Quantization of ecological interactions yields insights into food web assembly and dynamics

Justin D. Yeakel[a,b], **Mathias Pires**[c], **James O'Donnell**[d], **Marcus de Aguiar**[c], **Paulo R. Guimarães Jr.**[e], **Dominique Gravel**[f], **and Thilo Gross**[g]

[a]School of Natural Sciences, University of California Merced, Merced, CA 95343, USA; [b]Santa Fe Institute; [c]Universidade Estadual de Campinas; [d]University of Washington; [e]Universidade de São Paulo; [f]Universitè de Sherbrooke; [g]University of California Davis, Davis CA

In this paper we investigate the effects of heterogeneity in food available on the survival times of foraging individuals, using a simple model to incorporate the basic elements of time-stationary foraging process.

Ecosystem assembly | Ecological networks | Colonization | Extinction | Ecosystem engineering

Simplifying the abundant complexity and idiosyncrasies of nature is necessary to unravel their secrets. The layers of natural history giving rise to an ecological community can be simplified to the set of species from which it is composed and the multifarious set of interactions binding those species, and their populations, together. This minimalist perspective can provide insight into community structure, and consequently, the potential dynamics of species' populations and the fortunes of the system as a whole. Community structure can directly impact the ability of the system to absorb external perturbations, the existence and ubiquity of tipping points that can cause sudden changes in species' populations, and contributes to species' extinction risk. Importantly, this community structure is the result of a dynamical process over time, where species are added to or removed from the community. This process is generally referred to as community assembly, and though it is an engine for much that we observe in nature, it is not well understood.

Generally, a single interaction type is used to describe both the structure and dynamics of a community. For example, food webs are composed of trophic interactions between species, whereas mutualistic networks are composed of service-resource or service-service mutualisms between species. More recently, there has been a growing interest in understanding community structure by taking into account many types of interactions (**?**).

The structure of communities can be described by the number and types of organisms from which they are composed, organismal physiological constraints (Hempson, Science 2015) and functional traits (McGill et al. 2006), external environmental conditions (ref), and the structure of species interactions (Dunne 2002). Although these structures can be described statistically (Diamond, Williams and Martinez), such structures are products of a dynamic assembly process that includes but is not limited to colonization and extinction (MacArthur and Wilson, Simberloff), and evolutionary change (Gillespie, Rominger). Understanding the underlying mechanisms driving community assembly is necessary to know A dynamic perspective of communities, taking into account not only the structure of these systems, but the processes from which such structures arise (MacArthur and Wilson, Diamond), serves to provide a mechanistic understanding of ecological communities. Consequently, an understanding of community structure cannot be uncoupled from its assembly (Simberloff) and disassembly (Yeakel), via colonization and extinction, respectively.

Though theoretical examinations of community dynamics Many interaction: trophic to mutualistic.

Absent from ecological network theory is any consideration of how species presence within a community impacts the local environment. Engineering...

Here we...

Diverse interactions without engineers

Community assembly in the absence of engineers reveals the emergence of food web and mutualistic network properties consistent with observations of assembling and steady state empirical systems. Because only primary producers that do not have outgoing need interactions can colonize initially, a diverse base of autotrophic species typically constitutes the early assembly process. In order for communities to have > 1 pure autotroph, we do not consider competitive exclusion of the basal resource, such that all non-mutualistic pure autotrophs have the potential to co-exist. Following the establishment of a suite of autotrophs, both mixotrophs and lower trophic-level heterotrophs assemble into the community (Fig. **??**a). As species richness increases, available resources accumulate consumers and competitive exclusion leads to an increase in the extinction rate until a steady state is reached at $S_A^* = 130$ species. This community steady state increases as the number of mutualisms established in the source pool decreases (lower $\mathrm{E}\{f_n\}$) because mutualisms introduce dependencies that

www.pnas.org/cgi/doi/10.1073/pnas.XXXXXXXXXX

PNAS | February 5, 2019 | vol. XXX | no. XX | 1–4

inhibit colonization.

As the community assembles, we find that the connectance of trophic interactions ($C = L/S^2$) where in our case $L = \sum_{i,j} \mathrm{e}_{\boldsymbol{A}(i,j)}$ and $S = \mathcal{S}_{\boldsymbol{A}}$) follows a decay-like trajectory to values similar to – but on average 9% greater than – the connectance of the source pool $\boldsymbol{P}$ (Fig. **??**b). Decaying connectance has been documented in the assembly of mangrove communities (Piechnik), however this decay is a statistical inevitability, as a growing food web early in the assembly process inevitably has high link density (few species that are nearly fully connected), and over the course of increasing species richness and the establishment of different trophic levels and compartments, a decline in link density. That the connectance of assembled communities is greater than the source pool is due to the fact that only species connected by trophic interactions can enter the community to begin with, increasing expected link density compared to the overall pool.

Recent empirical work has indicated that generalist species may play an important role early in community assembly, whereas specialists tend to colonize after a diverse resource base has accumulated. Because the definition of a specialist or generalist to some degree depends on the size and connectance of the larger food web, trophic generality can be defined as $G_i = \sum_j \mathrm{e}_{\boldsymbol{A}(i,j)}(L/S)^{-1}$, such that the number of trophic interactions for a consumer is scaled by the average number of trophic interactions per species in the community L/S (Piechnik, others). A species is classified as a generalist if the number of its trophic interactions is greater than the average number of links per species, or $G_i > 1$, and a specialist if $G_i < 1$, where a community can be described by the proportion of specialists found therein.

For interaction networks that are assembling over time, generality can be scaled by a number of different measures of L/S, and this has a large effect on our interpretation of the role of generality in community assembly. For instance, L/S may be quantified by either including all autotrophic species or only autotrophic functional groups. Furthermore, the scaling of generality may be made with respect to the current state of the community at each point in time, or with respect to the community at steady state. For instance, in their investigation of assembling mangrove food webs (originally described by Simberloff, xxx), Piechnik et al. (2008) scaled trophic breadth to a standard steady state value of $L^*/S^* = 0.2$ averaged across 102 food webs. To examine how our assessment of the role of generalism over the course of assembly changes based on the application of different scalings, we employ three different measures of L/S to calculate G_i: 1) G_i^{all}, where L accounts for all links in the food web and S accounts for all species relative to each time interval in the assembly process (circles; Fig. 3b); 2) G_i^{hetero}, where we consider only the links and species richness of heterotrophs, excluding autotrophs (points; Fig. 3b); 3) G_i^*, where L and S are measured with respect to the communities at steady state, which is most similar to the
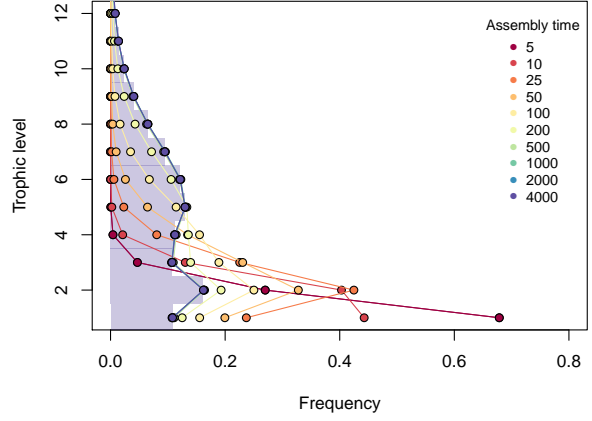

Fig. 1. The frequency distribution of trophic levels over the course of assembly taken across 10^4 replicates. Autotrophs occupy trophic level 1. We set $S = 200$, $p_e = 0.01$, $p_n = 0.002$, and $t = 4000$ (see methods).

measure used to evaluate assembling mangrove food webs (diamonds; Fig. 3b).

Whether trophic breadth is scaled to the current state of L/S or the steady state value of L^*/S^* has a large influence on the estimated proportion of generalists in the community, particularly when the size of the system is small. We observe that for G_i^{all}, the system is initially assembled by specialist species, though over the course of assembly the proportion of specialists relative to generalists declines to intermediate values (circles representing the average over replicates in Fig. 3). If only the trophic links between non-autotrophs are considered as in G_i^{hetero}, specialists still dominate early in assembly, but there is a greater range, such that some systems can be described by a mixed proportion of specialists and generalists (individual points representing independent replicates in Fig. 3). If generalism is measured with respect to the steady state L^*/S^* as in G_i^*, we observe that generalists dominate early in assembly, with an increase in specialists as assembly progresses (diamonds representing the average over replicates in Fig. 3). At steady state, all measures of L/S are approximately equivalent, and the proportion of specialists levels out at ca. 60%, which is similar to the empirical observations for Simberloff's mangrove communities in Piechnick et al. (2008).

The role of specialists early in assembly is primarily due to the accumulation of autotrophs specializing on the basal resource. This is evident when we observe that the trophic level distribution early in assembly is peaked at the lowest trophic level (trophic level 1). Four trophic levels are typically established by $t = 50$, where colonization is still the dominant dynamic, and by the time communities reach steady state the interaction networks are characterized by ca. 10 trophic levels (Fig. 1). The

distribution of trophic levels changes shape over the course of assembly: early on, we observe that the community exhibits a pyramidal structure, where the vast majority of species inhabit low-trophic positions. At steady state, we observe that intermediate trophic levels (2-7) dominate, with frequencies that reveal an hour-glass structure. We emphasize that these structures are diversity-weighted rather than biomass or abundance-weighted as is often the case (Trebilco et al. 2013, Gibert & Yeakel 2019). Trophic levels higher than 7 do occur, but are increasingly rare.

Because the ENIgMa framework includes multiple types of interactions, we must examine whether structures characteristic of mutualistic networks are observed. Empirical observations of mutualistic systems reveal that such interactions tend to be nested (where specialist interactions are subsets of generalist interactions). Increasing the frequency of need interactions increases the frequency of both service-resource ($e \leftrightarrow n$) and service-service ($n \leftrightarrow n$) mutualisms. In the ENIgMa framework, the probability of competitive exclusion is reduced with an increase in the prevalence mutualistic interactions, and this should lead more stable nested bipartite motifs over the course of assembly (see Supplementary Information). Our expectation then, is that nestedness should increase with the frequency of mutualisms, though this is difficult to predict a priori. As we increase the frequency of need interactions in the source pool, we indeed observe an increase in nestedness (measured as NODF; Fig. 2). That the absolute values of nestedness are low compared to those measured for empirical mutualistic networks is unsurprising: observations of mutualistic interactions are generally for bipartite networks and isolated to specific systems (e.g. ant-plant mutualisms). Here the NODF metric is taken across both eat and need interactions acroos the entire assembled community.

The role of engineers
Beyond engineering

Materials and Methods

We model an ecological system with a network where nodes represent *ecological entities* such as populations of species and or the presence of inanimate objects affecting species such as (examples). Following Pillai (REF), we do not track the abundances of entities but only track their presence or absence. The links of the network represent interactions between pairs of entities (x,y). We distinguish three types of such interactions: x eats y, x needs y to be present, x makes object y.

The model is initialized by creating S species and $O = \eta S$ objects, such that $N = S + O$ is the total number of entities and η is the number of objects per species in the system. For each pair of species (x,y) there is a probability p_e that x eats y and probability p_n that x needs y. For each pair of species x and object o, there is a probability q_e that x eats o and a probability q_n that species x needs object o.. Additionally, each species makes a number of objects that is drawn from a Poisson distribution with mean $\mu = \eta e(e-1)^{-1}$ where e is Euler's number. Once the number of objects per species is determined, each object is assigned to a species independently. This means that there multiple species may make the same object, and that may be some objects that are not made by any species.

In addition to interactions with ecosystem entities, there can be interactions with a basal resource, which is always present. The
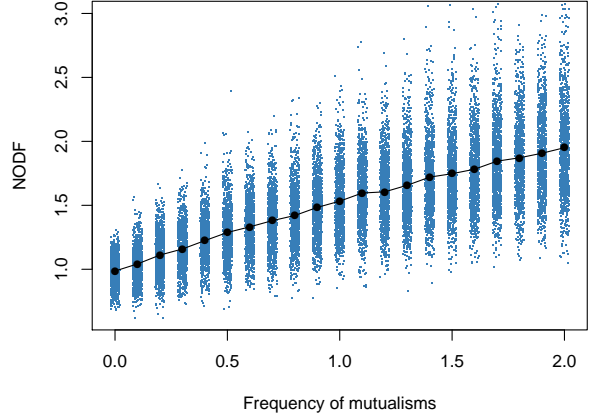
first species always eats this resource, such that there is always a primary producer in the pool. Other species eat the basal resource with probability p_e.

We then consider the assembly of a community which at any time will contain a subset of entities in the pool and always the basal resource. In time, the entities in the community are updated following a set of rules. A species from the pool can colonize the community if the following conditions are met: 1) all entities that a species needs are present in the community, and 2) at least one entity that a species eats is present in the community. If a colonization event is possible, it occurs stochastically in time with rate $r_{\rm c}$.

An established species is at risk of extinction if it is not the strongest competitor at least one of its resources that it eats. We compute the competitive strength of species i as

$$\sigma_i = c_{\rm n} n_i - c_{\rm e} e_i - c_{\rm v} v_i, \qquad [1]$$

where n_i is the number of entities that species i needs, e_i is the number of entities from the pool that species i can eat, and v_i is the number of species in the community that eat species i. This captures the ecological intuition that mutualisms provide a fitness benefit, specialists are stronger competitors than generalists, and many predators entail an energetic cost. The coefficients $c_{\rm n}$, $c_{\rm e}$, $c_{\rm v}$ describe the relative effects of these contributions to competitive strength. In the following, we use the values $c_{\rm n} = \pi$, $c_{\rm e} = \sqrt{2}$, $c_{\rm v} = 1$, such that the competitive benefit of adding an additional mutualism is greater than the detriment incurred by adding another prey or predator. A species at risk of extinction leaves the community stochastically in time at rate r_e.

An object is present in the community whenever at least one species that makes the object is present. If a species that makes an object colonizes a community, the object is created immediately, however objects may persist for some time after the last species that makes the object goes extinct. Any object that has lost all of its makers disappears stochastically in time at rate r_o.

The model described here can be simulated efficiently with an event-driven simulation utilizing a Gillespie algorithm. In these types of simulations, one computes the rates r_j of all possible events j in a given step. One then selects the time at which the next event happens by drawing a random number from an exponential distribution with mean $1/\sum_j r_j$. At this time, an event occurs that is randomly selected from the set of possible events such that the probability of event a is $r_a/\sum_j r_j$. Then the effect of the event is realized and the list of possible events is updated for the next step. This algorithm is known to offer a much better approximation to the true stochastic continuous time

Parameter	Definition	Value/Range
$\overrightarrow{a}$	assimilate	
$\overrightarrow{n}$	need	
$\overrightarrow{i}$	ignore	
$\overrightarrow{m}$	make	
$e \leftrightarrow i$	Asymmetric consumption	$p_{ei} = p_i(p_e/(p_e + p_n + p_i)) + p_e(p_i/(p_a + p_i + p_n))$
$e \leftrightarrow e$	Symmetric consumption	$p_{ee} = p_e(p_e/(p_i + p_n + p_e))$
$e \leftrightarrow n$	Trophic mutualism	$p_{en} = p_n(p_e/(p_e + p_n + p_i + p_m)) + p_e(p_n/(p_a + p_i + p_n))$
$n \leftrightarrow n$	Non-trophic mutualism	$p_{nn} = p_n(p_n/(p_e + p_n + p_i + p_m))$
$n \leftrightarrow i$	Commensalism	$p_{ni} = p_n(p_i/(p_e + p_n + p_i + p_m)) + p_i(p_n/(p_e + p_n + p_i))$
$m \leftrightarrow n$	Engineering	$p_{mn} = p_n(p_m/(p_e + p_n + p_i + p_m)) + p_m$
$i \leftrightarrow i$	Null	$p_{ii} = p_i(p_i/(p_e + p_n + p_i))$
$\mathcal{N}$	Number of species + objects	dyn.
$\mathcal{S}$	Number of species	dyn.
$\mathcal{O}$	Number of objects	dyn.

Table 1. Table of parameters, definitions, and assigned values or ranges.
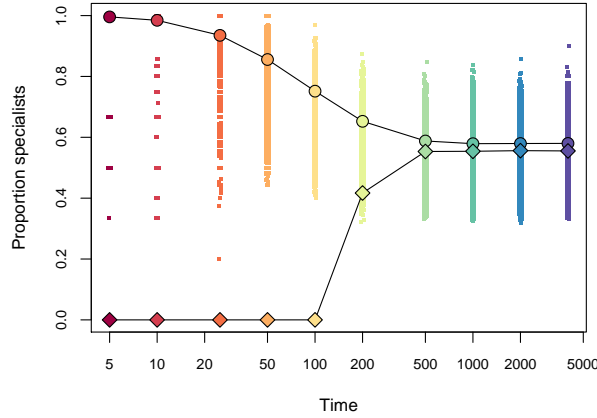

Fig. 3. The proportion of specialists as a function of assembly time, where a specialist is defined as a species with a generality index $G_i < 1$. All measures of G_i are scaled by the average number of links per species L/S, and we consider different values of L/S on G_i: Circles: G_i^{all} where L accounts for all links in the food web and S accounts for all species relative to each time interval in the assembly process (averaged across replicates); Points: G_i^{hetero}, where we consider only the links and species richness of heterotrophs, excluding autotrophs (each point shows an individual replicate); Diamonds: G_i^*, where L and S are measured with respect to the communities at steady state, which is most similar to the measure used to evaluate assembling mangrove food webs (averaged across replicates).

process than a simulation in discrete time steps, while providing a much higher numerical efficiency (Gillespie, 1974ish).
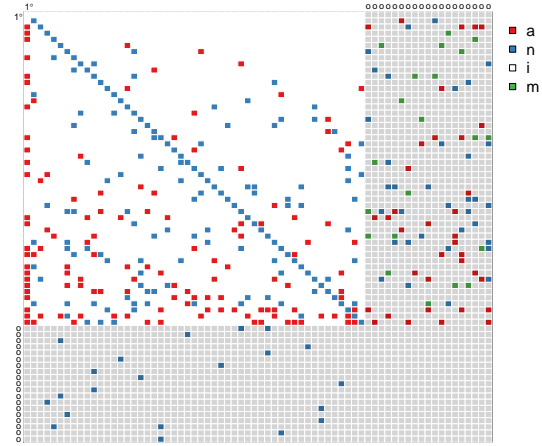

Fig. 4. An example of the source pool interaction matrix where $\mathcal{S} = 50$. Species and objects are aligned across the rows and columns; objects are shaded and labeled by 'o' to distinguish them from species. The interaction recorded in row i and column j describes the directed interaction from species/object i to species/object j. The first row/column represents the basal resource; species that assimilate the primary resource are capable of primary production. Species interact with other species and/or objects; objects only interact with their engineers by 'needing' them; objects do not interact with other objects.

4