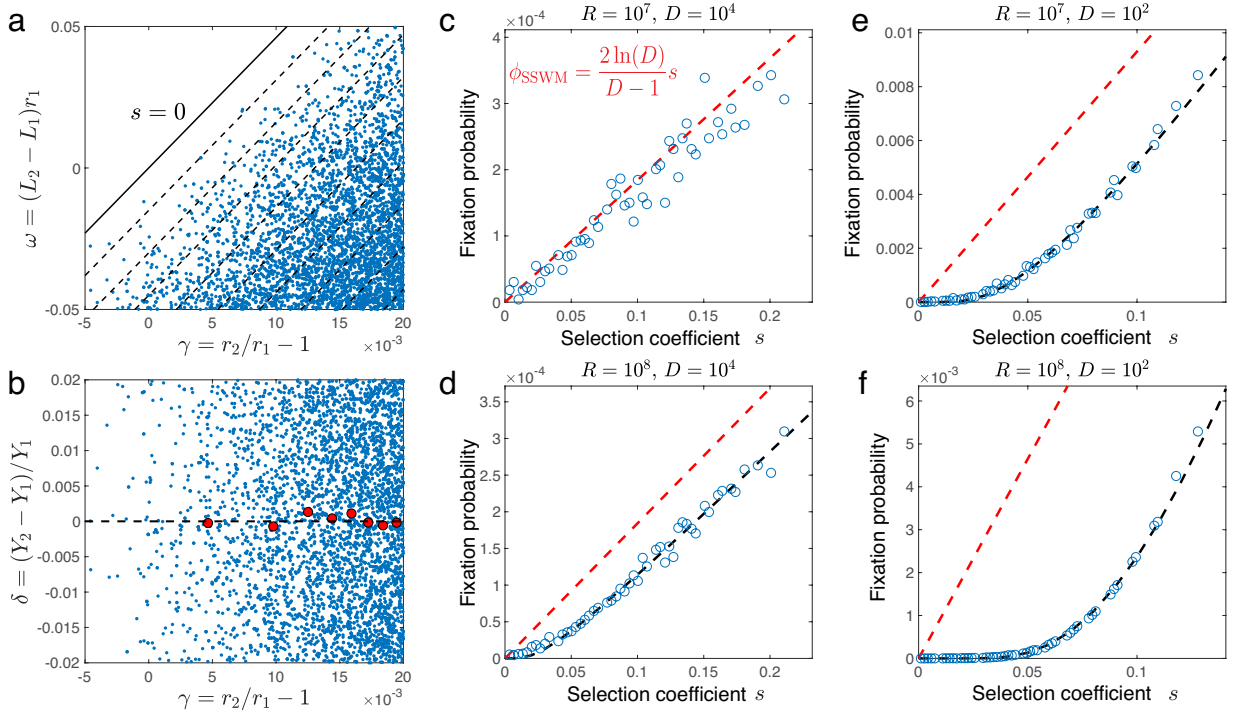


FIG. 2. Selection coefficient determines fixation probability. (a) The relative growth rates γ and the relative lag times ω of fixed mutations against their background strain. Dashed lines mark contours of constant selection coefficient with interval $\Delta s = 0.015$ while the solid line marks $s = 0$. (d) Same as (a) but for relative growth rate γ and the relative yield δ . The red dots mark the relative yield of fixed mutations averaged over binned values of the relative growth rate γ . In (a) and (d), $D = 10^2$ and $R = 10^7$. (b,c,e,f) Fixation probability of mutations against their selection coefficients for different amounts of resource R and dilution factors D as indicated in the titles. The red dashed line shows the fixation probability predicted in the SSWM regime (Eq. 5), while the black line shows a numerical fit of the data points to Eq. 6 with parameters $A = 0.1145$ and $B = 0.0801$ in (c), $A = 0.0017$ and $B = 0.0421$ in (e), and $A = 0.2121$ and $B = 0.2192$ in (f). In all panels mutations randomly arise from a uniform distribution p_{mut} with $-0.02 < \gamma < 0.02$, $-0.05 < \omega < 0.05$, and $-0.02 < \delta < 0.02$.

of population dynamics regimes by varying the dilution factor D and the amount of resources R .

For small populations, mutations generally arise and either fix or go extinct one at a time, a regime known as “strong-selection weak-mutation” (SSWM) [34]. In this case, we expect the fixation probability of a beneficial mutation with selection coefficient $s > 0$ to be [22, 35, 36]

$$\phi_{\text{SSWM}}(s) = \frac{2 \ln D}{D-1} s. \quad (5)$$

This is similar to the standard Wright-Fisher fixation probability of $2s$ [37], but with a different prefactor due to averaging over the different times in the exponential growth phase at which the mutation can arise (Supplementary Methods Sec. VIII). Indeed, we see this predicted dependence matches the simulation results for the small population size of $N_0 \sim R/D = 10^3$ (Fig. 2c).

For larger populations, multiple beneficial mutations will be simultaneously present in the population and interfere with each other, an effect known as clonal interference [38–43]. Our simulations show that, as for the SSWM case, the fixation probability depends only on the selection coefficient (Eq. 3) relative to the mutation’s im-

mediate ancestor and not on the individual combination of mutant traits (Fig. 2d,e,f), with all other population parameters held constant. Previous work has determined the dependence of the fixation probability on the selection coefficient under clonal interference using various approximations [38, 40–42]. Here, we focus on an empirical relation based on Gerrish and Lenski [38]:

$$\phi_{\text{CI}}(s) = A s e^{-B/s}, \quad (6)$$

where A and B are two constants that depend on other parameters of the population (D , R , and the distribution of mutational effects); we treat these as empirical parameters to fit to the simulation results, although Gerrish and Lenski [38] predicted $A = 2 \ln D / (D - 1)$, i.e., the same constant as in the SSWM case (Eq. 5). The $e^{-B/s}$ factor in Eq. 6 comes from the probability that no superior beneficial mutations appears before the current mutation fixes. Since the time to fixation scales as $1/s$, we expect the average number of superior mutations to be proportional to $1/s$ (for small s). This approximation holds only for selection coefficients that are not too small and therefore are expected to fix without additional beneficial mutations on the same background; Eq. 6 breaks

down for weaker beneficial mutations that typically fix by hitchhiking on stronger mutations [40]. Nevertheless, Eq. 6 matches our simulation results well for a wide range of selection coefficients achieved in our simulations and larger population sizes $N_0 \sim R/D > 10^4$ (Fig. 2d,e,f). Furthermore, the constant A we fit to the simulation data is indeed close to the predicted value of $2 \ln D / (D - 1)$, except in the most extreme case of $N_0 \sim R/D = 10^6$ (Fig. 2f).

Altogether Fig. 2 shows that mutations with different effects on cell growth — for example, a mutant that increases the growth rate and a mutant that decreases the lag time — can nevertheless have approximately the same fixation probability as long as their overall effects on selection are the same according to Eq. 3. To test the robustness of this result, we verify it for several additional distributions of mutational effects $p_{\text{mut}}(\gamma, \omega, \delta)$ in the Supplementary Methods: a Gaussian distribution of mutational effects, including the presence of correlated mutational effects (Fig. S1); a wider distribution of mutational effects with large selection coefficients (Fig. S2); and an empirical distribution of mutational effects estimated from single-gene deletions in *E. coli* (Fig. S3). In Fig. S4a we further test robustness by using the neutral phenotype (orthogonal to the selection coefficient) to quantify the range of γ and ω trait combinations that nevertheless have the same selection coefficient and fixation probability, and in Fig. S4b we show that the selection coefficient on growth alone is insufficient to determine fixation probability.

While the dependence of fixation probability on the selection coefficient is a classic result of population genetics [44], the existence of a simple relationship here is nontrivial since, strictly speaking, selection in this model is not only frequency-dependent [23] (i.e., selection between two strains depends on their frequencies) but also includes higher-order effects [24] (i.e., selection between strain 1 and strain 2 is affected by the presence of strain 3). Therefore in principle, the fixation probability of a mutant may depend on the specific state of the population in which it is present, while the selection coefficient in Eq. 3 only describes selection on the mutant in competition with its immediate ancestor. However, we see that, at least for the parameters considered in our simulations, these effects are negligible in determining the eventual fate of a mutation.

Adaptation of growth traits

As Fig. 3a shows, many mutations arise and fix over the timescale of our simulations, which lead to predictable trends in the quantitative traits of the population. We first determine the relative fitness of the evolved population at each time point against the ancestral strain by simulating competition between an equal number of evolved and ancestral cells for one cycle, analogous to common experimental measurements [1, 31]. The result-

ing fitness trajectories are shown in Fig. 3b. To see how different traits contribute to the fitness increase, we also calculate the average population traits at the beginning of each cycle; for instance, the average population growth rate at growth cycle n is $r_{\text{pop}}(n) = \sum_{\text{strain } k} r_k x_k(n)$. As expected from Eq. 3, the average growth rate increases (Fig. 3c) and the average lag time decreases (Fig. 3d) for all simulations. In contrast, the average yield evolves without apparent trend (Fig. 3e), since Eq. 3 indicates no direct selection on yield. We note that, while the cells do not evolve toward lower or higher resource efficiency on average, they do evolve to consume resources more quickly, since the rate of resource consumption (r_k/Y_k for each cell of strain k) depends on both the yield as well as the growth rate. Therefore the saturation time of each growth cycle evolves to be shorter, consistent with recent work from Baake *et al.* [45].

Figure 3 suggests relatively constant speeds of adaptation for the relative fitness, the average growth rate, and the average lag time. For example, we can calculate the adaptation speed of the average growth rate as the averaged change in the average growth rate per cycle:

$$W_{\text{growth}} = \langle r_{\text{pop}}(n+1) - r_{\text{pop}}(n) \rangle, \quad (7)$$

where the bracket denotes an average over replicate populations and cycle number. In the Supplementary Methods (Secs. IX and X) we calculate the adaptation speeds of these traits in the SSWM regime to be

$$\begin{aligned} W_{\text{growth}} &= \sigma_\gamma^2 r_0 (\ln D) \left(\frac{\mu R Y_0 \ln D}{D - 1} \right), \\ W_{\text{lag}} &= -\frac{\sigma_\omega^2}{r_0} \left(\frac{\mu R Y_0 \ln D}{D - 1} \right), \\ W_{\text{fitness}} &= \frac{W_{\text{growth}}}{r_0} \ln D - W_{\text{lag}} r_0, \end{aligned} \quad (8)$$

where σ_γ and σ_ω are the standard deviations of the underlying distributions of γ and ω for single mutations ($p_{\text{mut}}(\gamma, \omega, \delta)$), r_0 is the ancestral growth rate and Y_0 the ancestral yield (we assume the yield does not change on average according to Fig. 3e). Furthermore, the ratio of the growth adaptation rate and the lag adaptation rate is independent of the amount of resource and mutation rate in the SSWM regime:

$$\frac{W_{\text{growth}}}{W_{\text{lag}}} = -r_0^2 \frac{\sigma_\gamma^2}{\sigma_\omega^2} \ln D. \quad (9)$$

Equation 8 predicts that the adaptation speeds of the average growth rate, the average lag time, and the relative fitness should all increase with the amount of resources R and decrease with the dilution factor D (for large D); even though this prediction assumes the SSWM regime (relatively small $N_0 \sim R/D$), it nevertheless holds

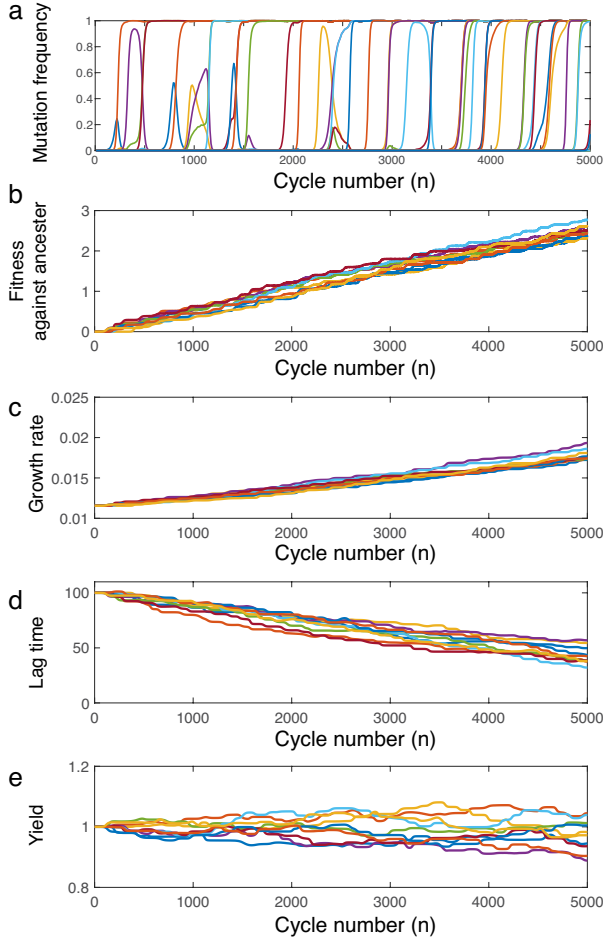


FIG. 3. Dynamics of evolving populations. (a) Frequencies of new mutations as functions of the number n of growth cycles. Example trajectories of (b) the fitness of the evolved population relative to the ancestral population, (c) the evolved average growth rate, (d) the evolved average lag time, and (e) the evolved average yield. In all panels the dilution factor is $D = 10^2$, the amount of resource at the beginning of each cycle is $R = 10^7$, and mutations randomly arise from a uniform distribution p_{mut} with $-0.02 < \gamma < 0.02$, $-0.05 < \omega < 0.05$, and $-0.02 < \delta < 0.02$.

across a wide range of R and D values (Fig. 4a,b,c), except for $R = 10^8$ where the speed of fitness increase is non-monotonic with D (Fig. 4c). The predicted adaptation speeds in Eq. 8 also quantitatively match the simulated trajectories in the SSWM case (Fig. 4d,e,f); even outside of the SSWM regime, the relative rate in Eq. 9 remains a good prediction at early times (Fig. S5).

Evolved covariation between growth traits

We now turn to investigating how the covariation between traits evolves. We have generally assumed that individual mutations have uncorrelated effects on different

traits. Campos *et al.* [33] recently systematically measured the growth curves of the single-gene deletions in *E. coli*. We compute the relative growth rate, lag time, and yield changes for the single-gene deletions compared with the wild-type and find that the resulting empirical distribution of relative growth traits changes shows very small correlations between these traits (Fig. S3b,c), consistent with our assumptions. We note that these measurements, however, are subject to significant noise (Supplementary Methods Sec. VI), and therefore any conclusions ultimately require verification by further experiments.

Even in the absence of mutational correlations, selection may induce a correlation between these traits in evolved populations. In Fig. 5a we schematically depict how the raw variation of traits from mutations is distorted by selection and fixation of multiple mutations. Specifically, for a single fixed mutation, selection induces a positive (i.e., antagonistic) correlation between the relative growth rate change and the relative lag time change. Figure 2a shows this for single fixed mutations, while Fig. 5b,c shows this positive correlation between the average growth rate and the average lag time across populations that have accumulated the same number of fixed mutations. For populations in the SSWM regime with the same number of fixed mutations, the Pearson correlation coefficient between the average growth rate and the average lag time across populations is approximately equal to the covariation of the relative growth rate change γ and the relative lag time change ω for a single fixed mutation:

$$\rho_{\text{fixed}} \approx \frac{\langle \gamma \omega \rangle_{\text{fixed}} - \langle \gamma \rangle_{\text{fixed}} \langle \omega \rangle_{\text{fixed}}}{\sqrt{(\langle \gamma^2 \rangle_{\text{fixed}} - \langle \gamma \rangle_{\text{fixed}}^2)(\langle \omega^2 \rangle_{\text{fixed}} - \langle \omega \rangle_{\text{fixed}}^2)}} \quad (10)$$

where $\langle \cdot \rangle_{\text{fixed}}$ is an average over the distribution of single fixed mutations (Supplementary Methods Sec. IX). We can explicitly calculate this quantity in the SSWM regime, which confirms that it is positive for uncorrelated mutational effects with uniform or Gaussian distributions (Supplementary Methods Sec. XI).

However, in evolution experiments we typically observe populations at a particular snapshot in time, such that the populations may have a variable number of fixed mutations but the same number of total mutations that arose and either fixed or went extinct (since the number of total arising mutations is very large, we neglect its fluctuation across populations). Interestingly, the variation in number of fixed mutations at a snapshot in time causes the distribution of growth rates and lag times across populations to stretch into a negative correlation; this is an example of Simpson's paradox from statistics [46]. Figure 5a shows this effect schematically, while Fig. 5d,e show explicit results from simulations. An intuitive way to understand the evolved negative correlation is to approximate the effects of all fixed mutations as deterministic, so that each fixed mutation increases the average growth rate and decrease the average lag time by the same amount. Therefore populations with a higher aver-

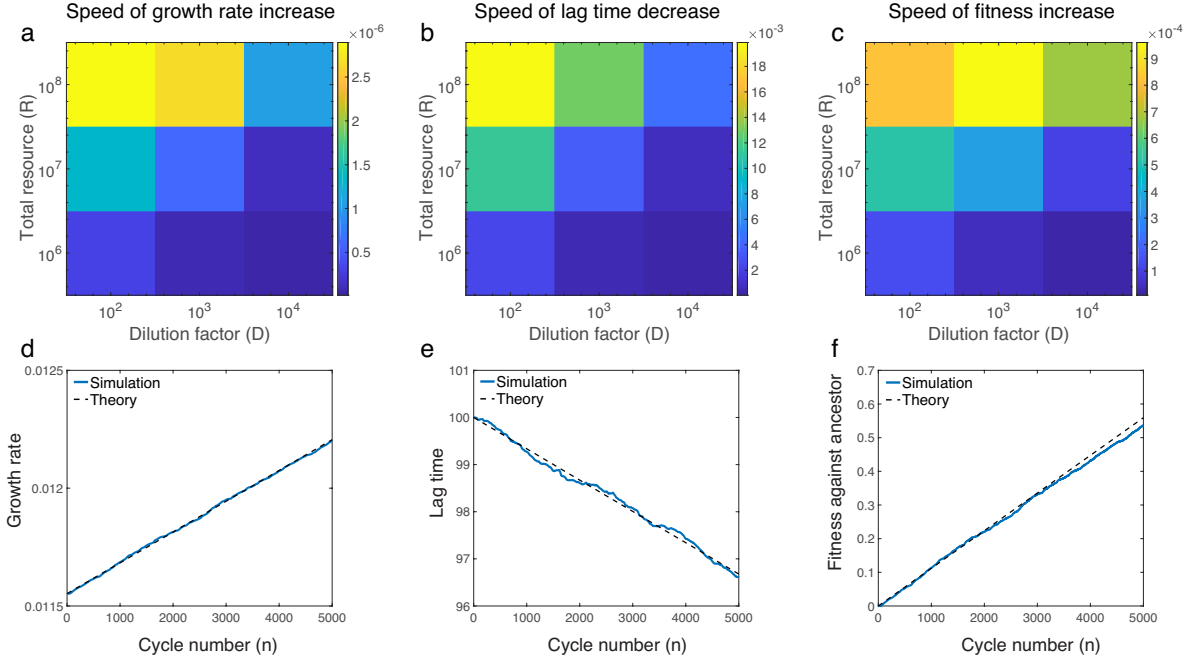


FIG. 4. Speed of adaptation. The average per-cycle adaptation speed of (a) the average growth rate, (b) the average lag time, and (c) the fitness relative to the ancestral population as functions of the dilution factor D and total amount of resources R . The adaptation speeds are averaged over growth cycles and independent populations. (d) The average growth rate, (e) the average lag time, and (f) the fitness relative to the ancestral population as functions of the number n of growth cycles. The dilution factor is $D = 10^4$ and the total resource is $R = 10^7$, so the population is in the SSWM regime. The blue solid lines are simulation results, while the dashed lines show the mathematical predictions in Eq. 8. All panels show averages over 500 independent simulated populations, with mutations randomly arising from a uniform distribution p_{mut} with $-0.02 < \gamma < 0.02$, $-0.05 < \omega < 0.05$, and $-0.02 < \delta < 0.02$.

age growth rate must have a larger number of fixed mutations and thus also a shorter average lag time, leading to a negative correlation between the average growth rates and the average lag times. In the Supplementary Methods (Sec. X), we calculate this evolved Pearson correlation coefficient across populations in the SSWM regime to be approximately

$$\rho_{\text{evo}} \approx \frac{\langle \gamma \omega \rangle_{\text{fixed}}}{\sqrt{\langle \gamma^2 \rangle_{\text{fixed}} \langle \omega^2 \rangle_{\text{fixed}}}}. \quad (11)$$

That is, the correlation of traits across populations with multiple mutations is still a function of the distribution of single fixed mutations, but it is not equal to the correlation of single fixed mutations (Eq. 10). In the Supplementary Methods (Sec. XI) we explicitly calculate ρ_{evo} in the SSWM regime for uncorrelated uniform and Gaussian distributions of mutational effects, which shows that it is negative. Furthermore, we prove that it must always be negative for any symmetric and uncorrelated distribution $p_{\text{mut}}(\gamma, \omega)$ (Supplementary Methods Sec. IX).

The predicted correlations in Eqs. 10 and 11 quantitatively match the simulations well in the SSWM regime (Fig. 5c,e). While they are less accurate outside of the SSWM regime, they nevertheless still produce the correct sign of the evolved correlation within the param-

eter regimes of our simulations (Fig. S6a,b,c). However, the signs of the correlations can indeed change depending on the underlying distribution of mutational effects $p_{\text{mut}}(\gamma, \omega, \delta)$. For example, in the Supplementary Methods we explore the effects of varying the mean mutational effects (Fig. S6d) — e.g., whether an average mutation has positive, negative, or zero effect on the growth rate — as well as the intrinsic mutational correlation between the relative growth rate change and the relative lag time change (Fig. S6e).

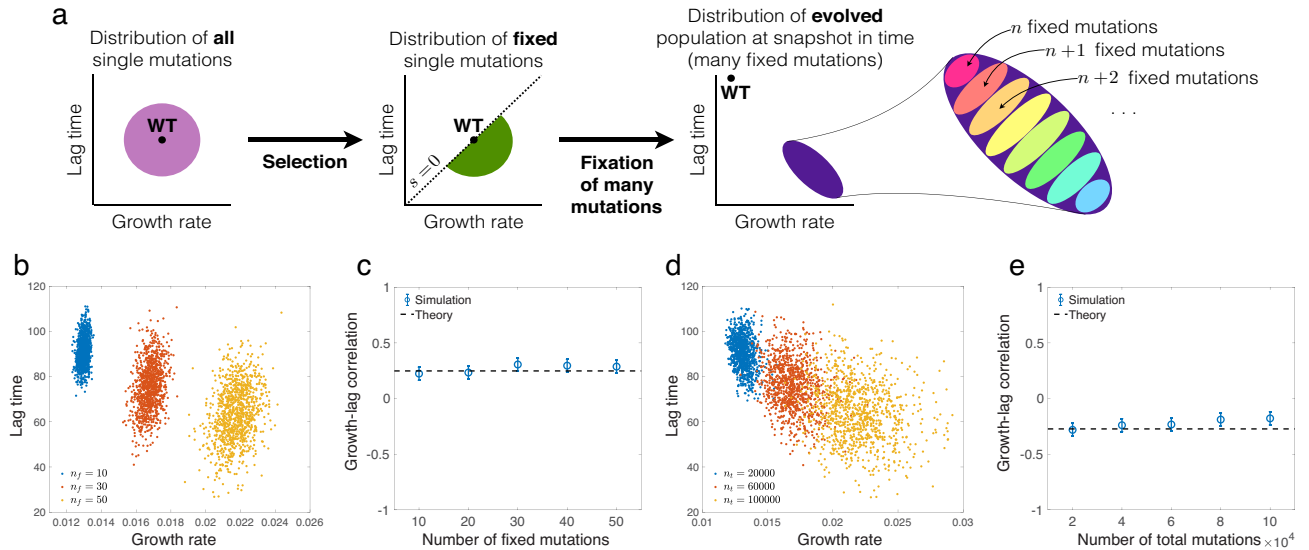


FIG. 5. Evolved patterns of covariation among growth traits. (a) Schematic of how selection and fixation of multiple mutations shape the observed distribution of traits. The sign of the Pearson correlation coefficient between the average growth rate and lag time depends on whether we consider an ensemble of populations with the same number of fixed mutations or the same number of total mutation events. (b) Distribution of average growth rate and lag time for 1000 independent populations with the same number of fixed mutations. Each color corresponds to a different number of fixed mutations (n_f) indicated in the legend. (c) Pearson correlation coefficient of growth rate and lag time for distributions in panel (b) as a function of the number of fixed mutations. The dashed line is the prediction from Eq. 10. (d) Same as (b) except each color corresponds to a set of populations at a snapshot in time with the same number of total mutation events. Each color corresponds to a different number of total mutations events (n_t) indicated in the legend. (e) Same as (c) but for the set of populations shown in (d). The dashed line is the prediction from Eq. 11. In (c) and (e) the error-bars represent 95% confidence intervals. In (b–e) we simulate the SSWM regime by introducing random mutations one-by-one and determining their fixation from Eq. 5 with $D = 10^3$.

III. DISCUSSION

We have investigated a model of microbial evolution under serial dilution, which is both a common protocol for laboratory evolution experiments [1, 6, 31, 47, 48] as well as a rough model of evolution in natural environments with feast-famine cycles. While there has been extensive work to model population and evolutionary dynamics in these conditions [2, 35, 36, 38, 45], these models have largely neglected the physiological links connecting mutations to selection. However, models that explicitly incorporate these features are necessary to interpret experimental evidence that mutations readily generate variation in multiple cellular traits, and that this variation is important to adaptation [17–20]. Wahl and Zhu [22] determined the relative fixation probabilities of mutations on different traits and the effects of mutation occurrence time and dilution factor, but the role of pleiotropy and evolutionary dynamics over many mutations were not considered.

In this paper, we have studied a model where mutations can affect three quantitative growth traits — the lag time, the exponential growth rate, and the yield (Fig. 1a) — since these three traits are widely measured for microbial populations. In particular, we have derived a simple expression (Eq. 3) for the selection coefficient of a mutation in terms of its effects on growth and lag and a single

environmental parameter, the dilution factor D . While previous work showed that this particular form of the selection coefficient determines the fixation probability of a single mutation in the SSWM regime [23], here we show that this holds even in the presence of clonal interference (Fig. 2c,d,e,f), which appears to be widespread in laboratory evolution experiments [9, 28, 49]. Our result is therefore valuable for interpreting the abundant experimental data on mutant growth traits. We have also calculated the adaptation rates of growth traits per cycle in the SSWM regime, which turn out to increase with the amount of resource R and decrease with the dilution factor D . These results are confirmed by numerical simulations and remain good predictions even outside of the SSWM regime. Furthermore, some of these results are independent of the specific form of the selection coefficient (Eq. 3), namely the fact that the fixation probability depends only on the selection coefficient (with other population parameters besides the mutant traits being held fixed) even in the clonal interference regime, and the expressions for the correlation coefficients of traits between populations (Eqs. 10 and 11).

An important difference with the previous work on this model is that here we used a fixed dilution factor D , which requires that the bottleneck population size N_0 fluctuates as the population evolves. In contrast, previous work used a fixed N_0 and variable D [23, 24]. We ob-

served two important differences between these regimes. First, in the case of fixed N_0 and variable D , the fold-change of the population during a single growth cycle, which is approximately $R\bar{Y}/N_0$ [23], determines the relative selection between growth and lag, since it determines how long the population undergoes exponential growth. Therefore one can experimentally tune this relative selection by varying either the total amount of resources R or the fixed bottleneck size N_0 . However, when the dilution factor D is fixed, the population fold-change is always constrained to exactly equal D and therefore D alone determines the relative selection on growth and lag (Eq. 3). The second difference is that, with fixed N_0 and variable D , the selection coefficient depends explicitly on the effective yield $\bar{Y}$ and is therefore frequency-dependent (Supplementary Methods Sec. II), which enables the possibility of stable coexistence between two strains [23, 24]. However, for the fixed D case, the frequency dependence of $\bar{Y}$ is exactly canceled by N_0 (Eq. 4). Therefore there is only neutral coexistence in this case, requiring the growth and lag traits of the strains to follow an exact constraint set by D (Supplementary Methods Sec. III).

A major result of our model is a prediction on the evolution of covariation between growth traits. In particular, we have shown that correlations between traits can emerge from selection and accumulation of multiple mutations even without an intrinsic correlation between traits from individual mutations (Figs. 5 and S6). We have also shown that selection alone produces no correlation between growth and yield, in the absence of correlated mutational effects (Figs. 2b and 3e). This is important for interpreting evolved patterns of traits in terms of selective or physiological tradeoffs. Specifically, it emphasizes that the evolved covariation between traits conflates both the underlying supply of variation from mutations as well as the action of selection and other aspects of population dynamics (e.g., genetic drift, spatial structure, recombination), and therefore it is difficult to make clear inferences about either aspect purely from the outcome of evolution alone. For example, simply observing a negative correlation between two traits from evolved populations is insufficient to infer whether that correlation is due to a physiological constraint on mutations (e.g., mutations cannot improve both traits simultaneously) or due to a selective constraint (e.g., selection favors specialization in one trait or another).

These questions, of course, have been the foundation of quantitative trait genetics [50]. Historically this field has emphasized polymorphic populations with abundant recombination as are applicable to plant and animal breeding. However, this regime is quite different from microbial populations which, at least under laboratory conditions, are often asexual and dominated by linkage between competing mutations [9, 28, 49]. We therefore need a quantitative description of both between-population as well as within-population covariation of traits of microbial populations in this regime. In the present study, we focus on between-population covariation in growth traits,

but recent work by Gomez *et al.* [51] provides insight into the case of within-population covariation. They showed that a tradeoff across individuals within a population evolves between two quantitative traits under positive, additive selection; this suggests that while growth rate and lag time will be negatively correlated across populations (Fig. 5d,e), they should be positively correlated within populations.

Microbial growth traits should indeed be an ideal setting for this approach due to abundant data, but conclusions on the nature of trait covariation have remained elusive. Physiological models have predicted a negative correlation between growth rate and lag time across genotypes [52, 53], while models of single-cell variation in lag times also suggests there should be a negative correlation at the whole-population level [54]. However, experimental evidence has been mixed, with some studies finding a negative correlation [13, 16], while others found no correlation [10, 11, 14]. Studies of growth-yield correlations have long been motivated by r/K selection theory, which suggests there should be tradeoffs between growth rate and yield [55]. For instance, metabolic models make this prediction [56–58]. However, experimental evidence has again been mixed, with some data showing a tradeoff [26, 59, 60], while others show no correlation [15, 18, 19, 61] or even a positive correlation [11, 47]. Some of this ambiguity may have to do with dependence on the environmental conditions [19] or the precise definition of yield. We define yield as the proportionality constant of population size to resource (Eq. 1) and neglect any growth rate dependence on resource concentration. Under these conditions, we predict no direct selection on yield, which means that the only way to generate a correlation of yield with growth rate is if the two traits are constrained at the physiological level, so that mutational effects are correlated. In such cases yield could evolve but only as a spandrel [62, 63]. Ultimately, we believe more precise single-cell measurements of these traits, both across large unselected mutant libraries as well as evolved strains, are necessary to definitively test these issues [33].

IV. ACKNOWLEDGMENTS

MM was supported by an F32 fellowship from the US National Institutes of Health (GM116217) and an Ambizione grant from the Swiss National Science Foundation (PZ00P3.180147). AA was supported by NSF CAREER grant 1752024 and the Harvard Dean’s Competitive Fund. AA and JL thank support from Harvard’s MRSEC (DMR-1420570). We thank Christine Jacobs-Wagner for kindly sharing the data of single-gene deletions of *E. coli*, and we thank Yipei Guo for suggesting the proof of the sign of the evolved growth rate and lag time correlation presented in the Supplementary Methods.

- [1] S. F. Elena and R. E. Lenski, "Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation," *Nat Rev Genet* **4**, 457–469 (2003).
- [2] M. M. Desai, "Statistical questions in experimental evolution," *J Stat Mech* **2013**, P01003 (2013).
- [3] J. E. Barrick and R. E. Lenski, "Genome dynamics during experimental evolution," *Nat Rev Genet* **14**, 827–839 (2013).
- [4] Michael J. Wiser, Noah Ribeck, and Richard E. Lenski, "Long-term dynamics of adaptation in asexual populations," *Science* (2013).
- [5] S. Kryazhimskiy, D. P. Rice, E. R. Jerison, and M. M. Desai, "Global epistasis makes adaptation predictable despite sequence-level stochasticity," *Science* **344**, 1519–1522 (2014).
- [6] S. F. Levy, J. R. Blundell, S. Venkataram, D. A. Petrov, D. S. Fisher, and G. Sherlock, "Quantitative evolutionary dynamics using high-resolution lineage tracking," *Nature* **519**, 181–186 (2015).
- [7] S. Wielgoss, J. E. Barrick, O. Tenaillon, S. Cruveiller, B. Chane-Woon-Ming, C. Médigue, R. E. Lenski, and D. Schneider, "Mutation rate inferred from synonymous substitutions in a long-term evolution experiment with *Escherichia coli*," *G3* **1**, 183–186 (2011).
- [8] J. E. Barrick, D. S. Yu, S. H. Yoon, H. Jeong, T. K. Oh, D. Schneider, R. E. Lenski, and J. F. Kim, "Genome evolution and adaptation in a long-term experiment with *Escherichia coli*," *Nature* **461**, 1243–1247 (2009).
- [9] G. I. Lang, D. P. Rice, M. J. Hickman, E. Sodergren, G. M. Weinstock, D. Botstein, and M. M. Desai, "Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations," *Nature* **500**, 571–574 (2013).
- [10] I. Levin-Reisman, O. Gefen, O. Fridman, I. Ronin, D. Shwa, H. Sheftel, and N. Q. Balaban, "Automated imaging with ScanLag reveals previously undetectable bacterial growth phenotypes," *Nat Methods* **7**, 737–739 (2010).
- [11] J. Warringer, E. Zörgö, F. A. Cubillos, A. Zia, A. Gjuvsland, J. T. Simpson, A. Forsmark, R. Durbin, S. W. Omholt, E. J. Louis, G. Liti, A. Moses, and A. Blomberg, "Trait variation in yeast is defined by population history," *PLOS Genet* **7**, e1002111 (2011).
- [12] M. Zackrisson, J. Hallin, L.-G. Ottosson, P. Dahl, E. Fernandez-Parada, E. Ländström, L. Fernandez-Ricaud, P. Kaferle, A. Skyman, S. Stenberg, S. Omholt, U. Petrovic, J. Warringer, and A. Blomberg, "Scan-omatic: High-resolution microbial phenomics at a massive scale," *G3* **6**, 3003–3014 (2016).
- [13] N. Ziv, B. M. Shuster, M. L. Siegal, and D. Gresham, "Resolving the complex genetic basis of phenotypic variation and variability of cellular growth," *Genetics* **206**, 1645–1657 (2017).
- [14] B. V. Adkar, M. Manhart, S. Bhattacharyya, J. Tian, M. Musharbash, and E. I. Shakhnovich, "Optimization of lag phase shapes the evolution of a bacterial enzyme," *Nat Ecol Evol* **1**, 0149 (2017).
- [15] J. M. Fitzsimmons, S. E. Schoustra, J. T. Kerr, and R. Kassen, "Population consequences of mutational events: effects of antibiotic resistance on the r/K trade-off," *Evol Ecol* **24**, 227–236 (2010).
- [16] N. Ziv, M. L. Siegal, and D. Gresham, "Genetic and Nongenetic Determinants of Cell Growth Variation Assessed by High-Throughput Microscopy," *Molecular Biology and Evolution* **30**, 2568–2578 (2013).
- [17] F. Vasi, M. Travisano, and R. E. Lenski, "Long-term experimental evolution in *Escherichia coli*. II. changes in life-history traits during adaptation to a seasonal environment," *Am Nat* **144**, 432–456 (1994).
- [18] M. Novak, T. Pfeiffer, R. E. Lenski, U. Sauer, and S. Bonhoeffer, "Experimental tests for an evolutionary trade-off between growth rate and yield in *E. coli*," *Am Nat* **168**, 242–251 (2006).
- [19] C. Reding-Roman, M. Hewlett, S. Duxbury, F. Gori, I. Gudelj, and R. Beardmore, "The unconstrained evolution of fast and efficient antibiotic-resistant bacterial genomes," *Nat Ecol Evol* **1**, 0050 (2017).
- [20] Y. Li, S. Venkataram, A. Agarwala, B. Dunn, D. A. Petrov, G. Sherlock, and D. S. Fisher, "Hidden complexity of yeast adaptation under simple evolutionary conditions," *Current Biology* **28**, 515–525 (2018).
- [21] H. L. Smith, "Bacterial competition in serial transfer culture," *Math Biosci* **229**, 149–159 (2011).
- [22] L. M. Wahl and A. D. Zhu, "Survival probability of beneficial mutations in bacterial batch culture," *Genetics* **200**, 309–320 (2015).
- [23] M. Manhart, B. V. Adkar, and E. I. Shakhnovich, "Trade-offs between microbial growth phases lead to frequency-dependent and non-transitive selection," *Proc R Soc B* **285**, 20172459 (2018).
- [24] M. Manhart and E. I. Shakhnovich, "Growth tradeoffs produce complex microbial communities on a single limiting resource," *Nat Commun* **9**, 3214 (2018).
- [25] J. Lin and A. Amir, "The effects of stochasticity at the single-cell level and cell size control on the population growth," *Cell Systems* **5**, 358–367 (2017).
- [26] J.-N. Jasmin and C. Zeyl, "Life-history evolution and density-dependent growth in experimental populations of yeast," *Evolution* **66**, 3789–3802 (2012).
- [27] L.-M. Chevin, "On measuring selection in experimental evolution," *Biol Lett* **7**, 210–213 (2011).
- [28] B. H. Good, M. J. McDonald, J. E. Barrick, R. E. Lenski, and M. M. Desai, "The dynamics of molecular evolution over 60,000 generations," *Nature* **551**, 45–50 (2017).
- [29] S. E. Finkel, "Long-term survival during stationary phase: evolution and the gasp phenotype," *Nature Reviews Microbiology* **4**, 113 (2006).
- [30] Sarit Avrani, Evgeni Bolotin, Sophia Katz, and Ruth Hershberg, "Rapid genetic adaptation during the first four months of survival under resource exhaustion," *Molecular Biology and Evolution* **34**, 1758–1769 (2017).
- [31] R. E. Lenski, M. R. Rose, S. C. Simpson, and S. C. Tadler, "Long-term experimental evolution in *Escherichia coli*. i. adaptation and divergence during 2,000 generations," *The American Naturalist* **138**, 1315–1341 (1991).
- [32] G. Chevereau, M. Dravecká, T. Batur, A. Guvenek, D. H. Ayhan, E. Toprak, and T. Bollenbach, "Quantifying the determinants of evolutionary dynamics leading to drug resistance," *PLoS Biol* **13**, e1002299 (2015).
- [33] Manuel Campos, Sander K Govers, Irnov Irnov, Genevieve S Dobihal, François Cornet, and Christine Jacobs-Wagner, "Genomewide phenotypic analysis

- of growth, cell morphogenesis, and cell cycle events in *Escherichia coli*,” *Mol Syst Biol* **14**, e7573 (2018).
- [34] J. H. Gillespie, “Molecular evolution over the mutational landscape,” *Evolution* **38**, 1116–1129 (1984).
- [35] Lindi M Wahl and Philip J Gerrish, “The probability that beneficial mutations are lost in populations with periodic bottlenecks,” *Evolution* **55**, 2606–2610 (2001).
- [36] Y. Guo, M. Vucelja, and A. Amir, “Stochastic tunneling across fitness valleys can give rise to a logarithmic long-term fitness trajectory,” *Sci Adv* **5**, eaav3842 (2019).
- [37] J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory* (Harper and Row, New York, 1970).
- [38] P. J. Gerrish and R. E. Lenski, “The fate of competing beneficial mutations in an asexual population,” *Genetica* **102**, 127 (1998).
- [39] Michael M Desai and Daniel S Fisher, “Beneficial mutation–selection balance and the effect of linkage on positive selection,” *Genetics* **176**, 1759–1798 (2007).
- [40] Stephan Schiffels, Gergely J. Szöllösi, Ville Mustonen, and Michael Lässig, “Emergent neutrality in adaptive asexual evolution,” *Genetics* **189**, 1361–1375 (2011).
- [41] 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**, 4950–4955 (2012).
- [42] Daniel S Fisher, “Asexual evolution waves: fluctuations and universality,” *Journal of Statistical Mechanics: Theory and Experiment* **2013**, P01011 (2013).
- [43] Benjamin H. Good and Michael M. Desai, “Deleterious passengers in adapting populations,” *Genetics* **198**, 1183–1208 (2014).
- [44] Daniel L Hartl, Andrew G Clark, and Andrew G Clark, *Principles of population genetics*, Vol. 116 (Sinauer associates Sunderland, 1997).
- [45] Ellen Baake, Adrin Gonzalez Casanova, Sebastian Probst, and Anton Wakolbinger, “Modelling and simulating lenski’s long-term evolution experiment,” *Theoretical Population Biology* **127**, 58 – 74 (2019).
- [46] E. H. Simpson, “The interpretation of interaction in contingency tables,” *Journal of the Royal Statistical Society: Series B (Methodological)* **13**, 238–241 (1951).
- [47] L. S. Luckinbill, “r and K selection in experimental populations of *Escherichia coli*,” *Science* **202**, 1201–1203 (1978).
- [48] Karin E. Kram, Christopher Geiger, Wazim Mohammed Ismail, Heewook Lee, Haixu Tang, Patricia L. Foster, and Steven E. Finkel, “Adaptation of *Escherichia coli* to long-term serial passage in complex medium: Evidence of parallel evolution,” *mSystems* **2** (2017).
- [49] G. I. Lang, D. Botstein, and M. M. Desai, “Genetic variation and the fate of beneficial mutations in asexual populations,” *Genetics* **188**, 647–661 (2011).
- [50] M. Lynch and B. Walsh, *Genetics and Analysis of Quantitative Characters* (Sinauer, Sunderland, MA, 1998).
- [51] K. Gomez, J. Bertram, and J. Masel, “Directional selection rather than functional constraints can shape the G matrix in rapidly adapting asexuals,” *Genetics* **211**, 715–729 (2019).
- [52] J. Baranyi and T. A. Roberts, “A dynamic approach to predicting bacterial growth in food,” *Int J Food Microbiol* **23**, 277–294 (1994).
- [53] Y. Himeoka and K. Kaneko, “Theory for transitions between exponential and stationary phases: Universal laws for lag time,” *Phys Rev X* **7**, 021049 (2017).
- [54] J. Baranyi, “Comparison of stochastic and deterministic concepts of bacterial lag,” *J Theor Biol* **192**, 403–408 (1998).
- [55] D. Reznick, M. J. Bryant, and F. Bashey, “r- and K-selection revisited: The role of population regulation in life-history evolution,” *Ecology* **83**, 1509–1520 (2002).
- [56] T. Pfeiffer, S. Schuster, and S. Bonhoeffer, “Cooperation and competition in the evolution of ATP-producing pathways,” *Science* **292**, 504–507 (2001).
- [57] R. C. MacLean, “The tragedy of the commons in microbial populations: insights from theoretical, comparative and experimental studies,” *Heredity* **100**, 471–477 (2007).
- [58] J. R. Meyer, I. Gudelj, and R. Beardmore, “Biophysical mechanisms that maintain biodiversity through trade-offs,” *Nat Commun* **6**, 6278 (2015).
- [59] J.-N. Jasmin, M. M. Dillon, and C. Zeyl, “The yield of experimental yeast populations declines during selection,” *Proc R Soc B* **279**, 4382–4388 (2012).
- [60] H. Bachmann, M. Fischlechner, I. Rabbers, N. Barfa, F. B. dos Santos, D. Molenaar, and Bas Teusink, “Availability of public goods shapes the evolution of competing metabolic strategies,” *Proc Natl Acad Sci USA* **110**, 14302–14307 (2013).
- [61] G. J. Velicer and R. E. Lenski, “Evolutionary trade-offs under conditions of resource abundance and scarcity: Experiments with bacteria,” *Ecology* **80**, 1168–1179 (1999).
- [62] S. J. Gould and R. C. Lewontin, “The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme,” *Proc R Soc B* **205**, 581–598 (1979).
- [63] A. Amir, “Point of view: Is cell size a spandrel?” *eLife* **6**, e22186 (2017).