Emergent neutrality in consumer-resource dynamics

Rafael D'Andrea [1], Theo Gibbs [1], James P. O'Dwyer [1]
1 Department of Plant Biology, University of Illinois, Urbana IL USA

Statement of Authorship: RD and JOD conceived the study. RD and TG performed analyses and wrote the paper. All authors contributed to the final version of the manuscript.

Abstract

Neutral theory assumes all species and individuals in a community are ecologically equivalent. This controversial hypothesis has been tested across many taxonomic groups and environmental contexts, and successfully predicts species abundance distributions across multiple high-diversity communities. However, it has been critiqued for its failure to predict a broader range of community properties, particularly regarding community dynamics from generational to geological timescales. Moreover, it is unclear whether neutrality can ever be a true description of a community given the ubiquity of interspecific differences, which presumably lead to ecological inequivalences. Here we derive analytical predictions for when and why non-neutral communities of consumers and resources may present neutral-like outcomes, which we verify using numerical simulations. Our results, which span both static and dynamical community properties, demonstrate the limitations of summarizing distributions to detect non-neutrality, and provide a potential explanation for the successes of neutral theory as a description of macroecological pattern.

Introduction

One of the central questions in community ecology is how species interactions affect community structure and ecological dynamics, and conversely whether summarized representations of the latter can be used to make inferences about the former. Macroecological patterns such as species abundance distributions in particular have been extensively studied, and often used in inference approaches (McGill, 2003; McGill et al., 2007; Harte, 2011).

However, these inference approaches have been called into question, especially following intense debate over the merits of the neutral theory of biodiversity (McGill, 2003; Harte, 2003; Gotelli and McGill, 2006; Alonso et al., 2006; Clark, 2009; Rosindell et al., 2011; Clark, 2012; Rosindell et al., 2012). Neutral theory assumes that all species and individuals in a community are ecologically equivalent (Caswell, 1976; Hubbell, 2001). This hypothesis has been heavily tested across many taxonomic groups and environmental contexts (Chave, 2004; Wootton, 2005; McGill et al., 2006; Volkov et al., 2007), and despite its radically simple assumptions, has experienced some success in describing macroecological patterns. Prominent among these successes is the prediction of a broad range of species abundances, matching the functional form for the species abundance distribution observed in multiple high-diversity communities (Hubbell, 2001; Bell, 2001; Volkov et al., 2003, 2007). On the other hand, the theory has been criticized on two fronts: first, it is unclear whether neutral theory could successfully predict a broader range of community properties, particularly regarding community dynamics from generational to geological timescales (Clark and McLachlan, 2003; Volkov et al., 2004; Pigolotti et al., 2005; Ostling, 2012; Wang et al., 2013; Kalyuzhny et al., 2014; O'Dwyer et al., 2015; Fung et al., 2016; D'Andrea and Ostling, 2017). Second, it has been argued that the patterns that neutral theory successfully fits do not uniquely reflect the underlying ecological properties of the species involved; in particular, it is possible that summarized indices of community structure and dynamics may look neutral even when species are not ecologically equivalent (Chave et al., 2002; McGill et al., 2007; Chisholm and Pacala, 2010; Al Hammal et al., 2015; D'Andrea and Ostling, 2017; Rael et al., 2018).

One earlier approach to investigating neutral-like outcomes under non-neutral dynamics has focused on phenomenological models of species interactions such as Lotka-Volterra competition (Haegeman and Loreau, 2011; Fisher and Mehta, 2014). These modeling approaches have required that the average behavior of a stochastic community model follows a deterministic Lotka-Volterra model. But this requirement does not completely determine the stochastic transition rates, and the ambiguity in translating a deterministic Lotka-Volterra model to a stochastic Lotka-Volterra model means that demographic noise enters into the model as a free parameter that can be tuned to adjust the balance of noise relative to deterministic ecological selection. As the strength of stochasticity is increased, species shift from a regime where they are pinned close to their deterministic equilibrium to one where they undergo long excursions away from equilibrium. This ap-

proach has successfully demonstrated that the stationary distribution of stochastic Lotka-Volterra models may be hard to distinguish from a neutral stochastic model where species abundances are stabilized by immigration from outside the system. On the other hand, these models do not clarify whether and when neutral behavior can emerge from competition models where the level of stochasticity is set endogenously, and not tuned to be high.

A second approach to understanding neutral-like phenomena has been in the context of consumer-resource models, where the mechanisms of competition are modeled explicitly in terms of overlapping resource requirements among groups of species. In this context, neutral behavior has been shown to arise from consumers competing for a small number of resources (Posfai et al., 2017). In this model, consumers satisfy metabolic tradeoffs, leading them to coexist on a neutrally stable manifold of fixed points on which drift subsequently determines the stationary distribution. On the other hand, if consumers violate the precise metabolic tradeoff constraints, then only a small number of consumers coexist at one fixed point, and the stationary abundance distribution of these consumers is no longer guaranteed to appear neutral. It is an open question, therefore, whether the species engaged in consumer-resource competition may appear neutral when their resource requirements are not fine-tuned in this way.

Here we present a model of stochastic, non-neutral consumer-resource dynamics, with which we investigate the potential for neutral-like outcomes from two perspectives: snapshots of community structure, and dynamics over multiple timescales. Using a deterministic version of our model, we derive predictions for a threshold between neutral and non-neutral behavior, which we test against numerical simulations of the full model. Specifically, we hypothesize that when the timescales of relaxation to equilibrium under non-neutral dynamics are commensurate with the timescales of drift to extinction under neutral dynamics, the system will appear neutral. We investigate two qualitatively different depatures from neutrality to test this hypothesis: a generalist scenario, where consumers have different preferences for resources but without a marked preferences for any one resource over others; and a specialist scenario, where each consumer has its own preferred resource, while consuming all other resources at a lower rate. Our analysis is based on these two particular ways to 'break' neutrality, but is not confined to them. However, we demonstrate that these cases support our hypotheses: consumer resource models with a degree of non-neutrality anywhere below our predicted threshold display neutral-like dynamics and patterns.

Methods

Consumer-Resource Model

We base our analysis on a stochastic model of S consumers, whose abundances we denote by N_i for $i \in \{1, ..., S\}$, that compete for K abiotic, substitutable resources, whose concentrations are R_k for $k \in \{1, ..., K\}$. Resources are externally supplied, and depleted only by consumption. Consumers grow purely through the consumption of these resources, and also undergo density-dependent mortality. While this is a challenging model to solve exactly, we can straightforwardly simulate a stochastic process with these three event types (resource inflow, consumption, and consumer mortality). In this simulation approach, we define $T\{X \to X+1\}$ as the transition rate at which the species with abundance X gains one individual, while all other abundances stay precisely the same. In this notation, the following transition rates represent the events that occur in our stochastic process:

$$\begin{aligned} T\{R_k \to R_k + 1\} &= \rho_k \\ T\{R_k \to R_k - 1\} &= (1-\epsilon) R_k \sum\nolimits_j C_{kj} N_j \\ T\{R_k \to R_k - 1, N_i \to N_i + 1\} &= \epsilon N_i C_{ki} R_k \\ T\{N_i \to N_i - 1\} &= \eta_i N_i \ . \end{aligned} \tag{1}$$

The basic event types that lead to the model in (1) are conceptually represented in Fig. 1. ρ_k is the inflow rate of the k-th resource. C_{ij} are non-negative coefficients which measure the rate of consumption of resource i by consumer j, forming a $K \times S$ matrix, C. ϵ measures the consumers' efficiency in converting resources into consumer biomass, and it is also the probability that a consumption event between resource k and consumer i will result in a new individual of consumer i. η_i is the per capita mortality rate for the i-th consumer. Every consumer extinct over long enough time scales, but in any real ecological system, speciation processes will tend to maintain species richness. Here we represent this balance in a simple way: every time a species with abundance $N_i = 1$ is selected for a death event (extinction), we keep it at $N_i = 1$, while logging the time it took for the death event to occur. This is akin to point speciation (Etienne et al., 2007), although for simplicity we introduce the 'new' species with the same resource profile as the 'extinct' species. (Alternatively, this process could be interpreted as very rare immigration events for that same species). While speciation and immigration can be modeled in increasing levels of realism (e.g. Schreiber and Tobiason 2003; Rosindell et al. 2010), our results are primarily based on the dynamics of extant species, and hence are (at least to a first approximation) insensitive to the details of this process.

In parallel, we consider a deterministic version of the same model:

$$\begin{aligned} \dot{R}_k &= \rho_k - R_k \sum\nolimits_j C_{kj} N_j \\ \dot{N}_i &= \epsilon N_i \sum\nolimits_j C_{ji} R_j - \eta_i N_i \ . \end{aligned} \tag{2}$$

Using this deterministic model, we consumer and resource abundances at equilibrium. While ecological disturbance or other perturbations can lead to important transient or cyclic behavior (Scheffer and van Nes, 2006; Jansen and Sigmund, 1998), stable equilibria attract nearby transient states, and therefore have a special role in determining the behavior of a dynamical system. Accordingly, they have been studied extensively in community ecology (Armstrong and McGehee, 1976, 1980; Barabás et al., 2014). Here, we will focus the dynamics nearby to feasible solutions, i.e. equilibria of (2) where all resources and consumers have positive equilibrium abundances (Saavedra et al., 2017). Let's denote the abundances by vectors $\vec{R} = \begin{bmatrix} R_1 & \dots & R_K \end{bmatrix}^T$ and $\vec{N} = \begin{bmatrix} N_1 & \dots & N_S \end{bmatrix}^T$ and the inflow and mortality rates by vectors $\vec{\rho} = \begin{bmatrix} \rho_1 & \dots & \rho_K \end{bmatrix}^T$ and $\vec{\eta} = \begin{bmatrix} \eta_1 & \dots & \eta_S \end{bmatrix}^T$. Then, given an invertible matrix C, we can choose positive inflow and mortality vectors $\vec{\rho}$ and $\vec{\eta}$ that lead to any feasible abundance vectors ($\vec{R}^*$ and $\vec{N}^*$) as equilibrium solutions of (2). For simplicity, we will choose $\vec{\rho}$ and $\vec{\eta}$ so that all resources converge to a single value r and all consumers converge to a value n. The values of the fixed points of (2) correspond approximately to the average values of the abundances in (1), so the stationary distribution for each consumer and resource will have a mean of n and r respectively.

In both (1) and (2), consumer i is determined by its consumption preferences (the i-th column of C) and mortality rate η_i. Fixing a given set of equilibrium abundances determines $\vec{\eta}$, leaving the principal object of biological interest as the set of consumption preferences in C. For example, if all entries in C are identical, we recover a neutral limit where all consumer species have the same resource preferences. In this manuscript we consider two different ways to move away from this neutral limit, shown in Fig. 1. The first, which we call the specialist scenario, is when C has all diagonal entries equal to one value C_d, while all off-diagonal entries are equal to another value C_o, where $C_d > C_o$. In this case, each consumer has a unique preferred resource that they consumer more quickly than all others. Other than this special resource, each species consumes all other resources at the same rate. If $C_d = C_o$, then all species are equivalent and we expect the stochastic version of this model to behave neutrally. In our second parametrization, called the generalist scenario, we sample each entry of C independently from a uniform probability distribution with mean μ and variance σ^2. To ensure that the C_{ij} are interpretable as rates of consumption, we require that $\mu > 0$ and choose σ^2 so that all the entries of C are positive. When $\sigma^2 = 0$, the entries of C are simply μ and we again expect the system to behave neutrally. In this way, the magnitude of the ratio C_d/C_o in the specialists scenario and the coefficient of variation $CV = \sigma/\mu$ in the generalists scenario are intuitive measures of the system's departure from neutral consumption

preferences. In the next section, we predict the threshold values of these two measures of non-neutrality, below which we predict the community dynamics will look neutral. In other words, we will make a prediction for how far from neutrality the underlying model can be, while still displaying neutral outcomes.

Neutral-Niche Threshold

The abundances in our stochastic process are constantly being perturbed away from the deterministic equilibrium by demographic noise. At the same time, consumer-resource feedbacks stabilize their abundances. Because it is difficult to calculate the rates at which the abundances return to equilibrium in the stochastic model, we will use the dynamics of the deterministic model near to equilibrium as a proxy for the stabilizing force that the consumers experience in the stochastic system. In our Supplementary Information, we characterize the spectrum of this deterministic model by deriving a linearized model. In each of the two ways for C to depart from neutrality, the spectrum of (2) has a characteristic structure – there is one bulk of real eigenvalues near zero, one bulk of real eigenvalues centered at a negative value, and a complex conjugate pair of eigenvalues. The eigenvalues near zero have eigenvectors that largely correspond to consumer dynamics, and the more negative eigenvalues have eigenvectors that largely correspond to resource dynamics (see Supplementary Figure S.2). In the specialist parametrization, the eigenvalues near zero are all equal, while in the generalist parametrization, they are described by the Marchenko-Pastur density from random matrix theory (Marčenko and Pastur, 1967). After any kind of perturbation, we therefore expect resource abundances to quickly decay back to equilibrium, while consumer abundances will decay more slowly, because the eigenvalues which dictate their dynamics are small in magnitude. This time-scale separation recapitulates the classic expectation of fast resource dynamics relative to consumer dynamics, but doesn't alone tell us whether the system will be neutral-like, or not.

We now let $Q_k(R_k|\vec{N},t)$ denote the probability that resource k has abundance R_k at time t in our stochastic model. As mentioned, consumer dynamics are slow, so that $\vec{N}$ is approximately constant in time while resource dynamics play out. The master equation for $Q_k(R_k|\vec{N},t)$ is then

$$\begin{aligned}\frac{dQ_k(R_k|\vec{N},t)}{dt} &= \rho_k Q_k(R_k-1|\vec{N},t) + \left((R_k+1)\sum_{j=1}^{S} C_{kj}N_j\right) Q_k(R_k+1|\vec{N},t) \\ &\quad - \left(\rho_k + R_k\sum_{j=1}^{S} C_{kj}N_j\right) Q_k(R_k|\vec{N},t)\ .\end{aligned} \tag{3}$$

Because of our timescale separation, the sums in (3) are constant, so the stationary distribution

is simply a Poisson distribution with rate $\frac{\rho_k}{\sum_j C_{kj} N_j}$. In our simulations, we find that the distribution of a single resource throughout time, and also the distribution of resource abundances across the community, quickly converge to Poisson distributions, in agreement with Equation (3) (see Supplementary Figs. S.6, S.7).

Next, we define $P_i(N_i|\vec{N},t)$, the distribution of consumer abundance values N_i over an ensemble of trajectories in our stochastic model. Since the resource abundances are tightly concentrated at their mean values, we assume that they are equal to their means, and we eliminate then from the master equation for consumer dynamics, leading to:

$$\begin{aligned}\frac{dP_i(N_i|\vec{N},t)}{dt} &= \epsilon(N_i-1)\left(\sum_{j=1}^{K}\frac{C_{ij}^T\rho_j}{C_{jj}(N_j-1)+\sum_{l\neq j}C_{jl}N_l}\right)P_i(N_i-1|\vec{N},t) \\ &\quad + \eta_i(N_i+1)P_i(N_i+1|\vec{N},t) - N_i\left(\eta_i+\epsilon\sum_{j=1}^{K}\frac{C_{ij}^T\rho_j}{\sum_l C_{jl}N_l}\right)P_i(N_i|\vec{N},t)\ .\end{aligned} \tag{4}$$

If the consumption coefficients are identical $C_{ij} = \mu$, then the sum $\sum_l C_{jl}N_l$ is approximately proportional to the average of the stationary distribution for consumers, which is constant in time at equilibrium. Equation (4) then becomes the master equation for a neutral birth-death process, with a log-series stationary distribution. As the coefficients C_{ij} become more heterogeneous, consumer abundances are more strongly affected by specific combinations of consumers, and we expect drift to fail to predict the resulting abundance patterns, because it does not incorporate these relationships. For example, if consumers are completely specialized and C is a diagonal matrix, then $\sum_l C_{jl}N_l = C_{jj}N_j$ is not a good estimate of the average value of the stationary distribution, since it is strongly affected by stochastic fluctuations in N_i at equilibrium. In fact, the stationary distribution for the master equation in (4) is a Poisson distribution for each consumer in the limit consumers are pure specialists, reflecting the Poisson distribution of their specialized resources. When instead we are in the neutral scenario, so that $C_{ij} = \mu$, we find that $\sum_l C_{jl}N_l = \mu\sum_l N_l$ is normally distributed with a mean of μSn and a variance determined by the variance of the resulting stationary distribution. This distribution is not very sensitive to abundance N_i and has relatively small fluctuations at equilibrium, so we expect it to be well approximated by simply replacing it by its mean value. In this way, we see how neutral drift can plausibly emerge from a consumer-resource model, and we want to derive predictions for when drift could be a good description of the equilibrium state.

In order to identify a transition between niche and neutral dynamics, we consider a neutral community where the average species abundance is n. Let T_n be the expected time to extinction (measured in generations) for a species undergoing drift from its mean abundance. We then use $T_n/\bar{\eta}$ as a characteristic timescale for drift in our model, where $\bar{\eta}$ is the average mortality rate—i.e. $T_n/\bar{\eta}$ sets a timescale over which drift is sufficient to take a species with average abundance all the

way to extinction. We propose that drift dynamics will dominate when this timescale is shorter than the timescale of the restoring, stabilizing forces due to niche specialization. To make this comparison, we will use $1/|\lambda_+|$ as the timescale for the stabilizing mechanism in the consumer-resource model, where λ_+ is the most negative eigenvalue corresponding to consumer dynamics in the deterministic model (i.e. λ_+ determines the fastest timescale among the consumer dynamics). If the characteristic time for consumer abundances to deterministically return to equilibrium is much longer than the timescale for drift to cause large abundance fluctuations (ie. if $1/|\lambda_+| \gg T_n/\bar{\eta}$), we expect that drift is the principal force governing consumer dynamics. This order-of-magnitude inequality suggests that there is a threshold value at which drift ceases to be a good description of the abundance patterns, namely when $1/|\lambda_+| \sim T_n/\bar{\eta}$. We should note here that our estimate of the characteristic timescale of the stabilizing mechanism assumes that the linearized system is a good approximation to the actual dynamics throughout a large range of abundances. In fact, the linearized system is only guaranteed to be a good description in the neighborhood of the fixed point, but we hypothesize that it yields a good order-of-magnitude estimate for this transition more generally.

In the Supplementary Information, we derive λ_+ in both of our parametrizations of C, thereby derive analytical expressions for when drift should be a good description of abundance patterns at equilibrium as a function of the other parameters in the model. We use the order-of-magnitude approximation that $T_n \approx n$ for simplicity, but using the full expression (8, see below) does not significantly change the functional form of the predicted thresholds. For the generalist parametrization, we find that the threshold for the coefficient of variation $CV = \sigma/\mu$ is

$$CV \sim \sqrt{\frac{S}{4n}\left(1 - \frac{\epsilon r}{n^2}\right)} \tag{5}$$

so the square of the threshold CV value should inversely depend on the mean consumer abundance n, if we disregard the term in parenthesis, which is small for our parameter choices. For the specialist parametrization, we derive a threshold scaling for the ratio between diagonal and off-diagonal elements of the consumption matrix, and find that

$$\frac{C_d}{C_o} \sim \frac{1 + \sqrt{\gamma}(N-1)}{1 - \sqrt{\gamma}} \tag{6}$$

where $\gamma = \frac{1}{n}\left(1 - \frac{\epsilon r}{n^2}\right)$. So, the specialist threshold is linear with the number of consumers in the community.

Testing for Neutrality

We now test our predicted threshold by comparing our observations to neutral predictions using both static and dynamical properties of the community. Under neutral dynamics, species abundances follow a distribution which, in high-diversity metacommunities, converges to Fisher's log-series (Fisher et al., 1943; Alonso and McKane, 2004). The probability that a species has abundance k in this case is

$$P_k = \alpha \frac{p^k}{k}, \tag{7}$$

where $\alpha = -1/\log(1-p)$, and the parameter p can be estimated from the mean abundance n by solving the equation $\alpha p/(1-p) = n$. We fit the log-series distribution to our simulated communities using the discrete Cramér-von Mises goodness-of-fit test (Arnold and Emerson, 2010), and consider the fit successful if the p-value exceeds 0.05.

To test our analytical predictions from the previous section, we define the neutrality threshold as the point of departure from neutral resource preferences at which the probability of a successful log-series fit drops below 50% (other choices for this probability cutoff do not change results qualitatively, see Supplementary Information). We determine this probability by fitting the log-series distribution onto ensembles of simulated communities, and running a logistic regression of the successful and rejected fits against the neutrality index of the different ensembles. In the generalist scenario, the neutrality index is the coefficient of variation in the consumption matrix, while in the specialist scenario it is the ratio between on- and off-diagonal entries.

We next define a test for whether long-timescale neutral dynamics fail to hold when the degree of non-neutrality of the consumer model passes through our predicted transition. Given a species current abundance k, its life expectancy (i.e. time to extinction) T_k under neutral dynamics has been shown to be (Chisholm and O'Dwyer, 2014)

$$T_k = \frac{1}{1-p}\left(p^{-k}B(p;1+k,0) + H_k + \log(1-p)\right), \tag{8}$$

where H_k is the k-th harmonic number (Abramowitz and Stegun, 1965), $B(z;a,b)$ is the incomplete beta function, and p is the log-series parameter (related to the speciation rate in neutral metacommunity models with speciation events). The extinction time in expression (8) is given in generation units, with a generation time defined as the inverse of the mean mortality rate of consumers, $\bar{\eta}$. In our numerical simulations, we obtain life expectancy estimates as follows. Over the course of one simulation, after a burn-in period to ensure stationarity, we record all species abundances every 30 minutes of simulation time, corresponding to roughly 54 (212) generations in the generalists (specialists) scenario, for a total of 30,000 generations of simulation runtime in either scenario. An extinction event occurs when a species with abundance 1 is chosen for a death event. Because in our model the state $N_i = 0$ is not allowed, multiple extinction events may

occur to the same species. Also, multiple abundance recordings may occur between successive extinction events, such that when an extinction eventually occurs, it will be linked to all previously recorded abundances since the last extinction. The time intervals between abundance recordings were sufficiently large to ensure independent data points, and the total runtime was long enough to observe extinctions in species with a wide range of initial abundances.

Finally, we test for neutral behavior at shorter timescales by comparing temporal fluctuations in species abundance against neutral predictions. For a subcritical stochastic birth-death process representing neutral dynamics, Chisholm and O'Dwyer (2014) showed that the probability that a species has abundance n at time t given that at time zero its abundance was n_0 is

$$\begin{aligned} P(n,t|n_0) = &\left(\frac{1}{\omega(t)\Omega(t)+1}\right)^{-n} \\ &\times (\Omega(t)(1+\omega(t)\Omega(t)))^{-n_0}\binom{n+n_0-1}{n} \\ &\times {}_2F_1(1-n_0,-n_0,1-n_0-n;(1-\Omega(t))(1+\omega(t)\Omega(t))), \end{aligned} \tag{9}$$

where ${}_2F_1$ is the hypergeometric function, $\omega(t) = \exp(-\nu t)$, and $\Omega(t) = \left(\frac{1}{\omega(t)-1}\right)\frac{1-\nu}{\nu}$, where ν is the speciation rate, equal to the difference between the per capita mortality and birth rates in the subcritical stochastic process. We use expression (9) to calculate the expected variance across histories of the stochastic process of the quantity $D(t) = \frac{n(t)-n_0}{\sqrt{n_0}}$ (Kalyuzhny et al., 2014), against which we compare our numerical observations. We use an ensemble of 100 histories, sampling the communities every generation for up to 100 generations. A species is considered to have zero abundance upon extinction, and we calculate the ensemble variance of $D(t)$ across species with the same n_0, and then average across all species with $n_0 \geq 20$ (the cutoff was necessary because low-n_0 species displayed noisy behavior at longer timescales, and would require a larger ensemble).

Results

In the generalist scenario, when the coefficient of variation (CV) in the consumption matrix is low (reflecting that consumers have similar preferences for all resources), the log-series distribution fits the SAD (Fig. 2A). However, as the CV of the consumption matrix increases, the probability that the SAD conforms to the LS distribution declines (Fig. 2B). For a community with 50 species and 50 resources, where the average species abundance is 500 and the average resource abundance is 100, a logistic regression indicates that the probability of a successful log-series fit drops below 50% when the CV of the consumption matrix is higher than 0.37. Defining this as the threshold for neutral-like patterns in the species abundance distribution, we note a power law be-

tween $CV^{\text{threshold}}$ and the mean species abundance n (Fig. 2C), with an exponent compatible with formula (5). This match supports the predictions of our threshold.

Results were analogous in the specialists scenario. Communities with low C_d/C_o ratio, indicating a small-magnitude difference between preferences for main and secondary resources, conform to the log-series distribution, while communities with sufficiently high C_d/C_o reject it (Fig. 2D). Indeed, the probability that the log-series distribution fits the SAD decreases as C_d/C_o increases (Fig. 2E). In communities with 50 species and 50 resources, with mean species and resource abundances at 100, the threshold ratio for a successful log-series fit is 4.0. This threshold increases linearly with the number of species in the community (Fig. 2F), in agreement with formula (6).

To compare the propensity for neutral-like abundance pattern across the generalist and specialist scenarios, we use a non-neutrality index applicable to both, based on the average cosine between the vectors representing the resource preference profiles of different species (i.e. the columns of the consumption matrix). The average cosine across all species pairs represents the average similarity between resource preference profiles. We therefore define the non-neutrality index as NNI $= 1 - \cos$. The limit of complete neutrality, i.e. identical resource preferences among all species, corresponds to a cosine of 1, and thus NNI $= 0$. The opposite limit of complete niche differentiation, i.e. where each species consumes a single resource with zero overlap, corresponds to a cosine of 0 (NNI $= 1$). Fig. 3 shows that in both the generalist and specialist scenarios, the probability of neutral-like abundance distribution is close to 100% in the neutral limit, as expected, but remains positive as we deviate from neutrality. The probability of rejecting the log-series only falls below 50% when the NNI is as high as 0.15 ± 0.05 in the specialist scenario, and 0.22 ± 0.03 in the generalist scenario (Fig. 3). For a given NNI, the abundance distribution in the generalist scenario typically appears more neutral-like than in the specialist scenario.

Extinction time results are shown in Fig. 4. As expected, life expectancy increases with species abundance. In the generalist scenario, species extinction times in communities with low CV closely match predictions from the neutral model, whereas communities with sufficiently high CV depart from neutral predictions (Fig. 4A). When there is a poor match, the neutral model underpredicts the extinction times, especially for species with high abundance (Fig. 4A). Plotting observed extinction times against neutral predictions in communities with different CV reveals that higher CVs lead to increasingly poor matches to neutrality, especially for high-abundance species (Fig. 4B). Interestingly, observed extinction times seem linearly related to predictions regardless of the CV. The slope of this relationship increases with the CV (Fig. 4C), being close to 1 at low CV, indicating a perfect match to neutral predictions, and > 1 at higher CV. This indicates that species of high mean abundance have particularly long life expectancy beyond neutral expectations, suggesting that niche differentiation has a disproportionate stabilizing effect on common species.

The specialist scenarios showed analogous results to the generalist scenarios regarding matches

to neutrality or lack thereof, except that extinction times of high-abundance species saturate. This could be because the restoring force in this scenario increases very quickly for large deviations from equilibrium, so a species that fluctuates to high abundance almost immediately returns to its mean abundance, thus not significantly increasing its life expectancy. By contrast, the neutral case lacks a strong stabilizing force, so excursions towards high abundance tend to substantially increase time to extinction.

Non-neutral community dynamics also displayed neutral-like behavior at shorter timescales (Fig. 5). Temporal fluctuations in species abundances in the specialist scenario were indistinguishable from neutrality for up to 10 generations, even when the non-neutrality index was as high as 0.5. At longer timescales, niche stabilization tended to reduce the ensemble variance of $D(t) = \frac{n(t)-n_0}{\sqrt{n_0}}$ relative to neutrality, as expected. However, it was only at maximum non-neutrality (NNI = 1), corresponding to fully specialized and therefore non-interacting species, that var(D) was immediately distinguishable from neutral, plateauing within a few generations.

Notably, the neutral-like behavior observed here applies to communities whose resource preference profiles would be easily distinguishable from true neutrality, if they were directly measured. In other words, our predictions and results are not solely for 'small' departures from neutrality, and yet we still observe neutral-like static and temporal patterns. To quantify this, we consider a neutral community of same richness and size as in Fig. 3. If all consumption events under true neutral dynamics are observed over the course of an average species lifetime (in this case, approximately 250 generations), the 95% confidence interval for the average pairwise cosine is estimated to lie between 0.992 and 0.996 (see Supplementary Fig. S.9), well above the threshold for 50% rejection of the log-series abundance distribution. This means communities that are unmistakably non-neutral under direct observations of consumption events would still look neutral under summary distributions like the SAD or extinction times. Of course, in most if not all cases it is unfeasible to directly observe consumption events, especially over such long timescales. This impracticality is one reason why ecologists have indeed tested theories of community assembly using summarizing distributions.

Discussion

We showed that communities of competing species may appear neutral despite differences in resource preferences that confer niche stabilization, and this behavior is quite predictable from summarized properties of the consumer-resource network. Neutral-like outcomes in non-neutral communities arise both in snapshots of community structure and species dynamics over short and long timescales. Viewed in one light, our results demonstrate that typical summarizing distribu-

tions are limited in their ability to detect non-neutrality. Viewed in another, our results show that neutral-like outcomes emerge from highly non-neutral consumer-resource dynamics, thus providing a potential explanation for the successes of neutral theory as a description of macroecological pattern.

Our simulation results qualitatively supported our hypothesis that a community undergoing non-neutral stochastic consumer-resource dynamics will be indistinguishable from a neutral community when the relaxation time to equilibrium in the stabilized system is comparable to the timescales of drift to extinction under neutral dynamics. Specifically, when differences in resource preferences fall below a threshold leading to commensurability of these timescales, one cannot easily reject the log-series abundance distribution, and species life expectancies match neutral predictions. Furthermore, at generational timescales, even full knowledge of species abundances over time may not suffice to reject neutrality. This has implications for inferring niche differentiation in real life systems, such as distinct metabolic syndromes in microbial communities, where the differences between consumers may not be large enough to overcome the influence of drift.

The neutrality threshold herein defined is a useful construct for testing our analytical predictions, not a methodological prescription for field ecologists. Indeed, the threshold's numerical value depends on the choice of cutoff for the probability of rejecting the log-series. In reality, this probability increases gradually rather than abruptly with increasing departures from neutral resource preferences. Furthermore, one's ability to reject neutral pattern in a community of interest depends not only on similarities between neutral and non-neutral outcomes, but also on the statistical power of goodness-of-fit tests. As such, in order to test our analytical predictions, the threshold's numerical value is less meaningful than how it scales with changes in parameters such as community size and richness. We verified that this scaling is robust to different cutoff choices, and presumably also to different tests and methodological approaches for rejecting neutrality.

The fact that abundances in specialist communities are easier to distinguish from neutrality than in generalist communities with similar niche overlap suggests that specializing towards a single resource has a stronger niche-like impact on abundance distributions than unstructured variation in resource preferences. It follows that departures from neutrality are not all equal with regards to impact on abundances, leaving open the possibility that some special structure in the consumption matrix may very quickly lead to detectable differences from neutral pattern. One method to test this possibility would be to infer the consumption matrix by fitting our consumer-resource model to experimental abundance measurements from monoculture experiments on each resource. Then, different consumption matrix structures could be investigated by choosing which consumers to include in community experiments. Although they did not infer a consumption network, Goldford et al. 2018's microbial community assembly experiments found that the resulting community structure was highly variable at the species level, while highly predictable at the family level. Our theoretical results suggest a possible interpretation of these findings. When viewed at the species

level, consumers do not differ enough to strongly affect the outcome of community assembly, suggesting that these species differences are below our analytical threshold and drift is the primary driver of abundance dynamics. On the other hand, consumer abundances converge to nearly deterministic outcomes at the family level, suggesting that between-family differences in resource profiles are well above our threshold.

The consumption matrix could be far more complex than the two cases considered here, because it captures the diverse metabolic strategies of consumers from different environments and evolutionary histories. For example, groups of the columns of C could be highly correlated when the corresponding groups of consumers perform similar functional roles in the community. The community as a whole could also exhibit a stronger preference for certain resources over others, introducing correlations between the rows of C. We ignored these complications for the sake of analytical tractability, but our results are likely generalizable: it is the bipartite structure of consumer-resource models, and not the specific structure of C, that separates the spectrum of our deterministic model into consumer and resource bulks. Because it is this separation that causes long excursions from the mean abundance in the stochastic model, our results suggest that a transition from neutral to niche dynamics will be a feature of consumer-resource models with more complex consumption preferences than considered here.

References

Abramowitz, M. and Stegun, I. A. (1965). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation.

Al Hammal, O., Alonso, D., Etienne, R. S., and Cornell, S. J. (2015). When Can Species Abundance Data Reveal Non-neutrality? *PLOS Computational Biology*, 11(3):e1004134.

Alonso, D., Etienne, R. S., and McKane, A. J. (2006). The merits of neutral theory. *Trends in Ecology & Evolution*, 21(8):451–7.

Alonso, D. and McKane, A. J. (2004). Sampling Hubbell's neutral theory of biodiversity. *Ecology Letters*, 7(10):901–910.

Armstrong, R. A. and McGehee, R. (1976). Coexistence of Species Competing for Shared Resources. *Theoretical Population Biology*, 9:317–328.

Armstrong, R. A. and McGehee, R. (1980). Competitive Exclusion. *The American Naturalist*, 115(2):151–170.

Arnold, T. B. and Emerson, J. W. (2010). *cvm.test: Discrete Cramer von Mises Goodness-of-Fit Tests*.

Barabás, G., Pásztor, L., Meszéna, G., and Ostling, A. M. (2014). Sensitivity analysis of coexistence in ecological communities: theory and application. *Ecology Letters*, 17:1479–1494.

Bell, G. (2001). Neutral macroecology. *Science (New York, N.Y.)*, 293(5539):2413–8.

Caswell, H. A. L. (1976). Community Structure: A Neutral Model Analysis. *Ecological Monographs*, 46(3):327–354.

Chave, J. (2004). Neutral theory and community ecology. *Ecology Letters*, 7(3):241–253.

Chave, J., Muller-Landau, H. C., and Levin, S. (2002). Comparing classical community models: theoretical consequences for patterns of diversity. *The American Naturalist*, 159(1):1–23.

Chisholm, R. A. and O'Dwyer, J. P. (2014). Species ages in neutral biodiversity models. *Theoretical Population Biology*, 93:85–94.

Chisholm, R. A. and Pacala, S. W. (2010). Niche and neutral models predict asymptotically equivalent species abundance distributions in high-diversity ecological communities. *Proceedings of the National Academy of Sciences of the United States of America*, 107(36):15821–5.

Clark, J. S. (2009). Beyond neutral science. *Trends in ecology & evolution*, 24(1):8–15.

Clark, J. S. (2012). The coherence problem with the Unified Neutral Theory of Biodiversity. *Trends in Ecology & Evolution*, 27(4).

Clark, J. S. and McLachlan, J. S. (2003). Stability of forest biodiversity. *Nature*, 423(6940):635–638.

D'Andrea, R. and Ostling, A. (2017). Biodiversity maintenance may be lower under partial niche differentiation than under neutrality. *Ecology*, 98(12):3211–3218.

Etienne, R. S., Alonso, D., and McKane, A. J. (2007). The zero-sum assumption in neutral biodiversity theory. *Journal of theoretical biology*, 248(3):522–36.

Fisher, C. K. and Mehta, P. (2014). The transition between the niche and neutral regimes in ecology. *Proceedings of the National Academy of Sciences*, 111(36):13111–13116.

Fisher, R., Corbet, A., and Williams, C. (1943). The Relation Between the Number of Species and the Number of Individuals in a Random Sample of an Animal Population. *Journal of Animal Ecology*, 12(1):42–58.

Fung, T., O'Dwyer, J. P., Kassim, A. R., Fletcher, C. D., and Chisholm, R. A. (2016). Reproducing static and dynamic biodiversity patterns in tropical forests : the critical role of environmental variance. *Ecology*, 97(5):1207–1217.

Goldford, J. E., Lu, N., Bajic, D., Estrela, S., Tikhonov, M., Sanchez-Gorostiaga, A., Segre, D., Mehta, P., and Sanchez, A. (2018). Emergent simplicity in microbial community assembly. *Science*, 361(August):469–474.

Gotelli, N. J. and McGill, B. J. (2006). Null versus neutral models: what's the difference? *Ecography*, 29(5):793–800.

Haegeman, B. and Loreau, M. (2011). A mathematical synthesis of niche and neutral theories in community ecology. *Journal of theoretical biology*, 269(1):150–65.

Harte, J. (2003). Tail of death and resurrection. *Nature*, 424(6952):1006–1007.

Harte, J. (2011). *Maximum entropy and ecology: a theory of abundance, distribution, and energetics*. OUP Oxford.

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, NJ.

Jansen, V. A. and Sigmund, K. (1998). Shaken not stirred: On permanence in ecological communities. *Theoretical Population Biology*, 54(3):195–201.

Kalyuzhny, M., Seri, E., Chocron, R., Flather, C. H., Kadmon, R., and Shnerb, N. M. (2014). Niche versus Neutrality: A Dynamical Analysis. *The American Naturalist*, 184(4):439–446.

Marčenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(457).

McGill, B. J. (2003). Strong and weak tests of macroecological theory. *Oikos*, 102(3):679–685.

McGill, B. J., Etienne, R. S., Gray, J. S., Alonso, D., Anderson, M. J., Benecha, H. K., Dornelas, M., Enquist, B. J., Green, J. L., He, F., Hurlbert, A. H., Magurran, A. E., Marquet, P. a., Maurer, B. a., Ostling, A. M., Soykan, C. U., Ugland, K. I., and White, E. P. (2007). Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology letters*, 10(10):995–1015.

McGill, B. J., Maurer, B. A., and Weiser, M. D. (2006). Empirical evaluation of neutral theory. *Ecology*, 87(6):1411–1423.

O'Dwyer, J. P., Kembel, S. W., and Sharpton, T. J. (2015). Backbones of evolutionary history test biodiversity theory for microbes. *Proceedings of the National Academy of Sciences*, 112(27):201419341.

Ostling, A. M. (2012). Large-scale spatial synchrony and the stability of forest biodiversity revisited. *Journal of Plant Ecology*, 5(1):52–63.

Pigolotti, S., Flammini, A., Marsili, M., and Maritan, A. (2005). Species lifetime distribution for simple models of ecologies. *Proceedings of the National Academy of Sciences of the United States of America*, 102(44):15747–15751.

Posfai, A., Taillefumier, T., and Wingreen, N. S. (2017). Metabolic Trade-Offs Promote Diversity in a Model Ecosystem. *Physical Review Letters*, 118(January):1–5.

Rael, R. C., D'Andrea, R., Barabás, G., and Ostling, A. (2018). Emergent niche structuring leads to increased differences from neutrality in species abundance distributions. *Ecology*.

Rosindell, J., Cornell, S. J., Hubbell, S. P., and Etienne, R. S. (2010). Protracted speciation revitalizes the neutral theory of biodiversity. *Ecology Letters*, 13(6):716–727.

Rosindell, J., Hubbell, S. P., He, F., Harmon, L. J., and Etienne, R. S. (2012). The case for ecological neutral theory. *Trends in Ecology & Evolution*, 27(4):204–209.

Rosindell, J., Hunter, M. O., and Etienne, R. S. (2011). The unified neutral theory of biodiversity and biogeography at age ten. *Trends in ecology & evolution*, 26(7):340–8.

Saavedra, S., Rohr, R. P., Bascompte, J., Godoy, O., Kraft, N. J. B., and Levine, J. M. (2017). A structural approach for understanding multispecies coexistence. *Ecological Monographs*, 87(3):470–486.

Scheffer, M. and van Nes, E. H. (2006). Self-organized similarity, the evolutionary emergence of groups of similar species. *Proceedings of the National Academy of Sciences of the United States of America*, 103(16):6230–6235.

Schreiber, S. J. and Tobiason, G. A. (2003). The evolution of resource use. *Journal of Mathematical Biology*, 47:56–78.

Volkov, I., Banavar, J. R., Hubbell, S. P., and Maritan, A. (2007). Patterns of relative species abundance in rainforests and coral reefs. *Nature*, 450(6197):45–49.

Volkov, I., Banavar, J. R., Hunter, M. O., and Maritan, A. (2003). Neutral theory and relative species abundance in ecology. *Nature*, 424(6952):1035–7.

Volkov, I., Banavar, J. R., Maritan, A., and Hubbell, S. P. (2004). The Stability of Forest Biodiversity - Reply. *Nature*, 427(6976):696–697.

Wang, S., Chen, A., Fang, J., and Pacala, S. W. (2013). Why abundant tropical tree species are phylogenetically old. *Proceedings of the National Academy of Sciences*, 110(40):16039–16043.

Wootton, J. T. (2005). Field parameterization and experimental test of the neutral theory of biodiversity. *Nature*, 433:309–312.

Figures

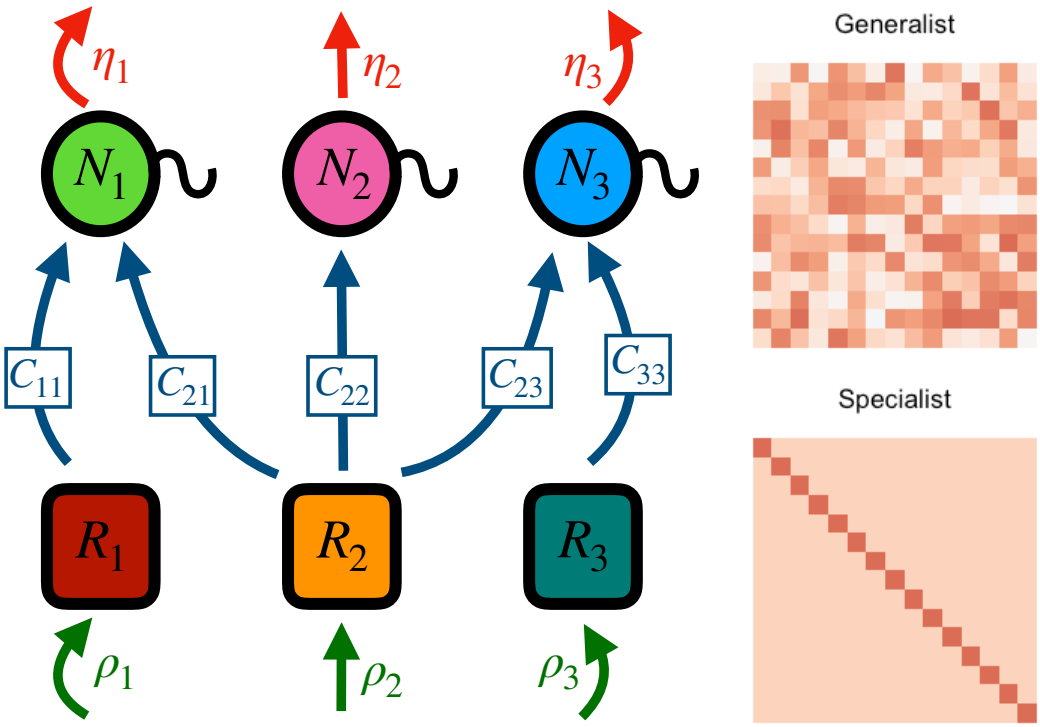

Figure 1: The left panel is a schematic of the models in (1) and (2). Resources flow into the system at rate ρ_i, while consumers undergo density dependent mortality at rates $\eta_i N_i$. The arrows connecting resources to consumers represent non-zero consumption coefficients C_{ij}. In this example, all consumers are capable of consuming resource 2, so there is community-wide competition for it. At the same time, consumer 1 is the only consumer that can utilize resource 1 and, similarly, consumer 3 is the only consumer that can deplete resource 3. The right panels are examples of the C matrices resulting from our two different parametrizations. In the generalist case, each entry is an independent sample from the same probability distribution with positive support. In the specialist case, consumer i can consumer resource i more quickly than any other resource. Other than this special resource, consumers consume every other resource at the same rate.

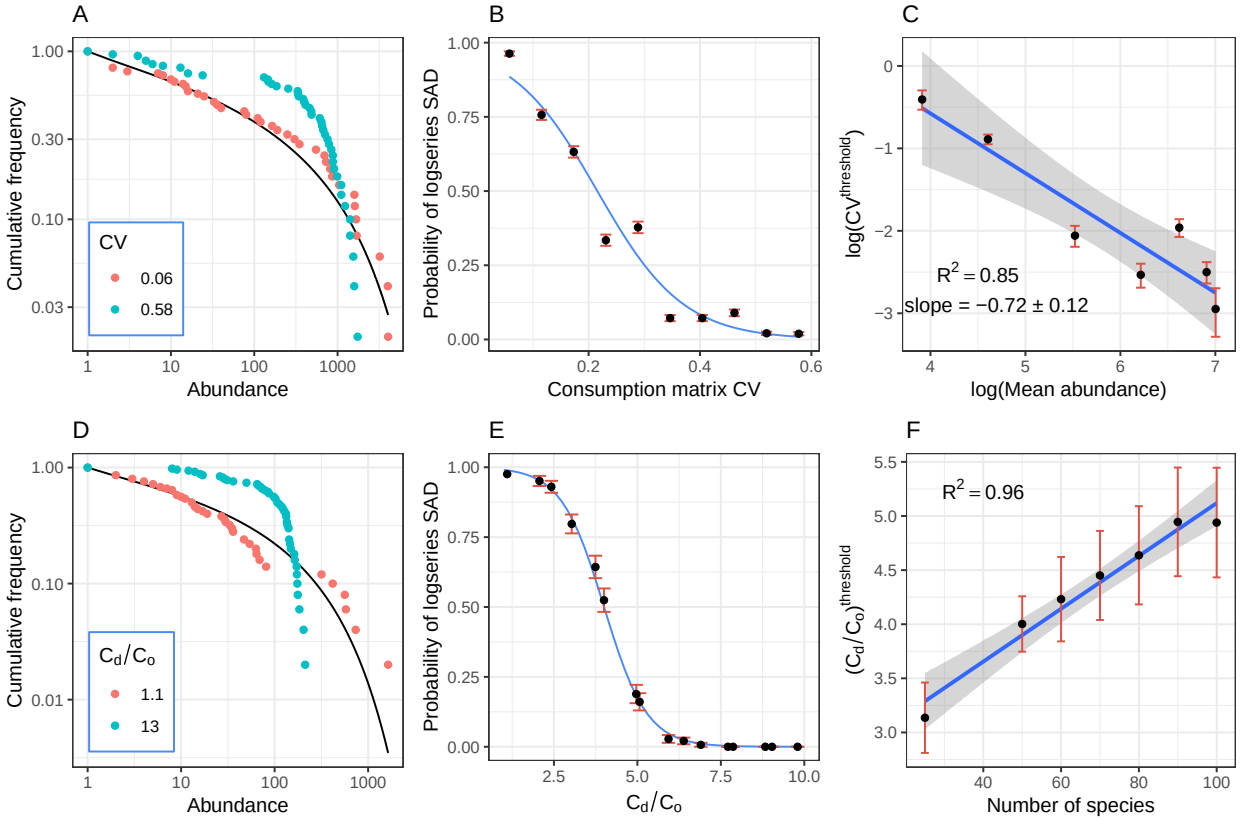

Figure 2: Species abundance distribution (SAD) results for the generalists scenarios (A-C) and specialist scenarios (D-F). **A**: In the generalist scenario, the log-series distribution fits the SAD when the coefficient of variation (CV) in the consumption matrix is low (red points; Cramér-von Mises goodness-of-fit test p-value 0.7), but is rejected when the CV is sufficiently high (blue points; CvM test p-value < 0.001). **B**: Probability that the logs-series distribution fits the SAD decreases with the CV of the consumption matrix. Points and error bars show the mean and standard error of the count of successful fits, out of an ensemble of 143 communities. Blue curve shows logistic regression. The threshold CV, defined as the point where the probability falls below 50%, is $CV^{\text{threshold}} = 0.37$. **C**: $\log(CV^{\text{threshold}})$ has a linear relationship with $\log(n)$, with slope -0.61 ± 0.09. This indicates a power law between $CV^{\text{threshold}}$ and n, with an exponent compatible with our analytic prediction of -0.5. Error bars show uncertainty propagated from the standard errors of the fitted parameters in the respective logistic regressions. Bands show the 95% CI of the linear regression. **D**: In the specialists scenario, communities with low C_d/C_o ratio (red points; CvM test p-value 0.267) conform to the LS distribution, while communities with sufficiently high C_d/C_o (blue points; CvM test p-value < 0.001) reject the LS distribution. **E**: The probability of the LS distribution fitting the SAD decreases as C_d/C_o increases, with the threshold at $C_d^{\text{threshold}} = 4.0$ (we set $C_o = 1$). Each data point summarizes an ensemble of 143 communities. **F**: $C_d^{\text{threshold}}$ increases linearly with the number of species in the community, in qualitative agreement with formula (6). Parameters: A-C: $K = S = 50$, $n = 500$, $r = 100$. D-F: $K = S = 50$, $n = r = 100$.

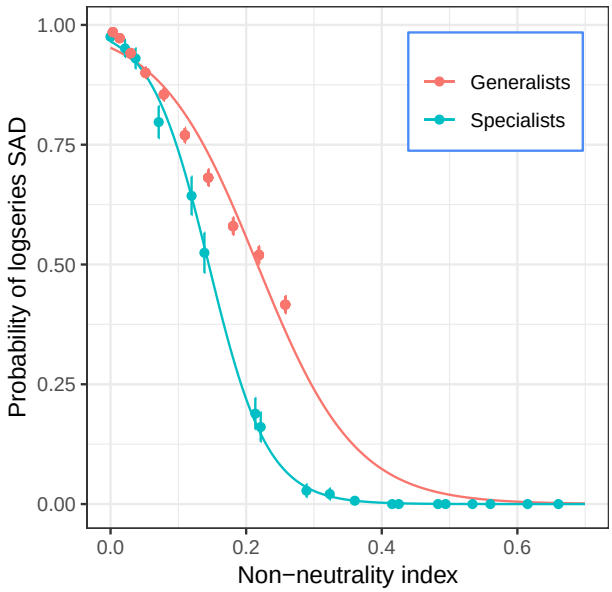

Figure 3: Probability of log-series SAD plotted against a non-neutrality index defined as NNI $= 1 - \cos$, where $\cos$ is the cosine between vectors representing species resource preferences, averaged across all species pairs in the community. Complete neutrality would correspond to NNI $= 0$ (i.e. $\cos = 1$), reflecting full overlap in resource preferences. For the same NNI value, communities in the generalist scenario are typically more likely to conform to the log-series distribution than communities in the specialist scenario. Parameters: $S = K = 50$, $n = r = 100$.

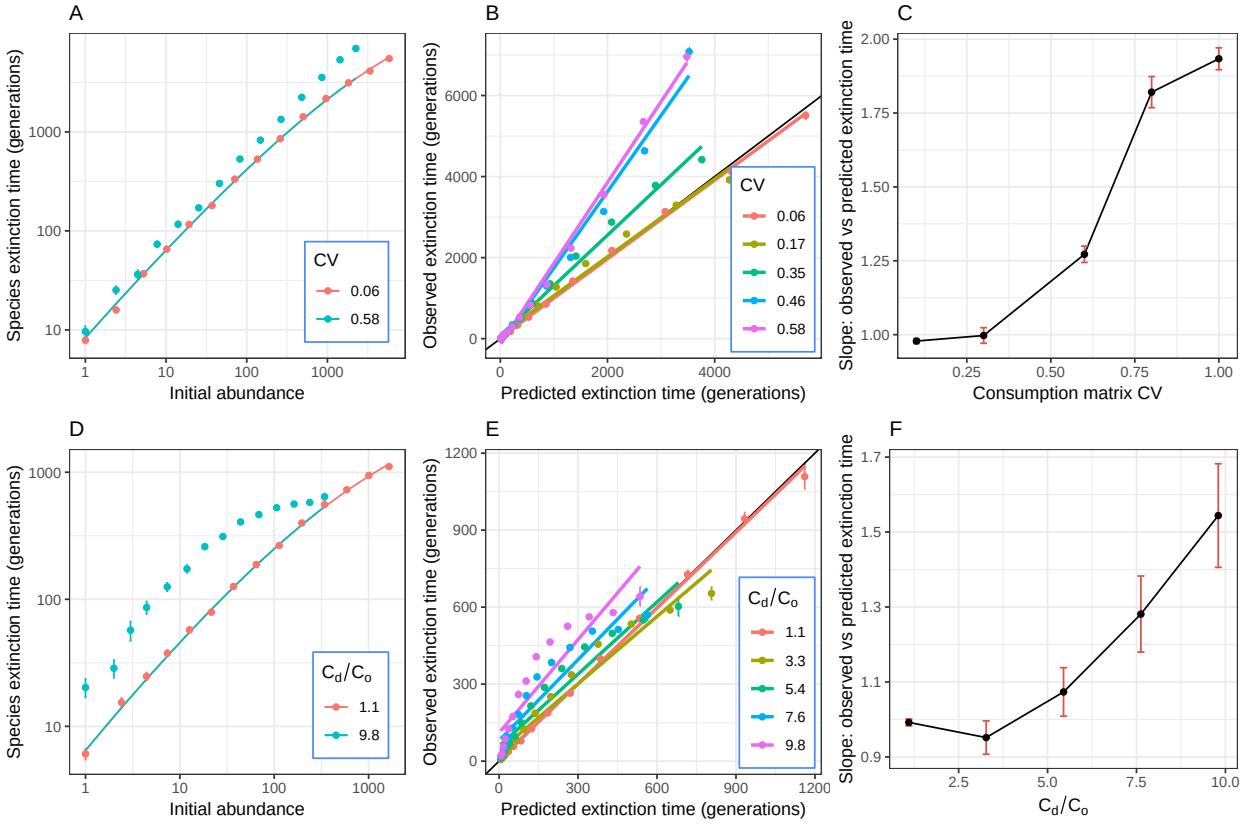

Figure 4: Extinction time results for the generalist (A-C) and specialist (D-F) scenarios. **A**: Points and error bars show average and standard errors of extinction times for species in logarithmically binned abundance categories. Curves show neutral predictions. In the generalist scenario, species extinction times match predictions from the neutral model in communities with low CV, but consistently exceed neutral predictions in communities with high CV, especially for high-abundance species. **B**: Plotting observed versus predicted extinction times in communities with different CV (colors) reveals that those with low CV conform closely to the neutral predictions (black line illustrates a perfect match), while higher CVs lead to increasingly poor matches to neutrality, especially for high-abundance species. Note that extinction times seem linearly related to predictions regardless of the CV. **C**: The slope of this relationship increases with the CV, being close to 1 (perfect match to neutral predictions) at low CV and > 1 at higher CV. **D-F**: Results in the specialist scenarios are analogous to the generalist scenarios, except that extinction times of high-abundance species saturate. Parameters: A-C: $K = S = 50$, $n = 500$, $r = 100$. D-F: $K = S = 50$, $n = r = 100$. Summary statistics were obtained from ca. 5,000 to 20,000 data points for each abundance bin in the generalists scenario, and 1,000 to 7,000 data points in the specialists scenario.

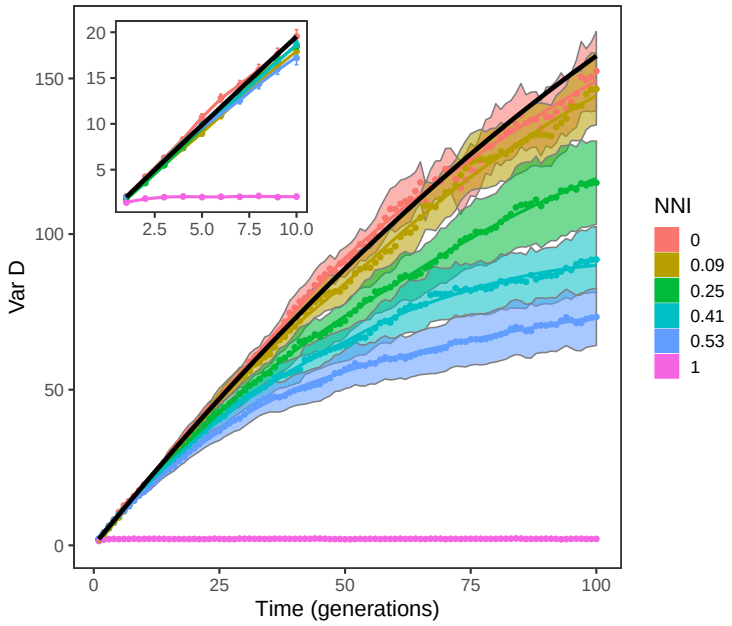

Figure 5: Variance in species abundances over time, for different parametrizations of the specialist scenario. Vertical axis plots the variance across histories of the stochastic process of $D(t) = \frac{n(t)-n_0}{\sqrt{n_0}}$, which is then averaged across species with different initial abundances n_0 (bands show standard error of the mean). Colors show parametrizations with increasing non-neutrality index (NNI), with lines showing the loess regression with smoothing parameter set to 1. Black line shows neutral prediction. Inset highlights similarity of all curves except NNI = 1 at timescales up to 10 generations.