bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Universality of evolutionary trajectories under arbitrary competition dynamics

Andrea Mazzolini[1] and Jacopo Grilli[2]

[1] *Laboratoire de physique de l'École normale supérieure (PSL University), CNRS, Sorbonne Université, and Université de Paris, 75005 Paris, France*

[2] *Quantitative Life Sciences, The Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste 34151, Italy.*

(Dated: May 6, 2022)

The assumption of constant population size is central in population genetics. It led to a large body of results, that are robust to modeling choices and that have proven successful to understand evolutionary dynamics. In reality, allele frequencies and population size are both determined by the interaction between a population and the environment. Relaxing the constant-population assumption have two big drawbacks. It increases the technical difficulty of the analysis, and it requires specifying a mechanism for the saturation of the population size, possibly making the results contingent on model details. Here, we develop a framework that encompasses a great variety of systems with an arbitrary mechanism for population growth limitation. By using techniques based on scale separation for stochastic processes, we are able to calculate analytically properties of evolutionary trajectories, such as the fixation probability. Remarkably, these properties assume a universal form with respect to our framework, which depends on only three parameters related to the inter-generation timescale, the invasion fitness, and the carrying capacity of the strains. In other words, different systems, such as Lotka-Volterra or a chemostat model (contained in our framework), share the same evolutionary outcomes after a proper re-mapping of their parameters. An important and surprising consequence of our results is that the direction of selection can be inverted, with a population evolving to reach lower values of invasion fitness.

I. INTRODUCTION

Competition is a fundamental agent of natural selection. The scarcity of resources (nutrient, water, space, ...) limits population growth, determining a selection pressure. Variants — strains in the following — that grow faster or more efficiently spread within the population. The long history of mathematical population genetics has produced several seminal results that quantitatively describe these processes. The vast majority of results are based on the assumption of constant population size [1–8]. The success of this assumption lies in the generality of the results. When the population size is large enough, the diffusion limit developed independently by Wright and Kolmogorov, pioneered in population genetics by Malécot[63] and Kimura [1, 9], is the convergence point of several alternative models. In fact, while different population genetics models (Wrigh-Fisher [10, 11] , Moran [12], conditional branching processes [13], and some Canning processes [14]) start from radically different assumptions about the genealogical and demographic structure of the population, they share the same predictions, up to a simple rescaling of timescales and parameters. For most of the theoretical advances in population genetics, the total population size is treated as an effective parameter of the model, which should be fitted from data, and the strain frequencies are the only dynamical degrees of freedom.

However, mechanistically, not only the strain frequencies but also the total population size is determined by the interaction between a population and the environment. Moreover, experimental works [15, 16] suggest that the

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

variation of the population size can play a role in the evolutionary process, implying that its dynamics should not be neglected in theoretical descriptions.

Non-stationary conditions are an important example where the dynamics of population size cannot be neglected to understand how evolution — here intended as the dynamics of relative frequencies of strains — unfolds. For instance, range expansions leave strong signature on the genetic diversity of a population [17–19].

A classical deterministic approach to couple the population growth and the frequencies of the strains is density-dependent selection [20–23]. Another approach to study evolution under varying total population sizes is to consider the dynamics of the latter as decoupled from the strain frequency, by assuming a priori the dynamics of the total population [24–27]. Here we instead focus on stochastic scenarios. This approaches require to model how environmental constraints limit population growth, for instance by considering a Lotka-Volterra dynamics [28–34] or generalizations of the Moran model [35]. In these kinds of models, the population size is a stochastic variable itself that can fluctuate. Those fluctuations are not independent of strain frequencies, leading to highly non-trivial evolutionary dynamics.

The increased realism of these models undermines, at least in principle, the generality of the standard population genetics results. To what extent do the details of the competition structure determine the evolutionary outcome? For instance, will a population whose growth is described by a logistic model differ in its evolutionary trajectory from a population described by Gompertz growth?

In this work, we address this question by introducing and considering a general ecological framework to describe the dynamics of a haploid asexual population. This approach encompasses several alternative models of competition — from Lotka-Volterra to consumer-resource models in a chemostat, to a Gompertzian growth — under the same mathematical set-up. In other words, this ecological framework defines the mathematical details of how the growth of a population is limited by competition for limited resources.

Given the generality of our framework, one could in principle expect that the evolutionary dynamics depend on the specific details of the model considered to describe population self-limitation. We show that evolutionary predictions are instead robust and insensitive to the details of population dynamics. Within our framework, evolutionary observables, such as the fixation probability and the fixation time, are universal and depend on only three parameters, related to the concepts of inter-generation timescale, invasion fitness, and carrying capacity.

We first present the general framework in the model section. We start to illustrate the deterministic limit, where basic ideas of the framework can be easily grasped. We then derive the stochastic dynamics, described as a system of Langevin equations. In the limit of small fitness differences, a time-scale separation allows us to calculate analytically the properties of the evolutionary dynamics. We describe the behavior of these properties over evolutionary timescales. We also illustrate an explicit example of the chemostat model for resource competition, showing how the previous general results can lead to an evolutionary trajectory with decreasing fitness over time.

II. MODEL

To help the reader to build its intuition on the presented framework, we first introduce our framework in the case of large population sizes (which allow neglecting stochasticity), in presence of a single clonal population. We then consider the case of finite population size and introduce the stochastic version for a clonal population. We then generalize our framework to multiple strains in the deterministic limit and, finally, multiple strains in the stochastic formulation,

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

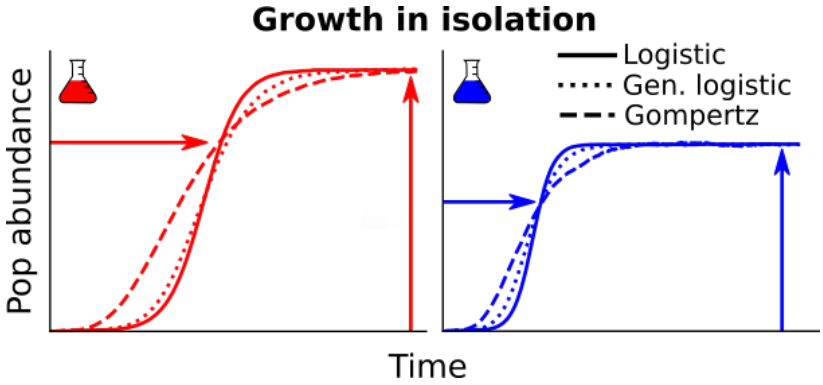

FIG. 1: Typical growth curve of a clonal population in isolation, characterized by a carrying capacity K_i and a timescale to reach the saturation T_i.

which is the framework considered for the rest of the paper.

A. A general model for the growth of a clonal population in presence of limiting factors

We consider a clonal population whose growth is limited by some resources. The growth curve of a clonal population displays a typical sigmoid shape, as depicted in Fig. 1: a rapid initial growth followed by a deceleration due to nutrient limitation, and, eventually, a convergence of the abundance to a carrying capacity, determined by the availability and quality of nutrients. Multiple models — such as Logistic, Gompertz, and Von Bertalanffy — capture this phenomenology, by explicitly quantifying the dependence of per-capita growth rates on population abundance. These models differ in the specific functional form of the population growth trajectory. For instance, in the Logistic model, the per-capita growth rate decreases linearly with population abundance, while in the Gompertz model it decreases logarithmically.

This large class of models can be captured, in full generality, by making the per-capita growth rate depend on the total population abundance N through generic functions:

$$\frac{1}{N}\frac{dN}{dt} = \frac{1}{T}\beta\left(u\frac{N}{K}\right)\left(1-\omega\left(\frac{N}{K}u\right)\right) \ . \tag{1}$$

The function $\beta(\cdot)$ and function $\omega(\cdot)$ determine the specific shape of growth curves, which distinguish between alternative models (see below). The value u simply equals to $\omega^{-1}(1)$, so that $\omega(uN/K) = 1$ for $N = K$. The two parameters T and K capture instead the dimensional components of these growth curves. The *carrying capacity* K defines the values of population abundance reached at large times. A *timescale* T sets the speed of convergence of population curves to carrying capacity. Fig. 1 shows the example of two populations with different strains, each characterized by different values of K and T.

Eq. 1 reduces to standard models under specific choices of the two functions $\beta(\cdot)$ and $\omega(\cdot)$. Logistic growth is obtained using $\beta(z) = 1$ and $\omega(z) = z$, the chemostat model [36] with $\beta(z) = 1/z$ and $\omega(z) = z$, Gompertz growth for $\beta(z) = 1$ and $\omega(z) = \log(z)$, and Von Bertalanffy model for $\beta(z) = z^{\alpha-1}$ and $\omega(z) = z^{1-\alpha}$ (see Appendix A).

To enforce the saturating phenomenology shown in Fig. 1, few constraints are required for the two generic functions. Specifically, that both are positive and that $\omega(\cdot)$ is monotonically increasing. This choice guarantees the existence of a

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

unique, globally stable, fixed point $N^* = K$, to which the population abundance converges for large times. These choices can be relaxed in presence of more complex biological and ecological mechanisms affecting population growth, that we will not consider in this manuscript. For instance, in presence of a strong Allee effect [37], the per-capita growth rate would be negative at small population abundance, corresponding to a non-monotonicity of $\omega(\cdot)$.

B. A general birth-death model for the growth of a clonal population

Eq. 1 can be seen as the continuous deterministic limit, obtained for large population sizes, of a discrete stochastic process that describes the dynamics of abundance of a finite number of individuals. It is natural to define such a model as a birth-death process with microscopic rates

$$\begin{aligned} N \to N+1 \ \text{ w.r. } \ b(N) &= \frac{N}{T}\beta\left(u\frac{N}{K}\right) \\ N \to N-1 \ \text{ w.r. } \ d(N) &= \frac{N}{T}\beta\left(u\frac{N}{K}\right)\omega\left(u\frac{N}{K}\right). \end{aligned} \tag{2}$$

This formulation further clarifies the interpretation of the parameters and the functions appearing in eq. 1. The value of T sets the unique timescale of the process, to which both the birth and death rate are inversely proportional, and is directly related to the generation time (see appendix C for more details). The function $\beta(\cdot)$ describes the density dependence of the birth rate, while the function $\omega(\cdot)$ sets the dependence of the ratio between birth and death rate.

The constraints on the two functions that have been introduced in the deterministic case have to be valid also in this context. First, $\beta(\cdot)$ and $\omega(\cdot)$ must be positive since the rates must be positive. Second, the monotonicity of $\omega(\cdot)$ enforces the population saturation: the death rate exceed the birth rate when $\omega\left(\frac{N}{K}u\right)$ becomes larger than 1, i.e., for $N > K$. The phenomenology of this class of stochastic models have been studied for specific choices of $\beta(\cdot)$ and $\omega(\cdot)$. In particular, extensive work has been done for logistic growth ($\beta(z) = 1$ and $\omega(z) = z$) [38].

The stochastic dynamics of N can be divided into three temporal phases. The first phase is dominated by a transient, which depends on the initial condition and directly corresponds to the transient of the deterministic dynamics depicted in Fig. 1. In the case of logistic growth, this first phase requires a time of the scale $T\log(K)$. The second phase corresponds to fluctuations around $N = K$ and displays a variability of the order of $\sqrt{K}$. Approximation for the quasi-stationary distribution of these fluctuations is well known [39, 40]. The timescale T determines the typical auto-correlation time of these fluctuations. The stability of the second phase is only apparent, as it is a meta-stable phase. The stochastic process defined in eq. 2 has an absorbing state in $N = 0$ and very rare fluctuations can drive the population to extinction, corresponding to the third phase. The rate of extinction is very small for large populations. The typical time to extinction — which, according to our definitions, corresponds to the duration of the second phase — scales exponentially with K. In the case of the logistic model [38] this time is already astronomical for relatively small population sizes. For instance, for $K = 100$ the average time to extinction is on the order of 10^{32} generation times.

C. Deterministic limit in the case of multiple strains

We want now to focus on the evolutionary dynamics of different strains that compete with each other, by studying how intra-population variability changes over time. We consider therefore A strains that differ in their growth rates

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

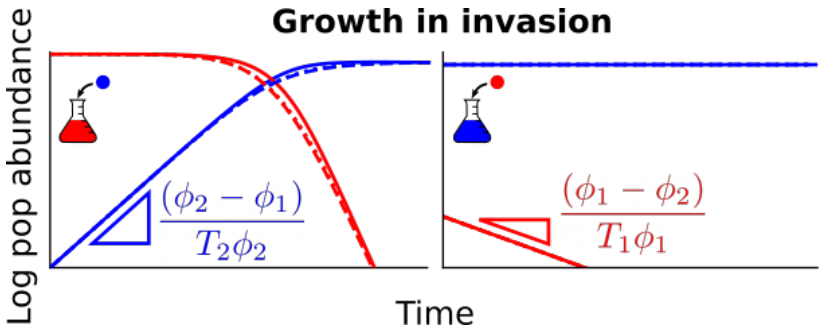

FIG. 2: Typical growth curves of two competitive strains, where one tries to invade tho other one at carrying capacity. The invasion success is determined by the invasion fitness ϕ_i. The lines are simulated with equation 3, for different choices of the generic functions as in figure 1.

through their demographic parameters. While K and T are enough to specify the scales of the growth curves obtained in isolation, they are in fact generally not enough to describe the dynamics in co-cultures. Fig. 2 depicts the setting of a pairwise competition experiment, where a single individual (or a small population) of a mutant is inoculated in a large resident population. In absence of multiple substitutable resources (or other factors giving rise to frequency-dependent selection), only one of the two strains — the one with the higher fitness — survives.

To include this feature in our model, the deterministic equation for the abundance of strain i reads

$$\frac{1}{n_i}\frac{dn_i}{dt} = \frac{1}{T_i}\beta\left(\sum_j \frac{n_j}{K_j}u_j\right)\left(1 - \frac{1}{\phi_i}\omega\left(\sum_j \frac{n_j}{K_j}u_j\right)\right) \tag{3}$$

where $u_i = \omega^{-1}(\phi_i)$. The generic functions $\beta(\cdot)$ and $\omega(\cdot)$ are independent of the strain and obey the same constraints as before. As described above, the values of T_i and K_i define the growth curve of that particular strain in a clonal population as shown in Fig. 1. The new crucial ingredient here is the parameter ϕ_i, which is related to the (invasion) fitness of strain i.

As shown in Appendix B, this system of non-linear differential equations admits a unique, globally stable, fixed point, where only one strain survives reaching a population abundance equal to its carrying capacity. The strain that survives is the one characterized by the largest value of ϕ, which we, therefore, identify as the fitness of a strain.

Consistently with the existence of a globally stable equilibrium, the outcomes of invasion experiments depend only on the fitness ϕ of the strains involved. In particular, the growth rate of the mutant strain (usually called "invasion fitness") is proportional to the difference between its fitness ϕ, and the one of the resident population (see Figure 2 and Appendix C). It is important to remark that, in general, the invasion fitness does not correspond to the Malthusian fitness [41] (i.e. the per-capita growth rate in exponential growth) as the latter refers only to growth in isolation. For instance, the Malthusian fitness in the logistic model is equal to $1/T_i$. Therefore the strain that is expected to out-compete the others (the one with the highest value of ϕ) is not necessarily the one with the higher growth rate in the exponential phase ($1/T$ for the logistic model).

A particularly simple and paradigmatic case occurs when $T_i = 1/(\tau\phi_i)$ [64] and $K_i = K$ for all i. In that case, in fact, the dynamics of the total population abundance has a globally stable fixed point $N^* = K$ independently of the relative abundances of individual strains. This allows to write explicitly the dynamics for the relative population

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

abundances $x_i = n_i/N$, which, by imposing $N = K$, turns out to correspond to the replicator equation

$$\frac{dx_i}{dt} = \beta\left(\bar{u}\right)\left(\phi_i - \bar{\phi}\right) \ , \tag{4}$$

where $\bar{\cdot}$ denotes the average over the population (e.g. $\bar{\phi} = \sum_i \phi_i x_i$).

D. Birth-death model for multiple strains

Finally, we generalize therefore eq. 2 to describe the coupled birth-death process of multiple strains, whose deterministic dynamics is eq. 3. The birth and death rates $b_i(\mathbf{n})$ and $d_i(\mathbf{n})$ depend on the abundance of all the other strains, $\mathbf{n} = n_1, \ldots, n_A$, as all the individual compete — even if with different ability — for the same limited resources. The rates read

$$\begin{aligned} b_i(\mathbf{n}) &= \frac{n_i}{T_i}\beta\left(\sum_j \frac{n_j}{K_j}u_j\right) \\ d_i(\mathbf{n}) &= \frac{1}{\phi_i}\frac{n_i}{T_i}\beta\left(\sum_j \frac{n_j}{K_j}u_j\right)\omega\left(\sum_j \frac{n_j}{K_j}u_j\right) . \end{aligned} \tag{5}$$

If $\phi_i = 1$, $T_i = T$, and $K_i = K$ for all i, the population abundance $N = \sum_j n_j$ is described by the rates of eq. 2.

It is important to notice that in this model, contrarily to standard models in population genetics that assume a fixed population size N, both the relative abundances of each strain n_i/N and the total population size $N = \sum_i n_i$ are changing stochastically over time in an interdependent manner. Even at stationarity, the population size N is not fixed but oscillates around a typical value, determined by the strains' parameters.

The presented framework includes all the works based on a competitive Lotka-Volterra like evolutionary dynamics [28–34] (for $\beta(z) = 1$ and $\omega(z) = z$), and generalization of the Moran model [35] (for $\beta(z) = 1$ and $\omega(z) = \alpha(1 + z)$).

III. STOCHASTIC DYNAMICS

While we want to include stochasticity in our description, taking into account the effect of finite population sizes, we are interested in the limit of large population sizes, which in our context corresponds to $K_i \gg 1$. In this limit classic models of population genetics (such as the Moran model or Wright-Fisher model) converge to the same effective diffusive limit, which corresponds to Kimura's equation.

The birth-death process defined in eq. 5 can be approximated by Itô stochastic differential equations as follows

$$\frac{dn_i}{dt} = b_i(\mathbf{n}) - d_i(\mathbf{n}) + \sqrt{b_i(\mathbf{n}) + d_i(\mathbf{n})}\xi_i \ , \tag{6}$$

where ξ_i is a Gaussian white noise, $\langle \xi_i(t) \rangle = 0$, $\langle \xi_i(t)\xi_j(t') \rangle = \delta_{ij}\delta(t - t')$ (see Appendix D for the derivation).

A. Separation of ecological and evolutionary timescales

Eq. 6 describes the stochastic dynamics in the limit of large population sizes ($K_i \gg 1$). In addition to this limit, we consider the limit of small fitness differences ($|\phi_i - \phi_j|/\phi_i \ll 1$). In this case, as depicted in Figure 3, the dynamics

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

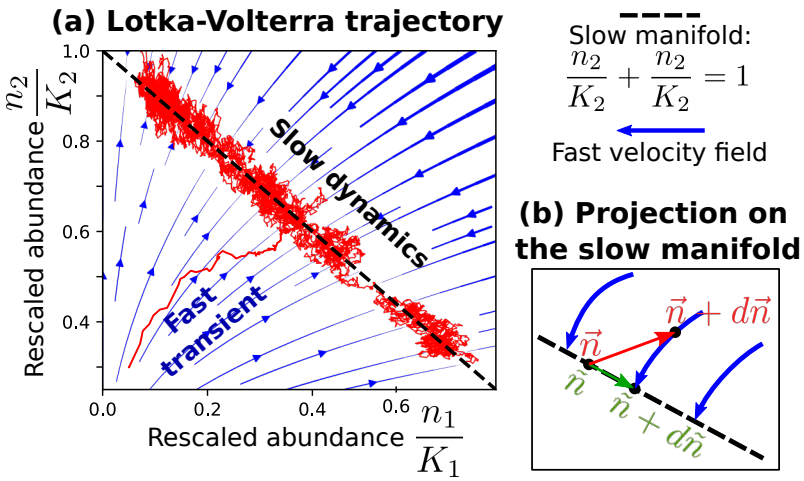

FIG. 3: The panel (a) shows an illustrative example of the time-scale separation in a Lotka-Volterra for two species: $\beta(z) = 1$, $\omega(z) = z$, $K_1 = K_2 = 2000$, $t_1 = T_2 = 1$, $\phi_1 = 1$, $\phi_2 = 1.002$. The variables n_i/K_i are the abundances re-scaled by the carrying capacity. The panel (b) sketches the idea underlying the projection of the dynamics onto the slow manifold: infinitesimal steps outside the manifold (in red) are mapped again on it by the fast velocity field. As a result, to correctly take into account these effects, the increment of the effective variable $\tilde{n}$ (in green) corresponds to the projection of $\vec{dn}$ along the velocity field.

separate in a fast transient followed by slow dynamics (see Appendix E). The initial trajectory drives the system to a slow manifold of solutions, corresponding to the equilibria of the deterministic dynamics in absence of fitness differences. What follows is a dynamic constrained on the slow manifold.

It is convenient to describe the dynamics on the slow manifold in terms of the relative abundances $x_i = n_i/\sum_j n_j$. The replicator equation 4 describes the deterministic dynamics of the relative abundances on the slow manifold only under the specific parameter choices discussed in sec. II C. Its derivation does not trivially generalize to the stochastic case and in presence of fully general parameter combinations. The "slow" dynamics is in fact determined by the combination of two forces (see Figure 3). Stochasticity (through genetic drift) pushes the system away from the manifold of solutions. The deterministic part of the dynamics is pushing the system back to the manifold. These two forces do not act orthogonally to the manifold, but the non-linearity of population dynamics and the multiplicative nature of demographic stochasticity result in a non-trivial combination with a net force, which corresponds to an effective frequency-dependent selection that drives the evolution of the system.

In the following we consider two strains, where $x \equiv x_1$ as the relative abundance of strain 1 and $x_2 = 1 - x$. The effective dynamics on the slow manifold reduces to

$$\frac{dx}{dt} = \gamma_{r,c}(x)\frac{x(1-x)}{T_1}\left(s - \frac{\sigma_{r,c}(x)}{K_2}\right) + \sqrt{\frac{x(1-x)}{K_2 T_1}\delta_{r,c}(x)} \cdot \eta(t), \tag{7}$$

where $\eta(t)$ is a white noise term (see Appendix E). In addition to K_2 and T_1, the dynamics of the relative abundance x depends on three parameters, s, r, c, through the three functions $\gamma_{r,c}(x)$, $\sigma_{r,c}(x)$, $\delta_{r,c}(x)$ specified below. These three quantities are related to the ratios between the parameters characterizing the population dynamics of the two strains. In particular, s is the selection coefficient $s = (\phi_1 - \phi_2)/\phi_1$ while $r = T_2/T_1$, and $c = K_2/K_1$. The three functions

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

read

$$\begin{aligned}\gamma_{r,c}(x) :=& \frac{1-x+xc}{1-x+xrc}\\ \sigma_{r,c}(x) :=& \frac{(1-r)(1-x+xc)}{(1-x+crx)^2}\Big(1-2c-\\ &-(3+c(cr-4))x+2(1-c)(1-cr)x^2\Big)\\ \delta_{r,c}(x) :=& \frac{(1-x+cx)^3\,(1-x+rx)}{(1-x+crx)^2}\ .\end{aligned} \tag{8}$$

The trajectories of the relative abundances of two strains growing together are therefore determined also by differences in inter-generation times (case $r \neq 1$) and in carrying capacities (case $c \neq 1$). If the inter-generation times T and the carrying capacities K are equal (i.e., if $r = c = 1$), eq. 7 reduces to Kimura's diffusion limit. In fact, $\gamma_{1,1}(x) = \delta_{1,1}(x) = 1$ and $\sigma_{1,1}(x) = 0$.

The sign of the deterministic term of equation 7 depends on the balance between two forces. The first term is the selection coefficient s. The second one (equal to $\sigma_{r,c}/K_2$) is absent in the classic Kimura's diffusion limit. It depends non-trivially on the time scales, the carrying capacities, and strain frequencies.

To better understand this dependence, it is instructive to consider the growth of a rare strain with relative abundance x. Since the strain is rare the initial growth (or decline) will be approximately described by a stochastic exponential growth, which can be obtained by expanding eq. 7 for $x \ll 1$. In particular, the average relative abundance $\langle x \rangle$ is determined by

$$\frac{d\langle x\rangle}{dt} = \frac{\langle x\rangle}{T_1}\left(s - \frac{(2c-1)(r-1)}{K_2}\right)\ . \tag{9}$$

If the carrying capacity of the resident population is very large, $K_2 \to \infty$, the growth rate converges to the deterministic limit s/T_1, depicted in Fig. 2. More precisely, the deterministic limit is correct whenever $sK_2 \gg (1-2c)(1-r)$.

On the other hand, the presence of large but finite population sizes and large enough trait variation drastically change the relative abundance trajectories. A particularly interesting scenario emerges when $\frac{(2c-1)(r-1)}{K_2} < s < 0$. In that case, accordingly to standard results strain 2 would a selective advantage over strain 1 ($s < 0$). However, the presence of the other extra-term counterbalances the positive value of the selection coefficient, making strain 1 able to invade. This implies that, under some choices of c and r, it can happen that an invader having smaller invasion fitness can grow on average within the "fitter" resident population. To better quantify this statement, in the next section we compare the fixation probabilities of the two strains.

IV. CONSEQUENCES ON EVOLUTIONARY DYNAMICS

A. Carrying capacity, inter-generation times, and invasion fitness control the evolutionary success

The evolutionary success of a strain in an environment described by 7 can be quantified by the fixation probability: the likelihood that one individual of a given strain is able to invade a population of a second strain. The ratio between the probability of fixation of one individual of the first strain $p_{1\text{fix}}$ in a resident population of strain 2 and the fixation of one individual of the second strain $p_{2\text{fix}}$ in a resident population of strain 1 can be expressed analytically (see

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

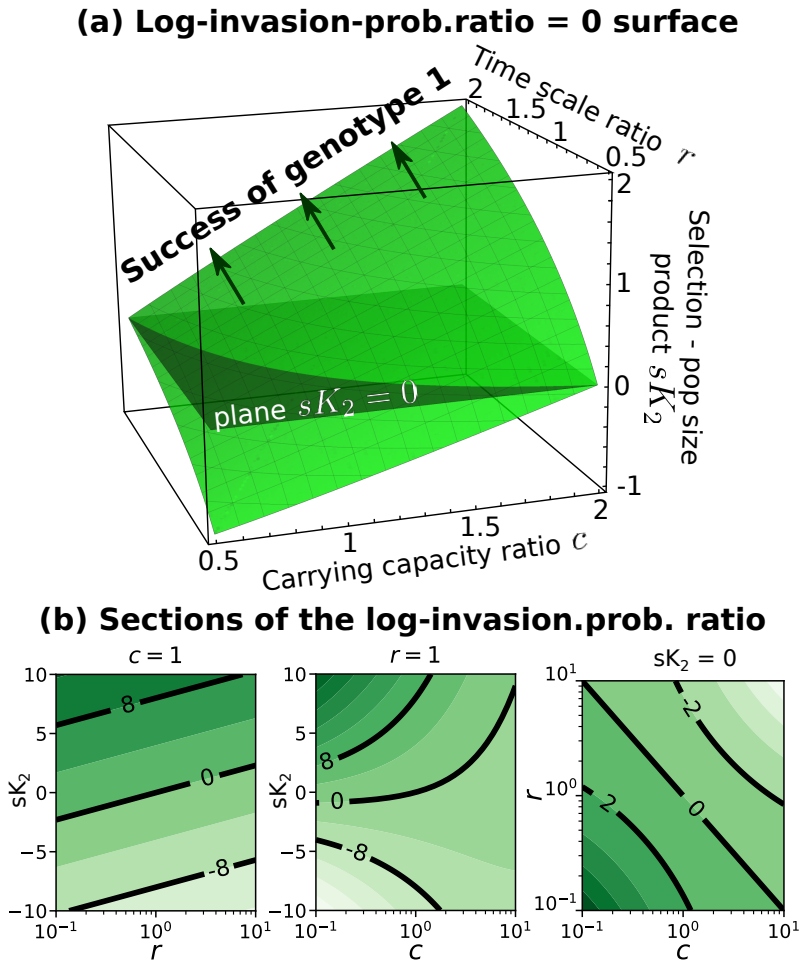

FIG. 4: How the fixation probability depends on the three parameters. Panel (a) shows the surface $\log(p_{1\text{fix}}/p_{2\text{fix}}) = 0$, panel (b) different sections of the fixation probability log-ratio. A darker shade of green represents a larger ratio.

Appendix F) and equals to

$$\log \frac{p_{1\text{fix}}}{p_{2\text{fix}}} = sK_2 \frac{r-1}{r-c} + \frac{sK_2 r(c-1) - (1-rc)(r-c)}{(r-c)^2} \log \frac{c}{r} \ . \tag{10}$$

which, in the singular case of $r = c$ becomes $\log(p_{1\text{fix}}/p_{2\text{fix}}) = \frac{r+1}{r}\left(\frac{sK_2}{2} + 1 - r\right)$.

Despite the generality of the modeling framework 5, this quantity depends on only three parameters: the inter-generation times ratio r, the ratio between the carrying capacities c, and the product between the selection coefficient and the size of the second population at carrying capacity sK_2.

Figure 4a shows the surface $\log(p_{1\text{fix}}/p_{2\text{fix}}) = 0$, which separates the regions of parameters between a more successful first strain, $p_{1\text{fix}} > p_{2\text{fix}}$, and the opposite scenario. Note that, differently from the deterministic approximation of eq. 3, invasion fitness does not unequivocally determine the most likely evolutionary outcome. It exists in fact a region of parameters within the volume below the $s = 0$ plane and above the $\log(p_{1\text{fix}}/p_{2\text{fix}}) = 0$ surface of the Figure 4 for which a the strain with negative selection coefficient is more likely to invade than the one with higher invasion fitness.

In absence of differences in the inter-generation time and carrying capacity (i.e., if $r = c = 1$), one recovers the Kimura's diffusion limit and the classical result obtained under a constant total population size K_2 [8]: $\log(p_{1\text{fix}}/p_{2\text{fix}}) =$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

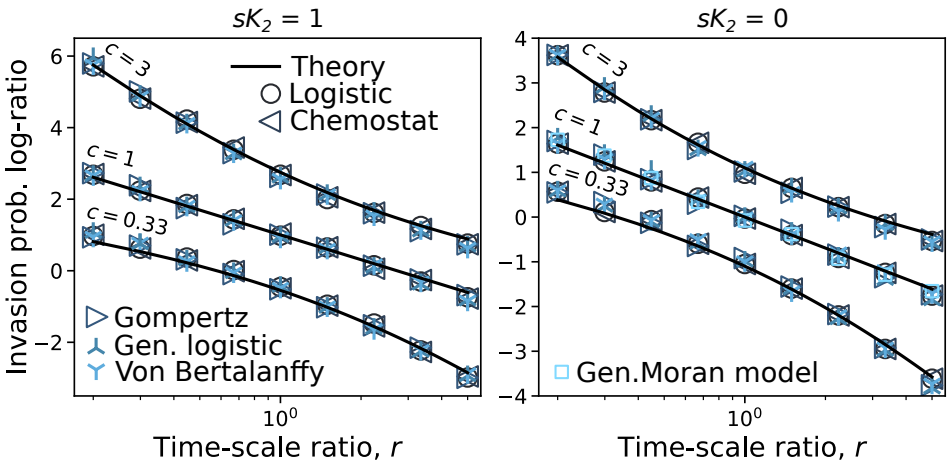

FIG. 5: Collapse of competitive models on the same fixation probability, given by equation 10, black lines. The simulations consider the competitive Lotka-Volterra, a chemostat model [42], the Gompertz growth dynamics [43, 44], a generalized Lotka-Volterra [45] with exponent $\nu = 1/2$, the Von Bertalanffy model [46] with $\alpha = 2/3$ and the Generalized Moran model [35] which is defined only for $sK_2 = 0$ and $c = 1$. For more details about the models see Appendix A

$K_2 s$.

It can also be instructive to isolate the dependency on the carrying capacities when intrinsic fitnesses and inter-generation times do not vary (i.e., if $s = 0$ and $r = 1$). In that case, $p_{1\text{fix}}/p_{2\text{fix}} = 1/c = K_1/K_2$. This implies that that larger carrying capacities, corresponding to lower demographic fluctuations and genetic drift, are favoured. Similarly, if the selection coefficient equals 0 and the carrying capacity are equal between the strains ($c = 1$), the ratio of the fixation probabilities reduces to $p_{1\text{fix}}/p_{2\text{fix}} = 1/r = T_1/T_2$. Also in this case, longer inter-generation times, which correspond to lower per-capita birth and death rates at carrying capacities, and therefore lower stochasticity, are favored.

B. The evolutionary outcome is independent of modeling details

One key result of our derivation is that the evolutionary trajectories, described in the diffusion limit by eq. 7, do not depend on the specific modeling choices. More precisely, eq. 7 does not depend on the shape of the functions $\beta(z)$ and $\omega(z)$.

Figure 5 shows numerical simulations of the fixation probability in alternative models, differing for the mathematical dependency of birth and death rates on strain abundances. The figure shows that once these different models are mapped onto our framework 5, all the invasion probabilities collapse onto the same curves as predicted by equation 10. The results of figure 4 are therefore general and do not depend on the specific model considered. For instance, for $c = 1$ and $s = 0$ one recovers the fixation probability of [35] and [33]. Details about how the simulations are performed are shown in Appendix G. The definitions of the simulated models and how to map them in our framework are in Appendix A.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

C. The asymmetry in total mutation rates induces an additional dependency on population size and inter-generation timescale

The expected number of mutations per generation depends on the population size. Since in our case the total population size can vary depending on which strains are present and in what abundance, also the total mutation rate depends on which strains are present and in which abundance. This dependency of the total mutation rate on population size induces an additional effect of the varying total population on the most likely evolutionary trajectory.

Let us consider two strains and a symmetric mutation probability $U \ll 1$. If we start with a clonal population of strain 2 at carrying capacity K_2, the birth and death rates of 5 are balanced by definition of steady state. The birth rate reads therefore $b_2 = \rho_2 \beta(\omega^{-1}(\phi))K_2$. By multiplying this rate by the mutation probability U one obtains the mutation rate, which, in this simple case, corresponds to the rate at which an individual of strain 1 appears. Therefore, one can compute the rate at which strain 1 substitutes the resident population composed of strain 2, which equals $m_{1\text{fix}} = Ub_2 p_{1\text{fix}}$. The ratio of this quantity for the two strains reads then

$$\frac{m_{1\text{fix}}}{m_{2\text{fix}}} = \frac{Ub_2 p_{1\text{fix}}}{Ub_1 p_{2\text{fix}}} = \frac{c}{r}\frac{p_{1\text{fix}}}{p_{2\text{fix}}} \ , \tag{11}$$

where the ratio $p_{1\text{fix}}/p_{2\text{fix}}$ is given by eq. 10. Including explicitely the effect of generation time and population size on mutation introduces an additional dependency on c and r.

In the periodic selection regimes, i.e., when the time between two successful mutations is much longer than the time to fixation of a successful mutation, this ratio is directly related to the relative probability of finding the population dominated by strain 1 vs strain 2.

Figure 6 shows the analytical behavior of this expression. Also in this case there is a large range of parameters for which the population evolves to lower values of invasion fitness.

D. Evolution in the chemostat model with metabolic trade-off can revert the direction of evolution

As an example for the application of the previous results, we consider a population growing in a chemostat on externally provided resources [36]. We consider a birth-death process where the birth-rate depends on the expected availability of resources, while the death rate is a constant factor. In the deterministic limit the equation reduces to

$$\frac{1}{n_i}\frac{dn_i}{dt} = \frac{\eta_i \alpha_i}{\frac{1}{M}\sum_j n_j \alpha_j} - \delta_i \ . \tag{12}$$

The parameter α_i represents the resource intake rate of strain i, δ_i is a death/dilution rate, while $M\eta_i$ is the efficiency of resource-to-biomass conversion (equivalent to an inverse yield), and $M \gg 1$ a parameter that sets the scale of the total population size. One can then map this model into our general framework, for the choice $\beta(z) = 1/z$ and $\omega(z) = z$, as described in Appendix A

A particularly interesting scenario appears when α, δ, and η are not independent, being subject to a tradeoff. For instance, yield (or, equivalently, efficiency) decreases during selection in experimental evolution [47–49], giving rise to a trade-off between growth rate and yield.

Specifically, we focus on mutations of the resource intake rate α and consider the case $\delta_i/\eta_i = \alpha_i + \alpha_0$, where $\alpha_0 \ll 1$ is a small parameter (in principle of the order of $1/M$). This choice implies that the population size at carrying

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

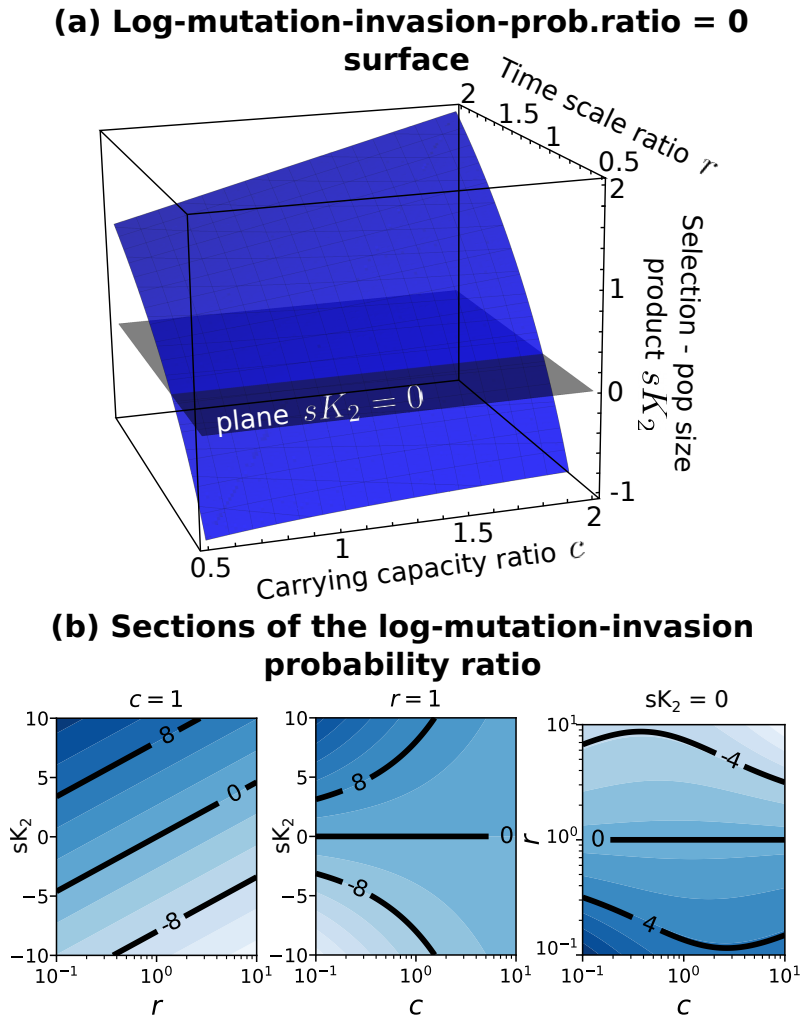

FIG. 6: Behavior of the invasion probability through mutations. The blue surface is $\log(m_{1\text{fix}}/m_{1\text{fix}}) = 0$ defined in eq. 11

capacity in a clonal population, i.e. $K_i = M\eta_i/\delta_i = M/(\alpha_i + \alpha_0)$, decreases with the intake rate α, giving rise to the growth-rate and yield trade-off.

In the deterministic limit, it is easy to see that strains with higher intake rates (i.e. with higher invasion fitness $\phi_i = \alpha_i\eta_i/\delta_i = \alpha_i/(\alpha_i + \alpha_0)$) always invade populations with lower values of α. By using using equations 10 and 11, we show that the presence of a large, but finite, strain-dependent total population size can invert this behavior. Moreover, the outcome of the evolutionary process depends on whether the dependency on α is included in the death rate or in the efficiency. We consider here the former case, while we discuss the latter, which nevertheless leads to similar results (see Appendix H).

If the intake conversion from the resource to the biomass is constant, η_0, the death rate per-capita can be expressed as $\delta_i = \eta_0(\alpha_i + \alpha_0)$. The computation of the time-scale ratio leads to $r = c$. In such a case, there is no difference between external invasions, eq. 10, or invasions through mutations, eq. 11. In particular, using eq. F9 (the special case $r = c$ of eq. 10) one obtains

$$\log \frac{p_{1\text{fix}}}{p_{2\text{fix}}} = \frac{(\alpha_1 - \alpha_2)(\alpha_1 + \alpha_2)(\alpha_0 M - 2\alpha_1\alpha_2)}{2\alpha_1^2\alpha_2^2} \ . \tag{13}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

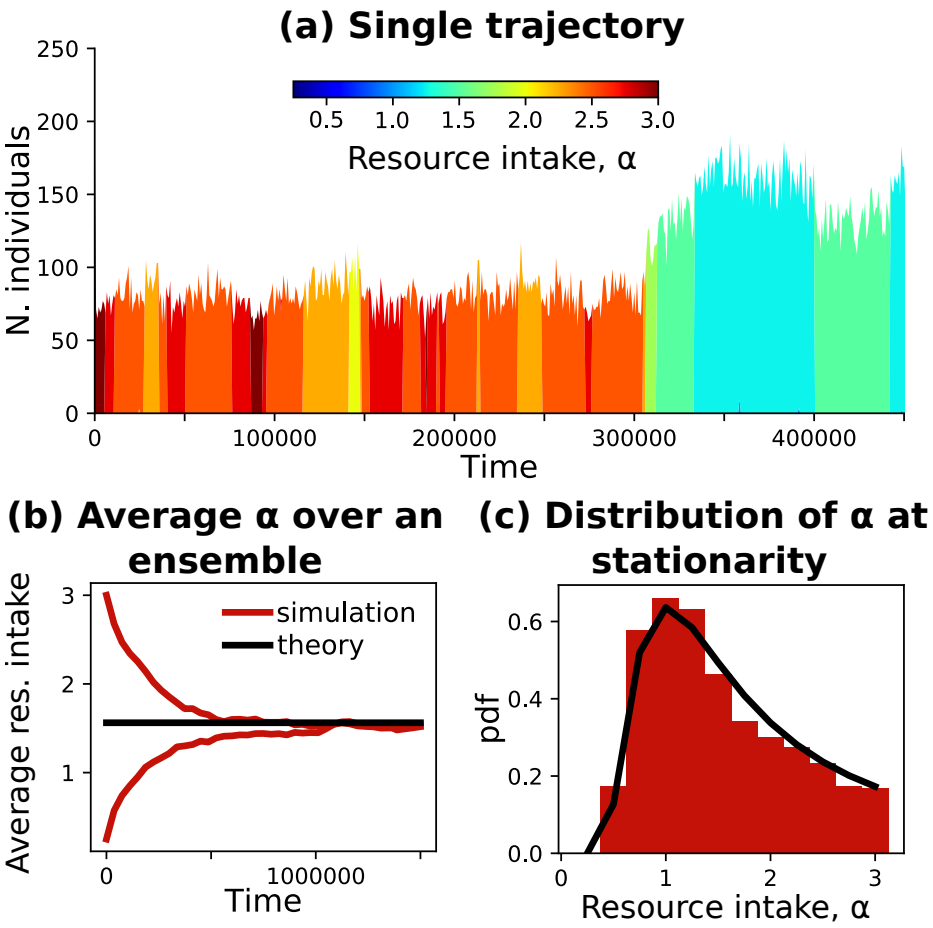

FIG. 7: Evolutionary dynamics of the chemostat with a constant intake rate $\eta_0 = 1$. The resource intake α can mutate to a neighbour $\alpha + \Delta\alpha$ or $\alpha - \Delta\alpha$ with rate $U = 2 \cdot 10^{-5}$. It is constrained to be in $[0, 3]$. The other parameters are: $\Delta\alpha = 0.25$, $\alpha_0 = 0.01$, $M = 200$. Panel (a) shows the number of individuals colored by their value of α of a Gillespie simulation. There is no clonal interference since, typically, there is no co-existence of sub-populations with different α. Panel (b) is the average resource intake over 300 realizations of two ensembles starting from different initial conditions. Panel (c) is the distribution of α at stationarity of the 300 realizations.

Strain 1 is advantaged ($p_{1\text{fix}}/p_{2\text{fix}} > 1$) only if $\alpha_0 M - 2\alpha_1\alpha_2 > 0$. If, without loss of generality, we assume $\alpha_1 > \alpha_2$, we obtain that strain 1 is advantaged over strain 2 in the case $\alpha_2 < \alpha_1 < M\alpha_0/(2\alpha_2)$. Therefore if a strain i has a value of intake rate α_i such that $\alpha_i = M\alpha_0/(2\alpha_i)$, no other strain can be advantaged over it. This indicates the existence of an optimal value of α, equal to $\sqrt{\alpha_0 M/2}$, to which the evolutionary trajectories converge.

Figure 7 shows a simulation of this scenario, where α can mutate taking discrete, uniformly spaced, values $\{\alpha_i\}$. The evolutionary trajectory, starting both above and below the predicted optimal value converges to it in the long term. Figure 7b shows that the expected value of the intake rate α (and therefore the invasion fitness) decreases over time if the population has initial values of α above the predicted value.

The presence of stochasticity implies that the optimal value is reached on average, with a fluctuating value of α around the optimal value. Under the assumption of periodic selection, one can predict the distribution $p(\alpha)$ at equilibrium by assuming that evolution follows a jump process which satisfies the detailed balance: $p(\alpha_i)p(\alpha_i \rightarrow$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$\alpha_{i+1}) = p(\alpha_{i+1})p(\alpha_{i+1} \to \alpha_i)$ (details in Appendix H). The prediction for the average and the full distribution is in agreement with the simulation in panels (b) and (c) of Figure 7.

V. DISCUSSION

Here we have studied the evolutionary trajectories of populations in a broad class of models characterizing population self-limitation. We introduced a population dynamics framework that encompasses several known models (Logistic, Gompertz, Chemostat, and others). In presence of large, but finite, population sizes, we have shown that the evolutionary trajectories depend only on three quantities (timescale, invasion fitness, and carrying capacity), irrespective of the models' specific form.

We obtained these results by assuming a time-scale separation between the total population size and the strain (or alleles) frequencies [28, 31, 33, 35]. This assumption allows us to write down an effective equation that describes the time evolution of strain frequencies, which depends only on the values of the three evolutionary relevant quantities. This effective description reduces to Kimura's diffusion limit when strains differ only in their invasion fitness or in the limit of strong selection. In that case, variation in timescales and/or carrying capacities become effectively irrelevant.

One especially interesting aspect of the effective equation we obtained is the role played by the (finite) population size. In Kimura's diffusion limit, population size influences only the strength of genetic drift. While in that context drift masks the effect of selection (e.g., a deleterious mutation has a non-zero fixation probability), it never alters its direction (i.e., a beneficial mutation has always a larger fixation probability than a deleterious one). In our setting, the existence of finite population size might alter the course of evolution: mutants with lower invasion fitness might be more likely to invade than strains with a higher invasion fitness.

We explicitly show this effect in the chemostat model in the presence of a metabolic tradeoff. In the deterministic limit, larger resource intakes correspond to higher fitness, and therefore evolution drives the population to higher and higher resource intakes. In presence of a large, yet finite, population size, the naive expectation obtained in the deterministic case is not realized. The evolutionary trajectory converges in fact to an optimal value of the intake rate, which we analytically predict. If a clonal population has an initial intake rate larger than the optimal one, evolution will drive the population to decrease the intake rate, and therefore invasion fitness.

Invasion fitness decreases over time because of the presence of a finite population size. This might be reminiscent of Muller's ratchet [50–52], but it has a radically different origin. In the case of Muller's ratchet, the decrease in fitness is determined by the fact that mutations only give rise to deleterious mutations, which can get fixed because of the finite population size. The presence of a small fraction of beneficial mutations is enough to balance the effect of deleterious mutations and lead to an overall increase in fitness [53].

In our case, the mechanism in play is radically different. The presence of variation in the timescale and/or in the carrying capacity creates an effective force driving the population to higher values of the carrying capacity and lower values of the inter-generation times. The parameter identified as invasion fitness in the deterministic dynamics does not predict the (likely) outcome of the evolutionary dynamics. Lower carrying capacities correspond to a higher level of genetic drift (as shown in eq. S12 of the Supplementary Materials). The intuition is that the evolutionary trajectory drives the population toward lower values of genetic drift. This effect is alike to thermophoresis [54] or to what is observed in Brownian motion in presence of a position-dependent diffusion: an effective force drives the trajectories to

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

lower values of noise [55]. In our case, the effective force can counterbalance the sign of the selection coefficient.

This result highlights the relevance of genetic drift in shaping evolutionary trajectories. In our case, drift not only affects the speed of evolution [56], but also its direction. One remarkable aspect is that this effect turns out to be model-independent and determined only by differences in the timescale and carrying capacity.

The three quantities (invasion fitness, carrying capacity, and timescale) that effectively determine the trajectory of an evolving population are not independent traits. Variation of a given trait can influence one, two, or all three of them. For instance, in the example of the chemostat model that we discussed, only one trait, the resource intake, is under selection pressure. But since it influences all three quantities, the result of evolution is far from trivial.

Our results could be reminiscent of r/K selection [57, 58], but the connection is not obvious. In saturated environments, i.e. when populations are close to carrying capacities, K-selection dominates, favoring strains with high-competitive abilities (which in our setting correspond to higher values of ϕ). In un-saturated environments, r-selection favors strains with higher fecundity (in our setting, lower values of T). We considered the case of a saturated environment, studying evolution when total population abundance was close to carrying capacity. The effect of demographic stochasticity, as opposed to environmental fluctuations at the core of r/K selection, was responsible for non-trivial evolutionary effects (e.g., the existence of an optimal intake rate in the chemostat model) which disappear in the limit of large population sizes.

Our results go in the direction of building a model-independent eco-evolutionary theory. The extent to which the specific details of ecological interactions influence the evolutionary outcome is a major limitation to the development of a comprehensive general understanding of eco-evolutionary trajectories. Our results present a first step in this direction, as they describe the evolution of a population limited by a single factor. The next step will be to generalize our framework by considering multiple limiting factors, giving rise to the coexistence of different populations [59]. As a further generalization, one can consider different interaction types, going beyond the competition.

One important hypothesis of our analysis was to consider the case of strain independent functional forms of $\omega(\cdot)$ and $\beta(\cdot)$. This is a key assumption to obtain model-independent evolutionary trajectories. One could consider the more general scenario where the functional forms $\beta_i(\cdot)$ and $\omega_i(\cdot)$ explicitly depended on strain i in a way that could not be recast to the one eq. 3. While this case could be interesting for future work from a theoretical standpoint, it is important to highlight that all the commonly used models could be written in terms of eq. 3 and did not require introducing strain-dependant functional forms. Another limitation of our framework, again related to the choice of $\omega(\cdot)$ and $\beta(\cdot)$, lies in the fact that we restrict our analysis to a class of models that admit a Lyapunov function, through the hypothesis of monotonic $\omega(\cdot)$. This framework does not allow to include mechanisms allowing multi-stability (e.g. Allee effect).

In this work, we have only considered haploid populations. The extension to diploid (or complex) traits is numerically easy to implement as our setting can be generalized fairly simply. It is much more complicated to obtain closed equations for the diffusive limit under a separation of time scales. Preliminary analysis of numerical examples, shows however that, also in the diploid case, evolutionary trajectories are model-independent.

To conclude, our work shows that several aspects of the population dynamics are details that do not influence the trajectories of an evolving population. Only three demographic parameters effectively matter. Their variation is subject to selection in a non-trivial, drift mediated, way, which might lead to a decrease in what is usually identified as fitness over time.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Appendix A: Definition of the ecological models

1. Logistic growth and competitive Lotka-Volterra

Logistic growth is recovered for $\omega(z) = z$ and $\beta(z) = 1$. In this case, the per-capita growth rate of a clonal population, obtain from eq. 1 reads

$$\frac{1}{N}\frac{dN}{dt} = \frac{1}{T}\left(1 - \frac{N}{K}\right) \ . \tag{A1}$$

In the case of multiple strains, eq. 3 we obtain

$$\frac{1}{n_i}\frac{dn_i}{dt} = \frac{1}{T_i}\left(1 - \frac{1}{\phi_i}\sum_j \phi_j \frac{n_j}{K_j}\right) \tag{A2}$$

2. Chemostat

To derive the chemostat dynamics, let us first consider a clonal population. Let us assume that the per-capita birth rate depends on the concentration R of some resource, $b(N)/N = \alpha\eta H(R)$, where $H(\cdot)$ is some monotonic function of the concentration, an intake-rate α, and a conversion factor between resource intake and biomass η. We assume that the per-capita death rate is independent of abundance $d(N)/N = \delta$. The resource concentration changes accordingly to a classic chemostat equation

$$\dot{R} = M - \alpha H(R) N \ , \tag{A3}$$

where the first term represents a constant influx of resources and the second term is consumption. We assume that the resource concentration R quickly reaches a stationary value, which leads to $H(R^*) = M/(N\alpha)$ at the stationary value. We obtain therefore that $b(N) = \eta M$, from which we get

$$\frac{1}{N}\frac{dN}{dt} = \left(\eta\frac{M}{N} - \delta\right) \ , \tag{A4}$$

which corresponds to the case $\beta(z) = 1/z$ and $\omega(z) = z$.

This can be easily extented to multiple strains, following [42] in the case of a single resource, by assuming that α_i, η_i and δ_i have strain-dependent values, while resource concentration changes as

$$\dot{R} = M - H(R)\sum_i \alpha_i n_i \ , \tag{A5}$$

from which, after a time-scale separation, we obtain eq 12.

This model maps in our general framework using $T_i = 1/(\alpha_i\eta_i\ )$, $\phi_i = \alpha_i\eta_i/\delta_i$, and $K_i = M\eta_i/\delta_i$.

3. Gompertz growth

Gompertz growth [44] of a clonal population is defined as

$$\frac{1}{N}\frac{dN}{dt} = \alpha \log\left(\frac{K}{N}\right) \ , \tag{A6}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

which corresponds to $\beta(z) = 1$ and $\omega(z) = \log(z)$.

We generalize Gompertz growth to the case of multiple strains as

$$\frac{1}{n_i}\frac{dn_i}{dt} = \alpha_i \log\left(\frac{k_i}{\sum_j \chi_j n_j}\right), \tag{A7}$$

which recovers the classical Gompertz growth dynamics in the case of a single strain and $\chi_1 = 1$. The mapping with the general framework is obtained uwing: $T_i = 1/(\alpha_i \log(k_i))$, $\phi_i = \log(k_i)$, and $K_i = k_i/\chi_i$.

4. Generalized Lotka-Volterra

The generalized logistic dynamics [45] in a clonal population is defined as

$$\frac{1}{N}\frac{dN}{dt} = \frac{1}{T_i}\left(1 - \left(\frac{N}{K}\right)^{\nu}\right), \tag{A8}$$

where ν is a positive exponent. The generalize logistic dynamics correspond to the choice $\beta(z) = 1$ and $\omega(z) = z^{\nu}$

The case of multiple strains can be written as follows:

$$\frac{1}{n_i}\frac{dn_i}{dt} = \frac{1}{T_i}\left(1 - \left(\frac{\sum_j \chi_j n_j}{k_i}\right)^{\nu}\right), \tag{A9}$$

whose mapping reads: $\phi_i = k_i^{\nu}$, and $K_i = k_i/\chi_i$. The simulations of figure 5 are obtained for $\nu = 1/2$.

5. Von Bertalanffy model

The Von Bertalanffy growth [46] is defined by the equation

$$\frac{1}{N}\frac{dN}{dt} = pN^{\alpha-1} - q\ , \tag{A10}$$

where $\alpha \in [0, 1)$. Von Bertalanffy growth corresponds to $\beta(z) = z^{\alpha-1}$ and $\omega(z) = z^{1-\alpha}$.

The natural extension to competing strains reads:

$$\frac{1}{n_i}\frac{dn_i}{dt} = p_i\left(\sum_j \chi_j n_j\right)^{\alpha-1} - q_i\ . \tag{A11}$$

It is included in the general framework by choosing $T_i = 1/p_i$, $\phi_i = p_i/q_i$, and $K_i = (q_i/p_i)^{1/(\alpha-1)}/\chi_i$. The simulations of figure 5 are obtained for $\alpha = 2/3$.

6. Generalized Moran model

The Generalized Moran model [35], is defined by the birth and death rates

$$\begin{aligned} n_i &\to n_i + 1 \ \text{ w. rate } \frac{n_i}{T_i} \\ n_i &\to n_i - 1 \ \text{ w. rate } \frac{n_i}{T_i}\alpha\left(1 + \sum_j \frac{n_j}{K}\right)\ , \end{aligned} \tag{A12}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

which corresponds to the case $\beta(z) = 1$ and $\omega(z) = \alpha(1 + z)$, where α is the same across strains. Note that, in this case, $K_i = K$ and $\phi_i = 1$ for all strains.

This corresponds in fact to the "quasi-neutral" case, where α, defined as the ratio of death and birth rate is constant for each genotype.

Appendix B: Deterministic dynamics of multiple strains

Equation 3 describes the deterministic limit of the dynamics in the case of multiple strains. A Lyapunov function for the dynamical system of eq.3 reads

$$L = -\sum_i \frac{n_i}{K_i} u_i \phi_i + \int_0^{\sum_j u_j n_j / K_j} dz\ \omega(z)\ , \tag{B1}$$

where $u_j = \omega^{-1}(\phi_j)$. One can in fact show that

$$\frac{dL}{dt} = \sum_i \frac{dn_i}{dt}\frac{\partial L}{\partial n_i} = -\frac{n_i}{K_i}\frac{u_i}{T_i \phi_i}\beta\left(\sum_j u_j \frac{n_j}{K_j}\right)\left(\phi_i - \omega\left(\sum_j \frac{n_j}{K_j} u_j\right)\right)^2 \leq 0\ , \tag{B2}$$

under the assumption that $\beta(\cdot)$ is positive. The second derivatives of L read

$$\frac{\partial^2 L}{\partial n_i \partial n_j} = \frac{1}{K_i K_j}\omega'\left(\sum_j \frac{n_j}{K_j} u_j\right)\ . \tag{B3}$$

The Hessian is therefore positive definite if and only if the function $\omega(\cdot)$ is monotonically increasing, i.e. if $\omega'(\cdot)$ is a positive function. Under this assumption, the function L is a convex function. The fact that the function $\omega(\cdot)$ is monotonically increasing also implies that the function defined in eq. B1 is bounded from below.

The existence of a convex Lyapunov function implies that, in the non-degenerate case (i.e., where multiple populations are characterized by the same value of ϕ) there is a unique, globally stable, fixed point of eq. 3. By minimizing eq. B1, one obtains that the globally stable fixed point is

$$n_i^* = \begin{cases} K_i, & \text{if } \phi_i = \max_j \phi_j \\ 0, & \text{otherwise} \end{cases} \tag{B4}$$

Unless multiple strains are characterized by the same values of ϕ_i, eq. 3 does not, therefore, allow the coexistence of multiple strains. For large times, the community is dominated by the strain with the largest value of ϕ_i.

Appendix C: Interpretation of parameters

The parameter K_i represent the carrying capacity of a strain in isolation. In absence of others, in fact, the birth and death rate of a strain (eq. 5) are equal to each other when $n_i = K_i$. In particular, they reads

$$\left.\frac{b_i(n_i)}{n_i}\right|_{n_i=K_i} = \left.\frac{d_i(n_i)}{n_i}\right|_{n_i=K_i} = \frac{\beta(\omega^{-1}(\phi_i))}{T_i}\ . \tag{C1}$$

Up to the factor $\beta(\omega^{-1}(\phi_i))$, the parameter T_i represents therefore the expected generation time, i.e. the average interval between the birth of an individual and the birth of its offspring, when the population is at carrying capacity.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

For instance, in the case of the logistic growth (when $\beta(z) = 1$), T_i exactly corresponds to the generation time at carrying capacity.

Moreover, when a clonal population is rare $n_i \ll K_i$, the average growth of strain i reads

$$\frac{dn_i}{dt} \approx \frac{n_i}{T_i}\beta(0)\left(1 - \frac{\omega(0)}{\phi_i}\right) . \tag{C2}$$

The population is therefore expected to grow exponentially, with growth rate proportional to $1/T_i$. The proportionality constant is model dependent, and in particular depends on the values $\omega(0)$ and $\beta(0)$. For instance, in the Logistic case, since $\omega(0) = 0$ and $\beta(0) = 1$, the growth rate is exactly equal to $1/T_i$. Note that for some of the models, $\beta(z)$ and/or $\omega(z)$ diverge in the limit $z \to 0$. In those cases, the growth at small populations abundances is not exponential.

As shown in Appendix B, the values of ϕ unequivocally determine which strain survives in the long term, and can therefore be naturally interpreted as fitness. Such an interpretation becomes more clear by considering a rare mutant/invader spreading in the population in presence of a large resident population at carrying capacity. Let the strain 1 be a rare invader (i.e. with initial population $n_1 \approx 1 \ll K_1$) and 2 the type of the resident population. Up to small terms (of the order $1/K_1$), the equation 3 for the invader reads:

$$\frac{dn_1}{dt} \approx \frac{n_1}{\phi_1 T_1}\beta(\omega(\phi_2)^{-1})(\phi_1 - \phi_2), \tag{C3}$$

where the growth rate is proportional to $(\phi_1 - \phi_2)/(\phi_1 T_1)$, as also depicted in Fig. 2.

Appendix D: Diffusive limit of the birth-death process

Here we are interested in the limit $K_i \to \infty$ of the birth-death process 5. In such a limit, the stochastic process can be approximated, via a diffusion limit, with a stochastic differential equation, which treats the abundances n_i as continuous variables. The diffusion limit can be formally obtained by truncating the Kramers-Moyal expansion to the leading terms. Here below we provide an intuitive derivation of the diffusion limit in the case of one strain. The generalization to an arbitrary number of strains does not present any further conceptual difficulty.

Let us consider a very small interval of time $\Delta t \ll 1/(b(n) + d(n))$ such that it is very unlikely that more than one birth and/or death events occur in Δt. This allows us to write down the probability of having n individuals at time $t + \Delta t$ as:

$$P(n, t + \Delta t) = b(n-1)\Delta t P(n-1, t) + d(n+1)\Delta t P(n+1, t) + (1 - \Delta t(b(n) + d(n)))\, P(n, t). \tag{D1}$$

Let us now consider the re-scaled variable $y = n/K$, where K is the carrying capacity of the strain. The functions that depended on $n \pm 1$, now depend on $y \pm 1/K$. Since $K \gg 1$ we can consider y as a continuous variable and $1/K$ as a small increment, which allows us to expand those functions in the Taylor series. By expanding up to the second term, we obtain

$$\partial_t P \left(b - \frac{\partial_y b}{K} + \frac{\partial_y^2 b}{2K^2}\right)\left(P - \frac{\partial_y P}{K} + \frac{\partial_y^2 P}{2K^2}\right) + \left(d + \frac{\partial_y d}{K} + \frac{\partial_y^2 d}{2K^2}\right)\left(P + \frac{\partial_y P}{K} + \frac{\partial_y^2 P}{2K^2}\right) + -(b+d)P, \tag{D2}$$

where we omitted the arguments (y, t) from all the functions for the sake of a simpler notation, and neglected the terms of order $1/K^3$ and higher. By regrouping the terms at different orders of $1/K$, one obtains

$$\partial_t P = -\frac{1}{K}\left((b-d)\partial_y P + P\partial_y(b-d)\right) + \frac{1}{2K^2}\left((b+d)\partial_y^2 P + 2\partial_y P \partial_y(b+d) + P\partial_y^2(b+d)\right), \tag{D3}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

which can be simplified to the following Fokker-Plank equation:

$$\partial_t P = \frac{1}{K}\partial_y\left((b-d)P)\right) + \frac{1}{2K^2}\partial_y^2\left((b+d)P\right). \tag{D4}$$

From this expression one can obtain the deterministic coefficient, $(b-d)/K$, and the diffusion coefficient, $(b+d)/(2K^2)$, and therefore the associated stochastic differential equation. In the case of A strains, the resulting Fokker-Plank reads

$$\partial_t P = \sum_i \frac{1}{K_i}\frac{\partial}{\partial y_i}\left((b-d)P)\right) + \sum_i \frac{1}{2K_i^2}\frac{\partial^2}{\partial y_i^2}\left((b+d)P\right) \ , \tag{D5}$$

which is equivalent to the following system of Itô SDEs:

$$\begin{aligned}\dot{y}_i =& \frac{b_i(\mathbf{n}) - d_i(\mathbf{n})}{K} + \sqrt{\frac{b_i(\mathbf{n}) + d_i(\mathbf{n})}{K^2}} \cdot \xi_i = \\ =& y_i \frac{1}{T_i}\beta\left(z(\mathbf{n})\right)\left(1 - \frac{\omega\left(z(\mathbf{n})\right)}{\phi_i}\right) + \sqrt{\frac{y_i}{K_i}\frac{1}{T_i}\beta\left(z(\mathbf{n})\right)\left(1 + \frac{\omega\left(z(\mathbf{n})\right)}{\phi_i}\right)} \cdot \xi_i,\end{aligned} \tag{D6}$$

where ξ_i is an uncorrelated normal random variable, $\langle \xi_i(t)\rangle = 0$, $\langle \xi_i(t)\xi_j(t')\rangle = \delta_{ij}\delta(t-t')$, and we defined $z(\mathbf{n}) = \sum_j u_j n_j / K_j$. From this expression one can easily obtain the equation 6 by simply by change of variable $n_i = y_i K_i$. In the limit $K \to \infty$, the fluctuations become negligible and one obtains the deterministic limit of equation 3.

Appendix E: Separation of Ecological and Evolutionary timescales

Let us consider the diffusive dynamics derived above $\dot{y}_i = a_{y_i} + \sqrt{2D_{y_i}} \cdot \xi_i$, whose coefficients read:

$$\begin{aligned} a_{y_i} &= \frac{y_i}{T_i}\beta\left(z\right)\left(1 - \frac{\omega\left(z\right)}{\phi_i}\right) \\ D_{y_i} &= \frac{y_i}{2K_iT_i}\beta\left(z\right)\left(1 + \frac{\omega\left(z\right)}{\phi_i}\right) \ , \end{aligned} \tag{E1}$$

where we have omitted the arguments $\mathbf{n}$ in $z(\mathbf{n})$.

When fitness differences are small (i.e., $|\phi_i - \phi_j|/\phi_j \ll 1$) the stochastic dynamics separates in two temporal phases, depicted in Fig. 3. The first phase corresponds to a fast transient: the total population size $N = \sum_i n_i$ changes to reach a value close to the carrying capacity. Fitness differences and stochasticity are negligible in this first transient but determine the dynamics of the second phase.

We introduce the selection coefficient $s_i = (\phi - \phi_i)/\phi$ of strain i relative to strain 1, where $\phi = \phi_1$. As we are interested in the limit $s_i \ll 1$ and $K_i \gg 1$, we can expand eq. E1. The expansion up to the first first order in s_i and $1/K_i$ reads,

$$\begin{aligned} a_{y_i} &= \underbrace{\frac{y_i}{T_i}\beta\left(z\right)\left(1 - \frac{\omega\left(z\right)}{\phi}\right)}_{\text{Fast dynamics}} - \underbrace{s_i\frac{y_i}{T_i}\beta\left(z\right)\frac{\omega\left(z\right)}{\phi}}_{\text{Slow dynamics}} \\ D_{y_i} &= \underbrace{\frac{y_i}{2K_iT_i}\beta\left(z\right)\left(1 + \frac{\omega\left(z\right)}{\phi}\right)}_{\text{Slow dynamics}} \ . \end{aligned} \tag{E2}$$

This expansion makes apparent the distinction between fast and slow dynamics. The latter terms are proportional to s_i or $1/K_i$, which are both small, while the former is not proportional to those factors. The fast dynamics (see

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 3(a) for an example with Lotka-Volterra dynamics) is determined by the first deterministic term in eq E2, independent of K_i and s_i. This fast dynamics drives the system to a manifold of solutions where $\omega(z) = \phi$, which implies $\sum_i n_i/K_i = \sum_i y_i = 1$.

The slow dynamics that follows is determined by the deterministic term linear in s_i and the stochastic term $1/K_i$. The time-scale separation allows us to consider the effective evolution along the slow manifold. Such effective dynamics is non-trivial (see Figure 3). The small stochastic terms of the original dynamics are no longer negligible and kick the variable outside the manifold. As soon as the abundances move away from the manifold, the fast dynamics become relevant again, and it pushes back — to a non-trivial position — the trajectory onto the manifold.

To correctly take into account this tension between the slow and fast terms, the increment of the effective variable constrained to move along the manifold, can be obtained by projecting the original increment dy onto the manifold following the flow field of the fast dynamics. To perform this calculation we take advantage of the fact that there exists a conserved quantity of the fast dynamics. This simplifies a lot the calculation of [60]. The derivation is similar to the computation proposed in [33].

1. Fast transient and conserved quantity

The deterministic equation describing the fast transient dynamics reads

$$\frac{dy_i}{dt} = \frac{y_i}{\phi T_i}\beta\left(\omega^{-1}(\phi)\sum_i y_i\right)\left(\phi - \omega\left(\omega^{-1}(\phi)\sum_i y_i\right)\right). \tag{E3}$$

This equation has a manifold of marginally stable fixed points, identified by the simplex $\sum_i y_i = 1$. Any point in this simplex is a fixed point of eq. E3. The dynamics drive the system to one of these solutions, which will differ depending on the initial condition.

The existence of a manifold of fixed points is a consequence of a symmetry in the dynamics of eq. E3 which corresponds to the existence of a conserved quantity, which reads

$$\alpha(\mathbf{y}) := \exp\left(\sum_{j=2}^{A}\left(\log y_j - \frac{T_1}{T_j}\log y_1\right)\right) . \tag{E4}$$

From which one obtains

$$\frac{d\alpha(\mathbf{y})}{dt} = \sum_i \frac{dy_i}{dt}\frac{\partial\alpha}{\partial y_i} = -\frac{dy_1}{dt}\frac{T_1}{y_1}\sum_j \frac{1}{T_j} + \sum_{j=2}^{A}\frac{1}{y_j}\frac{dy_j}{dt} = 0 . \tag{E5}$$

In the simple case of two strains, the conserved quantity reads:

$$\alpha(y_1, y_2) = y_1^{-T_1/T_2} y_2 \tag{E6}$$

2. Effective slow dynamics

The value of the function E4 does not change along the flow field of the fast dynamics and it allows us to correctly calculate the projection onto the slow manifold. In the following, we consider the case of two strains.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Due to the presence of the terms proportional to s_i and $1/K_i$ in eq. E2 the value of $\alpha(\mathbf{y})$ is not conserved during the slow dynamics. The time evolution of $\alpha(\mathbf{y})$ can be obtained from eq. E2 using Itô calculus

$$\frac{d\alpha}{dt} = \frac{\partial \alpha}{\partial y_1}\left(a_{y_1} + \sqrt{2D_{y_1}} \cdot \xi_1\right) + \frac{\partial \alpha}{\partial y_2}\left(a_{y_2} + \sqrt{2D_{y_2}} \cdot \xi_2\right) + \frac{1}{2}\frac{\partial^2 \alpha}{\partial y_1^2} 2D_{y_1}, \tag{E7}$$

where a and D are the deterministic and diffusion terms appearing in eq. E2: $dy_i/dt = a_{y_i} + \sqrt{2D_{y_i}} \cdot \xi_i$. The term with a second-order derivative of y_2 is equal to zero because α is linear in y_2. The term with mixed y_1 and y_2 derivatives equals zero because the two noises are uncorrelated. By writing explicitly all the terms, one obtains:

$$\frac{d\alpha}{dt} = \beta(z)\frac{\alpha(y_1, y_2)}{T_2}\left(\frac{\omega+\phi}{\phi}\frac{c(1+r)}{2K_2 y_1 r} - s\frac{\omega}{\phi}\right) + \alpha(y_1, y_2)\sqrt{\beta(z)\frac{r y_1 + c y_2}{2 y_1 y_2 r T_2 K_2}\frac{\omega+\phi}{\phi}} \cdot \eta \tag{E8}$$

where η is a white Gaussian noise, and $r = T_2/T_1$, $c = K_2/K_1$, $s = s_2 \ll 1$.

The key point of the derivation that follows is that the change of α outside the manifold is the same on the manifold for a variable projected through the flow field. Let us parametrize the position on the manifold as $\Gamma(y) = (\tilde{y}, 1-\tilde{y})$, where the variable $\tilde{y}$ corresponds to y_1 when constrained to be on the slow manifold. The change in α for the full dynamics is equivalent to the one obtained by constraining the dynamics on the slow manifold

$$d\alpha(y_1, y_2) = d\alpha(\tilde{y}, 1-\tilde{y}). \tag{E9}$$

This implies that the dynamics of α, evaluated for the variables on the manifold, follow equation E8

$$\left.\frac{d\alpha}{dt}\right|_{\Gamma} = \beta(\omega^{-1}(\phi))\frac{\alpha(\tilde{y})}{T_2}\left(\frac{c(1+r)}{K_2 \tilde{y}\, r} - s\right) + \alpha(\tilde{y})\sqrt{\beta(\omega^{-1}(\phi))\frac{r\tilde{y} + c(1-\tilde{y})}{\tilde{y}(1-\tilde{y}) r T_2 K_2}} \cdot \eta. \tag{E10}$$

In the following we will remove the term $\beta(\omega^{-1}(\phi))$ as it can always be reabsorbed in the definition of $\beta(\cdot)$.

One can then derive the expression for the re-scaled abundance on the slow manifold $\tilde{y}$ using the Itô formula from eq. E10:

$$\frac{d\tilde{y}}{dt} = \frac{d\tilde{y}}{d\alpha}\left(a_\alpha + \sqrt{2D_\alpha} \cdot \eta\right) + \frac{1}{2}\frac{d^2\tilde{y}}{d\alpha^2} 2D_\alpha. \tag{E11}$$

To perform the calculation, one needs to compute the first and second derivatives by using the derivatives of the inverse function $\alpha(y)$:

$$\frac{dy}{d\alpha} = \left(\frac{d\alpha}{dy}\right)^{-1} \qquad \frac{d^2 y}{d\alpha^2} = -\left(\frac{d\alpha}{dy}\right)^{-3}\frac{d^2\alpha}{dy^2}. \tag{E12}$$

Putting everything together, the final expression for the slow variable is the following:

$$\frac{d\tilde{y}}{dt} = \frac{\tilde{y}(1-\tilde{y})}{(1-\tilde{y}+r\tilde{y})\, T_1}\left[s + \frac{1}{K_2(1-\tilde{y}+r\tilde{y})}\left(1 - rc + \frac{r(1-c)}{(1-\tilde{y}+r\tilde{y})}\right)\right] + \sqrt{\frac{\tilde{y}(1-\tilde{y})}{K_2\,(1-\tilde{y}+r\tilde{y})^2\, T_1}(r\tilde{y} + c(1-\tilde{y}))} \cdot \eta. \tag{E13}$$

One can express the dynamics in terms of the frequency $x = n_1/(n_1+n_2)$, which is related with the re-scaled abundance $\tilde{y} = n_1/K_1$ as $x = \tilde{y}/(c + \tilde{y}(1-c))$. By using again Itô calculus, one finally obtains

$$\frac{dx}{dt} = \frac{x(1-x)\overline{c}}{T_1\overline{c}\overline{r}}\left(s - \frac{\overline{c}}{K_2\overline{c}\overline{r}}\left(rc - 1 - (1-c)\left(\frac{r\overline{c}}{\overline{c}\overline{r}} - 2\overline{r}\right)\right)\right) + \sqrt{\frac{x(1-x)\overline{c}^3\overline{r}}{T_1 K_2 \overline{c}\overline{r}^2}} \cdot \eta\ , \tag{E14}$$

where we used the notation $\overline{f} = xf + 1 - x$ This equation is equivalent to eq. 7.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Appendix F: Invasion probability

The question that we address now is whether a mutant strain that appears in a population composed of only one resident strain can invade and overcome the original population. To this end, we need the fixation probability $\pi_1(\tilde{y})$, i.e. the probability that, given the initial re-scaled abundance $\tilde{y}$, the species will reach $\tilde{y} = 1$ (notice that the same computation can be done starting from the frequencies (E14) leading to the same result).

The fixation probability can be obtained by solving the differential equation $L^+\pi_1(\tilde{y}) = 0$, where L^+ is the backward operator of the stochastic process, and the boundary conditions are $\pi_1(0) = 0$ and $\pi_1(1) = 1$. This leads to the following formula:

$$a_{\tilde{y}} \frac{d\pi_1}{d\tilde{y}} + D_{\tilde{y}} \frac{d^2\pi_1}{d\tilde{y}^2} = 0 \ , \tag{F1}$$

The solution of this equation reads:

$$\pi_1(\tilde{y}) = \frac{\mathcal{L}(\tilde{y})}{\mathcal{L}(1)} \tag{F2}$$

where

$$\mathcal{L}(\tilde{y}) = \int_0^{\tilde{y}} \exp\left(-I(y')\right) dy' \ , \tag{F3}$$

and

$$I(y) = \int^{y} \frac{a_y(y')}{D_y(y')} dy' \ . \tag{F4}$$

The invasion probability for the first species, $p_{1\ \mathrm{fix}}$, is the fixation probability starting from one individual $n_1 = 1$, then $\tilde{y} = 1/K_1$. Since $1/K_1 \ll 1$ we can try to compute it by expanding in the Taylor series the fixation probability around zero. In this way, we can simplify the integrals (which can be expressed exactly as a combination of hypergeometric functions):

$$p_{1\ \mathrm{fix}} = \pi_1\left(\frac{1}{K_1}\right) \simeq \pi_1(0) + \frac{1}{K_1}\frac{d\pi_1}{dy}(0) = \frac{\exp\left(-I(0)\right)}{K_1\mathcal{L}(1)} \tag{F5}$$

One can obtain also the invasion probability of the second species by noting that the fixation probability reads $\pi_2(1-\tilde{y}) = 1-\pi_1(\tilde{y})$. The fixation probability has to be evaluated at $\tilde{y} = 1-1/K_2$, when there is only one individual of the second species and the first one is at carrying capacity. Similarly as before, one can obtain the invasion probability by expanding for $1/K_2 \ll 1$:

$$p_{2\ \mathrm{fix}} \simeq 1 - \pi_1(1) + \frac{1}{K_2}\frac{d\pi_1}{dy}(1) = \frac{\exp\left(-I(1)\right)}{K_2\mathcal{L}(1)} \tag{F6}$$

If one considers the ratio between these two probabilities (and neglecting the terms $O(1/K_2)$) the dependency with the integral $\mathcal{L}$ disappears:

$$\frac{p_{1\ \mathrm{fix}}}{p_{2\ \mathrm{fix}}} = c\exp\left(I(1) - I(0)\right) \tag{F7}$$

which leads to

$$\log\frac{p_{1\ \mathrm{fix}}}{p_{2\ \mathrm{fix}}} = sK_2\,\frac{r-1}{r-c} + \frac{sK_2\,r(c-1) - (1-rc)(r-c)}{(r-c)^2}\log\frac{c}{r} \ . \tag{F8}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

One can notice that in the case $r = c$ this expression is odd-defined. The case $r = c$ can be obtained by repeating the full calculation for $r = c$. One obtains

$$\log \frac{p_{1\text{ fix}}}{p_{2\text{ fix}}}\bigg|_{r=c} = \frac{r+1}{r}\left(\frac{sK_2}{2} + 1 - r\right) . \tag{F9}$$

Appendix G: Simulation of the invasion probability

All the trajectories used in the simulations are generated through a standard Gillespie algorithm for the birth-death process under study.

A straightforward approach can be to generate R independent trajectories, where the first strain starts at $n_1 = 1$, and the abundance of the second strain is sampled from the stationary distribution of the process without type 1: $\tilde{n}_2 \sim P_{\text{stat }2}$. The invasion probability is then just

$$p_{1\text{ fix}} = \sum_{n_2} P_{\text{stat }2}(n_2) P(n_2 \to 0|1, n_2) \approx \sum_{\tilde{n}_2 \sim P_{\text{stat }2}} \frac{\mathbb{I}(n_2 \to 0|1, \tilde{n}_2)}{R} \tag{G1}$$

where $P(n_2 \to 0|n_1, n_2)$ represents the probability of the event $n_2 \to 0$, i.e. the second population gets extinct, given an initial number of individuals (n_1, n_2). $\mathbb{I}(n_2 \to 0|n_1, n_2)$ is the indicator function that is 1 if the event $n_2 \to 0$ for a trajectory starting from (n_1, n_2) does occur. However, this approach is very inefficient, since these counts in a simulation are very small (for the parameter setting typically considered, an invasion event is very rare), and a long computational time is needed to have a sufficient amount of samples.

We consider an alternative approach. First, one can observe that, using Bayes rule and the fact that the process is Markovian, the invasion probability can be rewritten by splitting the extinction of the second strain in the probability that the first species reaches a given threshold $n_1 = k$ or not:

$$P(n_2 \to 0|1, n_2) = P(n_2 \to 0 \ \ \& \ \ n_1 \to k|1, n_2) + P(n_2 \to 0 \ \ \& \ \ \neg(n_1 \to k)|1, n_2) \tag{G2}$$

where $\&$ is the logical and, and $\neg$ is the logical negation. The event $n_1 \to k$ means that the trajectory has reached (at least once in the past) the value k. One can then split the simulation in two parts. One simulates R trajectories starting from $n_1 = 1$ and n_2 sampled from $P_{\text{stat }2}$ as before. Then one counts the events of extinction of the second species when the first one does not cross k, giving an estimate of the second addendum above. Otherwise, if n_1 reaches k, the simulation is stopped, and the number of individuals of the second type is recorded in a set of values $\{n_2^k\}$. The second part of the simulation consists in generating new R trajectories starting from $n_1 = k$ and n_2 sampled at random from the list $\{n_2^k\}$. Note that R is greater (or equal) than the number of trajectories that in the previous simulation have reached $n_1 \to k$. However, the sampling from $\{n_2^k\}$ ensures that the statistics of n_2 when n_1 hits k is reliable, and, therefore, these new trajectories behave like continuations of the previous ones. Given this new set of R trajectories, the count of the events of the second species extinction estimates the first addendum of the equation above. Note that, in this way, when n_1 starts from k, the extinction probability is larger, and the counts for its computation increase, leading to better precision. In symbols:

$$p_{1\text{ fix}} \approx \sum_{\tilde{n}_2 \sim P_{\text{stat }2}} \frac{\mathbb{I}(n_2 \to 0 \ \ \& \ \ \neg(n_1 \to k)|1, \tilde{n}_2)}{R} + \frac{\mathbb{I}(n_2 \to 0 \ \ \& \ \ n_1 \to k|1, \tilde{n}_2)}{R} \sum_{\tilde{n}_2^k \sim \{n_2^k\}} \frac{\mathbb{I}(n_2 \to 0|k, \tilde{n}_2^k)}{R}. \tag{G3}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To further improve the statistics, one can add several increasing thresholds $(k_1, \dots k_K)$, and the invasion probability can be obtained with the following iterative formula:

$$p_{1\ \mathrm{fix}}(k_i) \approx \sum_{\tilde{n}_2^{k_i} \sim \{n_2^{k_i}\}} \frac{\mathbb{I}(n_2 \to 0\ \ \&\ \ \neg(n_1 \to k_{i+1})|k_i, \tilde{n}_2^{k_i})}{R} + \frac{\mathbb{I}(n_2 \to 0\ \ \&\ \ n_1 \to k_{i+1}|k_i, \tilde{n}_2^{k_i}}{R} p_{1\ \mathrm{fix}}(k_{i+1}), \tag{G4}$$

where $p_{1\ \mathrm{fix}}(k_0) = p_{1\ \mathrm{fix}}$.

Appendix H: Evolution in the chemostat model

We consider the chemostat model

$$\begin{aligned} n_i &\to n_i + 1 \ \text{ w. rate } \ n_i \frac{\eta_i \alpha_i M}{\sum_j \alpha_j n_j} \\ n_i &\to n_i - 1 \ \text{ w. rate } \ n_i \delta_i, \end{aligned} \tag{H1}$$

with one constraint to the parameters of the model: $\delta_i/\eta_i = \alpha_i + \alpha_0$, with $\alpha_0 \ll 1$. The question that we address is which evolutionary outcome one can expect when the parameter α can mutate. The answer depends on the details of how the two parameters δ_i and η_i depend on α, and, also, if one considers invasions through or without mutations.

Let us first map the chemostat parameters into the general framework. The carrying capacity ratio c and the parameter sK_2 read: $c = \alpha_1/\alpha_2$, $sK_2 \approx M\alpha_0(\alpha_1 - \alpha_2)/(\alpha_1\alpha_2^2)$. The timescale ratio r depends instead on how δ and η depends on α. In particular, we need to separate the cases in which the dependency is contained in the death rate or in the efficiency.

In the main text we consider the case $\delta_i = \eta_0(\alpha_i + \alpha_0)$ and $\eta_i = \eta_0$, which corresponds to, at the leading order in α_0, $r = c = \alpha_1/\alpha_2$, and leads to eq. 13

The other case considers a constant death rate δ_0, and $\eta_i = \delta_0/(\alpha_i - \alpha_0)$. This leads to $r = 1$ and $c = \alpha_1/\alpha_2$. Using equation 10 we obtain the ratio of the invasion probabilities is larger than one if $\alpha_2 < \alpha_1 < M\alpha_0/\alpha_2$, The evolutionary equilibrium situation is obtained when $\alpha^* = \sqrt{\alpha_0 M}$, which differs from the constant-efficiency case by a constant factor. A different scenario appears if one instead considers invasions through mutation, as described by eq 11. One can obtain that the invasion is always possible if $\alpha_1 > \alpha_2$, implying that evolution drives the system to larger and larger values of the intake rate α.

1. Deriving the distribution of intake rate

In the simulation of figure 4 of the main text, α can mutate taking discrete-equispaced values, $\{\alpha_i\}$, with a very small mutation probability that guarantees that the periodic selection regime holds. In this case, one can predict the distribution $p(\alpha)$ at equilibrium assuming a jump process which satisfies the detailed balance: $p(\alpha_i)p(\alpha_i \to \alpha_{i+1}) = p(\alpha_{i+1})p(\alpha_{i+1} \to \alpha_i)$. The transition probabilities are exactly the invasion probabilities through mutations of equation 11:

$$\frac{p(\alpha_i \to \alpha_{i+1})}{p(\alpha_{i+1} \to \alpha_i)} = \frac{m_{\ \mathrm{fix}}(\alpha_i, \alpha_{i+1})}{m_{\ \mathrm{fix}}(\alpha_{i+1}, \alpha_i)} \tag{H2}$$

where $m_{\ \mathrm{fix}}(\alpha_1, \alpha_2)$ is the invasion probability through mutations of a population having intake rate α_1 in a resident population having α_2. Since this ratio is given by equation 11 and 10, one can derive all the ratios between consecutive

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

rates $p(\alpha_i)/p(\alpha_{i+1})$. All these conditions plus the normalization allow finding all the values of the probability distribution $p(\alpha_i)$.

---

[1] Motoo Kimura. On the probability of fixation of mutant genes in a population. *Genetics*, 47(6):713, 1962.

[2] Warren J Ewens. The sampling theory of selectively neutral alleles. *Theoretical population biology*, 3(1):87–112, 1972.

[3] John Frank Charles Kingman. The coalescent. *Stochastic processes and their applications*, 13(3):235–248, 1982.

[4] Richard R Hudson, Martin Kreitman, and Montserrat Aguadé. A test of neutral molecular evolution based on nucleotide data. *Genetics*, 116(1):153–159, 1987.

[5] Fumio Tajima. Statistical method for testing the neutral mutation hypothesis by dna polymorphism. *Genetics*, 123(3):585–595, 1989.

[6] Stanley A Sawyer and Daniel L Hartl. Population genetics of polymorphism and divergence. *Genetics*, 132(4):1161–1176, 1992.

[7] Ziheng Yang and Joseph P Bielawski. Statistical methods for detecting molecular adaptation. *Trends in ecology & evolution*, 15(12):496–503, 2000.

[8] John H Gillespie. *Population genetics: a concise guide*. JHU Press, 2004.

[9] Warren J Ewens. *Mathematical population genetics 1: theoretical introduction*, volume 27. Springer Science & Business Media, 2012.

[10] Sewall Wright. Evolution in mendelian populations. *Genetics*, 16(2):97, 1931.

[11] Ronald Aylmer Fisher. *The genetical theory of natural selection*. 1958.

[12] Patrick Alfred Pierce Moran. Random processes in genetics. In *Mathematical proceedings of the cambridge philosophical society*, volume 54, pages 60–71. Cambridge University Press, 1958.

[13] Samuel Karlin and James McGregor. Direct product branching processes and related markov chains. *Proceedings of the National Academy of Sciences of the United States of America*, 51(4):598, 1964.

[14] Chris Cannings. The latent roots of certain markov chains arising in genetics: a new approach, i. haploid models. *Advances in Applied Probability*, pages 260–290, 1974.

[15] Jeffrey E Barrick and Richard E Lenski. Genome dynamics during experimental evolution. *Nature Reviews Genetics*, 14(12):827–839, 2013.

[16] Alvaro Sanchez and Jeff Gore. Feedback between population and evolutionary dynamics determines the fate of social microbial populations. *PLoS Biol*, 11(4):e1001547, 2013.

[17] Alan R Rogers and Henry Harpending. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular biology and evolution*, 9(3):552–569, 1992.

[18] Oskar Hallatschek, Pascal Hersen, Sharad Ramanathan, and David R Nelson. Genetic drift at expanding frontiers promotes gene segregation. *Proceedings of the National Academy of Sciences*, 104(50):19926–19930, 2007.

[19] Olivia J Burton, Ben L Phillips, and Justin MJ Travis. Trade-offs and the evolution of life-histories during range expansion. *Ecology letters*, 13(10):1210–1220, 2010.

[20] Robert H MacArthur. Some generalized theorems of natural selection. *Proceedings of the National Academy of Sciences of the United States of America*, 48(11):1893, 1962.

[21] Jonathan Roughgarden. Density-dependent natural selection. *Ecology*, 52(3):453–468, 1971.

[22] Bryan Clarke. Density-dependent selection. *The American Naturalist*, 106(947):1–13, 1972.

[23] Marjorie A Asmussen and Marcus W Feldman. Density dependent selection 1: A stable feasible equilibrium may not be attainable. *Journal of Theoretical Biology*, 64(4):603–618, 1977.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[24] Warren J Ewens. The probability of survival of a new mutant in a fluctuating environment. *Heredity*, 22(3):438–443, 1967.

[25] Sarah P Otto and Michael C Whitlock. The probability of fixation in populations of changing size. *Genetics*, 146(2):723–733, 1997.

[26] Lindi M Wahl and Philip J Gerrish. The probability that beneficial mutations are lost in populations with periodic bottlenecks. *Evolution*, 55(12):2606–2610, 2001.

[27] Zaheerabbas Patwa and Lindi M Wahl. The fixation probability of beneficial mutations. *Journal of The Royal Society Interface*, 5(28):1279–1289, 2008.

[28] Todd L Parsons and Christopher Quince. Fixation in haploid populations exhibiting density dependence i: The non-neutral case. *Theoretical population biology*, 72(1):121–135, 2007.

[29] Todd L Parsons and Christopher Quince. Fixation in haploid populations exhibiting density dependence ii: The quasi-neutral case. *Theoretical population biology*, 72(4):468–479, 2007.

[30] Yen Ting Lin, Hyejin Kim, and Charles R Doering. Features of fast living: on the weak selection for longevity in degenerate birth-death processes. *Journal of Statistical Physics*, 148(4):647–663, 2012.

[31] George WA Constable and Alan J McKane. Models of genetic drift as limiting forms of the lotka-volterra competition model. *Physical review letters*, 114(3):038101, 2015.

[32] Thiparat Chotibut and David R Nelson. Evolutionary dynamics with fluctuating population sizes and strong mutualism. *Physical Review E*, 92(2):022718, 2015.

[33] Thiparat Chotibut and David R Nelson. Population genetics with fluctuating population sizes. *Journal of Statistical Physics*, 167(3-4):777–791, 2017.

[34] Peter Czuppon and Arne Traulsen. Fixation probabilities in populations under demographic fluctuations. *Journal of mathematical biology*, 77(4):1233–1277, 2018.

[35] Todd L Parsons, Christopher Quince, and Joshua B Plotkin. Some consequences of demographic stochasticity in population genetics. *Genetics*, 185(4):1345–1354, 2010.

[36] Hal L Smith and Paul Waltman. *The theory of the chemostat: dynamics of microbial competition*, volume 13. Cambridge university press, 1995.

[37] Franck Courchamp, Ludek Berec, and Joanna Gascoigne. *Allee effects in ecology and conservation*. OUP Oxford, 2008.

[38] Johan Grasman. The expected extinction time of a population within a system of interacting biological populations. *Bulletin of mathematical biology*, 58(3):555–568, 1996.

[39] Ingemar Nåsell. On the quasi-stationary distribution of the stochastic logistic epidemic. *Mathematical biosciences*, 156(1-2):21–40, 1999.

[40] Otso Ovaskainen. The quasistationary distribution of the stochastic logistic model. *Journal of Applied Probability*, 38(4):898–907, 2001.

[41] Derek A Roff. Defining fitness in evolutionary models. *Journal of genetics*, 87(4):339–348, 2008.

[42] Anna Posfai, Thibaud Taillefumier, and Ned S Wingreen. Metabolic trade-offs promote diversity in a model ecosystem. *Physical review letters*, 118(2):028103, 2017.

[43] Benjamin Gompertz. Xxiv. on the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. in a letter to francis baily, esq. frs &c. *Philosophical transactions of the Royal Society of London*, (115):513–583, 1825.

[44] Charles P Winsor. The gompertz curve as a growth curve. *Proceedings of the National Academy of Sciences of the United States of America*, 18(1):1, 1932.

[45] FJ Richards. A flexible growth function for empirical use. *Journal of experimental Botany*, 10(2):290–301, 1959.

[46] Ludwig Von Bertalanffy. Quantitative laws in metabolism and growth. *The quarterly review of biology*, 32(3):217–231, 1957.

[47] Farida Vasi, Michael Travisano, and Richard E Lenski. Long-term experimental evolution in escherichia coli. ii. changes in

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.17.448795; this version posted May 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

life-history traits during adaptation to a seasonal environment. *The american naturalist*, 144(3):432–456, 1994.

[48] Maja Novak, Thomas Pfeiffer, Richard E Lenski, Uwe Sauer, and Sebastian Bonhoeffer. Experimental tests for an evolutionary trade-off between growth rate and yield in e. coli. *The American Naturalist*, 168(2):242–251, 2006.

[49] Jean-Nicolas Jasmin, Marcus M Dillon, and Clifford Zeyl. The yield of experimental yeast populations declines during selection. *Proceedings of the Royal Society B: Biological Sciences*, 279(1746):4382–4388, 2012.

[50] Hermann Joseph Muller. The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 1(1):2–9, 1964.

[51] Wilfried Gabriel, Michael Lynch, and Reinhard Bürger. Muller's ratchet and mutational meltdowns. *Evolution*, 47(6):1744–1757, 1993.

[52] Richard A Neher and Boris I Shraiman. Fluctuations of fitness distributions and the rate of muller's ratchet. *Genetics*, 191(4):1283–1293, 2012.

[53] Sidhartha Goyal, Daniel J Balick, Elizabeth R Jerison, Richard A Neher, Boris I Shraiman, and Michael M Desai. Dynamic mutation–selection balance as an evolutionary attractor. *Genetics*, 191(4):1309–1319, 2012.

[54] Roberto Piazza. Thermophoresis: moving particles with thermal gradients. *Soft Matter*, 4(9):1740–1744, 2008.

[55] Sydney Chapman. On the brownian displacements and thermal diffusion of grains suspended in a non-uniform fluid. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 119(781):34–54, 1928.

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

[57] Eric R Pianka. r and k selection or b and d selection? *The American Naturalist*, 106(951):581–588, 1972.

[58] Robert H MacArthur and Edward O Wilson. *The theory of island biogeography.* Princeton university press, 1967.

[59] Benjamin H Good, Stephen Martis, and Oskar Hallatschek. Adaptation limits ecological diversification and promotes ecological tinkering during the competition for substitutable resources. *Proceedings of the National Academy of Sciences*, 115(44):E10407–E10416, 2018.

[60] Todd L Parsons and Tim Rogers. Dimension reduction for stochastic dynamical systems forced onto a manifold by large drift: a constructive approach with examples from theoretical biology. *Journal of Physics A: Mathematical and Theoretical*, 50(41):415601, 2017.

[61] Thomas Nagylaki. Gustave Malécot and the transition from classical to modern population genetics. *Genetics*, 122(2):253–268, 1989.

[62] James F Crow and Motoo Kimura. *An introduction to population genetics theory.* Blackburn Press, West Caldwell, NJ, January 2009.

[63] For a historical account of the diffusion limit in population genetics, and the role of Gustave Malécot, see [61]

[64] Note that if $\omega(0) = 0$, as for logistic growth, the Malthusian fitness is proportional to T_i. Requiring $T_i = 1/(\tau\phi_i)$ is equivalent to ask the Malthusian fitness to be equal to the fitness measured in competititive experiments. See [62], ch. 1.6.