

# Supplementary Methods:

## Trade-offs between microbial growth phases lead to frequency-dependent and non-transitive selection

Michael Manhart, Bharat V. Adkar, and Eugene I. Shakhnovich  
*Department of Chemistry and Chemical Biology,  
 Harvard University, Cambridge, MA 02138, USA*

*Proceedings of the Royal Society B*, 2018. Main text doi:10.1098/rspb.2017.2459

### S1. MINIMAL THREE-PHASE MODEL OF POPULATION GROWTH

Let each strain  $i$  have lag time  $\lambda_i$ , growth rate  $g_i$ , and initial population size  $N_i(0)$ , so that its growth dynamics obey (figure 1a)

$$N_i(t) = \begin{cases} N_i(0) & 0 \leq t < \lambda_i, \\ N_i(0)e^{g_i(t-\lambda_i)} & \lambda_i \leq t < t_{\text{sat}}, \\ N_i(0)e^{g_i(t_{\text{sat}}-\lambda_i)} & t \geq t_{\text{sat}}. \end{cases} \quad (\text{S1.1})$$

The time  $t_{\text{sat}}$  at which growth saturates is determined by a model of resource consumption. Let  $R$  be the initial amount of resources. We assume that each strain consumes resources in proportion to its population size, for example, if the limiting resource is space. Let the yield  $Y_i$  be the number of cells of strain  $i$  per unit of the resource. Therefore the resources are exhausted at time  $t = t_{\text{sat}}$  such that

$$\sum_i \frac{N_i(t_{\text{sat}})}{Y_i} = R. \quad (\text{S1.2})$$

We can alternatively assume that each strain consumes resources in proportion to its total number of cell divisions, rather than its total number of cells. The number of cell divisions for strain  $i$  that have occurred by time  $t$  is  $N_i(t) - N_i(0)$ . Redefining the yield  $Y_i$  as the number of cell divisions per unit resource, saturation now occurs at the time  $t = t_{\text{sat}}$  satisfying

$$\sum_i \frac{N_i(t_{\text{sat}}) - N_i(0)}{Y_i} = R. \quad (\text{S1.3})$$

For simplicity we use the first model (equation (S1.2)) throughout this work, but it is straightforward to translate all results to the second model using the transformation  $R \rightarrow R + \sum_i N_i(0)/Y_i$ . This correction will generally be small, though, since  $\sum_i N_i(0)/Y_i$  is the amount of resources that the initial population of cells consume for their first divisions, and this amount will usually be much less than the total resources  $R$ . It is also straightforward to further generalize this model to include other modes of resource consumption, such as consuming the resource per unit time during lag phase.

### S2. DEFINITION OF SELECTION COEFFICIENT

The selection coefficient per unit time is

$$\sigma(t) = \frac{d}{dt} \log \left( \frac{N_2(t)}{N_1(t)} \right). \quad (\text{S2.1})$$

In the minimal three-phase growth model (equation (S1.1)), we can write the growth curve as  $N_i(t) = N_i(0)e^{g_i(t-\lambda_i)\Theta(t-\lambda_i)}$ , where  $\Theta(t)$  is the Heaviside step function. Then the instantaneous selection coefficient is:

$$\begin{aligned} \sigma(t) &= \frac{d}{dt} \left[ g_2(t - \lambda_2)\Theta(t - \lambda_2) - g_1(t - \lambda_1)\Theta(t - \lambda_1) \right] \\ &= g_2\Theta(t - \lambda_2) - g_1\Theta(t - \lambda_1), \end{aligned} \quad (\text{S2.2})$$

for  $t < t_{\text{sat}}$ , and  $\sigma(t) = 0$  for  $t > t_{\text{sat}}$ .

Since we are mainly concerned with how the mutant frequency changes over whole cycles of growth, it is more convenient to integrate this instantaneous selection coefficient to obtain the total selection coefficient per cycle:

$$\begin{aligned} s &= \int_0^{t_{\text{sat}}} dt \sigma(t) \\ &= g_2(t_{\text{sat}} - \lambda_2)\Theta(t_{\text{sat}} - \lambda_2) \\ &\quad - g_1(t_{\text{sat}} - \lambda_1)\Theta(t_{\text{sat}} - \lambda_1), \end{aligned} \quad (\text{S2.3})$$

which, using equation (S2.1), is equivalent to the definition in equation (2.2) from the main text. If we exclude the trivial case where the time to saturation is less than one of the lag times (so that one strain does not grow at all), the selection coefficient simplifies to

$$s = g_2(t_{\text{sat}} - \lambda_2) - g_1(t_{\text{sat}} - \lambda_1). \quad (\text{S2.4})$$

### S3. DERIVATION OF SELECTION COEFFICIENT EXPRESSION

To determine how  $s$  explicitly depends on the underlying parameters, we must solve the saturation condition

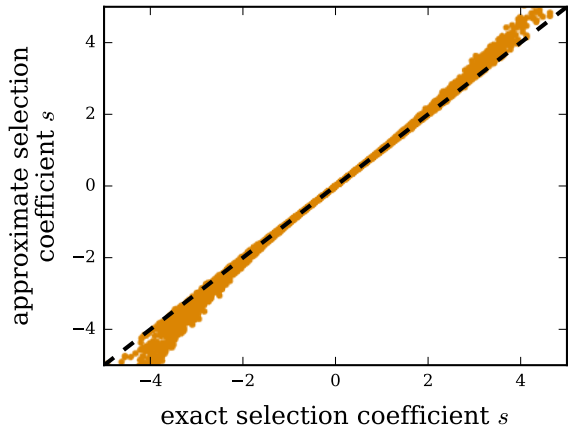


FIG. S1. **Test of selection coefficient approximation.** Comparison of the approximate selection coefficient formula (equation (3.1)) with the exact result obtained using the definition in equation (2.2) and a numerical solution to the saturation equation (equation (S3.1)). Each orange point corresponds to a different set of randomly-generated growth traits ( $\gamma$ ,  $\omega$ ,  $\nu_1$ ,  $\nu_2$ ; see equation (2.1)) and initial mutant frequencies  $x$ . The black dashed line is the line of identity.

in equation (S1.2) for  $t_{\text{sat}}$ :

$$R = \frac{N_0 x_1}{Y_1} e^{g_1(t_{\text{sat}} - \lambda_1)} + \frac{N_0 x_2}{Y_2} e^{g_2(t_{\text{sat}} - \lambda_2)}, \quad (\text{S3.1})$$

where  $N_0$  is the total initial population size and  $x_1, x_2$  are the initial frequencies of the wild-type and mutant. We ignore the trivial case where one strain saturates before the other starts to grow. While we cannot analytically solve this equation in general, we can obtain a good approximation in the limit of weak selection ( $|s| \ll 1$ ). We first rewrite the equation in terms of the selection coefficient using equation (S2.4):

$$R = N_0 e^{g_1(t_{\text{sat}} - \lambda_1)} \left( \frac{x_1}{Y_1} + \frac{x_2}{Y_2} e^s \right). \quad (\text{S3.2})$$

We then solve for  $t_{\text{sat}}$  and expand to first order in  $s$ :

$$t_{\text{sat}} \approx \lambda_1 - \frac{1}{g_1} \log \left[ \frac{N_0}{R} \left( \frac{x_1}{Y_1} + \frac{x_2}{Y_2} \right) \right] - \frac{x_2/Y_2}{g_1(x_1/Y_1 + x_2/Y_2)} s. \quad (\text{S3.3})$$

Self-consistency requires this expression for  $t_{\text{sat}}$  to be invariant under exchange of the mutant and wild-type indices and switching the sign of  $s$ ; equating these two equivalent expressions for  $t_{\text{sat}}$  allows us to solve for  $s$ , which gives the main result in equation (3.1).

In figure S1 we compare the selection coefficient calculated from this approximate expression with the exact

result obtained by numerically solving equation (S3.1) for  $t_{\text{sat}}$  and then directly calculating  $s$  using the definition of equation (2.2). This empirically shows that although the derivation relies on the approximation of weak selection ( $|s| \ll 1$ ), equation (3.1) is extremely accurate over a wide range of parameter values, even up to rather strong selection strengths  $|s| \sim 1$ . Furthermore, the expression is exact in two special cases: when the mutant and the wild-type are selectively neutral ( $s = 0$ ), and when the mutant and wild-type have equal growth rates ( $g_1 = g_2 = g$ ), since  $s = -(\lambda_2 - \lambda_1)g = -\omega$  according to equation (S2.4).

#### S4. SATURATION TIME AND TOTAL POPULATION SIZE

Here we derive expressions for the saturation time  $t_{\text{sat}}$  and the total population size at saturation

$$\begin{aligned} N_{\text{sat}} &= N_1(t_{\text{sat}}) + N_2(t_{\text{sat}}) \\ &= N_0 x_1 e^{g_1(t_{\text{sat}} - \lambda_1)} + N_0 x_2 e^{g_2(t_{\text{sat}} - \lambda_2)}. \end{aligned} \quad (\text{S4.1})$$

We again assume the nontrivial case of  $t_{\text{sat}} > \lambda_1, \lambda_2$ . First, if the growth rates are equal ( $g_1 = g_2 = g$ ), we can obtain exact solutions since the two-strain saturation condition (equation (S3.1)) is analytically solvable for  $t_{\text{sat}}$ :

$$\begin{aligned} t_{\text{sat}} &= \frac{1}{g} \log \left[ \frac{R}{2N_0} H \left( \frac{Y_1 e^{g\lambda_1}}{x_1}, \frac{Y_2 e^{g\lambda_2}}{x_2} \right) \right], \\ N_{\text{sat}} &= \frac{1}{2} (x_1 e^{-g\lambda_1} + x_2 e^{-g\lambda_2}) H \left( \frac{RY_1 e^{g\lambda_1}}{x_1}, \frac{RY_2 e^{g\lambda_2}}{x_2} \right). \end{aligned} \quad (\text{S4.2})$$

If the growth rates are unequal ( $g_1 \neq g_2$ ), then we must rely on the small  $s$  approximation. We can rearrange equation (S2.4) to obtain  $t_{\text{sat}}$  as a function of  $s$ :

$$t_{\text{sat}} = \frac{s + g_2 \lambda_2 - g_1 \lambda_1}{g_2 - g_1}. \quad (\text{S4.3})$$

We can then substitute the approximate expression for  $s$  (equation (3.1)) into equation (S4.3):

$$\begin{aligned} t_{\text{sat}} \approx & \left( \frac{x_1/Y_1 + x_2/Y_2}{g_1 x_1/Y_1 + g_2 x_2/Y_2} \right) \left( \log \left[ \frac{R}{2N_0} H \left( \frac{Y_1}{x_1}, \frac{Y_2}{x_2} \right) \right] \right. \\ & \left. - \frac{g_1 g_2 (\lambda_2 - \lambda_1)}{g_2 - g_1} \right) + \frac{g_2 \lambda_2 - g_1 \lambda_1}{g_2 - g_1}. \end{aligned} \quad (\text{S4.4})$$

To obtain an expression for  $N_{\text{sat}}$  in this approximation, we rewrite its definition (equation (S4.1)) in terms of  $s$  using equation (S4.3):

$$N_{\text{sat}} = N_0 e^{g_1 g_2 (\lambda_2 - \lambda_1) / (g_2 - g_1)} \times \left( x_1 e^{g_1 s / (g_2 - g_1)} + x_2 e^{g_2 s / (g_2 - g_1)} \right). \quad (\text{S4.5})$$

For small  $s$ , we can show from equation (3.1) that

---


$$e^{g_1 g_2 (\lambda_2 - \lambda_1) / (g_2 - g_1)} \approx \frac{1}{2N_0} H \left( \frac{RY_1}{x_1}, \frac{RY_2}{x_2} \right) \exp \left[ - \left( \frac{s}{g_2 - g_1} \right) \left( \frac{g_1 x_1 / Y_1 + g_2 x_2 / Y_2}{x_1 / Y_1 + x_2 / Y_2} \right) \right] \quad (\text{S4.6})$$

$$\approx \frac{1}{2N_0} H \left( \frac{RY_1}{x_1}, \frac{RY_2}{x_2} \right) \left[ 1 - \left( \frac{s}{g_2 - g_1} \right) \left( \frac{g_1 x_1 / Y_1 + g_2 x_2 / Y_2}{x_1 / Y_1 + x_2 / Y_2} \right) \right].$$


---

Substituting this into equation (S4.5) and expanding to first order in  $s$  shows that the saturation population size is approximately

$$N_{\text{sat}} \approx \frac{1}{2} H \left( \frac{RY_1}{x_1}, \frac{RY_2}{x_2} \right) \times \left[ 1 - \frac{x_1 x_2 (Y_2^{-1} - Y_1^{-1})}{x_1 / Y_1 + x_2 / Y_2} s \right]. \quad (\text{S4.7})$$

Therefore the saturation size in the neutral case ( $s = 0$ ) is

$$N_{\text{sat}} = \frac{1}{2} H \left( \frac{RY_1}{x_1}, \frac{RY_2}{x_2} \right) = H \left( \frac{N_{\text{sat},1}}{2x_1}, \frac{N_{\text{sat},2}}{2x_2} \right), \quad (\text{S4.8})$$

since  $RY_i = N_{\text{sat},i}$ , where  $N_{\text{sat},i}$  is the saturation population size of strain  $i$  if no other strains are present. So for a neutral pair of strains, the total population grows to the harmonic mean of the saturation population sizes of the individual strains; this shows that we can interpret the harmonic mean of both strains' yields as the effective yield for the combined population. When selection is nonzero, the effective yield is perturbed above this value if the strain with higher yield is also positively selected (e.g.,  $Y_2 > Y_1$  and  $s > 0$ ), while otherwise it is perturbed below the neutral value.

## S5. EFFECT OF CORRELATED PLEIOTROPY ON SELECTION

Mutational effects on growth traits may not only be pleiotropic, but they may also be correlated. The simplest case is a linear correlation between growth traits across many mutations or strains:

$$\lambda \approx \frac{a}{g} + \text{constant}, \quad \nu \approx bg + \text{constant}, \quad (\text{S5.1})$$

where  $a$  and  $b$  are proportionality constants. We take lag time to be linearly correlated with growth time (reciprocal growth rate), rather than growth rate, since then both traits have units of time and the constant  $a$  is dimensionless. Various models predict linear correlations of this form [1–6], which have been tested on measured distributions of traits [5, 7–12] (see section 4 in the main text).

We can combine this model with the selection coefficient in equation (3.1) to quantify how much selection is amplified or diminished by correlated pleiotropy. That is, if a mutation changes growth rate by a small amount  $\Delta g = g_2 - g_1$  from the wild-type, then according to equation (S5.1) it will also change lag time by  $\Delta \lambda \approx -a \Delta g / g^2$  and yield by  $\Delta \nu = b \Delta g$ , and hence the expected selection coefficient will be (using  $\gamma = \Delta g / g$ )

$$s \approx \gamma (\log \nu + a). \quad (\text{S5.2})$$

This shows that correlations between growth and yield have no effect on selection to leading order, since selection only depends logarithmically on yield. Correlations between growth and lag, however, can have a significant amplifying or diminishing effect. Since  $\log \nu > 0$  always, synergistic pleiotropy ( $a > 0$ ) will tend to increase the magnitude of selection, while antagonistic pleiotropy ( $a < 0$ ) will tend to reduce it. The significance of this effect depends on the relative value of  $a$  compared to  $\log \nu$ ; in general, the logarithm and the dimensionless nature of  $a$  suggest both should be order 1 and therefore comparable.

## S6. FREQUENCY DEPENDENCE OF SELECTION

The selection coefficient in equation (3.1) depends on the initial mutant frequency  $x$ . Here we show that  $s(x)$  is a monotonic function of the frequency  $x$ ; this is important because it means that conditional neutrality ( $s(\tilde{x}) = 0$ ) occurs at a unique frequency  $\tilde{x}$  (equation (3.5)). We use an exact argument starting from the original model because the approximate  $s(x)$  function in equation (3.1) has spurious non-monotonic behavior in some regimes. For simplicity we again use the dimensionless growth parameters defined in equation (2.1).

If the mutant and wild-type have equal growth rates ( $\gamma = 0$ ), then we have previously showed that  $s(x) = -\omega$ , so it is constant (and hence monotonic) in  $x$ . Now we consider  $\gamma \neq 0$ . In this case we can write the saturation condition in terms of  $s(x)$  by substituting equation (S4.3) for  $t_{\text{sat}}$  in equation (S3.1):

$$e^{\omega(1+\gamma)/\gamma} \left[ \frac{(1-x)}{\nu_1} e^{s(x)/\gamma} + \frac{x}{\nu_2} e^{(1+1/\gamma)s(x)} \right] = 1. \quad (\text{S6.1})$$

We can differentiate with respect to  $x$  and solve for  $ds/dx$  to obtain the differential equation

$$\frac{ds}{dx} = \frac{\gamma(1 - e^{s(x)}\nu_1/\nu_2)}{(1-x) + x(1+\gamma)e^{s(x)}\nu_1/\nu_2}. \quad (\text{S6.2})$$

The only way  $s(x)$  can be non-monotonic is if  $ds/dx = 0$  for some  $x$  without  $s(x)$  being constant. Since the denominator of equation (S6.2) is always positive,  $ds/dx = 0$  only if  $s(x) = \log(\nu_2/\nu_1)$  for some  $x$ . However, if  $s(x) = \log(\nu_2/\nu_1)$  for any  $x$ , then it must be constant at  $\log(\nu_2/\nu_1)$  for all  $x$ . We show this by substituting  $s(x) = \log(\nu_2/\nu_1)$  into the saturation equation (equation (S6.1)). The  $x$ -dependence drops out and we are left with

$$\frac{\nu_2^{1/\gamma}}{\nu_1^{1+1/\gamma}} e^{\omega(1+1/\gamma)} = 1. \quad (\text{S6.3})$$

Therefore if the parameters satisfy this condition, then  $s(x) = \log(\nu_2/\nu_1)$  for all  $x$ . Therefore  $ds/dx$  only equals zero when  $s(x)$  is constant, and so  $s(x)$  can never be a non-monotonic function of  $x$ .

Figure S2a shows the sign of  $ds/dx$  over growth-lag trait space for strains with equal yields ( $\nu_1 = \nu_2$ ); figure S2b shows the case of unequal yields ( $\nu_1 \neq \nu_2$ ). The boundaries between signs of  $ds/dx$  are where  $s(x)$  is a constant, and thus they are given by  $\gamma = 0$  and equation (S6.3). Note that for equal yields,  $s(x)$  is constant at zero along the neutral boundary (figure S2a), whereas for unequal yields there is a separate boundary, away from the conditionally-neutral region, where  $s(x)$  has a

constant but nonzero value (figure S2b).

Another way to measure the frequency dependence of selection is to consider its total variation across the whole range of frequencies. We define the relative variation of selection as  $|(s_{\text{max}} - s_{\text{min}})/s(1/2)|$ , where  $s_{\text{max}}$  and  $s_{\text{min}}$  are the maximum and minimum values of  $s(x)$  across all frequencies, and  $s(1/2)$  is selection at the intermediate frequency  $x = 1/2$ . Since  $s(x)$  is always a monotonic function of  $x$ , the maximum and minimum values are attained at the endpoints  $x = 0$  and  $x = 1$ . The selection coefficient is not technically defined for these values (since either the mutant or the wild-type is extinct), but we can determine its value in the limits  $x \rightarrow 0$  and  $x \rightarrow 1$ . In the limit of  $x \rightarrow 0$ , the saturation time must be the time for the wild-type alone to consume all the resources, and vice-versa for  $x \rightarrow 1$ :

$$\begin{aligned} \lim_{x \rightarrow 0} t_{\text{sat}}(x) &= \lambda_1 + \frac{1}{g_1} \log \left( \frac{RY_1}{N_0} \right), \\ \lim_{x \rightarrow 1} t_{\text{sat}}(x) &= \lambda_2 + \frac{1}{g_2} \log \left( \frac{RY_2}{N_0} \right). \end{aligned} \quad (\text{S6.4})$$

Using the relationship between  $s$  and  $t_{\text{sat}}$  in equation (S2.4) and converting to dimensionless parameters (equation (2.1)), we have

$$\begin{aligned} \lim_{x \rightarrow 0} s(x) &= \gamma \log \nu_1 - \omega(1 + \gamma), \\ \lim_{x \rightarrow 1} s(x) &= \left( \frac{\gamma}{1 + \gamma} \right) \log \nu_2 - \omega. \end{aligned} \quad (\text{S6.5})$$

Hence the total variation of selection coefficients is

$$\begin{aligned} |s_{\text{max}} - s_{\text{min}}| &= \left| \lim_{x \rightarrow 1} s(x) - \lim_{x \rightarrow 0} s(x) \right| \\ &= \left| \gamma \left( \frac{\log \nu_2}{1 + \gamma} - \log \nu_1 + \omega \right) \right|. \end{aligned} \quad (\text{S6.6})$$

This result is exact (no weak selection approximation), but the approximate  $s(x)$  expression in equation (3.1) gives an identical result.

Normalizing this total range of selection by its magnitude at some intermediate frequency, such as  $x = 1/2$ , measures the relative variation in  $s(x)$  over frequencies. For equal yields ( $\nu_1 = \nu_2$ ), the relative variation simplifies to

$$\left| \frac{s_{\text{max}} - s_{\text{min}}}{s(1/2)} \right| = \left| \frac{\gamma(2 + \gamma)}{2(1 + \gamma)} \right|. \quad (\text{S6.7})$$

It is small over a large range of the trait space (figure S2c), indicating that the frequency dependence of selection is relatively weak for equal yields. In contrast, when the yields are unequal ( $\nu_1 \neq \nu_2$ ), the variation becomes very large near the conditionally-neutral region

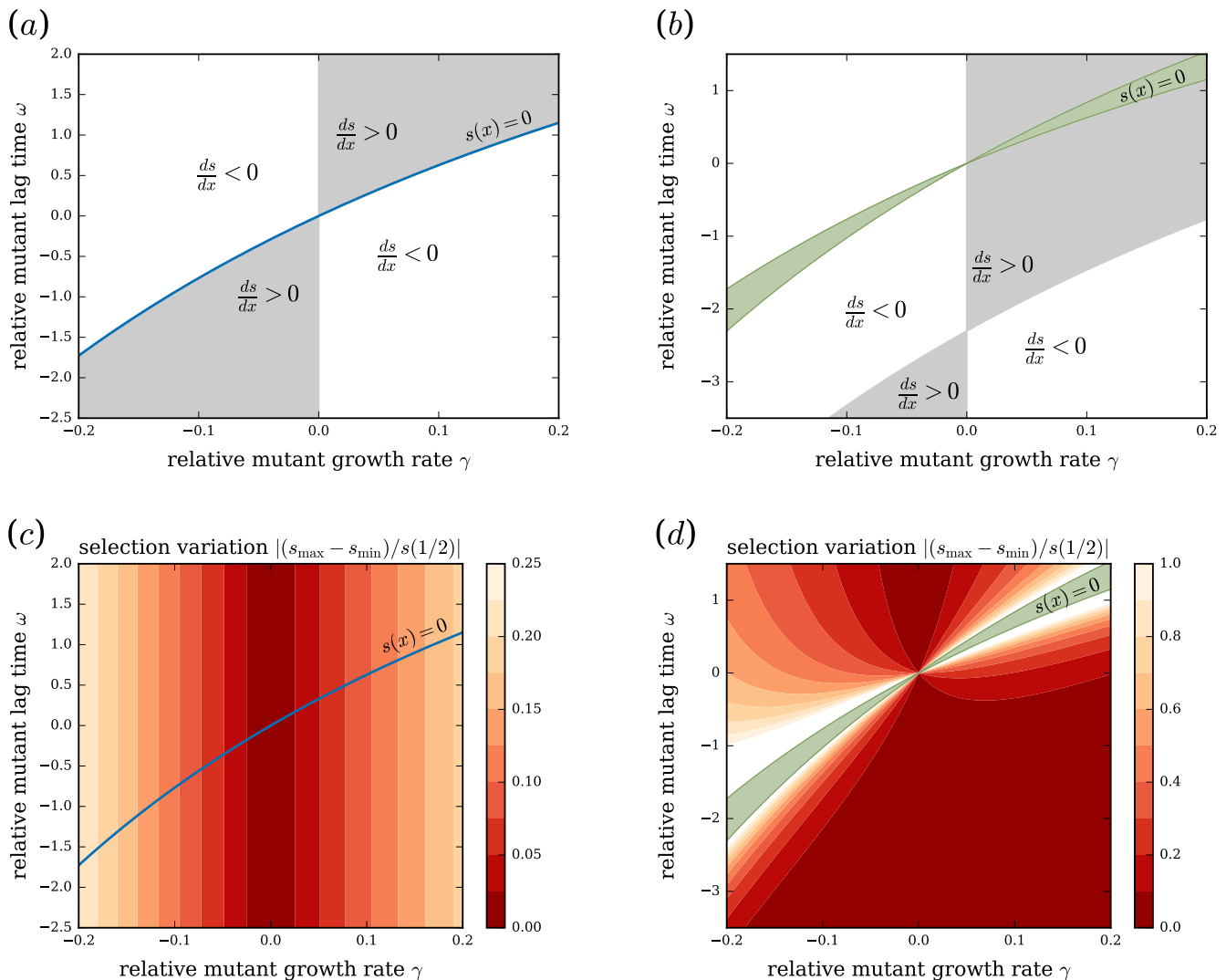


FIG. S2. **frequency dependence of the selection coefficient over growth-lag trait space.** (a) For a mutant and wild-type with equal yields ( $\nu_1 = \nu_2 = 10^3$ ), the gray and white regions indicate where the selection coefficient  $s(x)$  increases as a function of mutant frequency ( $ds/dx > 0$ ) or decreases ( $ds/dx < 0$ ). The neutral boundary is in blue. (b) Same as (a) but for a mutant and wild-type with unequal yields ( $\nu_1 = 10^3, \nu_2 = 10^4$ ). The conditionally-neutral region is shown in green. (c) Relative variation of the selection coefficient over mutant frequencies when the mutant and wild-type have equal yields. Yield values and the neutral boundary are the same as (a). (d) Same as (c) but for a mutant and wild-type with unequal yields; yield values and the conditionally-neutral region are the same as (b). The relative variation diverges in the conditionally-neutral region since  $s(1/2) = 0$  for some points.

(figure S2d). This is because  $s(1/2)$  goes to zero for some points in the conditionally-neutral region, while the total range  $|s_{\max} - s_{\min}|$  remains finite. Thus, the frequency dependence of selection is most significant for mutants in the conditionally-neutral region; this is expected since these are the mutants that can coexist or be bistable with the wild-type.

## S7. ROBUSTNESS OF COEXISTENCE TO GENETIC DRIFT

If the bottleneck population size  $N_0$  at the beginning of each round is small, then stochastic effects of sampling from round to round (genetic drift) may be significant. We can gauge the robustness of coexistence to these fluctuations by comparing the magnitude of those fluctuations, which is of order  $1/N_0$ , with  $ds/dx$  measured at the coexistence frequency  $\tilde{x}$  (equation (S6.2)), which estimates the strength of selection for a small change in frequency around coexistence. Coexistence will be robust

against fluctuations if

$$\left| \frac{\gamma(1 - \nu_1/\nu_2)}{(1 - \tilde{x}) + \tilde{x}(1 + \gamma)\nu_1/\nu_2} \right| > \frac{1}{N_0}. \quad (\text{S7.1})$$

This tells us the critical value of the bottleneck size  $N_0$ , which we can control experimentally, needed to achieve robust coexistence. For example, if the mutant has 10% slower growth rate ( $\gamma = -0.1$ ) but 10% higher yield ( $\nu_2/\nu_1 = 1.1$ ), and coexistence occurs at  $\tilde{x} = 1/2$ , then  $N_0$  must be greater than 100 for stabilizing selection at the coexistence frequency to be stronger than genetic drift.

### S8. FIXATION UNDER FREQUENCY-DEPENDENT SELECTION

If the population at the end of a competition round is randomly sampled to populate the next round, this is equivalent to a Wright-Fisher process with frequency-dependent selection coefficient  $s(x)$  and effective population size  $N_0$  [13]. In the limit of a large population ( $N_0 \gg 1$ ) and weak selection ( $|s(x)| \ll 1$ ), the fixation probability of a mutant starting from frequency  $x$  is

$$\phi(x) = \frac{\int_0^x dx' e^{2N_0 V(x')}}{\int_0^1 dx' e^{2N_0 V(x')}}, \quad (\text{S8.1a})$$

where  $V(x)$  is the effective selection ‘‘potential’’:

$$V(x) = - \int_0^x dx' s(x'). \quad (\text{S8.1b})$$

This is defined in analogy with physical systems, where selection plays the role of a force and  $V(x)$  is the corresponding potential energy function. The mean time (number of competition rounds) to fixation, given that fixation eventually occurs, is

$$\theta(x) = \int_x^1 dx' \psi(x') \phi(x') (1 - \phi(x')) + \left( \frac{1 - \phi(x)}{\phi(x)} \right) \int_0^x dx' \psi(x') (\phi(x'))^2, \quad (\text{S8.2a})$$

where

$$\psi(x) = \frac{2N_0 e^{-2N_0 V(x)}}{x(1-x)} \int_0^1 dx' e^{2N_0 V(x')}. \quad (\text{S8.2b})$$

These results assume that mutations are rare enough to neglect interference from multiple mutations simultaneously present in the population.

For simplicity we focus on the case of a single mutant cell (frequency  $1/N_0$ ) at the beginning of a competition round. To test the effect of frequency dependence on fixation, we compare the true fixation probabilities and times, calculated from equations (S8.1) and (S8.2) using  $s(x)$  (equation (3.1)), with the fixation probabilities and times predicted if selection has a constant value at  $s(1/2)$ , as is often measured in competition experiments [14]. When selection is a constant across frequencies, equation (S8.1) simplifies to Kimura’s formula [13]:

$$\phi(1/N_0) = \frac{1 - e^{-2s}}{1 - e^{-2N_0 s}}. \quad (\text{S8.3})$$

Deviations from this relationship between  $\phi$  and  $s(1/2)$  are therefore indicative of significant frequency dependence.

For several sets of mutants, figure S3a shows their selection coefficients  $s(1/2)$  versus their fixation probabilities  $\phi(1/N_0)$ . In orange are mutants obtained by uniformly scanning a rectangular region of growth-lag trait space (e.g., the trait space shown in figure 2a). The black line shows the prediction from Kimura’s formula (equation (S8.3)) assuming  $s = s(1/2)$  is a constant selection coefficient for all frequencies; this frequency-independent approximation appears to describe these mutants well. The mean fixation times  $\theta(1/N_0)$  (figure S3b) for these mutants are also well-described by assuming constant selection coefficient  $s(1/2)$ . This is because the frequency dependence for these mutants is weak, as shown in figure S2c,d. Therefore a single measurement of the selection coefficient for these mutants at any initial frequency provides an accurate prediction of the long-term population dynamics.

The plots of selection variation in figure S2c,d indicate that the most significant frequency dependence occurs for mutants in the conditionally-neutral region with unequal yields, i.e., mutants with coexistence or bistability. We thus calculate the fixation probabilities and times for mutants with neutrality at particular frequencies, and compare these statistics to their selection coefficients at  $x = 1/2$  as would be measured experimentally (figure S3a,b). As expected, the fixation statistics show significant deviations from the predictions for constant selection. In particular, mutants with neutrality at  $\tilde{x} = 1/2$  (equation (3.5)) have  $s(1/2) = 0$  by definition, but they nevertheless show a wide range of fixation probabilities and times, some above the neutral values ( $\phi = 1/N_0$ ,  $\theta = 2N_0$ ) and some below.

Figure S3c,d shows the fixation probabilities and times of conditionally-neutral mutants as functions of their relative growth rates  $\gamma$ , which separates mutants with coexistence from those with bistability: the mutant has higher yield than that of the wild-type in this example ( $\nu_2 > \nu_1$ ), so the mutants with worse growth rate ( $\gamma < 0$ ) have coexistence while the mutants with better growth rate ( $\gamma > 0$ ) are bistable. Bistable mutants with a neutral frequency of  $\tilde{x} = 1/2$  fix with lower probability than

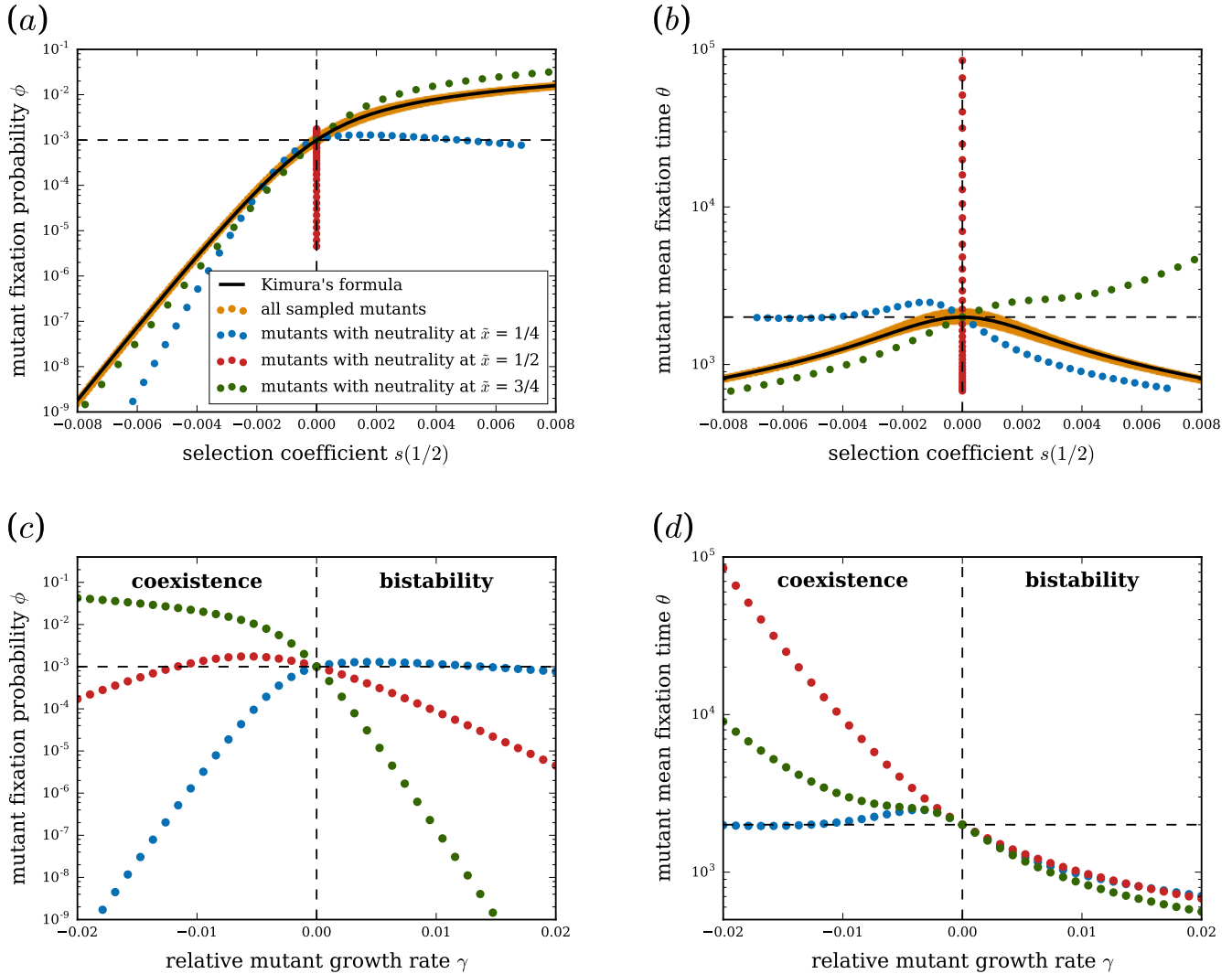


FIG. S3. **Fixation probabilities and times of a mutant.** (a) Fixation probability  $\phi(1/N_0)$  as a function of the selection coefficient at frequency  $x = 1/2$ . Orange points correspond to mutants uniformly sampled across a rectangular region of growth-lag trait space:  $(\gamma, \omega) \in [-10^{-3}, 10^{-3}] \times [-5 \times 10^{-3}, 5 \times 10^{-3}]$ . Other points correspond to mutants with neutrality at specific frequencies (blue for  $\tilde{x} = 1/4$ , red for  $\tilde{x} = 1/2$ , green for  $\tilde{x} = 3/4$ ). We calculate fixation probabilities using the frequency-dependent selection coefficient  $s(x)$  (equation (3.1)) and equation (S8.1); for comparison, the solid black line indicates the prediction from Kimura's formula (equation (S8.3)), assuming a frequency-independent selection coefficient  $s = s(1/2)$ . The horizontal dashed line marks the neutral fixation probability  $\phi = 1/N_0$ . (b) Same as (a), but with the mean fixation time  $\theta(1/N_0)$  (conditioned on eventual fixation) on the vertical axis. The solid black line marks the prediction for a frequency-independent selection coefficient (equation (S8.2)), and the horizontal dashed line marks the neutral fixation time  $\theta = 2N_0$ . (c) Fixation probability  $\phi(1/N_0)$  as a function of the relative growth rate  $\gamma$  for conditionally-neutral mutants. Colors indicate the same neutral frequencies as in (a) and (b). Mutants with  $\gamma < 0$  have coexistence, while mutants with  $\gamma > 0$  are bistable (since  $\nu_2 > \nu_1$  in this example). Dashed lines are the same as in (a). (d) Same as (c), but with the mean fixation time  $\theta(1/N_0)$  on the vertical axis. Dashed lines are the same as (b). In all panels, the relative yields are  $\nu_1 = 10^3$  and  $\nu_2 = 10^4$ , and the initial population size is  $N_0 = 10^3$ .

would a purely neutral mutant (figure S3c), but if they do fix, they do so in less time (figure S3d). We can understand this bistable case in analogy with diffusion across an energy barrier, using the effective selection potential defined in equation (S8.1b). The mutant starts at frequency  $1/N_0$ , and to reach fixation it must not only survive fluctuations from genetic drift while at low fre-

quency, but it also must cross the effective selection potential barrier at the neutral frequency  $\tilde{x}$ . Indeed, the mutant is actually deleterious at low frequencies (below the neutral frequency), and thus we expect the fixation probability to be lower than that of a purely neutral mutant. If such a mutant does fix, though, it will do so rapidly, since it requires rapid fluctuations from genetic

drift to cross the selection barrier. This effect is most pronounced for neutrality at relatively high frequencies; for low neutral frequencies, such as  $\tilde{x} = 1/4$ , the barrier is sufficiently close to the initial frequency  $1/N_0$  that it is easier to cross, and thus the fixation probability is closer to the neutral value (figure S3c).

Mutants with coexistence, on the other hand, are described by a potential well at the neutral frequency. The fixation of these mutants is determined by a tradeoff between the initial boost of positive selection toward the neutral frequency, which helps to avoid immediate extinction, and the stabilizing selection they experience once at coexistence. In particular, once at the neutral frequency, the mutant must eventually cross a selection barrier to reach either extinction or fixation. However, the barrier to fixation is always higher than the barrier to extinction, and thus the mutant has a greater chance of going extinct rather than fixing. As we see for mutants with coexistence at  $\tilde{x} = 1/2$ , decreasing  $\gamma$  from zero initially improves the probability of fixation over neutrality, but eventually it begins to decrease. Thus, the frequency dependence of mutants with coexistence or bistability plays a crucial role in shaping their fixation statistics, and their ultimate fates depend crucially on their individual trait values (i.e.,  $\gamma$ ).

### S9. ADDITIVITY OF THE SELECTION COEFFICIENT

The additivity condition (equation (3.6)) is approximately satisfied if strains  $i$ ,  $j$ , and  $k$  have only small differences in growth rates, lag times, and yields. Conceptually, this is because the saturation times  $t_{\text{sat}}$  for each binary competition between pairs of strains are all approximately equal, but we can also show this directly using the selection coefficient formula. Let  $\gamma_{ij} = (g_i - g_j)/g_j$ ,  $\omega_{ij} = (\lambda_i - \lambda_j)g_j$ , and  $\mu_{ij} = (\nu_i - \nu_j)/\nu_j$  be the relative differences in growth rate, lag time, and yield for strains  $i$  and  $j$ . If these relative differences are all small, then they each approximately obey the additivity condition across strains:

$$\begin{aligned} \gamma_{ik} &= (1 + \gamma_{ij})(1 + \gamma_{jk}) - 1 && \approx \gamma_{ij} + \gamma_{jk}, \\ \omega_{ik} &= \frac{\omega_{ij}}{1 + \gamma_{jk}} + \omega_{jk} && \approx \omega_{ij} + \omega_{jk}, \\ \mu_{ik} &= (1 + \mu_{ij})(1 + \mu_{jk}) - 1 && \approx \mu_{ij} + \mu_{jk}. \end{aligned} \quad (\text{S9.1})$$

In this same limit, the total selection coefficient for strains  $i$  and  $j$  is approximately

$$s_{ij} \approx \gamma_{ij} \log \nu_j - \omega_{ij}. \quad (\text{S9.2})$$

Note that, to leading order, the change in yield  $\mu_{ij}$  does not appear. Using equation (S9.1) and  $\nu_j = (1 + \mu_{jk})\nu_k$ , we have

$$\begin{aligned} s_{ij} + s_{jk} &\approx \gamma_{ij} \log \nu_j - \omega_{ij} + \gamma_{jk} \log \nu_k - \omega_{jk} \\ &\approx \gamma_{ik} \log \nu_k - \omega_{ik} \\ &\approx s_{ik}. \end{aligned} \quad (\text{S9.3})$$

Therefore the selection coefficient is approximately additive when differences between traits are small.

### S10. TRANSITIVITY OF THE SELECTION COEFFICIENT

Since we are only concerned with the sign of selection in determining transitivity, we focus on the signed component of the selection coefficient in equation (3.1). It is also more convenient to use growth times  $\tau_i = 1/g_i$  rather than growth rates, and the quantity  $h_{ij} = \log \left[ \frac{1}{2} H \left( \frac{\nu_j}{1-x}, \frac{\nu_i}{x} \right) \right]$  for the logarithm of the harmonic mean yield. We define the signed component of the selection coefficient for strain  $i$  over strain  $j$  to be

$$(\tau_j - \tau_i)h_{ij} + \lambda_j - \lambda_i. \quad (\text{S10.1})$$

That is,  $s_{ij}$  is proportional to this quantity up to an overall factor that is always nonnegative.

We first consider whether neutrality is a transitive property of strains. Three strains are all pairwise neutral if their traits satisfy

$$\begin{aligned} (\tau_1 - \tau_2)h_{21} + \lambda_1 - \lambda_2 &= 0, \\ (\tau_2 - \tau_3)h_{32} + \lambda_2 - \lambda_3 &= 0, \\ (\tau_3 - \tau_1)h_{13} + \lambda_3 - \lambda_1 &= 0. \end{aligned} \quad (\text{S10.2})$$

If all three strains have equal yields  $\nu_1 = \nu_2 = \nu_3$  ( $h_{21} = h_{32} = h_{13}$  for all frequencies), then any two of these equations imply the third (e.g., by adding them together), which means that neutrality is transitive when all strains have equal yields. If two of the yields are equal while the third is distinct, then transitivity only holds if two of the strains are identical (equal growth and lag times). For example, if  $\nu_1 = \nu_2 \neq \nu_3$ , then we can add together the last two equations in equation (S10.2) to obtain

$$(\tau_2 - \tau_1)h_{13} + \lambda_2 - \lambda_1 = 0, \quad (\text{S10.3})$$

(using  $h_{32} = h_{13}$ ), but this is only consistent with the first equation in equation (S10.2) if  $\tau_1 = \tau_2$  and  $\lambda_1 = \lambda_2$ , i.e., strains 1 and 2 are identical in all traits.

If all the yields have distinct values, then transitivity will generally not hold for arbitrary values of the growth traits. However, it is still *possible* for three strains with distinct yields to all be pairwise neutral, but only with very specific values of the traits. Note that with



unequal yields, neutrality at all frequencies is not possible, so pairs of strains are only conditionally neutral, where strain  $i$  is neutral at frequency  $\bar{x}_{ij}$  with strain  $j$  (equation (3.5)). These frequencies are encoded in the quantities  $h_{ij} = \log \left[ \frac{1}{2} H \left( \frac{\nu_j}{1-\bar{x}_{ij}}, \frac{\nu_i}{\bar{x}_{ij}} \right) \right]$ . We thus fix the yields and the desired neutral frequencies to arbitrary values, and without loss of generality, we can assume  $h_{21} < h_{13} < h_{32}$  (e.g., by putting the strains in order of increasing yields). We can also choose any values of  $\tau_1$  and  $\lambda_1$  since this amounts to a rescaling and shift of time units. Therefore we are left with three linear equations (equation (S10.2)) for four unknowns:  $\tau_2, \tau_3, \lambda_2, \lambda_3$ . If we choose any value of the strain 2 growth time that obeys

$$\tau_2 > \left( \frac{h_{13} - h_{21}}{h_{32} - h_{21}} \right) \tau_1 \quad (\text{S10.4})$$

(note the factor in parentheses is always positive by assumption), then equation (S10.2) has a unique solution for the remaining quantities:

$$\begin{aligned} \tau_3 &= \frac{\tau_2(h_{32} - h_{21}) - \tau_1(h_{13} - h_{21})}{h_{32} - h_{13}}, \\ \lambda_2 &= (\tau_1 - \tau_2)h_{21} + \lambda_1, \\ \lambda_3 &= (\tau_1 - \tau_2) \left( \frac{h_{32} - h_{21}}{h_{32} - h_{13}} \right) h_{13} + \lambda_1. \end{aligned} \quad (\text{S10.5})$$

The linear system actually has a unique solution regardless of equation (S10.4), but without that condition  $\tau_3$  may be negative. Therefore a set of three strains with unequal yields can all be pairwise conditionally neutral only if the growth traits for strains 2 and 3 satisfy equations (S10.4) and (S10.5). For example, in this manner one can construct three strains that all coexist in pairs.

We now turn to constructing sets of three strains such that there is a nontransitive cycle of selective advantage in binary competitions, i.e., strain 2 beats strain 1 in a binary competition, strain 3 beats strain 2, but strain 1 beats strain 3. Therefore the growth traits of the three strains must satisfy

$$\begin{aligned} (\tau_1 - \tau_2)h_{21} + \lambda_1 - \lambda_2 &> 0, \\ (\tau_2 - \tau_3)h_{32} + \lambda_2 - \lambda_3 &> 0, \\ (\tau_3 - \tau_1)h_{13} + \lambda_3 - \lambda_1 &> 0. \end{aligned} \quad (\text{S10.6})$$

All three yields cannot be equal; if they are, adding together any two of the inequalities in equation (S10.6) gives an inequality that is inconsistent with the third

one. Otherwise, the three yields can take arbitrary values, including two of them being equal. Since we can cyclically permute the strain labels, without loss of generality we assume strain 1 has the smallest yield ( $\nu_1 < \nu_2, \nu_3$ ). Therefore the harmonic mean logarithms obey  $h_{32} > h_{21}, h_{13}$ . We can also choose any values of  $\tau_1$  and  $\lambda_1$  as before.

We must now choose the growth traits of strains 2 and 3 ( $\tau_2, \tau_3, \lambda_2, \lambda_3$ ) to satisfy the inequalities of equation (S10.6). We use a geometrical approach to understand the available region of trait space for these strains. The lag time for strain 3 is bounded from above and below according to (combining the second and third inequalities in equation (S10.6))

$$(\tau_1 - \tau_3)h_{13} + \lambda_1 < \lambda_3 < (\tau_2 - \tau_3)h_{32} + \lambda_2. \quad (\text{S10.7})$$

The upper and lower bounds are both functions of  $\tau_3$ . The upper bound will be above the lower bound as long as  $\tau_3$  satisfies

$$\tau_3 < \frac{\tau_2 h_{32} - \tau_1 h_{13} + \lambda_2 - \lambda_1}{h_{32} - h_{13}}. \quad (\text{S10.8})$$

Since  $\tau_3$  must be positive, this upper bound of  $\tau_3$  must also be positive. The denominator of the right-hand side of equation (S10.8) is positive by assumptions about the yields, so therefore the numerator must be positive as well. This leads to a lower bound on the lag time  $\lambda_2$  of strain 2; we can combine this with an upper bound on  $\lambda_2$  from the first equation of equation (S10.6) (strain 2 beats strain 1) to obtain

$$\tau_1 h_{13} - \tau_2 h_{32} + \lambda_1 < \lambda_2 < (\tau_1 - \tau_2)h_{21} + \lambda_1. \quad (\text{S10.9})$$

Finally, the upper bound for  $\lambda_2$  will be above the lower bound as long as  $\tau_2$  satisfies

$$\tau_2 > \max \left( \left( \frac{h_{13} - h_{21}}{h_{32} - h_{21}} \right) \tau_1, 0 \right). \quad (\text{S10.10})$$

Altogether, we can construct a set of nontransitive strains by choosing any yields  $\nu_1, \nu_2, \nu_3$  satisfying  $\nu_1 < \nu_2, \nu_3$ , and any values for the growth traits  $\tau_1, \lambda_1$  of strain 1; we then choose  $\tau_2$  according to equation (S10.10) and  $\lambda_2$  according to equation (S10.9); finally, we choose  $\tau_3$  according to equation (S10.8) and  $\lambda_3$  according to equation (S10.7). These inequalities determine the shaded areas of trait space in figure 4b.

[1] Baranyi J, Roberts TA (1994) A dynamic approach to predicting bacterial growth in food. *Int J Food Microbiol* 23:277–294.

[2] Baranyi J (1998) Comparison of stochastic and deterministic concepts of bacterial lag. *J Theor Biol* 192:403–408.

[3] Pfeiffer T, Schuster S, Bonhoeffer S (2001) Cooperation

- and competition in the evolution of ATP-producing pathways. *Science* 292:504–507.
- [4] Swinnen IAM, Bernaerts K, Dens EJJ, Geeraerd AH, Impe JFV (2004) Predictive modelling of the microbial lag phase: a review. *Int J Food Microbiol* 94:137–159.
- [5] MacLean RC (2007) The tragedy of the commons in microbial populations: insights from theoretical, comparative and experimental studies. *Heredity* 100:471–477.
- [6] Himeoka Y, Kaneko K (2017) Theory for transitions between exponential and stationary phases: Universal laws for lag time. *Phys Rev X* 7:021049.
- [7] Novak M, Pfeiffer T, Lenski RE, Sauer U, Bonhoeffer S (2006) Experimental tests for an evolutionary trade-off between growth rate and yield in *E. coli*. *Am Nat* 168:242–251.
- [8] Fitzsimmons JM, Schoustra SE, Kerr JT, Kassen R (2010) Population consequences of mutational events: effects of antibiotic resistance on the r/K trade-off. *Evol Ecol* 24:227–236.
- [9] Warringer J et al. (2011) Trait variation in yeast is defined by population history. *PLoS Genet* 7:e1002111.
- [10] Jasmin JN, Zeyl C (2012) Life-history evolution and density-dependent growth in experimental populations of yeast. *Evolution* 66:3789–3802.
- [11] Jasmin JN, Dillon MM, Zeyl C (2012) The yield of experimental yeast populations declines during selection. *Proc R Soc B* 279:4382–4388.
- [12] Bachmann H et al. (2013) Availability of public goods shapes the evolution of competing metabolic strategies. *Proc Natl Acad Sci USA* 110:14302–14307.
- [13] Crow JF, Kimura M (1970) *An Introduction to Population Genetics Theory*. (Harper and Row, New York).
- [14] Elena SF, Lenski RE (2003) Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat Rev Genet* 4:457–469.