PNAS
www.pnas.org

Supplementary Information for

Adaptation limits ecological diversification and promotes ecological tinkering during the competition for substitutable resources

Benjamin H Good, Stephen Martis, Oskar Hallatschek

Benjamin H Good.
E-mail: benjamin.h.good@berkeley.edu

This PDF file includes:

Supplementary text
Figs. S1 to S4
References for SI reference citations

Supporting Information Text

1. Derivation of the model

In this section, we show how Eqs. (1) and (2) in the main text emerge from two different microscopic models. The first is a simplified class of consumer-resource models described in the main text. We also describe a second implementation that is a more direct extension of the traditional Wright-Fisher model. Motivated by results from population genetics, we then describe a special limit of Eq. (1) that is expected to describe the behavior of a much broader class of models, which differ in some of their microscopic details. To illustrate this limit, we show how it applies to a class of consumer-resource models analyzed by Refs. (1, 2).

1.1. Consumer-resource model. Our consumer-resource derivation closely follows the one described by Ref. (3), except that we now allow strains to vary in their total energy budget. In particular, we assume that all strains μ and resources i are present in a well-mixed volume V, which is diluted at rate D. In the consumer-resource framework, the per capita growth rate of each strain is mediated by the resource concentrations,

$$\partial_t n_\mu = g_\mu(\vec{c}) n_\mu - D n_\mu + \sqrt{n_\mu D} \cdot \eta_\mu(t)\,, \tag{S1}$$

where n_μ is the absolute number of individuals of strain μ, $g_\mu(\vec{c})$ is a strain-specific growth function, and $\eta_\mu(t)$ is a Brownian noise term (4) whose mean and covariance are given by

$$\langle \eta_\mu(t) \rangle = 0\,, \tag{S2}$$
$$\langle \eta_\mu(t) \eta_\nu(t') \rangle = \delta_{\mu,\nu}\delta(t-t')\,. \tag{S3}$$

The resource concentrations (in units of V^{-1}) obey a second set of equations,

$$\partial_t c_i = S_i - D c_i - \sum_\mu \frac{d_{\mu,i}(\vec{c}) n_\mu}{V} + \sqrt{\frac{c_i D}{V}} \cdot \eta_i(t)\,, \tag{S4}$$

where S_i is the input flux of resource i, $d_{\mu,i}(\vec{c})$ is the per capita depletion rate of resource i by strain μ, and $\eta_i(t)$ is an analogous set of noise terms that describe fluctuations in the resource concentrations. This general class of models has been studied previously by Refs. (5, 6), and others. Following Ref. (3), we consider a restricted subset of models where the growth and depletion functions take on a particularly simple form:

$$g_\mu(\vec{c}) = \sum_i b_{\mu,i} d_{\mu,i}(\vec{c})\,, \tag{S5}$$
$$d_{\mu,i}(\vec{c}) = r_{\mu,i}\lambda_i(\vec{c})\,. \tag{S6}$$

The first assumption states that the resources are *effectively substitutable*, i.e. biomass can be produced equally well from suitably normalized versions of any imported resource. The constant normalization factor $b_{\mu,i}^{-1}$ can be interpreted as the amount of imported resource i necessary to create one cell of strain μ. The second assumption states that the resource uptake rates can be factored into a species-and resource-specific (but concentration independent) factor $r_{\mu,i}$, and a species-independent (but resource and concentration-specific) function $\lambda_i(\vec{c})$. For example, $\lambda_i(\vec{c})$ could denote the uptake rate of a pathway that imports resource i, while $r_{\mu,i}$ denotes the constitutive expression of that pathway in an individual of strain μ. In this picture, strains can differ in their overall expression of a given pathway, but have limited ability to tune its biochemical properties. This should be a good approximation for strains that have recently descended from a common ancestor, though it may be violated for more rapidly evolving enzymes in distantly related species.

We assume that the resource fluxes and concentrations are both large, such that the dilution and noise terms can be neglected in Eq. (S4). Following Ref. (3), we also assume a separation of timescales between the dynamics of resource concentrations, such that the resource concentrations reach a quasi-equilibrium $S_i V \approx \sum_\mu d_{\mu,i}(\vec{c}) n_\mu$ before the strain abundances start to change significantly. Under these assumptions, we can eliminate the concentration variables entirely, and obtain a set of coarse-grained dynamics for the strain abundances:

$$\partial_t n_\mu = \left[-D + \sum_i \frac{S_i V b_{\mu,i} r_{\mu,i}}{\sum_\nu r_{\nu,i} n_\nu} \right] n_\mu + \sqrt{n_\mu D} \cdot \eta_\mu(t)\,. \tag{S7}$$

In this model, the dynamics of the total number of individuals, $\hat{N}(t) = \sum_\mu n_\mu(t)$, does not close, due to the μ dependence in the biomass conversion factor $b_{\mu,i}$:

$$\frac{\partial \hat{N}}{\partial t} = -D\hat{N} + \sum_i S_i V \left[\frac{\sum_\mu b_{\mu,i} r_{\mu,i} n_\mu}{\sum_\nu r_{\nu,i} n_\nu} \right] + \sqrt{\hat{N} D} \cdot \eta_\mu(t)\,. \tag{S8}$$

However, if we assume that the strains share similar biomass conversion factors $b_{\mu,i} \approx b_i$ (similar to our previous assumption that $\lambda(\vec{c})$ is independent of μ), then the equation for $\hat{N}(t)$ closes. We find that $\hat{N}(t)$ rapidly approaches a steady-state value

$N \equiv \sum_i S_i b_i V/D$ on a timescale of order $1/D$, with fluctuations of order $\sqrt{N}$. Such fluctuations become irrelevant in the large N limit, which suggests that we rewrite the dynamics in terms of the strain frequencies, $f_\mu = n_\mu / \sum_\nu n_\nu$ and measure time in units of D^{-1}. Following the derivation in Ref. (7), the dynamics of the frequencies f_μ can be shown to satisfy

$$\begin{aligned}\frac{\partial f_\mu}{\partial t} &= \left[-1 + \sum_i \frac{\beta_i \alpha_{\mu,i} e^{X_\mu}}{\sum_\nu \alpha_{\nu,i} e^{X_\nu} f_\nu}\right] f_\mu \\ &\quad + \sum_\nu \left[\delta_{\mu,\nu} - f_\mu\right] \sqrt{\frac{f_\nu}{N}} \eta_\nu(t)\,,\end{aligned} \tag{S9}$$

where we have defined the normalized parameters

$$\beta_i = \frac{S_i b_i}{\sum_j S_j b_j}\,, \tag{S10}$$

$$\alpha_{\mu,i} = \frac{r_{\mu,i}}{\sum_j r_{\mu,j}}\,, \tag{S11}$$

$$X_\mu = \log \sum_i r_{\mu,i}\,. \tag{S12}$$

The stochastic noise term $\xi_\mu(t)$ in Eq. (1) in the main text can therefore be identified with the linear combination

$$\xi_\mu(t) = \sum_\nu \left[\delta_{\mu,\nu} - f_\mu\right] \sqrt{f_\nu} \eta_\nu(t)\,, \tag{S13}$$

whose correlation structure ensures that $\sum_\mu f_\mu(t) = 1$ at all times.

1.2. Deterministic Lyapunov function. The deterministic part of Eq. (S9) possesses a Lyapunov function,

$$\Lambda(\vec{f}) = -\sum_\mu f_\mu + \sum_i \beta_i \log\left(\sum_\mu \alpha_{\mu,i} e^{X_\mu} f_\mu/\beta_i\right)\,, \tag{S14}$$

$$= -\sum_\mu f_\mu + \sum_i \beta_i \overline{X}_i\,, \tag{S15}$$

which is convex and bounded from above, and for which

$$\frac{d\Lambda}{dt} = \sum_\mu \frac{1}{f_\mu}\left(\frac{df_\mu}{dt}\right)^2 \geq 0\,. \tag{S16}$$

Among other things, this implies that the deterministic dynamics have a unique equilibrium that is approached at long times. We exploit this fact in the simulations in Appendix 5.

1.3. Subdivided environment model. The familiar form of Eq. (S9) suggests that these dynamics can also be obtained from a generalization of the standard Wright-Fisher model (8), in which the population is periodically subdivided into separate environments. In this model, the strains in environment i produce a number of gametes proportional to their Wrightian fitness, $W_{\mu,i}$. After a period of growth, $N\beta_i$ gametes are chosen from each environment and mixed together to obtain the next generation. The expected fraction of individuals in the next generation is

$$\langle f_\mu(t+\Delta t)\rangle = \sum_i \beta_i \left[\frac{W_{\mu,i} f_\mu}{\sum_\nu W_{\nu,i} f_\nu}\right]\,. \tag{S17}$$

When $\langle f_\mu(t+\Delta t) - f_\mu(t)\rangle$ is small, this update rule has the same continuum limit as Eq. (S9), with

$$X_\mu = \log\left(\sum_i W_{\mu,i}\right)\,, \tag{S18}$$

$$\alpha_{\mu,i} = \frac{W_{\mu,i}}{\sum_j W_{\mu,j}}\,. \tag{S19}$$

1.4. The near-ESS limit. Building on well-known results from population genetics, we expect that the model in Eq. (1) will attain its greatest generality in the limit that the intrinsic fitness differences X_μ and the resource-specific mean fitnesses $\overline{X}_i$ are both small compared to one (though still nonzero). This ensures a separation of timescales, in which the population- or ecosystem-level dynamics take place over times that are much longer than a single generation. Previous work has shown that the population-level dynamics in this case become insensitive to many assumptions about the underlying birth-death process (8).

The requirement that X_μ is small is familiar from the standard diffusion limit of the Wright-Fisher model (8). From the definition of the resource-specific mean fitness in Eq. (2), we see that a sufficient condition for $\overline{X}_i$ to be small is that the resource strategies $\alpha_{\mu,i}$ are close to β_i. Since we have previously identified $\alpha_{\mu,i} = \beta_i$ as a marginal evolutionarily stable state (ESS), we have termed this regime the *near-ESS limit*. Alternatively, it can be viewed as a generalization of the standard diffusion limit of population genetics.

We can access the near-ESS limit of Eq. (1) in several different ways. We can either work with the full model and take the near-ESS limit in the end, or else we can work directly with the near-ESS limit of Eq. (1). We have employed for the former strategy for most of this work. However, in certain cases, it can be more convenient take the near-ESS limit of Eq. (1) as our microscopic model (just as it is convenient to work directly with the Wright-Fisher diffusion process).

To obtain the near-ESS limit of Eq. (1), we first rewrite the resource uptake strategies in the form

$$\alpha_{\mu,i} = \beta_i \left(1 + \gamma_{\mu,i}\right) , \tag{S20}$$

for some rescaled vector $\gamma_{\mu,i}$. The normalization conditions for $\alpha_{\mu,i}$ and β_i yield a corresponding condition for $\gamma_{\mu,i}$,

$$\sum_i \beta_i \gamma_{\mu,i} = 0 . \tag{S21}$$

If we substitute Eq. (S20) into Eqs. (1) and (2), and expand to lowest order in X_μ, $\gamma_{\mu,i}$, and $1/N$, we obtain

$$\begin{aligned} \frac{\partial f_\mu}{\partial t} = &\left[X_\mu - \sum_\nu X_\nu f_\nu\right] f_\mu + \frac{\xi_\mu}{\sqrt{N}} \\ &- \sum_{i,\nu} \beta_i \gamma_{\mu,i} \gamma_{\nu,i} f_\nu f_\mu \\ &+ \left(\sum_{i,\nu,\sigma} \beta_i \gamma_{\nu,i} \gamma_{\sigma,i} f_\nu f_\sigma\right) f_\mu . \end{aligned} \tag{S22}$$

The first two terms coincide with the diffusion limit of the Wright-Fisher model, as expected, while the ecological interactions enter at $\mathcal{O}(\gamma^2)$ in the third term. These interactions take the form of a symmetric Lotka-Volterra model with a special interaction matrix formed by the outer product of the resource strategy vectors,

$$A_{\mu,\nu} = \sum_i \beta_i \gamma_{\mu,i} \gamma_{\nu,i} . \tag{S23}$$

The rank of this matrix is at most $\mathcal{R}$, regardless of the number of strains. The final term is analogous to the mean fitness term in the Wright-Fisher model, and ensures that $\sum_\mu f_\mu = 1$ at all times. Although this three-body term is formally outside of the Lotka-Volterra model, it is often small in practice, and can be neglected in many calculations.

1.5. Fitness differences arising from variable death rates. To illustrate the generality of the near-ESS limit, we show how it can apply to a separate class of consumer-resource models that have been studied in the literature (1, 2), in which the fitness differences arise from differences in the underlying death rate. In other words, we assume that the dilution rate D in Eq. (S1) can now vary between strains:

$$D \to D + m_\mu . \tag{S24}$$

Although this model is formally different than the one we consider in Eq. (1), it produces the same limiting behavior as Eq. (S22) when the fitness differences are small on the timescale of a single generation, i.e., when $m_\mu \ll D$. At lowest order, variation in the death rate will generate effective fitness differences of the form

$$X_\mu \to X_\mu - \frac{m_\mu}{D} . \tag{S25}$$

Even when m_μ/D is small, we have shown that these fitness differences can produce dramatic effects when integrated over many generations. Yet at lowest order, the effect of death rate variation is indistinguishable from the variation in growth rate that we have considered above.

Of course, the limiting behavior in Eq. (S22) will cease to apply if the fitness differences between strains are no longer small. In this regime, however, the dynamics will often depend on aspects of the birth-death process that are not captured by toy models like Eq. (S1) [e.g., cell-to-cell variation in growth rate (9), phenotypic delays in mutation penetrance (10), how genetic drift is implemented, etc.]. These would need to be carefully chosen to match the biological system of interest. To ensure the greatest generality, we therefore focus on the aspects of the model that can be captured by Eq. (S22). Enumeration of other universality classes is an interesting topic for future work.

2. Competition for two resources

In this section, we derive our main results for the two resource case. The major advantage of this limit is that the multidimensional resource space reduces to the scalar interval $(0, 1)$. Without loss of generality, we will write everything in terms of the first resource component, defining $\beta = \beta_1$ and $\alpha_{\mu,1} = \alpha_\mu$, with the remaining components $\beta_2 = 1 - \beta$ and $\alpha_{\mu,2} = 1 - \alpha_\mu$ fixed by the normalization condition. Following the description in the main text, we will begin by analyzing the competition between two strains, and then consider the effects of adding a third strain to a pair of previously coexisting strains.

2.1. Competition between two strains. To analyze the competition between two strains, we let α_1 and α_2 denote the strategy vectors of the two strains, and let $\Delta X = X_2 - X_1$ denote the fitness difference between them. We arbitrarily designate strain 1 as the "wildtype" and consider the frequency of the "mutant strain", $f \equiv f_2$. With these definitions, the stochastic term in Eq. (S13) can be written as

$$\begin{aligned}\xi_2(t) &= \sqrt{\frac{(1-f)^2 f}{N}}\eta_2(t) - \sqrt{\frac{f^2(1-f)}{N}}\eta_1(t)\,,\\ &= \sqrt{\frac{f(1-f)}{N}}\eta(t)\,,\end{aligned} \tag{S26}$$

where $\eta(t)$ is a third Brownian noise term with $\langle\eta(t)\rangle = 0$ and $\langle\eta(t)\eta(t')\rangle = \delta(t-t')$. Eq. (1) can then be rewritten in the familiar population genetic form (8),

$$\frac{\partial f}{\partial t} = s_e(f) f(1-f) + \sqrt{\frac{f(1-f)}{N}}\eta(t)\,, \tag{S27}$$

where the effective frequency-dependent selection coefficient, $s_e(f)$, is obtained from the deterministic portion of Eq. (1),

$$\begin{aligned}s_e(f) &\equiv \frac{1}{f(1-f)}\left(\frac{\partial f_2}{\partial t}\right)_{\text{deterministic}}\,,\\ &= \frac{\beta\left[\alpha_2 e^{\Delta x} - \alpha_1\right]}{\alpha_1 + \left[\alpha_2 e^{\Delta x} - \alpha_1\right] f}\\ &\quad + \frac{(1-\beta)\left[(1-\alpha_2)e^{\Delta X} - (1-\alpha_1)\right]}{1-\alpha_1 + \left[(1-\alpha_2)e^{\Delta X} - (1-\alpha_1)\right] f}\,.\end{aligned} \tag{S28}$$

Our main results can be derived from limiting versions of this basic model.

Invasion of a new strain. The invasion of a new strain corresponds to the $f \to 0$ limit, in which Eq. (S27) reduces to the linearized form,

$$\frac{\partial f}{\partial t} = S_{\text{inv}} f + \sqrt{\frac{f}{N}}\eta(t)\,, \tag{S29}$$

with an invasion fitness S_{inv} defined by

$$\begin{aligned}S_{\text{inv}} &\equiv \lim_{f\to 0} s_e(f)\,,\\ &= \left(e^{\Delta X} - 1\right) + e^{\Delta X}\left[\frac{(\beta-\alpha_1)(\alpha_2-\alpha_1)}{\alpha_1(1-\alpha_1)}\right]\,.\end{aligned} \tag{S30}$$

Eq. (S29) can be solved using standard methods (11). We will simply quote the relevant results here, while a more pedagogical exposition can be found in Chapter 1 of Ref. (12).

For initial frequencies small compared to the $1/NS_{\text{inv}}$, the genetic drift term dominates, and there is a high probability that the mutant will drift to extinction. However, with probability $p_{\text{est}} = 2NS_{\text{inv}}f(0)$, the mutant will drift to frequency $\sim 1/NS_{\text{inv}}$, after which point the selection term dominates over genetic drift. This "established" lineage will then grow deterministically as $f(t) = \frac{1}{2NS_{\text{inv}}}e^{S_{\text{inv}}t}$, which can be matched onto the full nonlinear (but deterministic) solution as f increases further. The full solution is somewhat unwieldy, but the first-order nature of the ODE shows that $f(t)$ cannot decrease as $t \to \infty$. Thus, once the mutant establishes, the deterministic dynamics will never drive the mutant close enough to the drift barrier that extinction becomes likely again. This suggests that the branching process description will be valid as long as $f(t)$ remains sufficiently small during the duration of the establishment process that $f(t) \ll 1$ and $s_e(f) \approx S_{\text{inv}}$. This will be true provided that these conditions are satisfied at the drift barrier, $1/NS_{\text{inv}}$, which leads to the conditions

$$NS_{\text{inv}} \gg 1\,, \tag{S31}$$

$$\frac{NS_{\text{inv}}\alpha_1}{\alpha_2 e^{\Delta X} - \alpha_1} \gg 1\,, \tag{S32}$$

$$\frac{NS_{\text{inv}}(1-\alpha_1)}{(1-\alpha_2)e^{\Delta X} - (1-\alpha_1)} \gg 1\,. \tag{S33}$$

These conditions can be satisfied simultaneously for sufficiently large N.

Stable coexistence. If $S_{\rm inv} > 0$ and the mutant is lucky enough to establish, then the frequency-dependent selection term will either drive the mutant to fixation ($f = 1$) or else stabilize at some intermediate frequency f^*. As described in the main text, stable coexistence requires that the reciprocal invasion fitness,

$$\begin{aligned} S^R_{\rm inv} &\equiv \lim_{f\to 1} -s_e(f)\,, \\ &= \left(e^{-\Delta X} - 1\right) + e^{-\Delta X}\left[\frac{(\beta-\alpha_2)(\alpha_1-\alpha_2)}{\alpha_2(1-\alpha_2)}\right], \end{aligned} \tag{S34}$$

is also positive. Solving this equation when $S^R_{\rm inv} = 0$ yields the critical fitness threshold

$$\Delta X_{\rm max} = \log\left(1 + \frac{(\alpha_1-\alpha_2)(\beta-\alpha_2)}{\alpha_2(1-\alpha_2)}\right), \tag{S35}$$

which reduces to Eq. (9) in the main text in the near-ESS limit. We might naively assume that this threshold would be equivalent to the fitness that gives strain 2 a higher uptake rate on *both* individual resources, i.e. $\alpha_2 e^{\Delta X} \geq \alpha_1$ and $(1-\alpha_2)e^{\Delta X} \geq 1-\alpha_1$. Although this is indeed a sufficient condition for strain 2 to fix, the true thresholds in Eqs. (8) and (9) are much weaker conditions, which depend on the environmental supply vector β. This means that in practice, stable coexistence will be disrupted long before one of the strains is uniformly better than the other.

When the conditions for stable coexistence are met, the equilibrium frequency f^* is obtained from the condition that $s_e(f^*) = 0$. From Eq. (S28), we see that this can only happen if $\alpha_2 e^{\Delta X} - \alpha_1$ and $(1-\alpha_2)e^{\Delta X} - (1-\alpha_1)$ have different signs, i.e. neither strain is uniformly better than the other. Solving for f^*, we find that

$$f^* = \frac{f_0^* + \left[f_0^* + \frac{\alpha_1(1-\alpha_1)}{\Delta\alpha^2}\right]\left(e^{\Delta X}-1\right)}{\left[1 + \frac{\alpha_2}{\Delta\alpha}\left(e^{\Delta X}-1\right)\right]\left[1 - \frac{(1-\alpha_2)}{\Delta\alpha}\left(e^{\Delta X}-1\right)\right]}, \tag{S36}$$

where $f_0^* = (\beta-\alpha_1)/\Delta\alpha$ is the equilibrium frequency in the absence of any fitness differences. This reduces to Eq. (10) in the main text in the limit that $\Delta X \to 0$ and $\Delta\alpha \to 0$.

When $f = f^*$, the resource-specific mean fitnesses $\overline{X}_i$ take on the values

$$\begin{aligned} \overline{X}_1 &= -\log\left[1 - \left(1-e^{-\Delta X}\right)\left(\frac{1-\alpha_2}{\Delta\alpha}\right)\right], \\ \overline{X}_2 &= -\log\left[1 + \left(1-e^{-\Delta X}\right)\left(\frac{\alpha_2}{\Delta\alpha}\right)\right], \end{aligned} \tag{S37}$$

which are independent of the resource supply vector β. This extends the "environmental shielding" behavior derived in the neutral limit by Ref. (3): when two strains coexist on two substitutable resources, the strain frequencies evolve so that the remaining selection pressures take on values that are independent of the environment, and depend only on the identities of the coexisting strains. We will revisit this behavior again in the multi-resource case below.

In the limit that fitness differences are small [specifically, when ΔX is small compared to 1, $\alpha_2/\Delta\alpha$, and $(1-\alpha_2)/\Delta\alpha$], Eq. (S37) reduces to the linearized version,

$$\overline{X}_1 = \frac{(1-\alpha_2)}{\Delta\alpha}\Delta X\,, \quad \overline{X}_2 = -\frac{\alpha_2}{\Delta\alpha}\Delta X\,, \tag{S38}$$

while Eq. (S36) reduces to the linear relation quoted in Eq. (10) in the main text. This defines a second fitness scale,

$$\begin{aligned} X_f &\equiv f^*(1-f^*)\left(\frac{\partial f^*}{\partial \Delta X}\right)^{-1}, \\ &= \frac{(\beta-\alpha_1)(\alpha_2-\beta)}{\beta(1-\beta) + (\beta-\alpha_1)(\alpha_2-\beta)}\,, \end{aligned} \tag{S39}$$

over which $f^*(\Delta X)$ changes significantly. Note that X_f has approximately the same scaling behavior for small and large $\beta-\alpha$ as the critical threshold $\Delta X_{\rm max}$ in Eq. (9).

For frequencies close to f^*, the selection term again grows small compared to the genetic drift term. Linearizing Eq. (S27) around $f \approx f^*$, the fluctuations $\delta f = f - f^*$ are described by

$$\frac{\partial \delta f}{\partial t} = -X_{\rm eq} f^*(1-f^*)\delta f + \sqrt{\frac{f^*(1-f^*)}{N}}\eta(t)\,, \tag{S40}$$

where we have defined the equilibrium restoring fitness

$$X_{\rm eq} \equiv -\left.\frac{\partial s_e(f)}{\partial f}\right|_{f=f^*} = \frac{1}{\beta(1-\beta)}\left(\frac{(\alpha_2 e^{\Delta X}-\alpha_1)[(1-\alpha_2)e^{\Delta X} - (1-\alpha_1)]}{\alpha_1[(1-\alpha_2)e^{\Delta X} - (1-\alpha_1)] - (1-\alpha_1)[\alpha_2 e^{\Delta X} - \alpha_1]}\right)^2. \tag{S41}$$

In the limit that $\Delta X \ll 1$, this becomes

$$X_{\rm eq} = \frac{\Delta\alpha^2}{\beta(1-\beta)}\left[1 + \frac{2\alpha_2 - 1}{\Delta\alpha}\Delta X\right]. \quad [S42]$$

This model can again be solved using standard methods (13). The stationary distribution of δf tends toward a normal distribution with mean zero and standard deviation $\sigma_f = 1/\sqrt{2NX_{\rm eq}}$, which decays on a timescale $\sim 1/X_{\rm eq} f^*(1-f^*)$. The quasi-deterministic model is therefore self-consistent provided that

$$\frac{\sigma_f}{f^*(1-f^*)} = \sqrt{\frac{\Delta\alpha^2}{2N(\beta-\alpha_1)^2(\alpha_2-\beta)^2}} \ll 1, \quad [S43]$$

which can be satisfied for sufficiently large N.

The fluctuations in f lead to similar fluctuations in the resource-specific mean fitnesses, $\overline{X}_i$, whose first order contribution is given by

$$\begin{aligned}
\delta\overline{X}_1 &= \frac{\Delta\alpha}{\beta}\left[1 + \frac{\alpha_2}{\Delta\alpha}(e^{\Delta X} - 1)\right] e^{-\overline{X}_1}\delta f, \\
\delta\overline{X}_2 &= -\frac{\Delta\alpha}{1-\beta}\left[1 - \frac{(1-\alpha_2)}{\Delta\alpha}\left(e^{\Delta X} - 1\right)\right] e^{-\overline{X}_2}\delta f.
\end{aligned} \quad [S44]$$

2.2. Competition between three strains. Having characterized the dynamics for a pair of strains, we next consider a scenario in which a third strain is introduced into a stable ecosystem where a pair of strains already coexist. Without loss of generality, we will assume that the third strain is a mutant version of the second strain, with fitness $X_3 = \Delta X + s$ and strategy vector α_3. From the definition of the model in Eq. (1), this mutant strain will have an invasion fitness

$$S_{\rm inv} = \alpha_3 \left(e^{\Delta X + s - \overline{X}_1} - 1 \right) + (1-\alpha_3) \left(e^{\Delta X + s - \overline{X}_2} \right) \tag{S45}$$

where the resource-specific mean fitnesses $\overline{X}_1(t)$ and $\overline{X}_2(t)$ are dictated by the two strain process in Eq. (S27). If the mutant was actually identical to its parent strain (i.e., if $\alpha_3 = \alpha_2$ and $s = 0$), it should never be favored to invade, since it can at best compete neutrally with its parent. This implies that

$$\alpha_2 \left(e^{\Delta X - \overline{X}_1} - 1 \right) + (1-\alpha_2) \left(e^{\Delta X - \overline{X}_2} - 1 \right) = 0 \,, \tag{S46}$$

when averaged over the timescales required for the mutation to invade. Multiplying this expression by e^s and subtracting it from Eq. (S45), we can then rewrite the general invasion fitness in the form

$$S_{\rm inv} = (e^s - 1) + (\alpha_3 - \alpha_2) \left(e^{-\overline{X}_1(t)} - e^{-\overline{X}_2(t)} \right) e^{\Delta X + s} \,, \tag{S47}$$

We consider the implications of this expression in various special cases below.

No fitness differences. In a completely neutral scenario ($\Delta X = s = 0$), the resource-specific mean fitnesses are solely determined by the fluctuations $\delta\overline{X}_1$ and $\delta\overline{X}_2$ from Eq. (S44), and Eq. (S47) reduces to

$$S_{\rm inv} = (\alpha_3 - \alpha_2) \left[\delta\overline{X}_2 - \delta\overline{X}_1 \right] = \frac{(\alpha_2 - \alpha_3)(\alpha_2 - \alpha_1)}{\beta(1-\beta)} \delta f(t) \,. \tag{S48}$$

Since $\langle \delta f(t) \rangle = 0$, this agrees with the deterministic results of Ref. (3), who found that all further invasion fitnesses vanish in a neutral population when the ecosystem is fully expoited. However, our stochastic analysis shows that fluctuations can induce momentary selection pressures of order

$$\delta S_{\rm inv} \sim \frac{(\alpha_2 - \alpha_3)(\alpha_2 - \alpha_1)}{\beta(1-\beta)} \frac{1}{\sqrt{N X_{\rm eq}}} \,, \tag{S49}$$

which can be large compared to $1/N$. However, these momentary selection pressures average out to zero over a timescale $1/X_{\rm eq} f^*(1-f^*)$. When N is large, this is much shorter than the timescale $\sim 1/\delta S_{\rm inv}$ required for the mutant lineage to escape the drift barrier. This shows that internal fluctuations cannot induce anomalous establishment events in our model. To leading order in N, ecological selection pressures vanish in a neutral population when two strains coexist on two substitutable resources.

Pure fitness mutations. In the case where the mutant lineage is created by a pure fitness mutation, $\alpha_3 = \alpha_2$, and the invasion fitness reduces to

$$S_{\rm inv} = e^s - 1 \approx s \,, \tag{S50}$$

which is identical to the standard Wright-Fisher model. This justifies our interpretation of X_μ as a fitness parameter. Eq. (S50) is a slightly stronger result, since it implies that pure fitness mutations continue to establish at the same rate, regardless of the structure of the ecosystem. When such a mutation establishes, it is guaranteed to displace its parent strain, resulting in a two-strain competition between strain 3 and strain 1, which now differ in fitness by an amount $\Delta X + s$. If $\Delta X + s \geq \Delta X_{\rm max}$ from Eq. (9), then stable coexistence will be disrupted, and strain 3 will take over the entire population. On the other hand, if $\Delta X + s < \Delta X_{\rm max}$, the mutant will only displace its parent strain, and will be prevented from sweeping through the entire population. Instead, the successful mutation will shift the equilibrium frequency by an amount

$$\begin{aligned} \Delta f &= f^*(\Delta X + s) - f^*(\Delta X) \,, \\ &\approx \frac{\beta(1-\beta) + (\beta - \alpha_1)(\beta - \alpha_2)}{\Delta\alpha^2} \cdot s \,, \end{aligned} \tag{S51}$$

where we have employed the linearized approximation for f^* from Eq. (10) in the main text.

Pure strategy mutations. If the mutant lineage is created by a pure strategy mutation ($s = 0$), then the invasion fitness reduces to

$$S_{\rm inv} = \frac{\alpha_3 - \alpha_2}{\alpha_1 - \alpha_2} \left(e^{\Delta X} - 1 \right) \,, \tag{S52}$$

where we have retained only the leading order contribution as $N \to \infty$. The $\Delta X \ll 1$ limit is listed in Eq. (15) in the main text. The interpretation of this expression, and the various scenarios that can arise after establishment, are described in the main text as well.

2.3. Evolution of a single-locus ecology. The results above allow us to analyze the effects of further evolution in our consumer resource model. As a first pass, we focus on a simplified scenario, in which strategy mutations switch between two fixed strategy vectors, α_1 and α_2, and occur at rate U_α. We assume that α_1 and α_2 span β, so that the strains can stably coexist. We also assume that α_1 and α_2 are sufficiently close to β that we can invoke the near-ESS limits of various expressions above. We note that while this assumption is also employed in the adaptive dynamics literature (14, 15), our model also differs from these results in a key way, as it includes α that go beyond the infinitesimal evolution assumption in adaptive dynamics.

Our model also differs from the canonical adaptive dynamics scenario in that it includes pure fitness mutations, which occur at rate $U_X \rho_X(s)$. We assume that the tails of $\rho_X(s)$ are sufficiently light that the distribution can be approximated by a characteristic beneficial fitness effect (16), which we will also denote by the generic variable s below. Our analysis here will focus on the strong-selection weak mutation (SSWM) regime that arises in the limit that $N \to \infty$ and $U_\alpha + U_X \to 0$. The first assumption guarantees that genetic drift is only relevant when mutations are sufficiently rare, so that the establishment process can be modeled by the branching process techniques above. The second assumption guarantees that all mutations establish or go extinct before the next mutation occurs, so that they can be described by the two- and three-strain competition processes above. Violations of this assumption are considered in more detail in a following section.

No strategy mutations. We first consider the dynamics under pure fitness mutations when $U_\alpha/U_X = 0$. We assume that the population has just diversified into a pair of coexisting strains with fitness difference $\Delta X = 0$, and equilibrium frequency f_0^*. Pure fitness mutations will occur in each clade at rate $NU_X f^*$ and $NU_X(1-f^*)$, respectively. According to Eq. (S50), these establish with probability $p_{\rm est} = 2s$, sweep through their parent clade, and result in a new fitness differential,

$$\Delta X = \begin{cases} +s & \text{with probability } f^*, \\ -s & \text{with probability } 1-f^*, \end{cases} \tag{S53}$$

which depends on the genetic background in which the mutation arose. This fitness differential will lead to a shift in the equilibrium frequency $\Delta f = \pm s/s_c$ described by Eq. (11) in the main text. If $\Delta f < -f^*$ or $\Delta f > 1 - f^*$, then by definition stable coexistence will be disrupted, since the new frequency would fall outside the interval from 0 to 1. [One can also see this directly from the fitness bounds $\Delta X_{\rm min}$ and $\Delta X_{\rm max}$ in Eqs. (8) and (9) in the main text.] Since the mutant has already swept through its parent clade, a disruption of coexistence implies that it will take over the entire population. In the near-ESS limit, we can combine these two conditions to obtain a convenient asymptotic condition for s:

$$s \gg s_c f_0^*(1-f_0^*) \equiv \frac{(\alpha_1-\beta)(\beta-\alpha_2)}{\beta(1-\beta)}\,. \tag{S54}$$

In this regime, the lifetime of coexistence is of order $\tau_{\rm collapse} \sim 1/NU_X s$ (the time that it takes for one fitness mutation to occur).

In the opposite regime, when $s \ll s_c f_0^*(1-f_0^*)$, individual fitness mutations lead to small shifts in f^*, and many such mutations must accumulate before stable coexistence is disrupted. In this case, we can model the changing equilibrium frequency using an effective diffusion process. In an interval of time δt, the fitness differential changes by $\delta\Delta X = s(k_2 - k_1)$, where k_2 and k_1 denote the number of fitness mutations that accumulate in the f^* and $1-f^*$ backgrounds, respectively. In the weak mutation limit, these occur as a Poisson process with rates $2NU_X s f^* \delta t$ and $2NU_X s(1-f^*)\delta t$, respectively, so that

$$\langle k_2 - k_1 \rangle = 2NU_X s(2f^* - 1)\delta t\,, \tag{S55}$$

$$\mathrm{Var}(k_2 - k_1) = 2NU_X s \delta t\,. \tag{S56}$$

The fitness difference ΔX can therefore be described by an effective diffusion process,

$$\frac{\partial \Delta X}{\partial t} = 2NU_X s^2[2f^*(\Delta X) - 1] + \sqrt{2NU_X s^3}\eta(t)\,, \tag{S57}$$

where the equilibrium frequency f^* itself depends on ΔX through Eq. (10) in the main text. Changing variables from ΔX to f^*, we obtain Eq. (12) in the main text. For our detailed calculations, it will be somewhat more convenient to work with the rescaled variables $Y = 2f^* - 1$ and $k = 2NU_X st$, which yields a related equation

$$\frac{\partial Y}{\partial k} = \frac{2s}{s_c} Y + \sqrt{\frac{4s^2}{s_c^2}} \cdot \eta(\tau)\,. \tag{S58}$$

This is similar to the equation for the drift-induced fluctuations in Eq. (S40), except that the bias is now a destabilizing force, rather than a stabilizing one. This reflects the fact that larger clades are more likely to acquire beneficial mutations in the weak mutation limit, which leads to further increases in frequency. We can quantify the strength of this snowballing effect by analyzing the ultimate fixation probability of strain 1 (i.e., the probability that $Y \to 1$) as a function of the current value of $Y = 2f^* - 1$. Eq. (S58) implies a corresponding backward equation for the fixation probability

$$\frac{2s}{s_c} Y \frac{\partial P_{\rm fix}}{\partial Y} + \frac{2s^2}{s_c^2} \frac{\partial P_{\rm fix}}{\partial Y} = 0\,, \tag{S59}$$

whose solution is given by

$$P_{\rm fix}(f^*) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{2f^*-1}{\sqrt{s/s_c}}} e^{-\frac{u^2}{2}}\, du\,. \tag{S60}$$

This function undergoes a sharp transition near $2f^* - 1 \sim \sqrt{s/s_c}$. When $|2f^* - 1| \ll \sqrt{s/s_c}$, fixation and extinction of the clade are equally likely, while for $2f^* - 1 \gg \sqrt{s/s_c}$, fixation is virtually guaranteed. This transition has a simple interpretation in terms of the relative strengths of the bias and noise terms in Eq. (12): $\sqrt{s/s_c}$ represents a critical frequency difference above which the bias dominates over the noise term. Since $\sqrt{s/s_c}$ is itself a small parameter in the $s \ll s_c$ regime, this implies that the random portion of the clade competition process is confined to frequencies near 50%. Reversals from frequencies near $f^* \approx 0$ or $f^* \approx 1$ are asymptotically unlikely.

To investigate the dynamics of this process, we analyze the mean squared frequency difference $\langle Y^2 \rangle$. Using Eq. (S58), we can derive a closed moment equation for $\langle Y^2 \rangle$

$$\frac{\partial \langle Y^2 \rangle}{\partial k} = \frac{4s}{s_c}\langle Y^2 \rangle + \frac{4s^2}{s_c^2}\,, \tag{S61}$$

whose solution is given by

$$\langle Y(k)^2 \rangle = Y(0)^2 e^{\frac{4sk}{s_c}} + \frac{s}{s_c}\left(e^{\frac{4sk}{s_c}} - 1\right)\,. \tag{S62}$$

Solving for k and converting back to units of time, we find that

$$t = \frac{s_c}{8NU_X s^2} \log\left(\frac{\langle Y(t)^2 \rangle + s/s_c}{Y(0)^2 + s/s_c}\right)\,. \tag{S63}$$

The behavior of this function has a simple heuristic interpretation based on the fundamental timescales of Eq. (12). These heuristics follow from standard arguments (17–19), so we will simply quote the relevant results here, while referring the interested reader to Chapter 1 of Ref. (12) for a more detailed exposition.

Starting from $|2f_0^* - 1| \ll \sqrt{s/s_c}$, the clade frequencies will wander diffusively for a time $\tau_{\rm drift} \sim \frac{s_c}{NU_X s^2}$ until the frequency difference reaches $\sqrt{s/s_c}$, after which point the major clade deterministically acquires mutations for $\tau_{\rm collapse} \sim \frac{s_c}{NU_X s^2} \log(s_c/s)$ more generations until it reaches fixation. On the other hand, if the clades start with a frequency difference $|2f_0 - 1| \gg \sqrt{s/s_c}$, then the major clade will deterministically fix within $\sim \frac{s_c}{NU_X s^2} \log\left(\frac{1}{|2f^*-1|}\right)$ generations

Including strategy mutations. We can use the results above to analyze the case where $U_\alpha/U_X > 0$. For very low values of U_α/U_X, strategy mutations will rarely occur before the ecosystem collapses according to the process described above. In this case, the main effect of strategy mutations is to re-diversify a population that consists of a single ecotype. The invasion fitness of such a mutation is therefore given by Eq. (5) in the main text, and will vary depending on which ecotype dominates the population.

We can therefore distinguish between two regimes. If $|2f_0^* - 1| \ll \sqrt{s/s_c}$, then both ecotypes are equally likely to fix, and the average invasion fitness is

$$\overline{X}_{\rm inv} = \frac{(\beta - \alpha_1)(\alpha_2 - \alpha_1)}{\beta(1-\beta)} + \frac{(\beta - \alpha_2)(\alpha_1 - \alpha_2)}{\beta(1-\beta)} = s_c\,, \tag{S64}$$

This leads to a diversification timescale $\tau_{\rm diversify} \sim 1/NU_\alpha s_c$, and the diversification-selection balance in Eq. (13) in the main text. On the other hand, if $|2f_0^* - 1| \gg \sqrt{s/s_c}$ then the clade with the larger initial frequency will typically be the one that fixes. Without loss of generality, we will relabel the strains so that f^* always represents this clade. In this scenario, the average invasion fitness is instead given by $\overline{X}_{\rm inv} \sim s_c(1 - f_0^*)$, which is strictly smaller than s_c. In this case, the diversification-selection balance is given by

$$\frac{\Pr[\mathcal{S}=2]}{\Pr[\mathcal{S}=1]} \approx \frac{\tau_{\rm collapse}}{\tau_{\rm diversify}} \sim \frac{U_\alpha}{U_X}\left(\frac{s_c}{s}\right)^2 (1 - f_0^*) \log\left(\frac{1}{2f_0^* - 1}\right)\,. \tag{S65}$$

For still larger values of U_α/U_X, strategy mutations will start to occur before one of the clades has fixed in the population. If the mutation occurs in the fitter clade, it will have an invasion fitness $S_{\rm inv} = |\Delta X|$, and will reset the fitness difference to zero if it establishes. On the other hand, if the mutation occurs in the less fit clade, it will have a negative invasion fitness and will not be able to establish. Thus, the net effect of these strategy mutations is to set $\Delta X = 0$ at a time-dependent rate

$$\lambda_0(t) = 2NU_\alpha f^* \Delta X \theta(\Delta X) - 2NU_\alpha(1 - f^*)\Delta X \theta(-\Delta X)\,, \tag{S66}$$

where $\theta(z)$ is the Heaviside step function, and $f^*(t)$ and $\Delta X(t)$ are determined by the effective diffusion process in Eq. (12) in the main text. The first successful strategy mutation will occur on a characteristic invasion timescale determined by the implicit relation

$$\int_0^{\tau_{\rm invade}} \lambda(t) dt \sim 1\,. \tag{S67}$$

Since the fitter strain will typically be the most abundant as well, Eq. (S66) will only differ by a factor of two from the much simpler expression

$$\lambda_0(t) \sim N U_\alpha |\Delta X|\,. \tag{S68}$$

Since Eq. (S67) is only accurate up to an order one factor, we will use this simpler approximation for $\lambda_0(t)$ instead.

Based on these definitions, we can obtain a self-consistent solution to Eq. (S67) in various regimes. If $U_\alpha \gg U_X$, then strategy mutations will arise much faster than individual mutations. In this case, a lucky fitness mutation will establish in one of the clades after a time of order $1/NU_X s$, so that

$$\int_0^t \lambda_0(t') dt' \sim N U_\alpha s \left(t - \frac{1}{N U_X s}\right)\,. \tag{S69}$$

This yields an invasion timescale

$$\tau_{\rm invade} \sim \frac{1}{N U_\alpha s} + \frac{1}{N U_X s} \sim \frac{1}{N U_X s}\,, \tag{S70}$$

which is self-consistent provided that $U_\alpha \gg U_X$.

If $\tau_{\rm invade} \gg 1/NU_X s$, then multiple fitness mutations will accumulate before the first successful strategy mutation arises. If $\tau_{\rm invade} \ll \tau_{\rm drift}$ then the fitness differential ΔX wanders diffusively as $|\Delta X| \sim \sqrt{NU_X s^3 t}$, and

$$\int_0^t \lambda(t') dt' \sim N U_\alpha \sqrt{N U_X} (st)^{3/2}\,. \tag{S71}$$

This leads to an invasion timescale of order

$$\tau_{\rm invade} \sim \frac{1}{N U_X s} \left(\frac{U_X}{U_\alpha}\right)^{2/3}\,, \tag{S72}$$

which is self consistent provided that $U_\alpha \ll U_X \ll U_\alpha \left(s_c/s\right)^{3/2}$.

If $\tau_{\rm invade} \gg \tau_{\rm drift}$, or if the initial frequency differential already exceeds the critical value $\sqrt{s/s_c}$, then the successful strategy mutation will occur when ΔX is growing deterministically as

$$\Delta X \sim s_c \sqrt{(2f_0^* - 1)^2 + \frac{s}{s_c}} \cdot e^{\frac{4NU_X s^2 t}{s_c}}\,, \tag{S73}$$

so that

$$\int_0^t \lambda_0(t') dt' \sim \frac{U_\alpha s_c^2}{U_X s^2} \sqrt{(2f_0^* - 1)^2 + \frac{s}{s_c}} \cdot e^{\frac{4NU_X s^2 t}{s_c}}\,. \tag{S74}$$

If $\tau_{\rm invade} \ll \tau_{\rm collapse}$, this leads to an invasion timescale,

$$\tau_{\rm invade} \sim \frac{s_c}{8NU_X s^2} \log\left(\frac{U_X^2}{U_\alpha^2} \frac{s^4}{s_c^4} \frac{1}{(2f_0^* - 1)^2 + \frac{s}{s_c}}\right)\,, \tag{S75}$$

which will be self-consistent provided that $U_\alpha \left(\frac{s_c}{s}\right)^{3/2} \ll U_X \ll U_\alpha \left(\frac{s_c}{s}\right)^2$. Finally, for $U_X \gg U_\alpha \left(\frac{s_c}{s}\right)^2$, strategy mutants are sufficiently rare that the ecosystem will typically collapse and re-diversify before invasion can occur. In this case, $\tau_{\rm invade}$ formally diverges. The various regimes for $\tau_{\rm invade}$ are summarized in Eq. (14) in the main text.

The effects of clonal interference. In our analysis above, we have focused on the weak mutation limit, in which only two or three strains exist within the population at any one time. While this enabled many analytical simplifications, it is also known that many microbial populations lie outside this regime. This is particularly true for many microbial evolution experiments in which stable coexistence has been observed to evolve spontaneously. While a thorough analysis of this regime is beyond the scope of the present work, we will summarize the key differences that are likely to arise in the effective diffusion process in Eq. (12) in the main text.

Outside of the weak-mutation limit, many established beneficial mutations will be driven to extinction due to clonal interference with other beneficial mutations that happen to segregate at the same time (20). In the limit that clonal interference is strong ($NU_X \gg 1$), this has two main consequences. First, the rate of adaptive substitution scales much more weakly with N than the linear expectation $NU_X s$ from the SSWM limit. In the case of coexisting strains, this will also apply to the subpopulations Nf^* and $N(1-f^*)$ that correspond to the two clades. As a result, the bias term in Eq. (12) will be significantly reduced (and essentially vanishes in the limit of strong clonal interference). Second, clonal interference causes the rate of adaptation to become more deterministic in addition to reducing it, since it is no longer limited by the supply of beneficial mutations. The dynamics of these fluctuations are poorly understood in the general case, though Ref. (21) has shown that they lead to a long-term diffusion constant,

$$D_X = \left[\frac{s}{\log\left(\frac{s}{U_X}\right)}\right]^3 . \qquad \text{[S76]}$$

for the total fitness gain in a model similar to ours. Thus, as long as $Nf^*(1-f^*)$ remains sufficiently large that clonal interference within each clade remains strong, we expect the effective diffusion model in Eq. (12) to be better approximated by the limiting form

$$s_c \frac{\partial f}{\partial t} \sim \sqrt{D_X} \cdot \eta(t) . \qquad \text{[S77]}$$

Due to the weaker bias term, we expect that the relative frequencies of the clades can undergo dramatic reversals before one or the other accumulates a fitness advantage that is large enough it to fix. Interestingly, such reversals have been observed in a long-term experiment in *E. coli* (22). However, a more thorough analysis of this clonal interference regime remains an interesting avenue for future work.

3. Competition for many resources

In this section, we show how many of the results derived in the two-resource case can be extended to systems with larger numbers of resources. Most of these results will apply for arbitrary values of $\mathcal{R}$, but we are particularly interested in the qualitative differences that arise in the many resource limit where $\mathcal{R} \gg 1$.

3.1. Invasion of a mutant strain. We begin by considering a mutation that occurs in an ecosystem with an arbitrary number of coexisting strains, with equilibrium resource-specific mean fitnesses, $\overline{X}_i$. Without loss of generality, we will assume that the mutation occurs in the $\mu = 1$ strain, and leads to a new phenotype $(X_\mu + s, \vec{\alpha}_\mu + \Delta\vec{\alpha})$, where the strategy perturbation must satisfy the normalization constraint $\sum_i \Delta\alpha_i = 0$. The invasion fitness for the resident strain μ must be zero, since it is by definition present at the ecological equilibrium. Using this fact, along with the normalization condition on $\Delta\vec{\alpha}$, one can show that the general invasion fitness for the mutant is given by

$$S_{\text{inv}} \equiv \sum_i (\alpha_{\mu,i} + \Delta\alpha_i) \left[e^{X_\mu + s - \overline{X}_i} - 1 \right] , \tag{S78}$$

$$= (e^s - 1) + e^s \sum_i \Delta\alpha_i \left(e^{X_\mu - \overline{X}_i} - 1 \right) , \tag{S79}$$

which generalizes the two-resource invasion fitness in Eq. (S47). In the near-ESS limit where s, X_μ, and $\overline{X}_i$ are all small compared to one, this expression reduces to Eq. (16) in the main text.

3.2. Ecological equilibria. The invasion fitness in Eq. (S78) depend on the structure of the stable ecosystem through the resource-specific mean fitnesses, $\overline{X}_i$, which depend on the equilibrium strain frequencies f^*_μ through the definition in Eq. (2). Compared to the two-resource case above, it is generally more difficult to calculate the ecological equilibrium for a set of strains when $\mathcal{R} > 2$. Part of this difficulty is caused by the vector nature of the resource space, which can no longer be projected down onto a single scalar dimension. However, this is more than just a book-keeping issue — there are also fundamentally new kinds of ecological equilibria that can arise when $\mathcal{R} > 2$. In a two-resource system, ecological equilibria are either monocultures (with $\mathcal{S} = 1$ resident strains), or else contain the maximum number of coexisting strains permitted by the environment ($\mathcal{S} = 2$). However, when $\mathcal{S} > 2$, one can also have stable coexistence at any intermediate value of $1 < \mathcal{S} < \mathcal{R}$, in addition to the saturated state with $\mathcal{S} = \mathcal{R}$. These two classes of equilibria turn out to have very different properties.

Saturated ecosystems. The saturated stable state ($\mathcal{S} = \mathcal{R}$) is the closest analogue of the two-resource equilibrium that we studied in Appendix 2. In this case, we can obtain an explicit solution for the strain frequencies, f^*_μ, and resource-specific mean fitnesses, $\overline{X}_i$, attained at equilibrium as a function of the phenotypes $(X_\mu, \vec{\alpha}_\mu)$ of the resident strains. By definition, the per capita growth rate $(\partial_t \log f_\mu)$ of each resident strain must vanish at equilibrium, which yields a system of $\mathcal{S}$ equations for the $\mathcal{R}$ resource-specific mean fitnesses:

$$\sum_i \alpha_{\mu,i} e^{-\overline{X}_i} = e^{-X_\mu} . \tag{S80}$$

When $k = p$, this system can be inverted to obtain

$$e^{-\overline{X}_i} = \sum_\mu \alpha^{-1}_{i,\mu} e^{-X_\mu} , \tag{S81}$$

where $\alpha^{-1}_{i,\mu}$ is the left inverse of $\alpha_{\mu,i}$. In the limit that $|X_\mu - X_\nu| \ll 1$, this reduces to Eq. (17) in the main text. Using the definition of $\overline{X}_i$ in Eq. (2) in the main text, we obtain a second system of $\mathcal{R}$ equations for the k equilibrium frequencies:

$$\beta_i = \sum_\mu \alpha_{\mu,i} e^{X_\mu - \overline{X}_i} f^*_\mu , \tag{S82}$$

which is the non-neutral generalization of Eq. (6) in the main text. Again, when $\mathcal{S} = \mathcal{R}$, we can invert this system to obtain

$$f^*_\mu = e^{-X_\mu} \sum_i \beta_i e^{\overline{X}_i} \alpha^{-1}_{i,\mu} = \sum_i \frac{\beta_i \alpha^{-1}_{i,\mu}}{\sum_\nu \alpha^{-1}_{i,\nu} e^{X_\mu - X_\nu}} , \tag{S83}$$

since the left and right inverses are equal in this case. In the limit that $|X_\mu - X_\nu| \to 0$, we obtain the leading order contribution

$$f^*_\mu \approx \sum_i \beta_i \alpha^{-1}_{i,\mu} - \sum_{i,\nu} \beta_i \alpha^{-1}_{i,\mu} \alpha^{-1}_{i,\nu} (X_\mu - X_\nu) . \tag{S84}$$

To gain intuition into these formulae, we consider a set of strains whose resource strategies are a mixture of specialist and generalist components:

$$\alpha_{\mu,i} = (1 - \epsilon)\beta_i + \epsilon\delta_{\mu,i} , \tag{S85}$$

where $0 \leq \epsilon \leq 1$ provides a measure of the "distance" between the resource strategies. In this case, the inverse matrix has the asymptotic limits

$$\alpha_{i,\mu}^{-1} \sim \begin{cases} \delta_{i,\mu} + \mathcal{O}(1-\epsilon) & \text{if } 1-\epsilon \ll 1, \\ \frac{\delta_{i,\mu}-\beta_\mu}{\epsilon} + \mathcal{O}(1) & \text{if } \epsilon \ll 1, \end{cases} \quad [S86]$$

so that

$$\overline{X}_i \sim \begin{cases} X_i & \text{if } 1-\epsilon \ll 1, \\ \frac{\sum_{j\neq i} \beta_j (X_i - X_j)}{\epsilon} & \text{if } \epsilon \ll 1, \end{cases} \quad [S87]$$

and

$$f_\mu^* \sim \begin{cases} \beta_\mu & \text{if } 1-\epsilon \ll 1, \\ \beta_\mu \left[1 - \frac{\sum_{\nu\neq\mu} \beta_\nu (X_\mu - X_\nu)}{\epsilon^2}\right] & \text{if } \epsilon \ll 1. \end{cases} \quad [S88]$$

Unsaturated ecosystems. In contrast to the saturated case, when the number of surviving species is less than the number of resources ($\mathcal{S} < \mathcal{R}$) the equations in Eq. (S80) underdetermine the resource-specific mean fitnesses, $\overline{X}_i$, so we must invoke the non-linear constraints in Eq. (S82) to jointly solve for $\overline{X}_i$ and f_μ^*. Alternatively Ref. (1) has shown that the equilibrium values of $\overline{X}_i$ can be obtained from the solution of a convex optimization problem, subject to the constraints in Eq. (S80). In particular, if we define the transformed variable $h_i = e^{-\overline{X}_i}$, then the equilibrium value of h_i is the solution to the convex optimization problem

$$\vec{h}^* = \text{argmax}_{\vec{h}} \left\{ \sum_i \beta_i \log h_i \; : \; \sum_i \alpha_{\mu,i} h_i = e^{-X_\mu} \; \forall \mu \right\}. \quad [S89]$$

In fact, this method yields a general solution for the equilibrium value of $\overline{X}_i$ for *any* initial collection of strains, provided that the equality constraints in Eq. (S89) are replaced by inequalities ($\leq$). Given the equilibrium values of $\overline{X}_i$, the surviving species correspond to the indices μ where the equality condition is satisfied. The corresponding values of f_μ^* satisfy the (generally overdetermined) set of equations in Eq. (S82), which can be inverted using constrained linear regression. We employ this technique to implement the SSWM simulations in Appendix 5 below.

We note that since the objective function in Eq. (S89) depends on β_i, the equilibrium values of $\overline{X}_i$ will also generally depend on the environmental supply vector in an unsaturated ecosystem, in contrast to the β-independent values obtained in the saturated case. Thus, the ecosystem is no longer able to dynamically adjust to "shield" the internal selection pressures from the current state of the environment (1–3). Shifts in β_i can therefore lead to new opportunities for evolutionary adaptation.

3.3. Evolution in a binary resource usage model.

Since there are few empirical constraints on the genetic architecture of resource strategies in the limit of many resources ($\mathcal{R} \gg 1$), we focused on a toy "binary usage" model similar to the one considered by Ref. (1). In this model, genomes can either encode the ability to utilize a given resource or not (e.g. through the presence or absence of a particular pathway), so that the resource strategy is of the form

$$\alpha_{\mu,i} = \frac{I_{\mu,i}}{\sum_i I_{\mu,i}}. \quad [S90]$$

where $I_{\mu,i} \in 0, 1$ is a binary indicator variables. Individuals can acquire loss-of-function mutations at rate $U_\alpha^- \sum_i I_{\mu,i}$, which cause one of the values of $I_{\mu,i} = 1$ to switch to $I_{\mu,i} = 0$. We also assume that they can acquire gain of function "mutations" (e.g. horizontal acquisition of a gene from the environment) at rate $U_\alpha^+ \mathcal{R}$, which force a randomly chosen uptake rate to the $I_{\mu,i} = 1$ state. Under these assumptions, the pure mutation dynamics will lead to an binomial ensemble of resource strategies, analogous to the one considered by Ref. (1), with a "success probability" of

$$\frac{\left\langle \sum_i I_{\mu,i} \right\rangle}{\mathcal{R}} = \frac{U_\alpha^+}{U_\alpha^-}. \quad [S91]$$

For simplicity, we will assume that the $\mathcal{R}$ resources are all supplied at nearly identical rates. Note that in the completely symmetric case ($\beta_i = 1/\mathcal{R}$), a "generalist" strain with $I_{\mu,i} = 1$ will constitute a marginal evolutionary stable state. To avoid this pathological behavior, we will consider small perturbations around the completely symmetric state:

$$\beta_i = \frac{1}{\mathcal{R}} \left(1 + \epsilon_i - \frac{1}{\mathcal{R}} \sum_j \epsilon_j \right), \quad [S92]$$

where the ϵ_i are small random perturbations drawn from some distribution, and sorted in descending order ($\epsilon_1 \geq \epsilon_2 \geq \ldots \epsilon_{\mathcal{R}}$). For simplicity, we will assume that the ϵ_i are i.i.d. Gaussian variables with scale $\epsilon \ll 1$. The steep tail ensures that the maximum perturbation scales as

$$\epsilon_1 \sim \sqrt{2\epsilon^2 \log \mathcal{R}}\,, \tag{S93}$$

and can be bounded to be sufficiently small for a suitable choice of ϵ. Under these assumptions, an ecosystem comprised of a single "generalist" strain will still have nonzero ecological selection pressures encoded by the resource-specific mean fitnesses,

$$\overline{X}_i \approx -\epsilon_i\,, \tag{S94}$$

so that some alternate resource strategies will be favored to invade.

By simulating evolutionary dynamics in this model in the weak mutation limit for various values of U_X and s (Appendix 5.2), we find that the long-term structure of the ecosystem tends toward a state in which there is a single generalist strain and $\mathcal{S}-1$ single loss-of-function variants that have recently descended from this strain (Figs. 5 and S2-S4). In the limit that $U_X/U_\alpha \to 0$, this state must also coincide with a saturated state ($\mathcal{S} = \mathcal{R}$). If we let f_1 denote the frequency of the generalist strain, then the equilibrium frequencies are given by

$$f_1 = 1 - (\mathcal{R}-1)\epsilon_1\,, \tag{S95}$$

$$f_i = \left(1 - \frac{1}{\mathcal{R}}\right)(\epsilon_1 - \epsilon_i)\,, \tag{S96}$$

where we have assumed that $\mathcal{R}\epsilon_1 \ll 1$. In other words, all resources except the one with the largest value of ϵ_i will have a loss-of-function strain. In the limit that $\epsilon_1\mathcal{R} \ll 1$, the loss-of-function strains will constitute a tiny fraction of the population, and most mutations will arise in the generalist strain. In particular, the accumulation of fitness mutations will cause the fitness of the generalist strain to grow as $X_1 \sim NU_X s^2 t$. We therefore wish to understand when and how this fitness differential drives some of the loss-of-function variants to extinction.

Due to the symmetry of the system, if j strains are driven to extinction, these must be strains with loss-of-function mutations in genes with the next j largest values of ϵ_i, i.e. $i = 2, \ldots, j+1$. Let $X_c(j)$ denote the critical value of X_1 required for these j strains to go extinct. Expanding Eq. (1) in the main text to lowest order in X_1, $(1-f_1)$, $1/\mathcal{R}$, and ϵ, the equilibrium frequencies satisfy

$$R - 1 = \sum_i (1-\delta_{\mu,i})\left[\frac{\mathcal{R}-1}{\mathcal{R}}(1 - X_1 f_1) + \epsilon_i + \frac{f_1}{\mathcal{R}} + f_i \delta_{i>j+1}\right]\,, \tag{S97}$$

or

$$f_i \approx \frac{1-f_1}{\mathcal{R}} - \epsilon_i - \mathcal{R} X_1 f_1\,. \tag{S98}$$

To self consistently solve for $1-f_1$, we sum over $i = j+2, \ldots, \mathcal{R}$ to obtain:

$$1 - f_1 \approx \mathcal{R} \cdot \frac{1}{j+1}\sum_{k=1}^{j+1}\epsilon_k - \frac{\mathcal{R}^2(\mathcal{R}-1-j)}{j+1} X_1\,. \tag{S99}$$

Substituting this back into our expression for f_i, we obtain:

$$f_i \approx \frac{1}{j+1}\sum_{i=1}^{j+1}\epsilon_k - \epsilon_i - \frac{\mathcal{R}^2}{j+1} X_1\,. \tag{S100}$$

We can then self-consistently solve for j by setting $f_{j+1} = 0$. For example, if $j = 1$, then we have

$$X_c(1) \sim \frac{\epsilon_1 - \epsilon_2}{\mathcal{R}^2}\,. \tag{S101}$$

This will be a quenched random variable, since we have assumed that the ϵ_i are randomly distributed. Given our Gaussian assumption, the typical value of $\epsilon_1 - \epsilon_2$ will occur for

$$\epsilon_1 - \epsilon_2 \sim \frac{\epsilon \log\log \mathcal{R}}{\sqrt{\log \mathcal{R}}}\,, \tag{S102}$$

which yields

$$X_c(1) \sim \epsilon \cdot \frac{\log\log \mathcal{R}}{\mathcal{R}^2\sqrt{\log \mathcal{R}}}\,. \tag{S103}$$

On the other hand, if $j \sim \mathcal{R}$, we have

$$X_c(j) \sim \epsilon \cdot \frac{1}{\mathcal{R}} \,. \tag{S104}$$

These two fitness scales are separated by a gap of order

$$\frac{X_c(\mathcal{R})}{X_c(1)} \sim \frac{\mathcal{R}\sqrt{\log \mathcal{R}}}{\log\log \mathcal{R}} \,, \tag{S105}$$

which grows increasingly large as $\mathcal{R} \gg 1$.

We can use these results to derive heuristic expressions for the number of species $\mathcal{S}$ at steady state as a function of U_α/U_X. We first consider the limit where $\mathcal{S} \ll R$. As mentioned above, the generalist strain comprises the vast majority of the population, so that to a first approximation, we can assume that all fitness and strategy mutations occur on this genetic background. Furthermore, since $\mathcal{S} \ll \mathcal{R}$, most loss of function mutations will target a resource i that does not already have a loss-of-function variant, where the resource-specific mean fitness is given by

$$\overline{X}_i \approx \log\left(\frac{\alpha_{1,i} e^{X_1}}{\beta_i}\right) \approx X_1 - \epsilon_i \,. \tag{S106}$$

According to Eq. (16), the invasion fitness for a loss-of-function variant that targets resource i is given by

$$S_{\text{inv}} \sim \frac{-\epsilon_i}{\mathcal{R}} \,. \tag{S107}$$

Since these loss of function mutations are produced from the generalist background at rate NU_α per resource, the number of coexisting strains increases at rate

$$\frac{d\mathcal{S}}{dt} = \sum_i NU_\alpha \cdot \frac{|\epsilon_i|}{R}\theta(-\epsilon_i) \sim NU_\alpha \epsilon \,. \tag{S108}$$

in the absence of fitness mutations.

However, as we mentioned above, the accumulation of fitness mutations will cause the fitness of the generalist strain to grow as $X_1(t) \sim NU_X s^2 t$. Since loss-of-function variants do not acquire further fitness mutations of their own, their fitness is frozen at whatever fitness the generalist strain had at the time that the mutation arose. The fitness difference between the mutant and the generalist therefore grows with time until it reaches a critical value $X_c(\mathcal{R}) \sim \epsilon/\mathcal{R}$, at which point the loss-of-function variant is driven to extinction. This gives rise to two characteristic dynamical regimes depending on whether $X_c(\mathcal{R})$ is larger or smaller than the effect s of a typical fitness mutation.

If $s \ll X_c(\mathcal{R})$, then the generalist lineage must acquire multiple fitness mutations to drive one of the loss-of-function variants to extinction. To a first approximation, the fitness difference between the generalist and the jth most-recently created loss-of-function variant in this regime is given by

$$\Delta X_j \sim \frac{j}{NU_\alpha \epsilon} \cdot NU_X s^2 \,. \tag{S109}$$

The number of coexisting ecotypes $\mathcal{S}$ at steady-state is therefore determined by the relation $\Delta X_{\mathcal{S}} \sim X_c(\mathcal{R})$, which reflects a balance between the elimination of the oldest loss-of-function variant due to the accumulation of fitness mutations and the production of new loss-of-function variants through strategy mutations. Solving for $\mathcal{S}$, we obtain the scaling relation,

$$\mathcal{S} \sim \frac{1}{R}\frac{U_\alpha}{U_X}\left(\frac{\epsilon}{s}\right)^2 \,, \tag{S110}$$

listed in Eq. (19) in the main text.

On the other hand, if $s \gg X_c(\mathcal{R})$, then a single fitness mutation in the generalist strain is sufficient to drive loss-of-function variants to extinction. Before this mutation arises, all the loss-of-function variants will share the same fitness difference ($\Delta X_j = 0$), and this value suddenly shifts to $\gg X_c(\mathcal{R})$ once the successful fitness mutation occurs, driving all of the existing loss-of-function variants to extinction. This leads to oscillations in $\mathcal{S}$ in the time between successive fitness mutations, which range from $\mathcal{S}_{\min} \sim 1$ immediately after the fitness mutation arises, to a maximum value of $\mathcal{S}_{\max} \sim U_\alpha \epsilon/U_X s$ right before the next mutation arises. Since the loss-of-function variants accumulate linearly with time, this leads to a time-averaged value

$$\mathcal{S} \sim \frac{U_\alpha \epsilon}{U_X s} \,. \tag{S111}$$

Both expressions should remain valid up to the point where there is an appreciable probability that new loss-of-function mutations target a resource that already has pre-existing variant ($\mathcal{S} \sim \mathcal{R}$). However, there can still be a broad intermediate regime between this point and the point where the ecosystem is completely saturated ($\mathcal{R} - \mathcal{S} \lesssim 1$). The saturated state will coincide with the evolutionary steady-state if the generalist strain is able to seed fitter loss-of-function variants into all the

relevant resource dimensions before $X_1(t)$ increases to the point $X_c(1)$, where the first strains start to go extinct. Once again, there are two characteristic timescales depending on whether $X_c(1)$ is large or small compared to s.

If $s \ll X_c(1)$, then the generalist lineage must acquire multiple fitness mutations to before the first loss-of-function variants are driven to extinction. This will happen over a timescale,

$$T_{\text{collapse}} \sim \frac{X_c(1)}{NU_X s^2} \sim \frac{\epsilon}{NU_X s^2} \cdot \frac{\log\log \mathcal{R}}{\mathcal{R}^2 \sqrt{\log \mathcal{R}}} \,. \tag{S112}$$

During this time, loss-of-function mutations will occur in the generalist background at rate $NU_\alpha \mathcal{R}$ and will establish with probability $\sim X_1(t)$. Since the loss-of-function mutations are chosen randomly, $\mathcal{R}\log\mathcal{R}$ such establishments are required to cover the total number of resource dimensions with high probability (23). This requires a timescale T_{div} that satisfies

$$\int_0^{T_{\text{div}}} NU_\alpha \mathcal{R} NU_x s^2 t \sim \mathcal{R}\log\mathcal{R} \,, \tag{S113}$$

or

$$T_{\text{div}} \sim \sqrt{\frac{\log \mathcal{R}}{NU_\alpha NU_X s^2}} \,. \tag{S114}$$

The ecosystem will remain saturated if $T_{\text{collapse}} \gg T_{\text{div}}$, which leads to the condition

$$U_X \ll \frac{U_\alpha}{\mathcal{R}^4 \log^2 \mathcal{R}} \left(\frac{\epsilon}{s}\right)^2 \,. \tag{S115}$$

We can compare this point to the transition to $\mathcal{R} - \mathcal{S} \sim \mathcal{R}$ from Eq. (19), which shows that these regimes are separated by a gap of order

$$\frac{U_X(\mathcal{R} - \mathcal{S} \sim 1)}{U_X(\mathcal{R} - \mathcal{S} \sim \mathcal{R})} = (\mathcal{R}\log\mathcal{R})^2 \,, \tag{S116}$$

while

$$\frac{U_X(\mathcal{R} - \mathcal{S} \sim \mathcal{R})}{U_X(\mathcal{S} \sim 1)} = \mathcal{R}^2 \,. \tag{S117}$$

3.4. Limits on the number of utilized resources. The fragility of the diversification-selection balance in Appendix 3.3 can be attributed in large part to the emergence of a fit generalist strain that is able to utilize all of the available resources. In practice, however, there might be biological constraints or other costs that limit the number of resources that a given strain can metabolize. This leads us to consider an extension of our binary resource usage model, in which the maximum number of utilized resources is capped at some value $\mathcal{R}_c \ll \mathcal{R}$. In this way, we can consider complex ecosystems ($\mathcal{R} \to \infty$) while restricting the metabolic repertoire of any given strain. A full analysis of this model is beyond the scope of the present work. Instead, we will outline a heuristic calculation that suggests that diversification-selection balance at large $\mathcal{R}$ is achieved for substantially higher values of $\mathcal{S}$ than in Appendix 3.3 above.

We first note that when $\mathcal{R}_c \ll \mathcal{R}$, multiple strains are required to cover the available resources. The minimum possible number of strains is $\mathcal{S} \sim \mathcal{R}/\mathcal{R}_c$, which is achieved when each of the strains specializes on a disjoint subset of $\mathcal{R}_c$ resources. To lowest order in ϵ, the frequencies of these strains are given by

$$f_\mu \approx \frac{\mathcal{R}_c}{\mathcal{R}} \,. \tag{S118}$$

With the same genetic architecture of strategy mutations that we assumed above, this state will form the basis of the new diversification-selection balance. Generalizing our analysis above, this state will consist of $\mathcal{R}/\mathcal{R}_c$ independent copies of the diversification-selection balance in Appendix 3.3, except with $\mathcal{R} \to \mathcal{R}_c$. The strains that utilize $\mathcal{R}_c$ resources will be prevented from branching into new resources because of the maximum resource capacity. Meanwhile, single loss-of-function mutants on these backgrounds will be too small to acquire a gain-of-function mutation before their parent acquires enough fitness differences to drive them to extinction.

However, this behavior is strongly dependent on the specific genetic architecture that we assumed, as well as our focus on the SSWM limit. In larger populations, there may be a substantial probability for strains to acquire multiple strategy mutations in a short period of time, which would allow them to break out of their resource neighborhood. To mimic this effect in the SSWM limit, we can introduce a new rate $U_\alpha^{(2)} \ll U_\alpha$ to represent the probability that two strategy mutations arise in the same lineage in a single generation. In particular, we will use this new rate to model *resource swap* events, in which one of the currently utilized resources is deleted and replaced with a randomly drawn resource. As above, we will assume that this mutation rate scales with the number of utilized resources, so that the net rate is given by $U_\alpha^{(2)} k_\mu$, where $k_\mu = 1/\sum_i \alpha_{\mu,i}^2$.

In this augmented model, if we start from the set of $\mathcal{R}/\mathcal{R}_c$ disjoint strains, then the fitnesses of these strains will wander diffusively as $X_\mu \sim NU_X f_\mu s^2 t \pm \sqrt{NU_X f_\mu s^3 t}$, so that the typical fitness differences between a pair of strains is of order

$\Delta X(t) \sim \sqrt{NU_X f_\mu s^3 t}$. These fitness differences will create a selection pressure for strategy mutations that swap a resource from one of the fitter strains with a resource from one of the less fit strains. Such a mutation will have an invasion fitness

$$S_{\rm inv} = \frac{e^{\Delta X(t)} - 1}{\mathcal{R}_c} \approx \frac{\Delta X(t)}{\mathcal{R}_c} \,. \tag{S119}$$

Successful swap mutations will be produced on a timescale $\tau_{\rm diversify}$ that satisfies

$$\int_0^{\tau_{\rm diversify}} NU_\alpha^{(2)} \mathcal{R}_c \cdot \frac{\sqrt{NU_X f_\mu s^3 t}}{\mathcal{R}_c}\, dt \sim 1 \,. \tag{S120}$$

Solving for $\tau_{\rm diversify}$, we obtain

$$\tau_{\rm diversify} \sim \frac{1}{s} \left(NU_\alpha^{(2)}\right)^{-2/3} \left(NU_X \cdot \frac{\mathcal{R}_c}{\mathcal{R}}\right)^{-1/3} , \tag{S121}$$

$$\Delta X(\tau_{\rm diversify}) \sim s \left(\frac{U_X \mathcal{R}_c}{U_\alpha^{(2)} \mathcal{R}}\right)^{1/3} . \tag{S122}$$

Once the successful swap mutation invades, it will create a new ecotype that coexists with the parent clade, as well as the ecotype that currently utilizes the new resource. To lowest order in $1/\mathcal{R}_c$, the equilibrium frequency is given by

$$f_\nu^* \sim \frac{1 - e^{-\Delta X(t)}}{2} \frac{\mathcal{R}_c}{\mathcal{R}} \,. \tag{S123}$$

When $\Delta X \ll 1$, this frequency will be small compared to the other dominant ecotypes. As above, fitness mutations will therefore preferentially accumulate in the dominant ecotypes, causing the fitness advantage, $\Delta X(t)$, of the swap mutant to decrease over time at rate $NU_X f_\mu^* s^2$. After a time of order

$$\tau_{\rm collapse} \sim \frac{\Delta X(\tau_{\rm diversify})}{NU_X s^2 \cdot \frac{\mathcal{R}_c}{\mathcal{R}}} \sim \frac{1}{s} \left(\frac{U_X \mathcal{R}_c}{U_\alpha^{(2)} \mathcal{R}}\right)^{1/3} \frac{\mathcal{R}}{NU_X \mathcal{R}_c} \,, \tag{S124}$$

the fitness of the less fit ecotype will have caught up to the swap mutant, and the latter will be driven to extinction. The ratio between $\tau_{\rm collapse}$ and $\tau_{\rm diversify}$ is therefore given by

$$\frac{\tau_{\rm collapse}}{\tau_{\rm diversify}} \sim \left(\frac{U_\alpha^{(2)} \mathcal{R}}{U_X \mathcal{R}_c}\right)^{1/3} . \tag{S125}$$

If this ratio is sufficiently large, then new swap mutants will typically establish before the fitness differences drive any of the existing swap mutants to extinction. For fixed $\mathcal{R}_c$, this will become increasingly true $\mathcal{R} \to \infty$.

On the other hand, if $\tau_{\rm collapse} \ll \tau_{\rm diversify}$, then a typical resource swap mutation will be driven to extinction before the next arises. However, because the timing of the swap mutations is a random process, anomalously late mutations may occur for which $\Delta X(t) \sim \mathcal{O}(1)$. In this case, the swap mutant is no longer rare compared to its parent, and there is strong selection pressure for loss- (and later gain-) of-function mutations to arise in this mutant background.

Together, these arguments suggest that the simplest generalization of the steady-state in Appendix 3.3 will generally be unstable whenever we impose a cap on the number of utilized resources, and that the corresponding diversification-balance will be attained for much higher values of $\mathcal{S}$ than we would expect based on our previous analysis. Further analysis of these dynamics are left for future work.

4. Connections to adaptive dynamics

Our model shares certain features associated with the traditional models studied in adaptive dynamics (14, 15), though it also differs from these models in several key ways. In this section, we attempt to make this connection more explicit, using the notation and terminology employed in the adaptive dynamics literature. As adaptive dynamics relies on the weak mutation limit, we will confine our discussion to this regime as well.

4.1. Two resources, no fitness differences. For simplicity, we will start by considering the two-resource case in the absence of fitness differences, where individuals are described by a scalar resource phenotype α. To make the connection with adaptive dynamics explicit, we will define a rescaled trait,

$$x = \frac{\alpha - \beta}{\sqrt{\beta(1-\beta)}}\,, \tag{S126}$$

such that $x \to 0$ as $\alpha \to \beta$. Following Ref. (14), we then let $s_x(y)$ denote the invasion fitness of a mutant of phenotype y in a monomorphic population of phenotype x. In the neighborhood of $x \to 0$, Eq. (5) shows that $s_x(y)$ takes on a simple quadratic form

$$s_x(y) = (x-y)x\,. \tag{S127}$$

Under the standard adaptive dynamics assumption that y is infinitesimally close to x, the phenotypes will evolve in the direction of the fitness gradient,

$$D(x) = \left.\frac{\partial s_x(y)}{\partial y}\right|_{y=x} = -x\,. \tag{S128}$$

This gradient vanishes for $x = 0$ which shows that $x^* = 0$ (or $\alpha{=}\beta$) is an evolutionarily singular strategy. This strategy is *convergence stable*, in that infinitesimal mutations drive the population toward $x = 0$ when $|x| > 0$. However, it is only marginally ESS-stable, since $\partial^2 s_x(y)/\partial y^2 = 0$ at $x = 0$. The second derivative classification in Ref. (14) also shows that stable dimorphisms can coexist in the neighborhood of x^*.

These two features combine to make our evolutionarily singular strategy behave as *both* an evolutionarily stable strategy (ESS) and an evolutionary branching point. On the one hand, $x^* = 0$ resembles an ESS because no mutant strains are favored to invade once the population reaches x^*. On the other hand, $x^* = 0$ resembles an evolutionary branching point because the population will typically branch into a stable dimorhpism once $x - x^*$ approaches the typical spacing between mutants. Thus, in practice, the population will always branch before it reaches the ESS, even if this is excluded under truly infinitesimal evolution. However, unlike a traditional branching point where $\partial^2 s^*_x(y)/\partial^2 y > 0$, there is no further selection to drive the branched phenotypes x_1 and x_2 away from each other once branching has occurred. We showed in the text that this can be viewed as a generic feature of a saturated ecosystem (where $\mathcal{S} = \mathcal{R}$) when there are no overall fitness differences between strains.

4.2. Resource continuum, no fitness differences. One might ask why the evolutionarily singular strategy is so peculiar in our model, given that consumer-resource theory is often touted as an example of evolutionary branching points in the adaptive dynamics literature (24). The key difference is that in this existing literature, the trait x does not usually refer to the uptake rate of a single resource, but instead is used to parameterize an entire curve of resource uptake rates for a continuum of different resources. To choose a simple example, one might imagine that the resources denote seeds of different sizes, which are indexed by a continuous parameter z. The function $\beta(z)$ then represents the distribution of seed sizes supplied by the environment, which is often assumed to have a Gaussian form

$$\beta(z) = \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\,, \tag{S129}$$

centered at some special value $z = 0$. Individual uptake rates are often assumed to have a similar Gaussian shape

$$\alpha(z|x) = \frac{e^{-\frac{(z-x)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\,, \tag{S130}$$

with a preferred value of $z = x$ and a characteristic width σ. The trait x is then subject to further evolution, rather than the individual uptake rates $\alpha(z)$. Substituting these functions into Eq. (1) (with $X_\mu = 0$), the invasion fitness for phenotype y in a monomorphic population of phenotype x is given by

$$s_x(y) = \exp\left(\frac{1}{\sigma^2}\left[\frac{x^2}{2}\left(1+\frac{1}{\sigma^2}\right) - \frac{y^2}{2}\left(1-\frac{1}{\sigma^2}\right) - \frac{2xy}{\sigma^2}\right]\right) - 1\,. \tag{S131}$$

The fitness gradient $\partial s_x(y)/\partial y$ vanishes when $x = 0$, showing that $x^* = 0$ is an evolutionarily singular strategy, as anticipated. The second derivative is given by

$$\frac{\partial s_x^*(y)}{\partial y^2} = \frac{1-\sigma^2}{\sigma^4}\,. \tag{S132}$$

For $\sigma > 1$, x^* is a true ESS, while for $\sigma < 1$, x^* is a true evolutionary branching point. In the neighborhood of x^*, selection will act to drive the phenotypes further apart from each other after branching has occurred.

We can understand this behavior using the intuition developed in the main text. Since there are an infinite number of resources in this model, the ecosystem is certainly not saturated when $\mathcal{S} = 2$. Thus, we can expect much of the selection pressure to focus on bringing the population averaged uptake rate $\overline{\alpha}(z)$ closer to the environmental supply rate $\beta(z)$. When the niche width σ is larger than the range of resources supplied by the environment, the best way to do this is with a single strain centered at $x = 0$. Branching is therefore not favored. On the other hand, if σ is smaller than the range of supplied resources, then the ecosystem as a whole can match the environmental supply rate better if there are two strains centered at intermediate locations on the real axis ($|x - y| > 0$).

In this way, we see that the $\mathcal{R} = 2$ resource case, far from being pathological, serves as a basic building block that allows us to understand more complex scenarios that are often considered in the literature. It also illustrates how the genetic architecture of the uptake rates (in this case, whether the $\alpha(z)$ can evolve independently or are restricted to the Gaussian family) can play a key role in determining the emergent dynamics of the model.

4.3. Directional selection as an intermediate asymptotic. We now return to the two-resource case above and examine how changes in the overall fitness (X) alter the adaptive dynamics analogy. Individuals are now described by a two-dimensional phenotype, (X, α). Generalizing our analysis above, we will now define a two-dimensional trait space:

$$x_1 = \frac{\alpha - \beta}{\sqrt{\beta(1-\beta)}}\,, \quad x_2 = X\,. \tag{S133}$$

In this notation, the invasion fitness in Eq. (7) becomes

$$s_{x_1,x_2}(y_1, y_2) = (y_2 - x_2) + (x_1 - y_1)x_1\,, \tag{S134}$$

whose fitness gradient is given by

$$\nabla_y s_{x_1,x_2}(y_1, y_2) = (-x_1, 1)\,. \tag{S135}$$

As expected, the overall fitness dimension always selects for phenotypes that increase X, regardless of the value of α. As a consequence, there are no longer any evolutionarily singular strategies in this model, so the formal classification such points in the adaptive dynamics framework does not apply any more. Nevertheless, we have seen that behaviors very similar to evolutionary branching still occur in our model if we project down onto the x_1 coordinate. Furthermore, the old evolutionarily singular strategy at $x_1^* = 0$ continues to play a key role in these dynamics. The major difference is that these ecologically stable polymorphisms are now only quasi-stable under evolutionary perturbations, as our analysis in the main text shows that further fitness evolution can drive one of the ecotypes to extinction (Fig. 2B). This behavior is consistent with observations from laboratory evolution experiments (22).

Although we have motivated this behavior with the abstract notion of overall fitness, our analysis suggests that similar behavior will generically arise in multi-dimensional phenotype spaces whenever one of the traits (i) approaches its marginal branching point x_i^*, while at least one of the other traits (j) remains far from x_j^* (Fig. S1). Previous work suggests that such highly asymmetric approaches to a stationary point may be common feature of gradient descent dynamics in high dimensional spaces (25). This suggests that the competition between resource and strategy mutations can be viewed as a more general intermediate asymptotic that describes the process of ecological diversification during the asymptotically long times required to approach an evolutionarily singular strategy in a high dimensional trait space.

Of course, our simplified consumer-resource model is peculiar in that it contains just a single marginal ESS ($\vec{\alpha} = \vec{\beta}$) even for large $\mathcal{R}$. For generic high-dimensional landscapes, in contrast, previous work suggests that the relevant critical points will often be more akin to saddle points (25). In this case, we conjecture that Fig. S1 would describe the approach to one such saddle point, until the positive eigenvalues combine to drive the population away from the current critical point and toward a saddle of lower index. This hypothesis suggests that repeated bouts of diversification and ecosystem collapse could occur even in glassy regimes where the population never reaches a true ESS. A detailed understanding of these dynamics remains an interesting topic for future work.

5. Simulations

5.1. Individual-based simulations. The simulations in Fig. 1 were carried out using an individual-based, discrete generation algorithm similar to the one employed in Refs. (16, 26) for a single resource. Each simulation starts with a clonal population of N individuals, and in each subsequent generation, the population undergoes a selection step followed by a mutation step. At each step, we keep track of the number of individuals n_μ with a given strategy vector $\alpha_{\mu,i}$ and overall fitness X_μ.

In the selection step, each lineage n_μ is assigned a new size from a Poisson distribution with mean

$$\lambda_\mu = C\left(\sum_{i=1}^{\mathcal{R}} \alpha_{\mu,i} e^{X_\mu - \overline{X}_i}\right), \tag{S136}$$

where

$$\overline{X}_i = \log\left[\sum_\mu \frac{\alpha_{\mu,i} e^{X_\mu}}{\beta_i} \cdot \left(\frac{n_\mu}{\sum_\nu n_\nu}\right)\right], \tag{S137}$$

and $C = N/\sum \lambda_\mu$ is a normalization constant chosen to ensure that the total population size remains near $N \pm \mathcal{O}(\sqrt{N})$.

In the mutation step, the new lineage size is pruned into multiple sublineages representing different mutations that occur on the original lineage background. With probability U_X, an individual founds a new sublineage ν is founded with fitness $X_\nu = X_\mu + s$, where s is drawn from the distribution of fitness effects $\rho_X(s)$. With probability U_α, an individual founds a new sublineage with a strategy vector $\vec{\alpha}_\nu$ drawn from the distribution $\rho_\alpha(\vec{\alpha}'|\vec{\alpha})$.

The simulations in Fig. 1 were carried out for $\mathcal{R} = 2$ with $\beta = 0.5$. We utilized a Gaussian distribution of fitness effects, $\rho_X(s) \propto \exp(-s^2/2s_0^2)$, for the pure fitness mutations. The distribution of strategy mutations, $\rho_\alpha(\alpha'|\alpha)$, was taken to be a beta distribution with mean α and coefficient of variation $\mathrm{Var}(\alpha')/E(\alpha')^2 = 0.05$, but with α' rounded to the nearest value of $1/5, \ldots, 4/5$. The initial resource strategy for the ancestral population was chosen uniformly at random from these discrete values.

Each simulation was performed for a total of $60,000$ generations with a population size of $N \sim 10^7$. Every 500 generations, we simulated a round of "metagenomic sequencing". We calculated the population frequencies of all mutations present in the population, and reported these values after binomial resampling at a depth of $D = 1000$.

A copy of our implementation in `C++` is available on Github (https://github.com/benjaminhgood/consumer_resource_simulations).

5.2. SSWM simulations. To simulate the long-term dynamics of the binary usage model in Fig. 4 (Appendix 3.3), we use an optimized simulation algorithm that is specifically designed for the strong-selection, weak mutation (SSWM) regime. Similar to traditional SSWM algorithms in population genetics (27), this algorithm gains an efficiency advantage by simulating only successful invasion events. In our case, however, the successful invasion events can now lead to non-trivial ecological equilibria, in addition to simple fixation.

Our simulations start with a collection of strains $(X_\mu, \vec{\alpha}_\mu)$ at time $t = 0$. To assess convergence to diversification-selection balance, we performed simulations for two initial conditions: (i) a single generalist strain with $\alpha_{\mu,i} = 1/\mathcal{R}$ (ii) a collection of $\mathcal{R}$ specialist strains with $\alpha_{\mu,i} = \delta_{\mu,i}$ and X_μ drawn from a Gaussian distribution with variance $\sigma = 10^{-7}$. Figs S3 and S4 show a comparison of these two initial conditions for $\mathcal{R} = 10$. Since the agreement is generally good, we utilized the more rapidly converging generalist initial conditions for the main simulations in Figs. 4 and S2.

After drawing the initial condition, we first calculate the ecological equilibrium, f_μ^*, for this collection of strains using the convex optimization procedure in Appendix 3.2, using the MOSEK software package (28). This algorithm yields the equilibrium values of $h_i^* = e^{-\overline{X}_i}$ and the set of ecotypes Σ^* that survive at equilibrium. Within this subset, the equilibrium frequencies are obtained from the solution of the linear system in Eq. (S82), which will be overdetermined when $\mathcal{S} < \mathcal{R}$. We obtain a solution to this system by solving the constrained least squares problem,

$$\vec{f}^* = \mathrm{argmin}_{\vec{f}}\left\{\sum_i \left|\sum_{\mu\in\Sigma^*} \alpha_{\mu,i} e^{X_\mu} f_\mu - \frac{\beta_i}{h_i^*}\right|^2 : \sum_{\mu\in\Sigma^*} f_\mu = 1\right\}, \tag{S138}$$

using the SciPy library (29).

Once the initial ecological equilibrium is obtained, the simulation proceeds via a series of virtual timesteps, each of which represents the successful invasion of a single mutation. In each step, we first enumerate the set of fitness and strategy mutants that are generated from mutations on each of the current strains μ, and calculate their corresponding invasion fitness from Eq. (S78). We use these values to calculate the net rate of successful invasions from each mutation type. We assume that fitness mutations confer a characteristic fitness benefit s, so that the rate of successful fitness mutations in strain μ is given by

$$R_\mu^X = NU_X f_\mu^* \left(e^s - 1\right). \tag{S139}$$

The rate of successful loss-of-function mutations for resource i is given by

$$R_{\mu,i}^- = \max\left\{0,\ NU_\alpha f_\mu^* \cdot \sum_{j\neq i} \frac{\alpha_{\mu,j}(h_j^* - h_i^*)}{k_\mu - 1} \cdot \theta(\alpha_{\mu,i} - \delta)\right\}, \tag{S140}$$

where $k_\mu = 1/\sum_i \alpha_{\mu,i}^2$ is the current number of resources utilized by strain μ, $\theta(z)$ is the Heavisde step function, and δ is an infinitesimally small positive number so that the step function is well-defined. The rate of successful gain-of-function mutations is given by an analogous expression,

$$R_{\mu,i}^{+} = \max\left\{0,\, NU_\alpha f_\mu^* \cdot \sum_{j\neq i} \frac{\alpha_{\mu,j}(h_i^* - h_j^*)}{k_\mu + 1} \cdot \theta(\delta - \alpha_{\mu,i})\right\}. \quad \text{[S141]}$$

Since these successful invasion events arise as a compound Poisson process, the time $T_{\rm est}$ to the next successful invasion event is exponentially distributed with rate

$$R_{\rm tot} = \sum_\mu \left[R_\mu^X + \sum_i \left(R_{\mu,i}^{+} + R_{\mu,i}^{-}\right)\right]. \quad \text{[S142]}$$

Using the Poisson thinning property, the identity of the invading mutation is chosen at random from the enumerated list with probability proportional to its corresponding R-value. Once the identity of the invading strain is determined, we find the new ecological equilibrium $\vec{h}^*$ and $\vec{f}^*$ using the constrained procedure above. By assumption, the time to reach this new equilibrium is negligible compared to $T_{\rm est}$ in the SSWM limit. The current time t is then incremented by $T_{\rm est}$, and the process repeats itself.

We repeated this process for a total of M successful invasion steps until the ecosystem converged to diversification-selection balance ($M \sim 100,000$). The simulations in Figs. 4 and S2-S4 were carried out for $\epsilon = 10^{-3}$ and $s = 10^{-7}$, scanning through different values of U_X/U_α.

A copy of our implementation in Python is available on Github (http://github.com/StephenMartis/consumer-resource-many-resources).

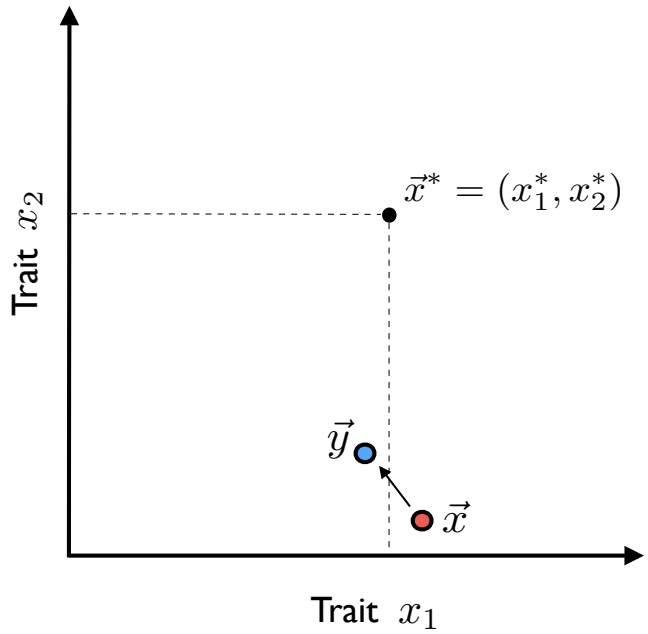

Fig. S1. An intermediate asymptotic of adaptive dynamics. In a multidimensional phenotype space, a population that is far from the evolutionarily singular strategy can display the quasi-stable branching behavior analyzed in the main text if one of the trait dimensions (x_1) is close to the singular coordinate (x_1^*). In the specific context of our consumer resource model, x_1 corresponds to the resource uptake strategy (α), while x_2 corresponds to the overall fitness (X).

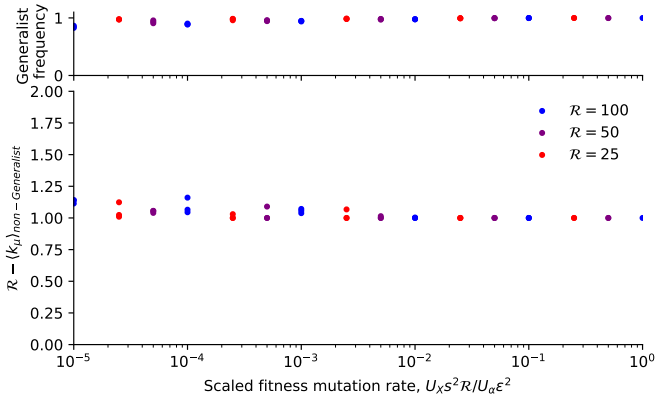

Fig. S2. Ecological structure at diversification-selection balance in a binary usage model. (Top) For each of the simulated populations in Fig. 4, the fraction of the population occupied by the generalist ecotype, $\alpha_{\mu,i} = 1/\mathcal{R}$. (Bottom) For the same populations, the frequency-weighted average of $k_\mu = 1/\sum_i \alpha_{\mu,i}$ (a measure of the number of utilized resources) for the remaining non-generalist ecotypes. A value of $\mathcal{R} - \langle k_\mu \rangle = 1$ indicates that the rest of the population consists of single loss-of-function mutants that descend from the generalist background.

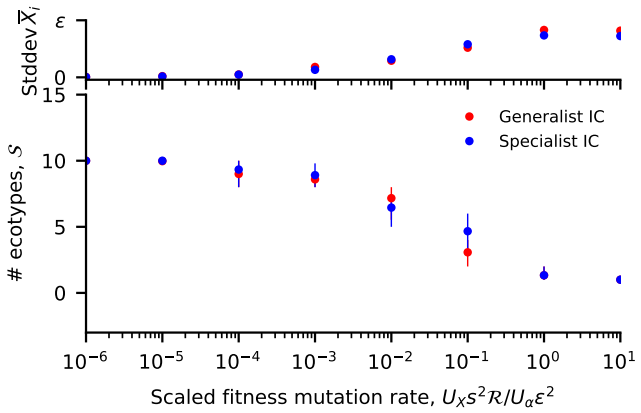

Fig. S3. Approach to diversification-selection balance from different initial conditions. An analogous version of Fig. 4 comparing specialist and generalist initial conditions for $\mathcal{R} = 10$.

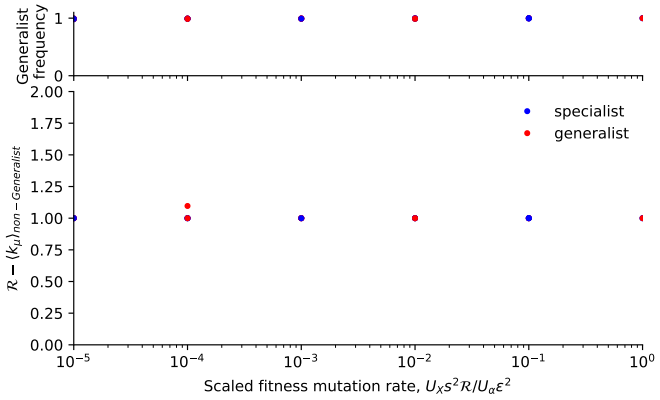

Fig. S4. Long-term ecological structure from different initial conditions. An analogous version of Fig. S2 comparing specialist and generalist initial conditions for $\mathcal{R} = 10$.

References

1. Tikhonov M, Monasson R (2017) Collective phase in resource competition in a highly diverse ecosystem. *Phys Rev Lett* 118(4):048103.
2. Advani M, Bunin G, Mehta P (2018) Statistical physics of community ecology: a cavity solution to macarthur's consumer resource model. *Journal of Statistical Mechanics: Theory and Experiment* 2018(3):033406.
3. Posfai A, Taillefumier T, Wingreen NS (2017) Metabolic trade-offs promote diversity in a model ecosystem. *Phys Rev Lett* 118(2):028103.
4. Gardiner C (1985) *Handbook of Stochastic Methods*. (Springer, New York).
5. Mac Arthur R (1969) Species packing, and what competition minimizes. *Proceedings of the National Academy of Sciences* 64(4):1369–1371.
6. Tilman D (1982) *Resource competition and community structure*. (Princeton University Press).
7. Good BH, Desai MM (2013) Fluctuations in fitness distributions and the effects of weak linked selection on sequence evolution. *Theor Pop Biol* 85:86–102.
8. Ewens WJ (2004) *Mathematical Population Genetics*. (Springer-Verlag, New York), second edition.
9. Levy SF, Ziv N, Siegal ML (2012) Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant. *PLoS biology* 10(5):e1001325.
10. Sun L, et al. (2018) Effective polyploidy causes phenotypic delay and influences bacterial evolvability. *PLoS biology* 16(2):e2004644.
11. Kendall DG (1948) On the generalized "birth-and-death" process. *The annals of mathematical statistics* 19(1):1–15.
12. Good BH (2016) Ph.D. thesis.
13. Kardar M (2007) *Statistical physics of fields*. (Cambridge University Press).
14. Geritz SA, Mesze G, Metz JA, , et al. (1998) Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary ecology* 12(1):35–57.
15. Doebeli M (2011) *Adaptive Diversification*. (Princeton University Press).
16. Good BH, Rouzine IM, Balick DJ, Hallatschek O, Desai MM (2012) Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. *Proc Natl Acad Sci USA* 109:4950–4955.
17. Desai MM, Fisher DS (2007) Beneficial mutation selection balance and the effect of genetic linkage on positive selection. *Genetics* 176:1759–1798.
18. Fisher DS (2011) Leading the dog of selection by its mutational nose. *Proc Natl Acad Sci USA* 108:2633–2634.
19. Cvijović I, Good BH, Jerison ER, Desai MM (2015) Fate of a mutation in a fluctuating environment. *Proceedings of the National Academy of Sciences* 112(36):E5021–E5028.
20. Gerrish P, Lenski R (1998) The fate of competing beneficial mutations in an asexual population. *Genetica* 127:127–144.
21. Fisher DS (2013) Asexual evolution waves: fluctuations and universality. *J Stat Mech* 2013:P01011.
22. Good BH, McDonald MJ, Barrick JE, Lenski RE, Desai MM (2017) The dynamics of molecular evolution over 60,000 generations. *Nature* 551(7678):45.
23. Feller W (2008) *An introduction to probability theory and its applications*. (John Wiley & Sons) Vol. 2.
24. Ackermann M, Doebeli M (2004) Evolution of niche width and adaptive diversification. *Evolution* 58(12):2599–2612.
25. Kurchan J, Laloux L (1996) Phase space geometry and slow dynamics. *Journal of Physics A: Mathematical and General* 29(9):1929.
26. Good BH, Desai MM (2015) The impact of macroscopic epistasis on long-term evolutionary dynamics. *Genetics* 199(1):177–190.
27. McCandlish DM, Stolzfus A (2014) Modeling evolution using the probability of fixation: history and implications. *Quarterly Review of Biology* 89:225–252.
28. Andersen ED, Andersen KD (2000) The mosek interior point optimizer for linear programming: an implementation of the homogeneous algorithm in *High performance optimization*. (Springer), pp. 197–232.
29. Jones E, Oliphant T, Peterson P, , et al. (2001–) SciPy: Open source scientific tools for Python.