The interplay of phenotypic variability and fitness in finite microbial populations

Ethan Levien[1], Jane Kondev[2], and Ariel Amir[3]

[1,3]John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
[1,2]Department of Physics, Brandeis University, Waltham, MA 02453 USA

Abstract

In isogenic microbial populations, phenotypic variability is generated by a combination of stochastic mechanisms, such as gene expression, and deterministic factors, such as asymmetric segregation of cell volume. Here we address the question: how does phenotypic variability of a microbial population affect its fitness? While this question has previously been studied for exponentially growing populations, the situation when the population size is kept fixed has received much less attention, despite its relevance to many natural scenarios. We show that the outcome of competition between multiple microbial species can be determined from the distribution of phenotypes in the culture using a generalization of the well-known Euler-Lotka equation, which relates the steady-state distribution of phenotypes to the population growth rate. We derive a generalization of the Euler-Lotka equation for finite cultures, which relates the distribution of phenotypes among cells in the culture to the exponential growth rate. Our analysis reveals that in order to predict fitness from phenotypes, it is important to understand how distributions of phenotypes obtained from different subsets of the genealogical history of a population are related. To this end, we derive a mapping between the various ways of sampling phenotypes in a finite population and show how to obtain the equivalent distributions from an exponentially growing culture. Finally, we use this mapping to show that species with higher growth rates in exponential growth conditions will have a competitive advantage in the finite culture.

1 Introduction

A central problem in microbiology is understanding how measurable phenotypic traits such as a cell's size, growth rate, or expression of a gene affect population dynamics [1, 2, 3, 4]. This problem is made difficult by the fact that there is rarely a well-defined mapping from genotype to phenotype, instead, a single genotype can give rise to a distribution of phenotypic traits throughout a population's history. This distribution can be generated at the cellular level due to *intrinsic* factors, such as the stochasticity of biochemical reactions [5, 6, 7, 8, 9, 10], or due to *environmental* fluctuations that affect the entire population [11, 12, 13, 14, 15]. It is now well established that a population's long term growth rate cannot be determined by the average values of single-cell traits, such as the growth rates or generation times of cells, but is instead determined by the variability and heritability of phenotypes measured throughout a population. The discrepancy is an example of *non-ergodicity* and results from epigenetic heritability of phenotypic traits that is present in all populations. It is therefore necessary to quantify the distribution of a phenotypic trait, and not just its average, in order to deduce how phenotypes affect fitness.

In the setting of exponential growth, a proxy for fitness is the population growth rate Λ, defined by the relation $N \sim \mathrm{e}^{\Lambda t}$, where N is the number of cells at time t. The problem of how Λ is related to the distribution of generation times has a rich history in the field of mathematical population dynamics, starting with Euler in the 17th century. Euler derived an implicit relationship between Λ and the birth and death rates of individuals in a growing population [16]. Euler's result was rediscovered by Lotka [16] and later adapted to microbial populations by Powell [17]. In its modern form, this result, known as the *Euler-Lotka equation*, states that for an exponentially growing population in which each individually has a generation time, τ, drawn independently from a distribution $f(\tau)$, Λ satisfies the implicit relationship

$$\frac{1}{2} = \int_0^\infty f(\tau) e^{-\Lambda\tau} d\tau. \tag{1}$$

Note that when generation times are deterministic ($f(\tau) = \delta(\tau - \tau_0)$), this reduces to relationship between the population's doubling time and growth rate, $\Lambda = \ln(2)/\tau_0$. Equation (1) illustrates the importance of considering variation in phenotypic traits, not just the averages, as it can clearly be seen that the entire

distribution of generation times shapes Λ. However, this formula has limited applicability to understanding fitness in microbial populations because generation times are typically not drawn independently, but rather they are correlated between mother and daughter cells. For example, negative correlations can arises due to cell-size regulation: if a cell grows abnormally large, its daughter cell will need to have a shorter generation time to account for the additional size inherited from its mother [18].

The relationship between the distribution of generation times and fitness in the presence of mother-daughter correlations is given by a more general equation, derived by Lebowitz [19]. This states that

$$\frac{1}{2} = \int_0^\infty \mathrm{e}^{-\Lambda\tau}\psi_{\mathrm{tree}}(\tau)d\tau, \tag{2}$$

where $\psi_{\mathrm{tree}}(\tau)$ is the distribution of generation times over the entire history of a growing population; see Figure 1 (a). Note that in generating this distribution, it is crucial to include the generation times of cells *currently* in the culture. The distribution $\psi_{\mathrm{tree}}(\tau)$ is, in general, distinct from the distribution of generation times along an isolated lineage. If $f(\tau|\tau')$ is the distribution of generation times conditioned on the mother cell's generation time, it was shown that $\psi_{\mathrm{tree}}(\tau)$ satisfied the self-consistent equation [19]

$$\psi_{\mathrm{tree}}(\tau) = 2\int_0^\infty f(\tau|\tau')\psi_{\mathrm{tree}}(\tau')\mathrm{e}^{-\Lambda\tau'}d\tau'. \tag{3}$$

In both natural and experimental settings, populations cannot grow exponentially for very long. For example, in rich growth conditions *Escherichia coli* has a doubling time of about 20 minutes and the weight of an individual bacteria is on the order of 1 picograms, hence a single day of continuous exponential growth starting from one cell would yield a colony that weighs around $2^{24\times 3} \approx 4.7\times 10^{21}$ picograms, or around 5 tons. Another 24 hours of doubling would yield a colony roughly 4 times the mass of the earth. In reality, populations of microbes are periodically diluted giving a population whose size is fixed, at least on average. For example, in a chemostat cells are continuously flushed out as new media is pumped into the culture. Analogous processes occur in natural environments, such as the digestive system of a host organism where waste is periodically disposed of. In both examples, exponential growth is an inaccurate model of the population dynamics. Instead, one could imagine a fixed population of N cells. One way to model such a population is with a *Moran process*, where N is kept fixed by selecting a cell to remove from the population at every division event. In the variant of the Moran process that we will adopt, the cell to remove is randomly selected from all of the $N+1$ cells in the population after every division (which includes the $N-1$ that did not divide and the two newborn daughter cells), thus reverting the population size back to N immediately after any cell divides. The assumption of this model, which we will study in the large N limit, is that the population is well-mixed and hence the probability of any given cell being removed is independent of its age or genealogical history. If two species are present in such a population, the principle of *competitive exclusion* state that one of them will eventually dominate the population.

What does the distribution of phenotypes tell us about the fate of a species in a finite population? The standard theory [20] tells us that if two species with growth rates Λ_1 and Λ_2 compete in a Moran process, the species with a larger population growth rate will eventually dominate. However, equation (2) tells us that the exponential growth rates are determined by the entire history of cells in a growing population, yet in the Moran process, most of the cells are flushed out. This raises the question: How is the outcome of a competition experiment related to the distribution of generation times, or any other measurable trait affecting growth, among the cells *in* the culture? In this paper, we will address this question by analyzing the dynamics of phenotypically heterogeneous populations. This will lead to a generalization of equation (2) for finite populations, equation (20), which relates the population growth rate to the distribution of generation times sampled over cells that remain in the culture. We find subtle distinctions between the phenotype distributions in the exponentially growing population and the finite population. However, it will be shown that the ultimate fate of species in the competition setting is independent of whether the population is growing or kept at a fixed size. In particular, one of our main results is that species with higher values of Λ_i in the exponentially growing culture will always dominate the finite culture, regardless of the initial conditions.

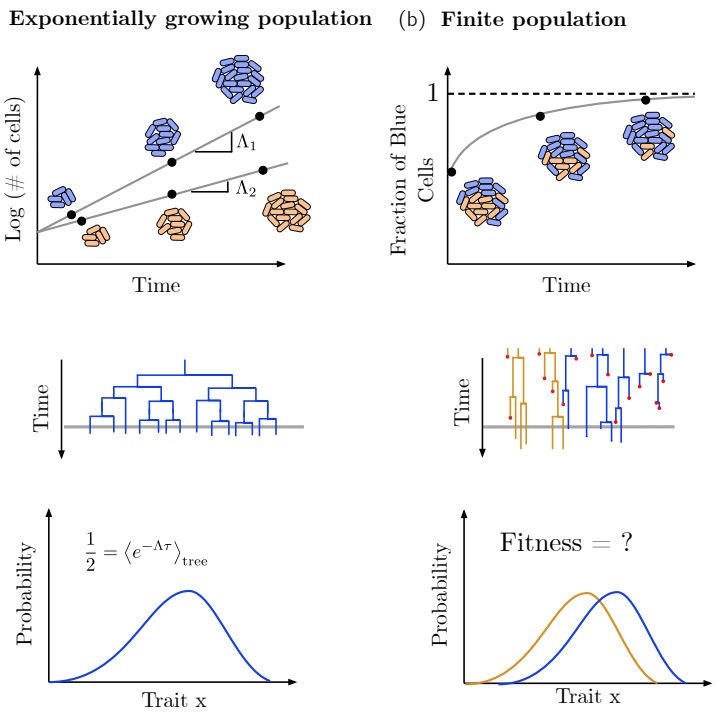

Figure 1: (a) (top) In a population growing exponentially the population growth rate, Λ, which is a proxy for fitness, can be computed from temporal dynamics of the number of cells. (middle) A cartoon of genealogies obtained from an exponentially growing population. Each vertical line represents a cell and the length of the line its generation time. (bottom) A histogram of some phenotypic trait that determines generation times obtained from the genealogical tree can be used to determine the fitness according to the Euler-Lotka equation. The precise procedure for sampling the phenotypes will be discussed in Section 3. (b) (top) In a finite population containing non-interacting genotypes, the principle of competitive exclusion says that one genotype will eventually dominate the culture. (middle) A cartoon of genealogies obtained from the two populations competing in the finite culture. (bottom) How do the distributions of phenotypes obtained from the genealogies in the finite culture relate to fitness?

2 A general model of phenotypic variability

We start by introducing a general modeling framework allowing us to study intrinsic variability in an arbitrary phenotypic trait, which will be represented by the vector $\mathbf{x} = (x_1, \ldots, x_l) \in \mathbb{R}_+^l$. We assume each cell is born with a value of $\mathbf{x}$ drawn from a distribution $f(\mathbf{x}|\mathbf{x}')$ depending on the phenotype $\mathbf{x}'$ of the mother cell. Importantly, $\mathbf{x}$ may not change over the cell-cycle. For example, $\mathbf{x}$ could represent the initial concentration of a protein, or some macroscopic phenotype, such as the average growth rate over the cell cycle. However, it could not represent the age or volume of a cell, since these change over the course of the cell cycle. After a time $\tau(\mathbf{x})$, the cell divides and produces two new cells with phenotypes drawn from $f(\mathbf{x}|\mathbf{x}')$ and generation times depending on these phenotypes. This framework can capture any possible form of *intrinsic* (as opposed to environmental) generation time variability through the appropriate choices of the phenotype $\mathbf{x}$ and distribution f. Therefore, existing models can be derived as specific cases of this framework. These include models where division is asymmetric [21, 22, 23, 24] and where the relationship between successive generation times is non-monotonic and nonlinear, such as the kicked cell-cycle model used to describe circadian rhythms [25].

Example: cell-size regulation model

As a concrete example, we will consider a model of cell-size regulation, which has been used extensively in the literature to understand how cells maintain homeostasis of their sizes [1]. This model will serve as a useful illustration of how our formalism can capture the effects of stochasticity in multiple phenotypic traits. The most important assumption of the cell-size regulation model is that cells grow exponentially at the single cell level and divide upon reaching a size depending on their size at birth, v_{birth}. We will assume that cells divide symmetrically so that v_{birth} is obtained by dividing the cells mother's size at division by 2. Phenotypic variability is introduced by adding noise to both the single-cell growth rate as well as the volume at division. To this end, we take the growth rate λ and division volume v_{div} of a cell to obey

$$\begin{aligned} \ln \lambda &= \ln\langle\lambda\rangle + \eta_\lambda \\ v_{\text{div}} &= 2(1-\alpha)v_{\text{birth}} + 2\alpha v_0 + \eta_v \end{aligned} \tag{4}$$

where η_λ and η_v are independent normally distributed random variables. The first equation captures the fact that growth rates are positive and approximately follow a Gaussian distribution for small noise, while the second equation tells us how cells decide when to divide based on their volume. Here, the cell-size regulation strategy is controlled by the parameter α. $\alpha = 1$ corresponds to cells dividing at a critical size (known as a "sizer"), while $\alpha = 1/2$ corresponds to cells adding a constant size v_0 (known as an "adder"). We refer to [1] for an in-depth discussion of the cell-size control model and its implications for population growth. Within this model, the generation time of a cell can be expressed as a deterministic function of the growth rate noise. In particular,

$$\begin{aligned} &\tau(\eta_\lambda, \eta_v, v_{\text{birth}}) \\ &= \frac{1}{\langle\lambda\rangle e^{\eta_\lambda}} \ln\left(\frac{2(1-\alpha)v_{\text{birth}} + 2\alpha v_0 + \eta_v}{v_{\text{birth}}}\right). \end{aligned} \tag{5}$$

It follows that in the notation of our general modeling framework the phenotype that controls generation times is $\mathbf{x} = (\eta_\lambda, \eta_v, v_{\text{birth}})$. The conditional distribution of phenotypes is given by

$$\begin{aligned} f(\mathbf{x}|\mathbf{x}') = \delta&\left(v_{\text{birth}} - \frac{v'_{\text{birth}}}{2} e^{\lambda' \tau(\eta'_\lambda, \eta'_v, v'_{\text{birth}})}\right) \\ &\times \left(\frac{1}{\sqrt{2\pi\sigma_\lambda^2}} e^{-\eta_\lambda^2/2\sigma_\lambda^2}\right)\left(\frac{1}{\sqrt{2\pi\sigma_v^2}} e^{-\eta_v^2/2\sigma_v^2}\right), \end{aligned} \tag{6}$$

where σ_λ^2 and σ_v^2 are the variances of η_λ and η_v, respectively Note that in the special case where $\alpha = 1$ and $\eta_v = 0$, the phenotype is given by $\mathbf{x} = \ln(\lambda)$ and the generation time is simply $\tau = \ln(2)/\lambda$.

3 Relationship between phenotypic variability and fitness in a finite culture

We now consider a model where N, the number of cells in the population, is fixed. In order to achieve this, every division event must correspond to a cell being expelled from the culture. As discussed in the introduction, this can be implemented by a *Moran process* [26]. Since the population is not growing, we can no longer obtain Λ from the exponential growth curves (as shown in Figure 1 (a)). Instead, one can measure growth in terms of the rate per cell at which cells are expelled from the population, or the *dilution rate* Λ_D.

Intuitively, if N is sufficiently large, the continuous expulsion of cells from the culture should have no influence on the relative frequencies of the phenotypes, and therefore the instantaneous rate at which cells are accumulated should be equal to ΛN. On the other hand, since the size of the culture remains fixed, this must be balanced by the rate at which cells are expelled, implying $\Lambda \approx \Lambda_D$ for sufficiently large N. It follows that Λ_D should satisfy the same relation as Λ. To make this intuition mathematically precise, we can derive a transport equation for the joint distribution of phenotypes ($\mathbf{x}$) and ages (u) at a given time, denoted $\psi(t, \mathbf{x}, u)$, which is given by (see Appendix A),

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial u}\right)\psi(t, \mathbf{x}, u) = -\Lambda_D \psi(t, \mathbf{x}, u). \tag{7}$$

This equation is the Von Foerster equation for an age structured population [16], although the role of the "death rate" is played by Λ_D, which is no longer an input to the model. Instead, Λ_D must be derived from a boundary condition to ensure the population remains finite. This boundary condition is

$$\psi(t, \mathbf{x}, 0) = 2 \int_0^\infty \cdots \int_0^\infty f(\mathbf{x}|\mathbf{x}')\psi(t, \mathbf{x}', \tau(\mathbf{x}'))dx_1' \cdots dx_l'. \tag{8}$$

The term $f(\mathbf{x}|\mathbf{x}')\psi(t, \mathbf{x}', \tau(\mathbf{x}'))dx_1' \cdots dx_l'$ in the integrand represents the probability of a cell with phenotype $\mathbf{x}'$ dividing to produce a cell with phenotype $\mathbf{x}$. The factor of 2 comes from the fact that each cell produces two daughters. Importantly, the probability that any specific cell is expelled is $1/N$, which vanishes in the limit $N \to \infty$; therefore, equation (8) does not need to account for the fact that one of the daughter cells could be expelled. Since each cell produces two offspring, integrating over all phenotypes and multiplying by a factor of 2 yields the distribution of phenotypes and ages evaluated at age $u = 0$. Note that this is proportional to the distribution of phenotypes among newborn cells, denoted $\psi^F_{\text{birth}}(\mathbf{x})$.

We now look at equation (7) in steady-state by setting the time derivative to zero,

$$\frac{\partial}{\partial u}\psi(\mathbf{x}, u) = -\Lambda_D \psi(\mathbf{x}, u). \tag{9}$$

Solving equation (9) yields

$$\psi(\mathbf{x}, u) = \psi(\mathbf{x}, 0)e^{-\Lambda_D u}, \quad u < \tau(\mathbf{x}). \tag{10}$$

As shown in Appendix A, $\psi(\mathbf{x}, 0)$ is proportional to the *tree distribution*, denoted $\psi^F_{\text{tree}}(\mathbf{x})$. In order to obtain the tree distribution, we sample the phenotype of every cell throughout the entire history of the population, including those that are expelled from the culture, and make a histogram of their phenotypes. (In general we will use the superscript F to indicate a distribution in the finite culture, but we omit this when there is no ambiguity.) After inserting (10) into (8), we find

$$\psi^F_{\text{tree}}(\mathbf{x}) = 2 \int_0^\infty \cdots \int_0^\infty f(\mathbf{x}|\mathbf{x}')\psi^F_{\text{tree}}(\mathbf{x}')\mathrm{e}^{-\Lambda_D \tau(\mathbf{x}')}dx_1' \cdots dx_l'. \tag{11}$$

When phenotypes are not correlated across generations, $f(\mathbf{x}|\mathbf{x}') = f(\mathbf{x})$, hence it follows from (11) that $\psi^F_{\text{tree}}(\mathbf{x}) = f(\mathbf{x})$. Otherwise, $\psi^F_{\text{tree}}(\mathbf{x})$ will be distinct from the distribution of generation times obtained from a single, isolated lineage. We may integrate with respect to $\mathbf{x}$ and replace Λ_D with Λ in order to obtain

$$\frac{1}{2} = \int_0^\infty \cdots \int_0^\infty \psi^F_{\text{tree}}(\mathbf{x})e^{-\Lambda\tau(\mathbf{x})}dx_1' \cdots dx_l'. \tag{12}$$

This generalizes equation (3) in two ways: first, it extends equation (3) to the context of a finite culture, and second, it relates Λ to a distribution of a more general phenotype, not only the generation times. Using this formalism, it is possible to derive the self-consistent equation for $\psi_{\text{tree}}(\tau)$ in the exponentially growing culture (equation (3)) in terms of $\psi^F_{\text{tree}}(\tau)$, the tree distribution of generation times in the finite culture. To begin, we note that $\psi^F_{\text{tree}}(\tau)$ can be obtained by marginalizing over $\psi^F_{\text{tree}}(\mathbf{x})$:

$$\psi^F_{\text{tree}}(\tau) = \int_{\{\mathbf{x}:\tau(\mathbf{x})=\tau\}} \psi^F_{\text{tree}}(\mathbf{x})dx_1 \cdots dx_l. \tag{13}$$

Here, $\psi^F_{\text{tree}}(\tau)$ is the distribution of generation times over the history of the culture, including the generation times of those cells that are expelled [1]. Applying the same marginalization procedure to the right-hand side of equation (11) yields

$$\begin{aligned}\psi^F_{\text{birth}}(\tau) &= 2 \int_0^\infty \left[\int_{\{\mathbf{x}':\tau(\mathbf{x}')=\tau'\}} \left[\int_{\{\mathbf{x}':\tau(\mathbf{x}')=\tau'\}} f(\mathbf{x}|\mathbf{x}')dx_1 \cdots dx_l\right] \psi^F_{\text{tree}}(\mathbf{x}')\mathrm{e}^{-\Lambda_D\tau'}dx_1' \cdots dx_l'\right] d\tau' \\ &= 2 \int_0^\infty f(\tau|\tau')\psi^F_{\text{tree}}(\tau')\mathrm{e}^{-\Lambda_D\tau'}d\tau'.\end{aligned} \tag{14}$$

[1] We emphasize that when computing $\psi^F_{\text{tree}}(\tau)$ one needs to know the generation times of the cells that are expelled from the culture, not the age at which they are expelled.

This equation is identical to equation 3, implying $\psi^F_{\text{tree}}(\tau)$ obeys the same self-consistent equation as $\psi_{\text{tree}}(\tau)$. It follows that the tree distribution in the finite culture is identical to the tree distribution in the exponentially growing culture. Integrating both sides of equation (14) with respect to τ yields the Euler-Lotka equation in the finite culture:

$$\frac{1}{2} = \int_0^\infty \psi^F_{\text{tree}}(\tau) \mathrm{e}^{-\Lambda_D \tau} d\tau. \tag{15}$$

A similar procedure can used to obtain a simple equation for the population growth rate in terms of the distribution of any single-cell observable. For example, consider the cell-size regulation model with $\alpha = 1$ and $\eta_v = 0$. In this case, the generation time is a deterministic function of the growth rate: $\tau = \ln(2)/\lambda$. It follows that the marginalization approach above can be used to obtain

$$\frac{1}{2} = \int_0^\infty 2^{-\Lambda/\lambda} f(\lambda) d\lambda. \tag{16}$$

Here we have used that growth rates are uncorrelated to replace $\psi^F_{\text{tree}}(\lambda)$ with $f(\lambda)$. Within this model, $f(\lambda)$ is simply a log-normal distribution with mean $\langle\lambda\rangle$ and variance σ^2_λ. This equation has previously been analyzed in [27], where a saddle point approximation was carried out to obtain

$$\Lambda \approx \langle\lambda\rangle \left[1 - \left(1 - \frac{\ln(2)}{2}\right) \mathrm{CV}^2_\lambda\right], \tag{17}$$

where CV_λ represent the coefficient of variation (the standard deviation over the mean) of growth rates. This shows that within the context of this model, growth rate variation tends to decrease population growth [27].

4 Computing the population growth rate from cells in the culture

The tree distribution, which is required to compute the population growth rate using equation (2), is only one of a number of ways to quantify phenotypic variability in a lineage tree and it is not always the most convenient to work with. For example, it is often the case that one does not have access to the cells that are expelled from the culture. An alternative distribution is given by the expression $2\psi^F_{\text{tree}}(\mathbf{x})\mathrm{e}^{-\Lambda_D \tau(\mathbf{x})}$; this term is twice the integrand of equation (12), which implies it must be a probability distribution, as it is normalized. As was previously noted by Powell [17] in the context of a model with uncorrelated generation times, this expression can be interpreted as the distribution of $\mathbf{x}$ over all cells that divide before being expelled from the culture; see Figure 2. Following Powell's terminology, we define the *carrier distribution* as

$$\psi^F_{\text{car}}(\mathbf{x}) = 2\psi^F_{\text{tree}}(\mathbf{x})\mathrm{e}^{-\Lambda_D \tau(\mathbf{x})}. \tag{18}$$

To derive Powell's interpretation of the carrier distribution in the more general setting, first note that $\mathrm{e}^{-\Lambda_D \tau(\mathbf{x})}$ is the probability that a cell stays in the culture until it divides, given that is has phenotype $\mathbf{x}$. Also, the probability that any cell throughout the history of the population divides in the culture must be $1/2$ in order for the number of cells to remain fixed. The interpretation of $\psi^F_{\text{car}}(\mathbf{x})$ then follows from an application of Bayes' Theorem where $\psi^F_{\text{tree}}(\mathbf{x})$ and $\psi^F_{\text{car}}(\mathbf{x})$ play the role of the *prior* and *posterior*, respectively.

In practice the distribution $\psi^F_{\text{car}}(\mathbf{x})$ may be easier to obtain than $\psi^F_{\text{tree}}(\mathbf{x})$ because sampling $\psi^F_{\text{tree}}(\mathbf{x})$ requires knowledge of generation times of cells that are expelled from the culture. In certain experimental settings, these cells may be difficult to track, and additionally, their growth may be not be statistically identically to those cells that remain in the culture due to the different conditions in and out of the device where cells are observed [28]. This motivates us to develop a method for estimating the population growth rate, Λ, from the available lineage data. To this end, we utilize the normalization of $\psi^F_{\text{tree}}(\mathbf{x})$ to obtain

$$2 = \int_0^\infty \cdots \int_0^\infty e^{\Lambda \tau(\mathbf{x})} \psi_{\text{car}}(\mathbf{x}) dx_1 \cdots dx_l. \tag{19}$$

As we have done above in equation (13), we can marginalize over the phenotype $\mathbf{x}$ to recast this equation in terms of the corresponding generation time distribution:

$$2 = \int_0^\infty e^{\Lambda \tau} \psi_{\text{car}}(\tau) d\tau. \tag{20}$$

Equation (20) addresses the question of how Λ is related to the distribution generation times among cells *in* the culture.

The ancestral and division distributions

In the exponentially growing culture the distribution $2\psi_{\text{tree}}(\mathbf{x})e^{-\Lambda\tau(\mathbf{x})}$ has a different interpretation; it is equal to the distribution over all the current cell's ancestors, denoted $\psi_{\text{ans}}(\mathbf{x})$ (this has been referred to as the *branch distribution* in [27]). This distribution is obtained by sampling all cells in the tree excluding those that are currently alive, as shown in Figure 2 (a). The fact that $\psi^F_{\text{car}}(\mathbf{x}) = \psi_{\text{ans}}(\mathbf{x})$ suggests there may be an alternative interpretation of these distributions that is common to both types of cultures. In fact, in both the continuous culture and exponentially growing culture, the carrier (or branch) distribution can be interpreted as the distribution of phenotypes of cells that have just divided, or the *division distribution* $\psi_{\text{div}}(\mathbf{x})$. We refer to Appendix B for the derivation.

Now consider the the distribution over the ancestral cells in the finite culture, denoted $\psi^F_{\text{ans}}(\mathbf{x})$. In order to obtain $\psi^F_{\text{ans}}(\mathbf{x})$, we take the phenotypes of all cells currently alive in the culture and trace back through their ancestors, collecting their phenotypes. How might this distribution be related to $\psi_{\text{ans}}(\mathbf{x})$ in the exponential culture (or equivalently, the carrier distribution in the finite culture, $\psi^F_{\text{car}}(\mathbf{x})$)? Because all cells have a common ancestor we can compute $\psi^F_{\text{ans}}(\mathbf{x})$ by taking only the ancestor cells of this common ancestor – as we trace back in time far enough the more recent cells make up a negligible fraction of the observed phenotypes. $\psi_{\text{ans}}(\mathbf{x})$ is therefore the distribution of phenotypes obtained by looking *backwards* along a lineage beginning from an arbitrary cell in the culture. If phenotypes are not correlated across generations, then the phenotypes sampled along the lineage are independent and this is therefore equivalent to sampling cells in the population that have divided in the culture; as we saw before, this is $\psi_{\text{car}}(\mathbf{x})$. However, if there are correlations across generations, we need to consider how the ancestor's phenotypes along a lineage are biased by the descendants' phenotypes. This is captured by the recursive formula for $\psi_{\text{ans}}(x)$ (derived in Appendix B):

$$\psi^F_{\text{ans}}(\mathbf{x}) = \psi^F_{\text{car}}(\mathbf{x}) \int_0^\infty \cdots \int_0^\infty f(\mathbf{x}'|\mathbf{x}) \frac{\psi^F_{\text{ans}}(\mathbf{x}')}{\psi^F_{\text{tree}}(\mathbf{x}')} dx'_1 \cdots dx'_l. \tag{21}$$

From this equation it is easy to see that $\psi^F_{\text{ans}}(\mathbf{x}) = \psi^F_{\text{car}}(x)$ precisely when $f(\mathbf{x}|\mathbf{x}') = f(\mathbf{x})$, as expected from the argument above.

Our results concerning the relationship between phenotype distributions in exponential and finite cultures are summarized in Figure 2 (c). To explicitly demonstrate these relationships we performed simulations of the cell-size regulation model. Results of these simulations are shown in Figure 2 (d). For realistic parameter values, we see that the differences between the distributions are quite small.

5 Competition between multiple species in a finite population

Having established that the Euler–Lotka equation can be generalized to the setting of the a finite culture in steady-state, we now consider a model in which two cells compete in a finite culture. We will once again assume that the number of cells (N) is sufficiently large so that any stochastic fluctuations at the population level may be ignored. Before considering these competition dynamics directly, we suppose that the two species are grown separately in environments with unlimited space and nutrients. Under these conditions, the number N_i of cells in species i will then grow exponentially at some rate Λ_i. It is not difficult to see that the fraction of species 1, $\phi_1 = N_1/(N_1 + N_2)$, evolves according to the logistic growth equation

$$\frac{d}{dt}\phi_1 = \phi_1(1-\phi_1)\Delta\Lambda, \quad \Delta\Lambda = \Lambda_1 - \Lambda_2. \tag{22}$$

The quantity $S = \Delta\Lambda/\Lambda_2$ is known as the *selective advantage* of species 1 over species 2 and its sign determines which species eventually takes over the culture.

We now imagine placing both species in a finite culture that can hold exactly N cells; see Figure 1 (b). As before, this is implemented by removing a randomly selected cell from the culture every time a cell division occurs. For the moment, we will assume that nutrients are not limiting growth and the environment is constant. Under these assumptions, we can heuristically obtain the dynamics of the frequency

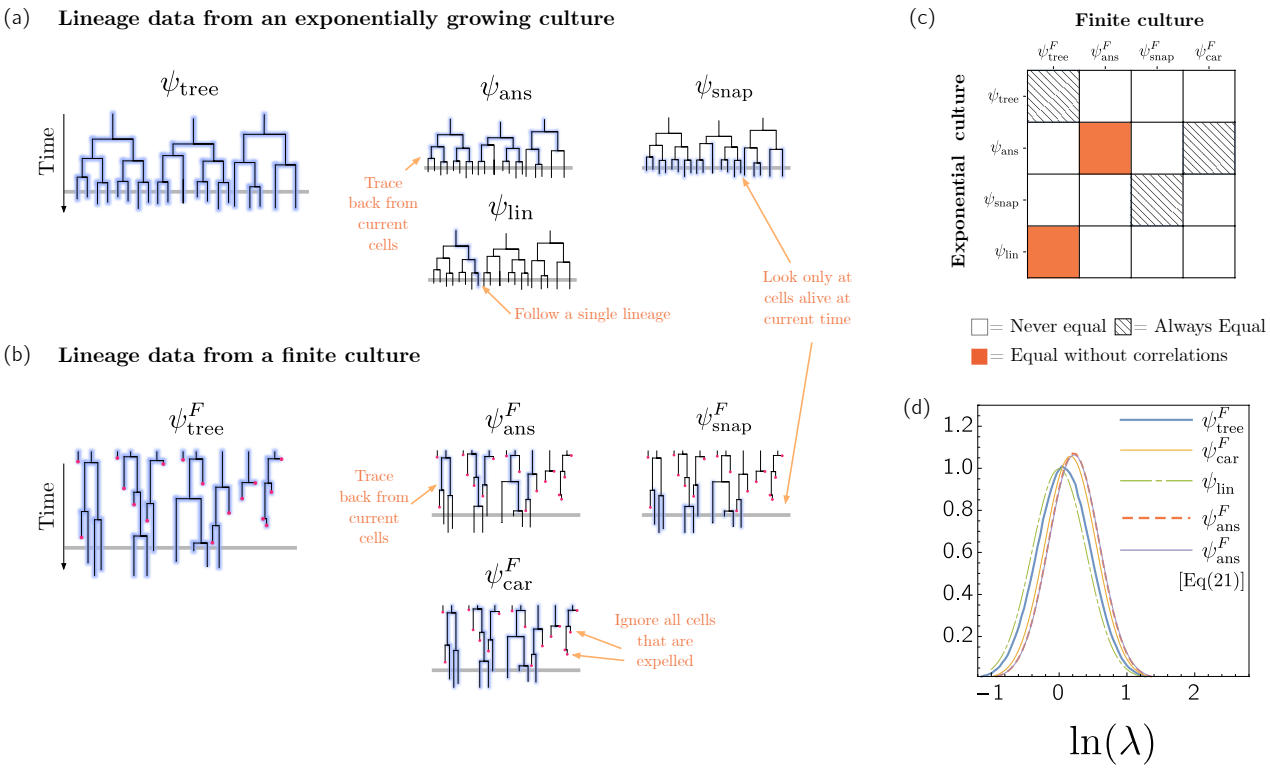

Figure 2: (a) (left) The various ways of sampling phenotypes from a lineage tree obtained in an exponential culture. The gray horizontal line on the right panel indicates the current time of observation. The blue highlighted lines correspond to cells that are sampled to obtain the distribution indicated above the lineage tree. For example, to obtain ψ_{ans} we sample all the cell in the tree excluding those that are currently alive. (b) The same image for a finite culture. The red dots represent cells that are expelled from the culture before they divide. (c) The relationship between distributions in the finite and exponentially growing culture. (d) Plots of the various distributions of single-cell growth rates obtained from simulations of the cell-size regulation model with $N = 10^4$, $CV^2_\lambda = 0.25$, $\langle\lambda\rangle = 1$, $v_{\text{div}} = 2$, $\alpha = 1$ and $\rho_\lambda = 0.2$. The lineage distribution, which can be computed analytically for this model, is also shown. For the ancestral distribution, we have plotted both the histogram of $\ln\lambda$ over all ancestral cells (red dashed line) as well as the result of evaluating the right hand side of equation (21), using the carrier and tree distributions (purple line).

of species 1 in the finite setting, $\phi_1(t)$, from the exponentially growing population by uniformly sampling N cells from the exponentially growing population at each time; this would suggest that the outcome of a competition experiment is also determined by $\Delta\Lambda$. However, we know that the steady-state phenotype distribution determines the population growth rates. Therefore, in order to justify the heuristic argument mathematically, it must be established that when two species are placed in a finite culture, their phenotype and age distributions converge to their steady-states. If this condition is not met, then there is a possibility that the outcome of the competition depends on the initial phenotype distribution. From a biological perspective, it is important to understand what happens when the initial distribution is not in steady-state, as it is rarely the case that ecological dynamics begin with each species in steady-state. For example, two species from different environments may meet after being consumed by the same host organism. Unless the conditions within the host are identical to the conditions outside, the populations will not be in steady-state at the onset of the competition.

In Appendix C, we have derived the dynamics of $\phi_1(t)$ within the context of the two-species Moran process model, along with the generalization to the case of M species. There, we show that the joint density

of phenotypes and ages among species i, denoted $\psi_i(t, \mathbf{x}, u)$, obeys

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial u}\right) \psi_i(t, \mathbf{x}, u) = -\Lambda_i(t)\psi_i(t, \mathbf{x}, u). \tag{23}$$

Here $\Lambda_i(t)$ is the per capita division rate of cells of species i, which may depend on time if the system has not reached its steady-state. In terms of ψ_i,

$$\Lambda_i(t) = \int_0^\infty \cdots \int_0^\infty \psi_i(t, \mathbf{x}, \tau_i(t, \mathbf{x}))dx_1 \cdots dx_l. \tag{24}$$

We have allowed for both the phenotype distribution and $\Lambda_i(t)$ to be time-dependent, since the culture may not be in steady-state. Equation (23) holds for $u < \tau_i(\mathbf{x})$ and is supplemented by the boundary conditions prescribing how the phenotypes are selected at birth,

$$\psi_i(t, \mathbf{x}, 0) = 2\int_0^\infty \cdots \int_0^\infty f(\mathbf{x}|\mathbf{x}')\psi_i(t, \mathbf{x}', \tau_i(\mathbf{x}'))dx_1' \cdots dx_l'. \tag{25}$$

Importantly, equations (23) and (25) are completely decoupled from the phenotype distributions of other species and $\phi_i(t)$. In other words, the distributions of phenotypes evolve completely independently from the frequencies of species in the culture. Some of our main results follow from this observation: First, in a competition experiment, the phenotype distributions will converge to their steady-states regardless of the initial conditions and relative fitnesses of each species. Second, the species with a higher value of Λ_i as determined by equation (2) will eventually dominate the culture, regardless of the initial state of the phenotype and age distributions. Finally, in the special case where the initial phenotype distributions are in their steady-states satisfying equation (8) and there are two species, the frequency $\phi_1(t)$ will obey the logistic growth equation (22).

In Figure 3 we have shown simulations for competition between two species modeled by the random growth rate model compared with the prediction of the logistic growth equation. When the competition begins with the phenotype distributions in their steady-state, we see that the simulations agree exactly with the prediction. If the initial phenotype distributions are not in steady-state, we find that the transient dynamics differ from the logistic growth equation; however, as predicted, they converge to the trajectory of the logistic equation within a few generations.

6 Discussion

In this paper, we have studied the dynamics of competition between phenotypically heterogeneous populations in finite cultures. While a number of studies have explored the dynamics of phenotypically heterogeneous populations under exponential growth conditions and shown how to relate lineage statistics to the population growth rate (e.g., [1, 27, 29, 30, 31, 32, 33]), little work has been done concerning the dynamics of multiple species in finite populations. In exponentially growing populations, the population growth rate is related to the steady-state phenotype distribution. When studying competition dynamics, it is not clear that the transient dynamics of the phenotype and age distributions do not affect the fate of species. We have shown that even if the initial populations are not initially in steady-state, the end result of a competition experiment can be predicted by the steady-state dynamics, which are the same for both exponential and finite cultures.

We have also clarified the relationship between the different distributions in both the exponentially growing and finite culture, as well as derived a simple mapping between the various distributions in the two types of cultures. Importantly, distributions obtained in equivalent ways do not always have the same meaning in different types of cultures. In particular, the ancestral distribution in a finite culture is not equivalent to the ancestral distribution of the exponentially growing culture due to a bias towards surviving cells, yet both are obtained by looking at the history of all cells in the culture. A corollary of our mapping between the lineage statistics in different cultures is a formula for the population growth rate (given by equation (20)) that requires knowledge only of the lineage data obtained in a sufficiently large finite culture. In addition to relating the lineage statistics to the population growth rate, we have shown that the population growth rate predicts the fitness of a species in a competition experiment.

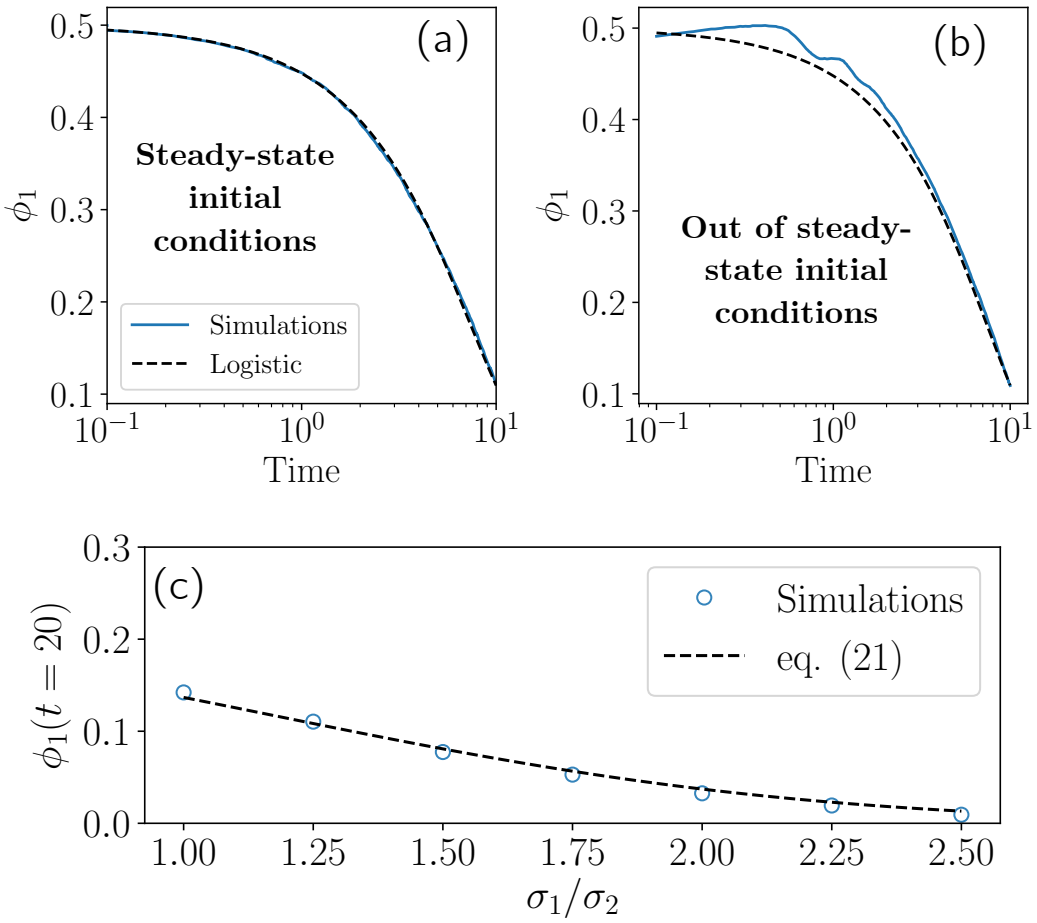

Figure 3: (a) and (b) Simulations of two species, each modeled by the random growth rate model, competing in the Moran process compared with the prediction of the logistic growth equation (22). The blue lines represent the average over 1000 simulations. In (a), the phenotype distributions are initially taken to be in steady-state, satisfying the self-consistent equation (3). In (b), the phenotype distributions are initially out of steady-state. This is achieved by drawing the initial growth rates from Gaussian distributions with different means. In both cases, we have used $\sigma_\lambda = 0.3$ for both species $\langle\lambda\rangle_1 = 1$ and $\langle\lambda\rangle_2 = 1.1$ (hence species 2 has an advantage). We have used the sizer model for cell-size regulation and neglected any size additive noise. (c) $\phi_1(t)$ at $t = 20$ generations for two populations with the same average growth rate and different values of σ_i^2, the variance in growth rates. We have used $\langle\lambda\rangle_1 = 1$ and $\langle\lambda\rangle_2 = 1.1$, and we have fixed $\sigma_2 = 0.2$ and increased σ_1. For both species growth rates are uncorrelated across generations. Each circle represents the average of 1000 simulations for the same parameter values.

By demonstrating that the outcomes of ecological competition are independent of the initial conditions, our results shed light on what is important in ecological dynamics of microbial populations. In particular, they suggest that, in the context of the Moran process, all memory of the transient phenotypic distributions is lost over the course of a few generations. This is important, as many natural populations do not live in constant environments, but are repeatedly transplanted to new environments where they are forced to compete (in a finite population) with new sets of species. For example, infectious bacteria may travel between many different hosts, each one containing distinct populations of microbes.

We emphasize that the simple Moran process in which species are non-interacting is a very simplified picture of ecological competition dynamics. However, by establishing the basic principles of competition in this context, we believe our results can help guide future studies seeking to understand higher order effects.

An interesting extension of our work could involve investigating species interactions. When species interact, the principle of competitive exclusion breaks down, and it is possible that two species can coexist indefinitely [20]. How does variability within species affect coexistence? A second extension would be to understand how demographic fluctuations (fluctuations due to the finite size of the population) influence the dynamics. If one allows for demographic fluctuations in the model, it is possible that even the species with a larger population growth rate can go extinct. Therefore, the more relevant quantity to consider is the *fixation probability*, defined as the probability that one species will eventually completely dominate the culture. How are the fixation probabilities related to the population growth rates? Do the initial conditions significantly affect the fixation probabilities? To our knowledge, these questions have not been answered in the existing literature.

Acknowledgements

We acknowledge Jie Lin for helpful discussions related to this work. We thank Farshid Jafarpour for offering helpful feedback on this manuscript. We acknowledge funding support from National Science Foundation under Award DMS-1902895 (EL), National Science Foundation grant DMR-1610737 (JK), MRSEC at Brandeis University DMR-1420382 (JK), the Simons Foundation (JK), MRSEC at Harvard University under grant DMR-1420570 (AA) and NSF CAREER 1752024 (AA).

A Derivation of Euler-Lotka equation

Here we give a more detailed derivation of equation 12. Let $\psi(t, \mathbf{x}, u)$ denote the joint distribution of species with phenotypes $\mathbf{x}$ and ages u in an instantaneous observation of a population with $N \gg 1$ microbes. We are mostly interested in the steady-state distribution $\psi(\mathbf{x}, u)$. In this section, we omit the superscript F because we are only concerned with the finite culture. We proceed by considering the change in the number of cells with phenotypes in the hypercube $\prod_{i=1}^{l}[x_i', x_i + dx_i)$ and ages between u and $u + du$ in an interval dt. This change is caused both by the aging of cells (not be confused with the way this term is used in [22], where it corresponds to the deterioration of growth capacity) and the expulsion of cells from the culture. Since each cell is chosen to be expelled with equal probability whenever a cell divides, the probability that any given cell is expelled is simply the fraction of cells that divide per unit time multiplied by dt, or $\Lambda_D dt$. It follows that

$$\begin{aligned}
&\underbrace{N\psi(t+dt, x, u)dx_1 \cdots dx_l dt - N\psi(t, \mathbf{x}, u)dx_1 \cdots dx_l dt}_{\text{change in number of cells with age } u \text{ and phenotype } \mathbf{x}} \\
&\approx \underbrace{N\psi(t, \mathbf{x}, u - dt)dx_1 \cdots dx_l dt}_{\text{cells gained from aging}} - \underbrace{N\psi(t, \mathbf{x}, u)dxdt}_{\text{cells lost from aging}} \\
&- \underbrace{N\psi(t, \mathbf{x}, u)\Lambda_D dx_1 \cdots dx_l dt}_{\text{cells expelled from the culture}}.
\end{aligned} \tag{26}$$

This equation is valid for cells that have not yet divided; hence, $\mathbf{x}$ and u are such that $u < \tau(\mathbf{x})$. Taking the limit $dt \to 0$ and dividing by dt gives the Von Foerster equation (7). Consider the distribution of phenotypes among cells that are just born. This is given by $\psi_{\text{birth}}(\mathbf{x}) = \psi(\mathbf{x}, 0)\Psi_0^{-1}$ where Ψ_0 is a normalization constant obtained by integrating $\psi(\mathbf{x}, 0)$; $\Psi_0 = \int_0^\infty \psi(\mathbf{x}, 0)dx_1 \cdots dx_l$. Empirically, $\psi_{\text{birth}}(\mathbf{x})$ is obtained by taking an instantaneous observation of the population and sampling only the cells that have just been born. By solving equation (9) and inserting the solution into (8), it is straightforward to obtain

$$\psi_{\text{birth}}(\mathbf{x}) = 2\int_0^\infty f(\mathbf{x}|\mathbf{x}')\psi_{\text{birth}}(\mathbf{x}')\mathrm{e}^{-\Lambda_D \tau(\mathbf{x}')}dx_1 \cdots dx_l. \tag{27}$$

Because this distribution is time invariant, sampling those cells that have just been born at any sequence of times $t_1, t_2, \ldots t_n$ yields the same distribution. Carrying out this sampling procedure at an infinite sequence of times samples every cell only once, and therefore $\psi_{\text{birth}}(\mathbf{x}) = \psi_{\text{tree}}(\mathbf{x})$. Integrating both sides of equation (27) results in equation (12).

B Relationship between phenotype distributions

The age distribution

The distribution of ages in the culture $g(u)$ is given by

$$g(u) = \int_{x:\tau(x)>u} \psi(\mathbf{x}, 0) \mathrm{e}^{-u\Lambda_D} dx_1 \cdots dx_l. \tag{28}$$

Another way to derive this expression is by noting that it is the probability a cell is not flushed out by age u, times the probability that it has not yet divided by age u. In the specific context of the random generation time model ($\tau = \mathbf{x}$), we obtain the result found by Powell [17] without generation time correlations:

$$g(u) = \Psi_0 \mathrm{e}^{-u\Lambda_D} \int_u^\infty \psi_{\text{birth}}(\tau) d\tau. \tag{29}$$

To find Ψ_0, note that the number of cells that divide in a time dt is the number which have ages equal to their generation time, hence

$$N\Lambda_D dt = N dt \int_0^\infty \psi(s, s) ds = N dt \Psi_0 \int_0^\infty \psi_{\text{birth}}(s) \mathrm{e}^{-s\Lambda_D} ds = \frac{N\Psi_0}{2} dt.$$

It follows that $\Psi_0 = 2\Lambda_D$, which generalizes the conclusion Powell came to in the uncorrelated setting. Note that for $\psi_{\text{birth}}(\tau) = \delta(\tau - \tau_0)$, we obtain an exponential age distribution.

The ancestral distribution

In the finite culture $\psi_{\text{ans}}(\mathbf{x})$ is the distribution over all cells that have descendents in the current culture. Let $\tilde{f}(\mathbf{x}'|\mathbf{x}) dx_1 \cdots dx_l$ be the probability that a cell's mother has phenotype in the hypercube $\prod_{i=1}^l [x_i', x_i' + dx_i)$ given that its phenotype is $\mathbf{x}$. Bayes' theorem tells us that

$$\tilde{f}(\mathbf{x}'|\mathbf{x}) = f(\mathbf{x}|\mathbf{x}') \frac{\psi_{\text{car}}(\mathbf{x}')}{\psi_{\text{tree}}(\mathbf{x})}. \tag{30}$$

It follows that

$$\psi_{\text{ans}}(\mathbf{x}') = \int_0^\infty \cdots \int_0^\infty \tilde{f}(\mathbf{x}'|\mathbf{x}) \psi_{\text{ans}}(\mathbf{x}') dx_1 \cdots dx_l = \psi_{\text{car}}(\mathbf{x}') \int_0^\infty f(\mathbf{x}|\mathbf{x}') \frac{\psi_{\text{ans}}(\mathbf{x})}{\psi_{\text{tree}}(\mathbf{x})} dx_1 \cdots dx_l. \tag{31}$$

Division distribution

In both the continuous culture and exponentially growing culture, the carrier (or branch) distribution can be interpreted as the distribution of cells that have just divided, or the *division distribution* $\psi_{\text{div}}(\mathbf{x})$. In terms of the joint density of phenotypes and ages throughout the population, $\psi(\mathbf{x}, u)$, the number of cells with phenotypic trait in $\prod_{i=1}^l [x_i, x_i + dx_i)$ that have just divided is

$$N\psi(\mathbf{x}, \tau(\mathbf{x})) dx = N\psi(\mathbf{x}, 0) \mathrm{e}^{-\Lambda_D \tau(\mathbf{x})} dx_1 \cdots dx_l = N 2\psi_{\text{tree}}(\mathbf{x}) \mathrm{e}^{-\Lambda_D \tau(x)} dx_1 \cdots dx_l. \tag{32}$$

Dividing by $N dx_1 \cdots dx_l$ gives the distribution of phenotypes among recently divided cells, which is $\psi_{\text{div}}(\mathbf{x}) = 2\psi_{\text{tree}}(x) \mathrm{e}^{-\Lambda_D \tau(\mathbf{x})}$. As this argument works for both a finite and exponentially growing culture, the carrier distribution, branch distribution and division distributions are identical.

Summary of distributions in the finite culture

To summarize, we have discussed the following ways to sample cells in the finite culture:

- *Tree distribution* ($\psi_{\text{tree}}^F(\mathbf{x})$): Sample all the cells throughout the history of the culture, including or excluding the current cells, but including those that are expelled before they divide.

- *Carrier distribution* ($\psi^F_{\text{car}}(\mathbf{x})$): Sample all the cells throughout the culture, excluding those that are expelled before they divide.
- *Ancestral distribution* ($\psi^F_{\text{ans}}(\mathbf{x})$): Sample all the cells in the history of the culture that have ancestors that are currently in the culture. Equivalently, select a random cell from the culture at any time and trace along a lineage backwards from this cell.
- *Current cells distribution* ($\psi^F_{\text{snap}}(\mathbf{x})$): Sample all the cells that are currently in the culture.
- *Birth distribution* ($\psi^F_{\text{birth}}(\mathbf{x})$): From all the cells that are in the culture at time t, sample only those that have just been born (meaning they have age ≈ 0), then repeat this for all t and average the results.
- *Division distribution* ($\psi^F_{\text{div}}(\mathbf{x})$): From all the cells that are in the culture at time t, sample only those whose age is approximately equal to their generation time, then repeat this for all t and average the results.

C Derivation of logistic growth equation

Here we provide a more systematic derivation of the logistic growth equation. Let n_i denote the number of cells of species i in a culture containing M species. Let $\phi_i = n_i/N$ be the fraction of cells of species i and $\psi_i(t, \mathbf{x}, u)dudx_1 \cdots dx_l$ be the fraction of cells of species i with ages between u and du and phenotype x_k between x_k and $x_k + dx_k$. When a cell of species i reaches age $\tau_i(t, \mathbf{x})$ it divides, hence over an interval dt, the number of cells in species i that divide is the product of the total number of cells in species i, or $\phi_i N$, and the fraction of the these cells that have reached age $u = \tau_i(t, \mathbf{x})$. It follows that the total number of cells dividing in an interval dt is

$$\sum_{i=1}^{M} \underbrace{\phi_i N}_{\substack{\text{number}\\ \text{of cells}\\ \text{in species } i}} \underbrace{\int_0^\infty \cdots \int_0^\infty \psi_i(t, \mathbf{x}, \tau_i(t, x))dx_1 \cdots dx_l dt}_{\text{fraction of species } i \text{ cells that are dividing}}. \tag{33}$$

The per capita dilution rate, $\Lambda(t)$, is obtained by dividing this expression by N. That is,

$$\Lambda(t) = \sum_{i=1}^{M} \phi_i \Lambda_i(t), \tag{34}$$

where $\Lambda_i(t)$ is the contribution of each species to the total dilution rate:

$$\Lambda_i(t) = \int_0^\infty \cdots \int_0^\infty \psi_i(t, \mathbf{x}, \tau_i(t, \mathbf{x}))dx_1 \cdots dx_l. \tag{35}$$

At a given time, the number of bacteria of species i in the hypercube $\prod_{i=1}^{l}[x'_i, x'_i + dx_i)$ is

$$\phi_i(t)N\psi_i(t, \mathbf{x}, u)dx_1 \cdots dx_l du. \tag{36}$$

Just as in the single species case, the change in the number of individuals of a given phenotype and age is caused by aging of cells and expulsion from the population. Therefore, over an interval dt, the change in the number of microbes of species i with age $u < \tau_i(\mathbf{x})$ is

$$\begin{aligned} &\underbrace{\phi_i(t + dt)N\psi_i(t + dt, \mathbf{x}, u)dx_1 \cdots dx_l dt - \phi(t)N\psi_i(t, \mathbf{x}, u)dx_1 \cdots dx_l dt}_{\text{change in number of cells with age } u \text{ and phenotype } \mathbf{x} \text{ in species } i} \\ &\approx \underbrace{\phi_i(t)N\psi_i(t, \mathbf{x}, u - dt)dx_1 \cdots dx_l dt}_{\text{cells gained from aging}} - \underbrace{\phi_i(t)N\psi_i(t, \mathbf{x}, u)dx_1 \cdots dx_l dt}_{\text{cells lost from aging}} \\ &- \underbrace{\phi_i(t)N\Lambda(t)dx_1 \cdots dx_l dt}_{\text{cells expelled from the culture}}. \end{aligned} \tag{37}$$

For each i, this equation is almost the same as equation (26); however, we have the additional factors of ϕ_i which account for the fact that only a fraction of the cells are in species i. Taking the limit $dt \to 0$ gives

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial u}\right)\psi_i(t,\tau,u) = -\left(\Lambda + \frac{1}{\phi_i}\frac{d}{dt}\phi_i(t)\right)\psi_i(t,\tau,u). \tag{38}$$

Compared to equation (9), we have picked up an additional term depending on ϕ_i, so that we need to consider the evolution of ϕ_i in order to completely specify the dynamics. The instantaneous growth of ϕ_i is the difference between the rate at which cells of species i are born ($\Lambda_i(t)$) and flushed out ($\Lambda(t)$). Hence,

$$\frac{d}{dt}\phi_i = (\Lambda_i(t) - \Lambda(t))\phi_i. \tag{39}$$

Note that with two species, equation (39) becomes the standard logistic growth equation:

$$\begin{aligned}\frac{d}{dt}\phi_1 &= (\Lambda_1(t) - (\phi_1\Lambda_1(t) + (1-\phi_1)\Lambda_2(t)))\phi_1 \\ &= \phi_1(1-\phi_1)(\Lambda_1(t) - \Lambda_2(t))\end{aligned} \tag{40}$$

By combing equation (39) with equation (38), we obtain equation (23). Along with an initial distribution $\psi_i(0,\mathbf{x},u)$, equations (23) and (25) describe how the distribution of phenotypes and ages evolves in the mixed culture. Importantly, these equations are decoupled from $\phi_i(t)$ and each other. Moreover, they turn out to be identical to the transport equations for a homogenous culture. This implies that the dynamics of the phenotype distribution are unchanged by the competition, and the fitnesses Λ_i are therefore determined by the population growth rates of the individual species in exponentially growing cultures.

References

[1] Ho PY, Lin J, Amir A. 2018 Modeling Cell Size Regulation: From Single-Cell-Level Statistics to Molecular Mechanisms and Population-Level Effects. *Annual Review of Biophysics* **47**, 251–271.

[2] Eldar A, Elowitz MB. 2010 Functional roles for noise in genetic circuits. *Nature* **467**, 167.

[3] Wang Z, Zhang J. 2011 Impact of gene expression noise on organismal fitness and the efficacy of natural selection. *PNAS; Proceedings of the National Academy of Sciences* **108**, E67–E76.

[4] Duveau F, Hodgins-Davis A, Metzger BP, Yang B, Tryban S, Walker EA, Lybrook T, Wittkopp PJ. 2018 Fitness effects of altering gene expression noise in Saccharomyces cerevisiae. *eLife* **7**.

[5] Elowitz MB, Levine AJ, Siggia ED, Swain PS. 2002 Stochastic Gene Expression in a Single Cell. *Science* **297**.

[6] Tsimring LS. 2014 Noise in biology. *Reports on Progress in Physics* **77**, 026601.

[7] Thattai M, Van Oudenaarden A. 2004 Stochastic gene expression in fluctuating environments. *Genetics* **167**, 523–530.

[8] Roux J, Hafner M, Bandara S, Sims JJ, Hudson H, Chai D, Sorger PK. 2015 Fractional killing arises from cell-to-cell variability in overcoming a caspase activity threshold. *Molecular Systems Biology* **11**, 803–803.

[9] Hilfinger A, Paulsson J. 2011 Separating intrinsic from extrinsic fluctuations in dynamic biological systems. *Proceedings of the National Academy of Sciences* **108**, 12167–12172.

[10] Paulsson J, Berg OG, Ehrenberg M. 2000 Stochastic focusing: Fluctuation-enhanced sensitivity of intracellular regulation. *Proceedings of the National Academy of Sciences* **97**, 7148–7153.

[11] Claudi B, Spröte P, Chirkova A, Personnic N, Zankl J, Schürmann N, Schmidt A, Bumann D. 2014 Phenotypic variation of Salmonella in host tissues delays eradication by antimicrobial chemotherapy. *Cell* **158**, 722–733.

[12] Manina G, Dhar N, McKinney JD. 2015 Stress and host immunity amplify Mycobacterium tuberculosis phenotypic heterogeneity and induce nongrowing metabolically active forms. *Cell host & microbe* **17**, 32–46.

[13] Sughiyama Y, Kobayashi TJ. 2017 Steady-state thermodynamics for population growth in fluctuating environments. *Physical Review E* **95**.

[14] Leibler S, Kussell E. 2010 Individual histories and selection in heterogeneous populations. *Proceedings of the National Academy of Sciences* **107**, 13183–13188.

[15] Vilar JMG, Rubi JM. 2018 Determinants of population responses to environmental fluctuations. *Scientific Reports* **8**.

[16] Bacaër N. 2011 A Short History of Mathematical Population Dynamics. .

[17] Powell EO. 1956 Growth rate and generation time of bacteria, with special reference to continuous culture. *Microbiology* **15**, 492–511.

[18] Amir A. 2014 Cell Size Regulation in Bacteria. *Physical Review Letters* **112**.

[19] Lebowitz JL, Rubinow SI. 1974 A theory for the age and generation time distribution of a microbial population. *Journal of Mathematical Biology* **1**, 17–36.

[20] Nowak MA. 2006 *Evolutionary dynamics*. Harvard University Press.

[21] Aldridge BB, Fernandez-Suarez M, Heller D, Ambravaneswaran V, Irimia D, Toner M, Fortune SM. 2012 Asymmetry and Aging of Mycobacterial Cells Lead to Variable Growth and Antibiotic Susceptibility. *Science* **335**, 100–104.

[22] Chao L, Rang CU, Proenca AM, Chao JU. 2016 Asymmetrical Damage Partitioning in Bacteria: A Model for the Evolution of Stochasticity, Determinism, and Genetic Assimilation. *PLOS Computational Biology* **12**, e1004700.

[23] Min J, Lin J, Amir A. 2019 Optimal Segregation of Proteins: Phase Transitions and Symmetry Breaking. *Physical Review Letters* **122**.

[24] Jafarpour F. 2018 Bridging the Timescales of Single-Cell and Population Dynamics. *Physical Review X* **8**.

[25] Mosheiff N, Martins BMC, Pearl-Mizrahi S, Grünberger A, Helfrich S, Mihalcescu I, Kohlheyer D, Locke JCW, Glass L, Balaban NQ. 2018 Inheritance of Cell-Cycle Duration in the Presence of Periodic Forcing. *Physical Review X* **8**.

[26] Moran PAP. 1958 Random processes in genetics. In *Mathematical Proceedings of the Cambridge Philosophical Society* vol. 54 pp. 60–71. Cambridge University Press.

[27] Lin J, Amir A. 2017 The Effects of Stochasticity at the Single-Cell Level and Cell Size Control on the Population Growth. *Cell Systems* **5**, 358 – 367.e4.

[28] Hashimoto M, Nozoe T, Nakaoka H, Okura R, Akiyoshi S, Kaneko K, Kussell E, Wakamoto Y. 2016 Noise-driven growth rate gain in clonal cellular populations. *PNAS; Proceedings of the National Academy of Sciences* **113**, 3251–3256.

[29] Thomas P. 2017a Single-cell histories in growing populations: relating physiological variability to population growth. *bioRxiv*.

[30] Thomas P. 2017b Making sense of snapshot data: ergodic principle for clonal cell populations. *Journal of The Royal Society Interface* **14**, 20170467.

[31] Jafarpour F. 2019 Cell Size Regulation Induces Sustained Oscillations in the Population Growth Rate. *Physical Review Letters* **122**.

[32] Lambert G, Kussell E. 2015 Quantifying Selective Pressures Driving Bacterial Evolution Using Lineage Analysis. *Physical Review X* **5**.

[33] Wakamoto Y, Grosberg AY, Kussell E. 2012 Optimal lineage principle for age-structured populations. *Evolution* **66**, 115–134.