

Diverse interactions and ecosystem engineering stabilize community assembly

Justin D. Yeakel,^{1,2} Mathias M. Pires,³ Marcus A. M. de Aguiar,³ James L. O'Donnell,⁴ Paulo R. Guimarães Jr.,⁵ Dominique Gravel,⁶ and Thilo Gross⁷

¹University of California, Merced, Merced, CA 95340, USA

²Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

³Universidade Estadual de Campinas, Campinas - SP, Brazil

⁴University of Washington, Seattle, WA 98195, USA

⁵Universidade de São Paulo, São Paulo, Brazil

⁶Université de Sherbrooke, Sherbrooke, QCJ1K0A5, Canada

⁷University of California, Davis, Davis, CA 95616, USA

The complexity of an ecological community can be distilled into a network, where diverse interactions connect species in a web of dependencies. Species interact not only with each other but indirectly through environmental effects, however the role of these ecosystem engineers has not yet been considered in models of ecological networks. Here we explore the dynamics of ecosystem assembly, where the colonization and extinction of species within a community depends on the constraints imposed by trophic, service, and engineering dependencies. We show that our assembly model reproduces many key features of ecological systems, such as the role of generalists during assembly, realistic maximum trophic levels, and increased nestedness with higher frequencies of mutualisms. We find that ecosystem engineering has large and nonlinear effects on extinction rates, facilitating robustness by creating niche space, but at the same time increasing the magnitude of extinction cascades. We emphasize the importance of redundancies in engineered effects and show that such redundancy lowers the barriers to colonization, promoting community diversity. Together, our results suggest that ecological engineers may enhance community diversity while increasing persistence by facilitating colonization and limiting competitive exclusion.

To unravel nature's secrets we must simplify its abundant complexities and idiosyncrasies. The layers of natural history giving rise to an ecological community can be distilled – among many forms – into a network, where nodes represent species and links represent interactions between them. Networks are generally constructed for one type of interaction, such as food webs capturing predation^{1–3} or pollination networks capturing a specific mutualistic interaction⁴, and continues to lead to significant breakthroughs in our understanding of the dynamical consequences of community structure^{5–7}, assembly⁸, and coevolution⁹. Recent interest in ‘multilayer networks’ comprising multiple interaction types (multiplex interactions) may provide additional insight into these processes^{10,11}. However, interactions where species affect others by altering the environment in a lasting way have not yet been incorporated into models of ecological networks. These interactions, known as ecosystem engineering^{12,13} or more generally niche construction^{14,15}, are quite common in nature and exist in almost every ecosystem.

Diverse interactions occur not only between species but indirectly through the effects that species have on their environment^{13,16,17}. Elephants root out large saplings and small trees, enabling the formation and maintenance of grasslands¹⁸ and creating habitat for smaller vertebrates¹⁹. Burrowing rodents create shelter and promote primary production by aerating the soil²⁰, salmon and aquatic invertebrates create freshwater habitats by changing stream morphology²¹, and leaf-cutter ants alter microclimates, influencing seedling survival and plant growth²². These examples illustrate ecosystem engineer-

ing, where the engineering organism alters the environment on timescales longer than its own²³.

Ecosystem engineering not only impacts communities on ecological timescales, but has profoundly shaped the evolution of life on Earth²⁴. For example, the emergence of multicellular cyanobacteria fundamentally altered the atmosphere during the Great Oxidation Event of the Proterozoic roughly 2.5 Byrs BP^{24,25}, paving the way for the biological invasion of terrestrial habitats. In the oceans it is thought that rRNA and protein biogenesis of aquatic photoautotrophs drove the nitrogen:phosphorous ratio (the Redfield Ratio) to ca. 16:1 matching that of plankton²⁶, illustrating that engineering clades can have much larger, sometimes global-scale effects.

The effect of the environment on species is commonly included in models of ecological dynamics^{27–29} due to its acknowledged importance and because it can – to first approximation – be easily systematized. By comparison the way in which species engineer the environment defies easy systematization due to the multitude of mechanisms by which engineering occurs. While interactions between species and the abiotic environment have been conceptually described^{17,30}, the absence of engineered effects in network models was addressed by Odling-Smeet et al.¹³, where they outlined a conceptual framework that included both species and abiotic compartments as nodes of a network, with links denoting both biotic and abiotic interactions.

Here we model the assembly of an ecological network where nodes represent ecological entities, including engineering species, non-engineering species, and the effects of the former on the environment, which we call abiotic

75 *modifiers*. The links of the network that connect both 133 species and modifiers represent trophic (*eat* interactions), 134 service (*need* interactions), and engineering dependen- 135 cies, respectively (Fig. 1; see Materials and Methods for 136 a full description). Trophic interactions represent both 137 predation as well as parasitism, whereas service interac- 138 tions account for all non-trophic interactions such as pol- 139 lination or seed dispersal. In our framework a traditional 140 mutualism (such as a plant-pollinator interaction) con- 141 sists of a service (need) interaction in one direction and 142 a trophic (eat) interaction in the other. These multitype 143 interactions between species and modifiers thus embed 144 multiple dependent ecological sub-systems into a single 145 network (Fig. 1). Modifiers in our framework overlap 146 conceptually with the ‘abiotic compartments’ described 147 in Odling-Smee et al.¹³. Following Pillai et al.³¹, we 148 do not track the abundances of biotic or abiotic enti- 149 ties but only track their presence or absence. We use 150 this framework to explore the dynamics of ecosystem as- 151 sembly, where the colonization and extinction of species 152 within a community depends on the constraints imposed 153 by the trophic, service, and engineering dependencies. 154 We then show how observed network structures emerge 155 from the process of assembly, compare their attributes 156 with those from empirical systems, and examine the ef- 157 fects of ecosystem engineers.

158 Our results offer four key insights into the roles of 159 multitype interactions and ecosystem engineering in 160 driving community assembly. First, we show that the 161 assembly of communities in the absence of engineering 162 reproduces many features observed in empirical systems. 163 These include changes in the proportion of generalists 164 over the course of assembly that accord with measured 165 data and trophic diversity similar to empirical observa- 166 tions. Second, we show that increasing the frequency 167 of mutualistic interactions leads to the assembly of 168 ecological networks that are more nested, a common 169 feature of diverse mutualistic systems³², but are also less 170 robust. Our third key result shows that increasing the 171 proportion of ecosystem engineers within a community 172 has nonlinear effects on observed extinction rates. While 173 we find that a low amount of engineering increases 174 extinction rates, a high amount of engineering has the 175 opposite effect. Finally we show that redundancies 176 in engineered effects promote community diversity by 177 lowering the barriers to colonization.

178 **Assembly without ecosystem engineering.** Com- 179 munities assemble by random colonization from a source 180 pool. A species from the source pool can colonize if it 181 finds at least one resource that it can consume (one *eat* 182 interaction is satisfied; cf. Ref. 33) and all of its non- 183 trophic needs are met (all *need* interactions are satisfied). 184 As such, the service interactions are assumed to be obli- 185 gate, whereas trophic interactions are flexible. Following 186 the establishment of an autotrophic base, the arrival of 187 mixotrophs and lower trophic heterotrophs create oppor- 188 tunities for organisms occupying higher trophic levels to 189

190 invade. This expanding niche space initially serves as an 191 accelerator for community growth.

192 Following the initial colonization phase, extinctions 193 begin to slow the rate of community growth. Primary 194 extinctions occur by the competitive exclusion of 195 species sharing similar resources. A species’ competition 196 strength is determined by its interactions: competition 197 strength is enhanced by the number of need interactions 198 and penalized by its trophic generality (number of prey) 199 and vulnerability (number of predators). Secondary ex- 200 tinctions occur when species lose its last trophic or any 201 of its service requirements. See Fig. 1D, E for an il- 202 lustration of the assembly process. As the colonization 203 and extinction rates converge, the community reaches a 204 steady state around which it oscillates (Fig. 2A). See 205 Materials and Methods and Supplementary Appendix 1 206 for a complete description of the assembly process.

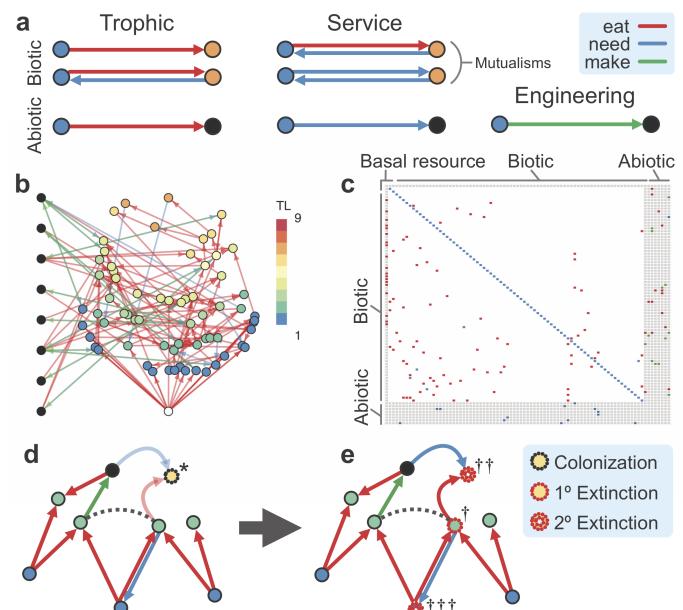


Figure 1. **a**, Multitype interactions between species (colored nodes) and abiotic modifiers (black nodes). **b**, An assembling food web with species (colored nodes; color denotes trophic level, TL) and modifiers (black nodes). The basal resource is the white node rooted at the bottom of the network. **c**, The corresponding adjacency matrix with colors denoting interactions between biotic (species) and abiotic (modifiers) entities. **d**, A species (*) can colonize a community when a single trophic and all service requirements are met. **e**, Greater vulnerability increases the risk of primary extinction via competitive exclusion (competition denoted by dashed line) to species (†). The extinction of species (†) will cascade to affect those connected by trophic (††) and service (†††) dependencies.

193 Assembly of ecological communities in the absence of 194 engineering results in interaction networks with struc- 195 tures consistent with empirical observations. As the com- 196 munity reaches steady state, we find that the connectance 197 of trophic interactions ($C = L/S^2$, where S is species 198 richness and L is the number of links) decays to a value 199

similar to that of the source pool (Fig. S1). Decay-
ing connectance has been documented in the assembly of
mangrove communities³⁴, however this decay is a statis-
tical inevitability, as early in the assembly process small
food webs will have high link density, from which it can
only decline. Compared to trophic networks constructed
using the Niche model³⁵ given similar species richness
and connectance, our framework results in networks with
degree distributions of similar means but with reduced
variance (Supplementary Appendix 2).

Recent empirical work has suggested that general-
ist species may predominate early in assembly, whereas
specialists colonize after a diverse resource base has
accumulated^{33,34}. Here the trophic generality of species
 i is defined $G_i = k_i^{\text{in}}/(L/S)$ where k_i^{in} is the in-degree, or
number of resources consumed, by species i ³⁵. A species
is classified as a generalist if $G_i > 1$ and a specialist if
 $G_i < 1$. If generalism is scaled to the steady state link
density (see Supplementary Appendix, section III), we
observe that generalists dominate early in assembly, with
an increase in specialists as assembly progresses (Fig.
2B). This confirms expectations from the trophic theory
of island biogeography³³, where early communities with
lower richness are less likely to support specialist con-
sumers than late, species-rich communities. At steady
state the proportion of specialists is ca. 56%, similar to
empirical observations of assembling food webs³⁴.

The role of specialists early in assembly is primarily
due to the accumulation of autotrophic specialists. This
is evident when we observe that the trophic level (TL)
distribution early in assembly ($t = 5$) has an average
 $\text{TL} = 1.6$. Four trophic levels are typically established
by $t = 50$, where colonization is still dominant, and by
the time communities reach steady state the interaction
networks are characterized by an average $\text{TL}_{\text{max}} (\pm \text{stan-}$
 $\text{dard deviation}) = 11 \pm 2.8$ (Fig. 2C). While the maxi-
mum trophic level is higher than that measured in most
predator-prey systems³⁶, it is not unreasonable if par-
asitic interactions (which we do not differentiate from
other consumers) are included³⁷. Overall, the most com-
mon trophic level among species at steady state is ca.
 $\text{TL} = 4.75$.

The distribution of trophic levels changes shape over
the course of assembly. Early in assembly, we observe
a skewed pyramidal structure, where most species feed
from the base of the food web. At steady state, we
observe that intermediate trophic levels dominate, with
frequencies taking on an hourglass structure (purple
bars, Fig. 2C). Compellingly, the trophic richness
pyramids that we observe at steady state follow closely
the hourglass distribution observed for empirical food
webs and are less top-heavy than those produced by
static food web models³⁸.

Structure and dynamics of mutualisms. Nested in-
teractions, where specialist interactions are subsets of
generalist interactions, are a distinguishing feature of
mutualistic networks³². A nested structure has been

shown to maximize the structural stability of mutu-
alistic networks³⁹, emerge naturally via adaptive for-
aging behaviors^{40,41} and neutral processes⁴², and pro-
mote the influence of indirect effects in driving coevo-
lutionary dynamics⁹. While models and experiments of
trophic networks suggest that compartmentalization con-
fers greater stabilizing properties^{43,44}, interaction asym-
metry among species may promote nestedness in both
trophic⁴⁵ and mutualistic systems⁴⁶. Processes that op-
erate on different temporal and spatial scales may have a
significant influence on these observations⁴⁷. For exam-
ple, over evolutionary time, coevolution and speciation
may degrade nested structures in favor of modularity⁴⁸,
and there is some evidence from Pleistocene food webs
that geographic insularity may reinforce this process⁴⁹.

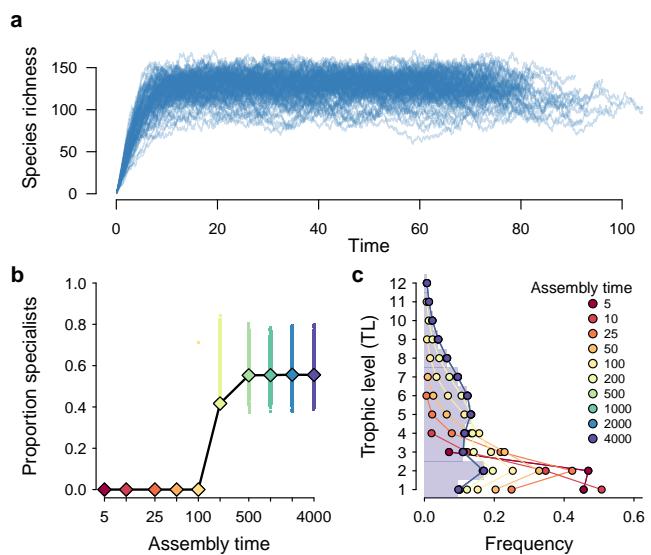


Figure 2. **a**, Assembling communities over time from a pool of 200 non-engineering species. Steady state species richness is reached by $t = 250$. **b**, The proportion of specialists as a function of assembly time (iterations), where a specialist is defined as a species with a generality index $G_i < 1$. All measures of G_i are scaled by the average number of links per species where L and S are measured at steady state. **c**, The frequency distribution of trophic levels as a function of assembly time (iterations). Autotrophs occupy $\text{TL} = 1$. Measures were evaluated across 10^4 replicates; see Materials and Methods for parameter values.

Does the assembly of ecological networks favor nest-
edness when mutualistic interactions are frequent? In-
creasing service dependencies (*need* interactions; see Fig.
1) promotes both service-resource and service-service de-
pendencies. Consider how species with more service in-
teractions compare to those with fewer. More service
i) increase a species' competition strength,
lowering its primary extinction risk while also *ii*) increas-
ing inter-species dependencies and its secondary extinc-
tion risk. While mutualisms convey fitness advantages
in order to evolve, the latter highlights the potential risk

associated with losing mutualistic partners^{50,51}. Indeed, the balance that mutualists must maintain with their partners may have large implications for the future of global biodiversity⁵².

We find that as we increase the frequency of mutualistic interactions, the assembled community at steady state becomes more nested (Fig. 3). In this case, nestedness emerges from the assembly process and provides structural robustness. The robustness can be observed by examining the exclusionary differences between species in a simple nested motif (Fig. 3, inset). In the trophic motif, species with high vulnerability (multiple predators) are at greater risk of primary extinction via competitive exclusion. This will result in the secondary extinction of the specialist consumer, rendering the nested structure prone to change. In our framework mutualistic networks are generally formed by composite interactions, where the consumer species is engaged in a trophic interaction while the resource species is engaged in a service interaction. As such, the consumer species becomes a trophic partner and the resource species gains the competitive advantages of the service. If the competitive advantages of services are greater than the costs of vulnerability (see Materials and Methods), it is the low vulnerability species with fewer trophic partners that is at greater risk of exclusion (Fig. 3, inset). Because its elimination will not cascade, the nested structure will be more resistant to change.

Our results also suggest that the addition of mutualistic interactions comes at a cost to the assembling community. Because mutualisms increase dependencies between species, they also increase the frequency of secondary extinctions (Fig. 3). Measuring persistence in terms of the proportion of time species are established in the network reveals that more frequent mutualisms leads on average to lower persistence. At the community scale, lower average persistence implies greater species turnover. Observations of empirical systems appear to support our model predictions. For example, assembling plant-pollinator systems have demonstrated high rates of species and interaction turnover, both during the assembly process and at the steady state⁸.

We emphasize that we have restricted ourselves to examining the effects of obligate mutualisms, although the importance of non-obligate mutualisms has long been recognized^{8,40,41,54,55}. While the inclusion of non-obligate mutualists will lower the likelihood of cascading effects in systems with higher frequencies of service interactions, the loss of obligate mutualistic partners will have larger dynamic consequences than the loss of more flexible non-obligate mutualistic partners. As such, we do not expect inclusion of non-obligate mutualisms to alter the qualitative nature of our findings.

Assembly with ecosystem engineering. The concept of ecosystem engineering, or more generally niche construction, has both encouraged an extended evolutionary synthesis⁵⁶ while also garnering considerable controversy^{57,58}. Models that explore the effects of

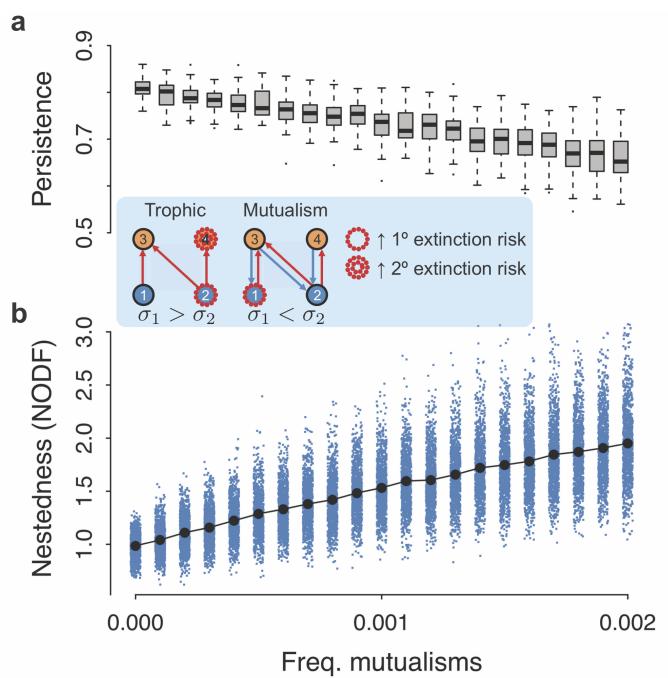


Figure 3. **a**, Species persistence with increasing frequencies of mutualistic (service) interactions without ecosystem engineers. **b**, Structural nestedness of communities, measured as NODF (Nestedness based on Overlap and Decreasing Fill)⁵³. Measures were evaluated across 10^4 replicates; see Materials and Methods for parameter values. Inset: A trophic and mutualistic nested motif for resource species 1, 2 and consumer species 3, 4. Trophic motif: the vulnerable species 2 is subject to primary extinction because it has a lower competition strength σ , resulting in an extinction cascade of species 4. Mutualistic motif: the least vulnerable species 1 with fewer mutualistic partners is subject to primary extinction without cascading effects.

ecosystem engineering are relatively few, but have covered important ground^{13,23}. For example, engineering has been shown to promote invasion⁵⁹, alter primary productivity⁶⁰, and change the selective environment over eco-evolutionary timescales^{61,62} which can lead to unexpected outcomes such as the fixation of deleterious alleles⁶³. On smaller scales, microbiota construct shared metabolic resources that have a significant influence on microbial communities⁶⁴, the dynamics of which may even serve as the missing ingredient stabilizing some complex ecological systems⁶⁵.

We next explore the effects of ecosystem engineering by allowing species to produce abiotic modifiers as additional nodes in the ecological network (Fig. 1). These modifier nodes produced by engineers can serve to fulfill resource or service requirements for other species. The parameter η defines the mean number of modifiers produced per species, drawn from a Poisson distribution (see Materials and Methods for details). Increasing the frequency of engineering interactions both increases the

318 number of engineering species (those species making ≥ 1 modifier) and the number of modifiers per species. There
 319 are two characteristics of engineering that have particular
 320 relevance for community assembly: modifiers can linger
 321 in the community even after the species that produce
 322 them have been excluded, and more than one engineer
 323 can produce the same modifier such that engineering re-
 324 dundancies increase with η (Supplementary Appendix 1).
 325

326 Increasing engineering has significant consequences for
 327 community robustness, but these effects also are sensi-
 328 tive to the frequency of service interactions within the
 329 community. We measure community robustness by *i*)
 330 rates of primary versus secondary extinctions, *ii*) species
 331 persistence, and *iii*) steady state community diversity.
 332 All measures were averaged over each species within the
 333 community across assembly time.

334 As the number of engineers increase, mean rates of
 335 primary extinction are first elevated and then decline
 336 (Fig. 4A). This nonlinear effect of engineering on rates
 337 of primary extinction results from two competing forces.
 338 Increased production of abiotic modifiers supplies con-
 339 sumers with additional resources, limiting secondary ex-
 340 tinctions and promoting persistence (Fig. 4B, C). How-
 341 ever, the stabilization of consumers ultimately results in
 342 increased vulnerability of prey. Engineering dependen-
 343 cies are considered rare when the ratio of modifier nodes
 344 per species is $0 < \eta \leq 0.5$. The cumulative effect in these
 345 species-rich/modifier-poor systems is increased compet-
 346 itive exclusion of prey and higher rates of primary ex-
 347 tinction (Fig. 4A). Notably the presence of even a small
 348 number of engineers serves to limit the magnitude of sec-
 349 ondary extinction cascades. Higher rates of primary ex-
 350 tinction coupled with lower rates of secondary extinction
 351 mean that extinctions are common, but of limited magni-
 352 tude such that disturbances are compartmentalized. As
 353 engineering becomes common ($\eta > 0.5$) the available
 354 niche space expands, lowering competitive overlap and
 355 suppressing both primary and secondary extinctions.

356 Increasing the frequency of service interactions pro-
 357 motes service interactions between species and engi-
 358 neered modifiers (Fig. 1). A topical example of the latter
 359 is the habitat provided to invertebrates by the recently
 360 discovered rock-boring teredinid shipworm (*Lithoredo*
 361 *abatanica*)⁶⁶. Here, freshwater invertebrates are serviced
 362 by the habitat modifications engineered by the shipworm,
 363 linking species indirectly via an abiotic effect (in our
 364 framework via a modifier node). As the frequency of
 365 service interactions increases, the negative effects asso-
 366 ciated with rare engineers is diminished (Fig. 4A). In-
 367 creasing service interactions both elevates the competi-
 368 tive strength of species receiving services (from species
 369 and/or modifiers), while creating more interdependen-
 370 cies between and among species. As trophic interactions
 371 are replaced by service interactions, previously vulnera-
 372 ble species gain a competitive foothold and persist (Fig.
 373 3, inset), lowering rates of primary extinctions (Fig. 4A).
 374 The costs of these added services to the community are
 375 an increased rate of secondary extinctions (Fig. 4B) and

376 higher species turnover (Fig. 4C). Low rates of primary
 377 extinction coupled with high rates of secondary extinc-
 378 tion mean that extinctions are less common but lead to
 379 larger cascades.

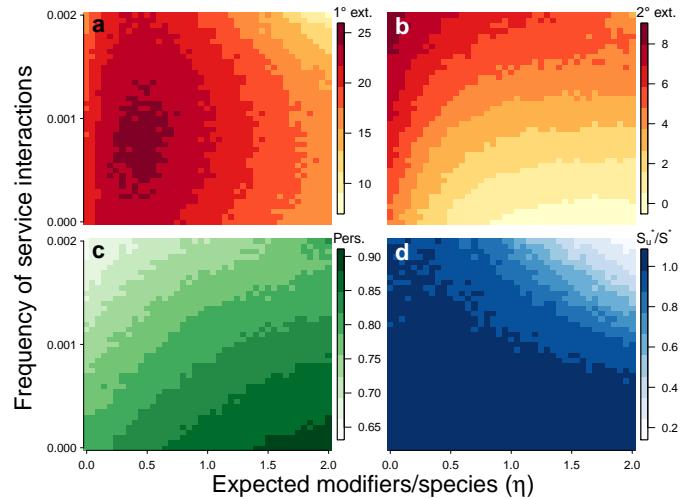


Figure 4. Community robustness as a function of the frequency of service interactions and modifiers per species. **a**, Mean rates of primary extinction, where primary extinctions occur from competitive exclusion of consumers over shared resources. **b**, Mean rates of secondary extinction, which cascade from primary extinctions. **c**, Mean species persistence. **d**, The ratio S_u^*/S^* , where S_u^* denotes steady states for systems where all engineered modifiers are unique to each engineer, and S^* denote steady states for systems with redundant engineering. Lower values of S_u^*/S^* mean that systems with redundant engineers have higher richness at the steady state than those without redundancies. Measures were evaluated across 50 replicates; see Materials and Methods for parameter values.

380 While the importance of engineering timescales has
 381 been emphasized previously²³, redundant engineering
 382 has been assumed to be unimportant¹². We argue that
 383 redundancy may be an important component of highly
 384 engineered systems, and particularly relevant when there
 385 exists a positive feedback between the effects of engi-
 386 neers on their fitness⁵⁹. The vast majority of contem-
 387 porary ecosystem engineering case studies focus on sin-
 388 gle taxa, such that redundant engineers appear rare¹².
 389 However if we consider longer timescales, increasing di-
 390 versity of engineering clades may promote redundancy,
 391 and in some cases this may feed back to accelerate
 392 diversification¹⁴. Such positive feedback mechanisms
 393 likely facilitated the global changes induced by cyanobac-
 394 teria in the Proterozoic^{24,25} among other large-scale engi-
 395 neering events in the history of life²⁴. Engineering redun-
 396 dancies are likely important on shorter timescales as well.
 397 For example, diverse sessile epifauna on shelled gravels in
 398 shallow marine environments are facilitated by the engi-
 399 neering of their ancestors, such that the engineered effects
 400 of the clade determine the future fitness of descendants⁶⁷.
 401 In the microbiome, redundant engineering may be very
 402 common due to the influence of horizontal gene transfer

in structuring metabolite production⁶⁸. In these systems, redundancy in the production of shared metabolicic reources may play a key role in community structure and dynamics^{64,65}.

When there are few engineers, each modifier in the community tends to be unique to a particular engineering species. Engineering redundancies increase linearly with η (Supplementary Appendix, section I; Fig. S5), such that the loss of an engineer will not necessarily lead to the loss of engineered modifiers. We examine the effects of this redundancy by comparing our results to those produced by the same model, but where each modifier is uniquely produced by a single species. Surprisingly, the lack of engineering redundancies does not alter the general relationship between engineering and measures of community robustness (Fig. S6). However we find that redundancies play a central role in maintaining species diversity. When engineering redundancies are allowed, steady state community richness S^* does not vary considerably with increasing service interactions and engineering (Fig. S7A). In contrast, when redundant engineering is not allowed, steady state community richness S_u^* declines sharply (Figs. 4D, S7B).

Communities lacking redundancy have lower species richness because sparse interdependencies preclude colonization (Fig. S7C,D). Colonization occurs only when trophic and service dependencies are fulfilled. A species requiring multiple engineered modifiers, each uniquely produced, means that each required entity must precede colonization. This magnifies the role of priority effects in constraining assembly order¹⁵, precluding many species from colonizing. In contrast, redundancy increases the niche space available to species while minimizing priority effects by allowing multiple engineers to fulfill dependencies. Our results thus suggest that redundant engineers may play important roles in assembling ecosystems by lowering the barriers to colonization thereby promoting community diversity.

Together, the results of our model point to the importance of considering multitype interactions both between species and as mediated through changes to the environment via engineering. We suggest that including the effects of engineers, either explicitly as we have done here, or otherwise, is vital for understanding the inter-dependencies that define ecological systems. As past ecosystems have fundamentally altered the landscape on which contemporary communities interact, future ecosystems will be defined by the influence of engineering today. Understanding the role of ecosystem engineers is thus tantamount to understanding our own.

Methods

We model an ecological system with a network where nodes represent *ecological entities* such as populations of species and/or the presence of abiotic modifiers affecting species such as (examples). Following Pilai et al.³¹, 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 modifier y.

The assembly process entails two steps: first a source pool of species is created, followed by colonization/extinction into/from a local community. The model is initialized by creating S species and $M = \eta S$ modifiers, such that $N = S + M$ is the average total number of entities and η is the average number of modifiers 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 modifier m, there is a probability q_e that x eats m and a probability q_n that species x needs modifier m. Additionally, each species makes a number of modifiers that is drawn from a Poisson distribution with mean $\mu = \eta e / (e - 1)$ where e is Euler's number. Once the number of modifiers per species is determined, each modifier is assigned to a species independently. This means that multiple species may make the same modifier, and that there may be some modifiers that are not made by any species, which are eliminated from the pool.

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 . Species with zero assigned trophic interactions are assumed to be primary producers. See Supplementary Appendix 1 for additional details on defining the source pool.

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_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

$$\sigma_i = c_n n_i - c_e e_i - c_v v_i, \quad (1)$$

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_n , c_e , c_v describe the relative effects of these contributions to competitive strength. In the following, we use the values $c_n > c_e > c_v$, 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 .

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

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 $r_a / \sum_j r_j$. The effect of the event is then 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 process than a simulation in discrete time steps, while providing a much higher numerical efficiency⁶⁹. Simulations described in the main text have default parameterizations of $S = 200$, $p_e = 0.01$, $c_n = \pi$, $c_e = \sqrt{2}$, $c_v = 1$, and 4000 iterations.

532 Data availability

533 The study is theoretical; no new empirical data were generated.

534

535 Code availability

536 The simulation code supporting this work is available for download
537 from <https://github.com/jdyeakel/Lego>.

538

- 539** [1] Paine, R. T. Food web complexity and species diversity. **595** **158**, 99 – 107 (2012).
- 540** *Am. Nat.* **100**, 65–75 (1966).
- 541** [2] Dunne, J. A., Williams, R. J. & Martinez, N. D. Food-**596** **19**] Pringle, R. M. Elephants as agents of habitat creation for
542 web structure and network theory: the role of con-**597** small vertebrates at the patch scale. *Ecology* **89**, 26–33
543 nectance and size. *Proc Natl Acad Sci USA* **99**, 12917–**598** (2008).
- 544** 12922 (2002).
- 545** [3] Pascual, M. & Dunne, J. *Ecological Networks: Linking Structure to Dynamics in Food Webs* (Ox-**601** **20**] Reichman, O. & Seabloom, E. W. The role of pocket go-
546 ford University Press, Oxford, UK, 2006). URL **602** phers as subterranean ecosystem engineers. *Trends Ecol
547 <http://www.google.com/search?client=safari&rls=en-us&q=Ecological+Networks&ie=UTF-8&oe=UTF-8>. **603** Evol **17**, 44 – 49 (2002).*
- 548** [4] Bascompte, J. & Jordano, P. *Mutualistic Networks*. **604** **21**] Moore, J. W. Animal Ecosystem Engineers
549 Monographs in Population Biology (Princeton Univer-**605** in Streams. *BioScience* **56**, 237–246 (2006).
550 sity Press, Princeton, NJ, 2013). URL <https://books.google.com/books?id=xVWzAAAAQBAJ>. **606** arXiv:<https://oup.prod.sis.lan/bioscience/article-pdf/56/3/237/26897844/56-3-237.pdf>.
- 551** [5] May, R. M. Will a large complex system be stable? **607** **22**] Meyer, S. T., Leal, I. R., Tabarelli, M. & Wirth, R.
552 *Nature* **238**, 413–414 (1972).
- 553** [6] Gross, T., Levin, S. A. & Dieckmann, U. Generalized **608** **23**] Hastings, A. *et al.* Ecosystem engineering in space and
554 models reveal stabilizing factors in food webs. *Science* **609** time. *Ecol Lett* **10**, 153–164 (2007).
- 555** **325**, 747–750 (2009).
- 556** [7] Allesina, S. & Tang, S. Stability criteria for complex **610** **24**] Erwin, D. H. Macroevolution of ecosystem engineering,
557 ecosystems. *Nature* **483**, 205–208 (2012).
- 558** [8] Ponisio, L. C., Gaiarsa, M. P. & Kremen, C. Opportunis-**611** niche construction and diversity. *Trends Ecol Evol* **23**,
559 tic attachment assembles plant–pollinator networks. **612** 304 – 310 (2008).
- 560** *Ecol Lett* **20**, 1261–1272 (2017).
- 561** [9] Guimarães Jr, P. R., Pires, M. M., Jordano, P., Bas-**613** **25**] Schirrmeyer, B. E., de Vos, J. M., Antonelli,
562 compte, J. & Thompson, J. N. Indirect effects drive co-**614** A. & Bagheri, H. C. Evolution of multicel-
563 evolution in mutualistic networks. *Nature* **18**, 586 (2017).**615** lularity coincided with increased diversification

564 [10] Kéfi, S., Miele, V., Wieters, E. A., Navarrete, S. A. &**616** of cyanobacteria and the great oxidation event.
565 Berlow, E. L. How structured is the entangled bank?**617** *Proc Natl Acad Sci USA* **110**, 1791–1796 (2013).
566 the surprisingly simple organization of multiplex ecologi-**618** arXiv:<https://www.pnas.org/content/110/5/1791.full.pdf>.

567 cal networks leads to increased persistence and resilience.**619** **26**] Loladze, I. & Elser, J. J. The origins of the Red-
568 *PLoS Biol* **14**, e1002527 (2016).

569 [11] Pilosof, S., Porter, M. A., Pascual, M. & Kéfi, S. The **620** field nitrogen-to-phosphorus ratio are in a homoeostatic
570 multilayer nature of ecological networks. *Nature Ecol
571 Evol* **1**, 0101 (2017).

572 [12] Lawton, J. H. What do species do in ecosystems? **621** **27**] Woodward, G., Perkins, D. M. & Brown, L. E. Climate
573 *Oikos* **71**, 367–374 (1994).

574 [13] Odling-Smee, J., Erwin, D. H., Palkovacs, E. P., Feld-**622** change and freshwater ecosystems: impacts across mul-
575 man, M. W. & Laland, K. N. Niche construction theory:**623** tiple levels of organization. *Philos T Roy Soc B* **365**,
576 a practical guide for ecologists. *Q Rev Biol* **88**, 4–28 **624** 2093–2106 (2010).

577 (2013).

578 [14] Odling-Smee, F., Laland, K. & Feldman, M. *Niche Con-**625** **28**] Brose, U. *et al.* Climate change in size-structured ecosys-
579 struction: The Neglected Process in Evolution (MPB-37).**626** tems. *Philos T Roy Soc B* **367**, 2903–2912 (2012).*

580 Monographs in Population Biology (Princeton Univer-**627** [29] Gibert, J. P. Temperature directly and indirectly influ-
581 sity Press, Princeton, NJ, 2013). URL <https://books.google.com/books?id=Ok-NujtWMZOC>.**628** ences food web structure. *Sci Rep-UK* **9**, 5312 (2019).

582 [15] Fukami, T. Historical Contingency in Community As-**629** [30] Getz, W. M. Biomass transformation webs provide a
583 sembly: Integrating Niches, Species Pools, and Priority **630** unified approach to consumer-resource modelling. *Ecol
584 Effects. *Annu Rev Ecol Evol Syst* **46**, 1–23 (2015). **631** Lett **14**, 113–124 (2011).*

585 [16] Jones, C. G., Lawton, J. H. & Shachak, M. Organisms **632** [31] Pillai, P., Gonzalez, A. & Loreau, M. Metacommunity
586 as ecosystem engineers. *Oikos* **69**, 373–386 (1994). **633** theory explains the emergence of food web complexity.

587 [17] Olff, H. *et al.* Parallel ecological networks in ecosystems. **634** *Proc Natl Acad Sci USA* **108**, 19293–19298 (2011).

588 *Philos T Roy Soc B* **364**, 1755 (2009).

589 [18] Haynes, G. Elephants (and extinct relatives) as earth-**635** [32] Bascompte, J., Jordano, P., Melián, C. J. & Olesen,
590 movers and ecosystem engineers. *Geomorphology* **157**–**636** J. M. The nested assembly of plant-animal mutualis-
591 **650** **641**tic networks. *Proc Natl Acad Sci USA* **100**, 9383–9387
592 (2003).

593 [19] Gravel, D., Massol, F., Canard, E., Mouillot, D. & Mou-**642** [33] Piechnik, D. A., Lawler, S. P. & Martinez, N. D.
594 quet, N. Trophic theory of island biogeography. *Ecol Lett*
595 **14**, 1010–1016 (2011).

596 [20] Food-web assembly during a classic biogeographic study:
597 species “trophic breadth” corresponds to colonization order.
598 *Oikos* **117** (2008).

599 [21] Williams, R. J. & Martinez, N. D. Simple rules yield
600 complex food webs. *Nature* **404**, 180–183 (2000).

- 651 [36] Williams, R. & Martinez, N. Limits to trophic 715
 652 levels and omnivory in complex food webs: The 716
 653 ory and data. *Am Nat* **163**, 458–468 (2004). 717
 654 arXiv:<https://doi.org/10.1086/381964>. 718
- 655 [37] Lafferty, K. D., Dobson, A. P. & Kuris, A. M. Parasites 719
 656 dominate food web links. *Proc Natl Acad Sci USA* **103**, 720
 657 11211–11216 (2006). 721
- 658 [38] Turney, S. & Buddle, C. M. Pyramids of species rich- 722
 659 ness: the determinants and distribution of species diver- 723
 660 sity across trophic levels. *Oikos* **125**, 1224–1232 (2016). 724
- 661 [39] Rohr, R. P., Saavedra, S. & Bascompte, J. On the 725
 662 structural stability of mutualistic systems. *Science* **345**, 726
 663 1253497–1253497 (2014). 727
- 664 [40] Valdovinos, F. S. *et al.* Niche partitioning due to adaptive 728
 665 foraging reverses effects of nestedness and connectance on 729
 666 pollination network stability. *Ecol Lett* **19**, 1277–1286 730
 667 (2016).
- 668 [41] Valdovinos, F. S. Mutualistic networks: moving closer to 732
 669 a predictive theory. *Ecol Lett* **0** (2019).
- 670 [42] Krishna, A., Guimarães Jr, P. R., Jordano, P. & Bas- 734
 671 compte, J. A neutral-niche theory of nestedness in mu- 735
 672 tualistic networks. *Oikos* **117**, 1609–1618 (2008). 736
- 673 [43] Stouffer, D. B. Compartmentalization increases food-web 737
 674 persistence. *Proc Natl Acad Sci USA* **108**, 3648–3652 738
 675 (2011).
- 676 [44] Gilarranz, L. J., Rayfield, B., Liñán-Cembrano, G., Bas- 740
 677 compte, J. & Gonzalez, A. Effects of network modular- 741
 678 ity on the spread of perturbation impact in experimental 742
 679 metapopulations. *Science* **357**, 199–201 (2017).
- 680 [45] Araújo, M. S. *et al.* Nested diets: a novel pattern of 744
 681 individual-level resource use. *Oikos* **119**, 81–88 (2010). 745
- 682 [46] Pires, M. M., Prado, P. I. & Guimarães Jr, P. R. Do 746
 683 food web models reproduce the structure of mutualistic 747
 684 networks? *PLoS ONE* **6**, e27280 (2011).
- 685 [47] Massol, F. *et al.* Linking community and ecosystem dy- 749
 686 namics through spatial ecology. *Ecol Lett* **14**, 313–323 750
 687 (2011).
- 688 [48] Ponisio, L. C. *et al.* A network perspective for community 752
 689 assembly. *Front Ecol Evol* **7**, 103 (2019).
- 690 [49] Yeakel, J. D., Guimarães Jr, P. R., Bocherens, H. & 754
 691 Koch, P. L. The impact of climate change on the 755
 692 structure of Pleistocene food webs across the mammoth 756
 693 steppe. *Proc Roy Soc B* **280**, 20130239–20130239 (2013).
- 694 [50] Bond, W. J., Lawton, J. H. & May, R. M. Do mutualisms 758
 695 matter? Assessing the impact of pollinator and disperser 759
 696 disruption on plant extinction. *Phil Trans Roy Soc B* **344**, 760
 697 83–90 (1994).
- 698 [51] Colwell, R. K., Dunn, R. R. & Harris, N. C. Coextinction 762
 699 and persistence of dependent species in a changing world. 763
Ann Rev Ecol Evol Sys **43**, 183–203 (2012).
- 700 [52] Dunn, R. R., Harris, N. C., Colwell, R. K., Koh, L. P. 764
 702 & Sodhi, N. S. The sixth mass coextinction: are most 765
 703 endangered species parasites and mutualists? *Proc Roy 766
 704 Soc B* **276**, 3037–3045 (2009).
- 705 [53] Almeida-Neto, M., Guimarães, P., Guimarães Jr, P. R., 767
 706 Loyola, R. D. & Ulrich, W. A consistent metric for nest- 768
 707 edness analysis in ecological systems: reconciling concept 769
 708 and measurement. *Oikos* **117**, 1227–1239 (2008).
- 709 [54] Ramos-Jiliberto, R., Valdovinos, F. S., Moisset de Es- 770
 710 panA's, P. & Flores, J. D. Topological plasticity in- 771
 711 creases robustness of mutualistic networks. *J Anim Ecol* 772
 712 **81**, 896–904 (2012).
- 713 [55] Vieira, M. C. & Almeida Neto, M. A simple stochastic 773
 714 model for complex coextinctions in mutualistic networks: 774
 robustness decreases with connectance. *Ecol Lett* **18**, 715
 144–152 (2015).
- 716 [56] Laland, K. N. *et al.* The extended evolutionary synthesis: 717
 718 its structure, assumptions and predictions. *Proc Roy Soc 719
 720 B* **282**, 20151019 (2015).
- 721 [57] Gupta, M., Prasad, N., Dey, S., Joshi, A. & Vidya, T. 722
 722 Niche construction in evolutionary theory: the construc- 723
 723 tion of an academic niche? *J Gen* **96**, 491–504 (2017).
- 724 [58] Feldman, M. W., Odling-Smee, J. & Laland, K. N. Why 725
 725 Gupta et al.'s critique of niche construction theory is 726
 off target. *J Gen* **96**, 505–508 (2017).
- 727 [59] Cuddington, K. Invasive engineers. *Ecol Model* **178**, 728
 335–347 (2004).
- 729 [60] Wright, J. P. & Jones, C. G. Predicting effects of ecosys- 730
 730 tem engineers on patch-scale species richness from pri- 731
 731 mary productivity. *Ecology* **85**, 2071–2081 (2004).
- 732 [61] Kylafis, G. & Loreau, M. Ecological and evolutionary 733
 733 consequences of niche construction for its agent. *Ecol 734
 734 Lett* **11**, 1072–1081 (2008).
- 735 [62] Krakauer, D. C., Page, K. M. & Erwin, D. H. Diversity, 736
 736 dilemmas, and monopolies of niche construction. *Am Nat* 737
 737 **173**, 26–40 (2009).
- 738 [63] Laland, K. N., Odling-Smee, F. J. & Feldman, M. W. 739
 739 Evolutionary consequences of niche construction and 740
 740 their implications for ecology. *Proc Natl Acad Sci USA* 741
 741 **96**, 10242–10247 (1999).
- 742 [64] Kallus, Y., Miller, J. H. & Libby, E. Paradoxes in leaky 743
 743 microbial trade. *Nat Commun* **8**, 1361 (2017).
- 744 [65] Muscarella, M. E. & O'Dwyer, J. P. Ecological insights 745
 745 from the evolutionary history of microbial innovations. 746
bioRxiv 220939 (2017).
- 747 [66] Shipway, J. R. *et al.* A rock-boring and rock-ingesting 748
 748 freshwater bivalve (shipworm) from the philippines. *Proc 749
 749 Roy Soc B* **286**, 20190434 (2019).
- 750 [67] Kidwell, S. M. Taphonomic feedback in Miocene assem- 751
 751 blages: testing the role of dead hardparts in benthic com- 752
 752 munities. *Palaios* **1**, 239–255 (1986).
- 753 [68] Polz, M. F., Alm, E. J. & Hanage, W. P. Horizontal 754
 754 gene transfer and the evolution of bacterial and archaeal 755
 755 population structure. *Trends Genet* **29**, 170 – 175 (2013).
- 756 [69] Gillespie, D. T. Exact stochastic simulation of coupled 757
 757 chemical reactions. *J Phys Chem* **81**, 2340–2361 (1977).
- 758 [70] Williams, R. J. & Purves, D. W. The probabilistic niche 759
 759 model reveals substantial variation in the niche structure 760
 760 of empirical food webs. *Ecology* **92**, 1849–1857 (2011).
- 761 [71] Warren, C. P., Pascual, M., Lafferty, K. D. & Kuris, 762
 762 A. M. The inverse niche model for food webs with para- 763
 763 sites. *Theor Ecol* **3**, 285–294 (2010).
- 764 [72] **Acknowledgements**
- 765 We would like to thank Uttam Bhat, Irina Birskis Barros, Emmet 766
 766 Brickowski, Ashkaan Fahimipour, Jean Philippe Gibert, Chris P 767
 767 Kempes, Taran Rallings, Samuel Scarpino, Megha Suswaram, and 768
 768 Ritwika VPS for insightful discussions and comments throughout 769
 769 the lengthy gestation of this manuscript. The original idea 770
 770 was conceived at the Networks on Networks Working Group in 771
 771 Göttingen, Germany (2014) and the Santa Fe Institute (2015). 772
 772 This work was formerly prepared as a part of the Ecological 773
 773 Network Dynamics Working Group at the National Institute 774
 774 for Mathematical and Biological Synthesis (2015–2019), spon- 775
 775 sored by the National Science Foundation through NSF Award 776
 776 DBI-1300426, with additional support from The University of 777
 777 Tennessee, Knoxville. Infinite revisions were conducted at the 778
 778 Santa Fe Institute made possible by travel awards to JDY and TG. 779
 779 Additional support came from UC Merced startup funds to JDY, 780
 the International Centre for Theoretical Physics ICTP-SAIFR,

⁷⁸⁰ FAPESP (2016/01343-7) and CNPq (302049/2015-0) to MAMA,
⁷⁸¹ CNPq and FAPESP (2018/14809-0) to PRG, and DFG research
⁷⁸² unit 1748 and EPSRC (EP/N034384/1) to TG.

⁷⁸³

⁷⁸⁴

⁷⁸⁵ **Author contributions**

⁷⁸⁶ JDY and TG conceived of the model framework. JDY, MMP,
⁷⁸⁷ MAMA, and TG designed the analyses. JDY, MMP, MAMA,
⁷⁸⁸ JLOD, PRG, DG, and TG analyzed the results and contributed
⁷⁸⁹ to multiple versions of the manuscript.

⁷⁹⁰

⁷⁹¹ **Competing interests**

⁷⁹² The authors declare no competing interests.

793

SUPPLEMENTARY METHODS

794

Appendix 1: Building the source pool

795 Here and henceforth, we refer to the assembly model presented
 796 in the main text as the ENIGMa model (E:eat, N:need, Ig:ignore,
 797 Ma:make). To initiate the ENIGMa assembly model, we must first
 798 construct the source pool, where each ecological entity (species and
 799 modifiers) is defined by its potential interactions with each other.
 800 The source pool interaction matrix P is generated by first setting
 801 the number of species in the pool S_P and determining the number
 802 of modifiers M_P that are made by ecosystem engineers. The result-
 803 ing matrix is $N_P \times N_P$ where $N_P = S_P + M_P$, and is subdivided
 804 into four quadrants, only two of which play a role here: species-
 805 species interactions and species-modifier interactions (see Fig. 1).
 806 In these two quadrants, the expected frequency of eat interactions
 807 $E\{p_e\}$ and the expected frequency of need interactions $E\{p_n\}$ are
 808 free parameters, as is the expected number of modifiers made per
 809 species $E\{M_i\} = \eta$. Here and throughout, we simplify this pa-
 810 rameter space by assuming that the frequency of eat and need in-
 811 teractions for species-species (SS) interactions and species-modifier
 812 (SM) interactions are equivalent, such that $E_{SS}\{p_e\} = E_{SM}\{p_e\}$
 813 and $E_{SS}\{p_n\} = E_{SM}\{p_n\}$. For each species, a set number of modi-
 814 fiers is drawn from Poiss(η), such that the expected proportion of
 815 species that are engineers (species that make modifiers) is $1 - e^{-\eta}$.
 816 If a particular modifier is randomly and independently drawn for
 817 a given engineer from a complete list of all possible modifiers, such
 818 that multiple species – with some probability – can make the same
 819 modifier, the expected number of modifiers is

$$E\{M_P\} = S_P \eta \left(1 - \frac{1}{e}\right), \quad (S1)$$

820 where e is Euler's number. The frequency of engineering (make)
 821 interactions is then calculated as

$$E\{p_m\} = \frac{\eta}{S_P \left(1 + \eta - \frac{\eta}{e}\right)^2}. \quad (S2)$$

822 Finally the frequency of the non-interaction is calculated as
 823 $E_{SS}\{p_\emptyset\} = 1 - E_{SS}\{p_e\} + E_{SS}\{p_n\}$ and $E_{SM}\{p_\emptyset\} = 1 - E_{SM}\{p_e\} +$
 824 $E_{SM}\{p_n\} + E_{SM}\{p_m\}$ for species-species and species-modifier in-
 825 teractions, respectively. Pairwise interactions are assigned ran-
 826 domly from these probabilities between species-species and species-
 827 modifiers independently in both quadrants, such that the source
 828 pool matrix has no imbued structure apart from the number of
 829 species, the number of modifiers, and the frequency of each direc-
 830 tional interaction type. Each source pool is provided a *basal re-*
 831 *source* (the first row/column). A species with a trophic interaction
 832 to this resource is identified as an autotroph (or mixotroph depend-
 833 ing on its other trophic interactions). If they do not have service
 834 dependencies with other species/modifiers, it is these species that
 835 are uniquely able to initiate assembly.

836 We can determine analytically the expected number of unique
 837 versus redundant modifiers in the source pool. As the total number
 838 of modifiers is given in Eq. S1, the number of unique modifiers is
 839 given by $E\{M_P\}_{\text{unique}} = S_P \eta e^{-1}$. The number of redundant
 840 modifiers is then given as

$$E\{M_P\}_{\text{redundant}} = \eta S_P \frac{e - 2}{e}, \quad (S3)$$

841 such that the proportion of redundant modifiers ϕ is

$$\phi = \frac{e - 2}{e - 1} \approx 0.418. \quad (S4)$$

842 Accordingly, we find that the number of redundant modifiers in-
 843 creases linearly with η , while the proportion of modifiers that are
 844 redundant is fixed. Figure S5A,B shows both analytical expecta-
 845 tions and numerically-derived measures for $E\{M_P\}_{\text{redundant}}$ and
 846 ϕ , respectively.

847

Appendix 2: Comparison to Niche Model

848 We compared certain structural features of ENIGMa at steady
 849 state to those of the Niche Model³⁵. Comparisons were restricted
 850 to networks constructed in the absence of engineering because engi-
 851 neers introduce indirect effects that are not considered in static food
 852 web models, and may make such comparisons irrelevant. While we
 853 there are many similarities, there are also some important differ-
 854 ences, some of which are highlighted in the main text. While we
 855 consider a comparison of our framework with other food web mod-
 856 els such as the Niche Model relevant, we emphasize that the mo-
 857 tivations underlying both are distinct. Our approach is intended
 858 to provide a deeper understanding into how multitype depen-
 859 dencies between species and the environment impact the dynamics of
 860 community assembly. While capturing general qualitative features
 861 of empirical systems demonstrates that the dynamics we consider
 862 are ecologically relevant, the goal of our approach is distinct from
 863 that of static food web models, which aim to maximize structural
 864 similarities between model and empirical systems^{35,70}.

865 We compared steady state ecological networks that emerge
 866 from ENIGMa (described in Materials and Methods, main text)
 867 with food webs constructed from the Niche Model³⁵ with simi-
 868 lar species richness and connectance. Because species richness and
 869 connectance of the Niche Model are often altered by eliminating dis-
 870 connected species, we compared *i*) species richness, *ii*) connectance,
 871 *iii*) mean species degree, *iv*) standard deviation of out-degree dis-
 872 tributions, and *v*) standard deviation of in-degree distributions av-
 873 eraged across 1000 replicates for each model.

874 We found that all measures resulted in fairly similar values be-
 875 tween ENIGMa and the Niche Model food webs with a some im-
 876 portant differences (Figs. S2,S3). While similar, ENIGMa pro-
 877 duces consistently lower values of connectance, mean species de-
 878 gree, as well as standard deviations of the in- and out-degree dis-
 879 tributions. This means that the food webs produced by ENIGMa
 880 are more sparsely connected with less variance between species.
 881 These results were expected, as the Niche Model assumes system-
 882 atically increasing dietary ranges with higher niche values, whereas
 883 the trophic interactions assigned to species in the source pool of
 884 ENIGMa are drawn independently. An important difference be-
 885 tween the Niche Model and ENIGMa is that we do not distinguish
 886 between predators and parasites. A different framework known
 887 as the Inverse Niche Model⁷¹ has been proposed to address par-
 888 asitic interactions. The Inverse Niche Model assumes increasing
 889 specialization with feeding hierarchies, which would serve to lower
 890 the average generality of species (lower degree). In addition, the
 891 Inverse Niche model outputs lower standard deviations of in- and
 892 out-degree distributions. Together these trends suggest that the
 893 qualitative structural differences that we observe for the assembly
 894 and Niche model may reflect an important structural distinction
 895 between food webs that do and do not include parasitic species.

896

Appendix 3: Measures of generality

897 The trophic breadth of potential colonizers is thought to play an
 898 important role in community assembly. The definition of a special-
 899 ist or generalist to some degree depends on the size and connectance
 900 of the larger food web. Trophic generality for a species i is defined
 901 $G_i = k_i^{\text{in}} / (L/S)$, where k_i^{in} is the in-degree, or number of resources
 902 consumed by species i ³⁵. A species is classified as a generalist if
 903 the number of its trophic interactions is greater than the average
 904 number of links per species, or $G_i > 1$, and a specialist if $G_i < 1$,
 905 where a community can be described by the proportion of special-
 906 ists found therein. For interaction networks that are assembl-
 907 ing over time, generality can be scaled by a number of different mea-
 908 sures of L/S , and this has a large effect on our interpretation of the
 909 role of generality in community assembly. For instance, L/S may
 910 be quantified by either including all autotrophic species or only au-
 911 totrophic 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, Piechnik et al.³⁴ 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. S4b); 2) G_i^{hetero} , where we consider only the links and species richness of heterotrophs, excluding autotrophs (points; Fig. S4b); 3) 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 (diamonds; Fig. S4b). 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. S4). 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. S4).

The different normalizations by which generality is measured will impact the interpretation of both empirical and model systems alike. In our framework, species colonizing early in the assembly process are generalists compared to how the term is defined at the steady state, but they are functionally specialists with respect to the assembling community. For example, a species that is trophically connected to 10 resource species in the source pool may colonize a community where it is consuming a small subset of its potential range. As the community grows, that species may realize more of its trophic niche if those resource species subsequently colonize the system. To what end we label this species a generalist or specialist relative to the assembling community is thus subject to multiple interpretations.

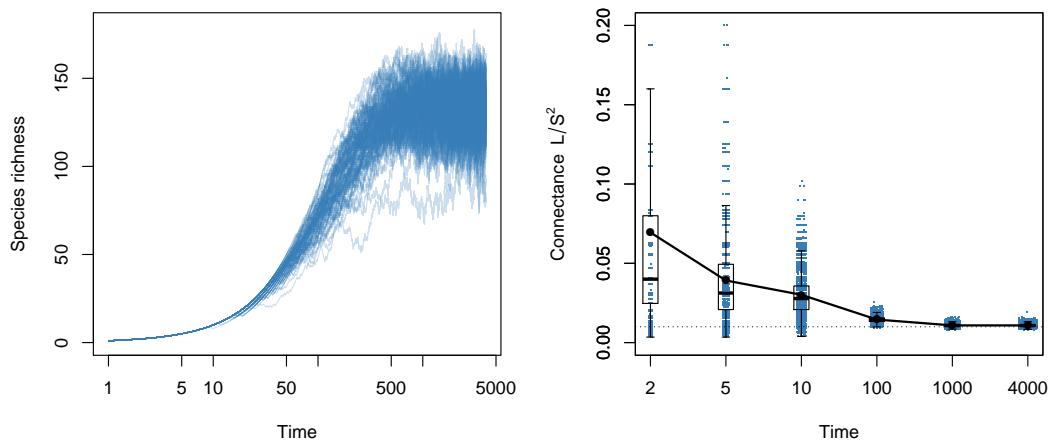


Figure S1. Left: Assembly of communities over time results in steady state species richness by ca. time-step 250. Right: Trophic connectance early in assembly is high because few species are tightly connected. Over time, connectance decays as species richness increases, and the density of trophic interactions declines.

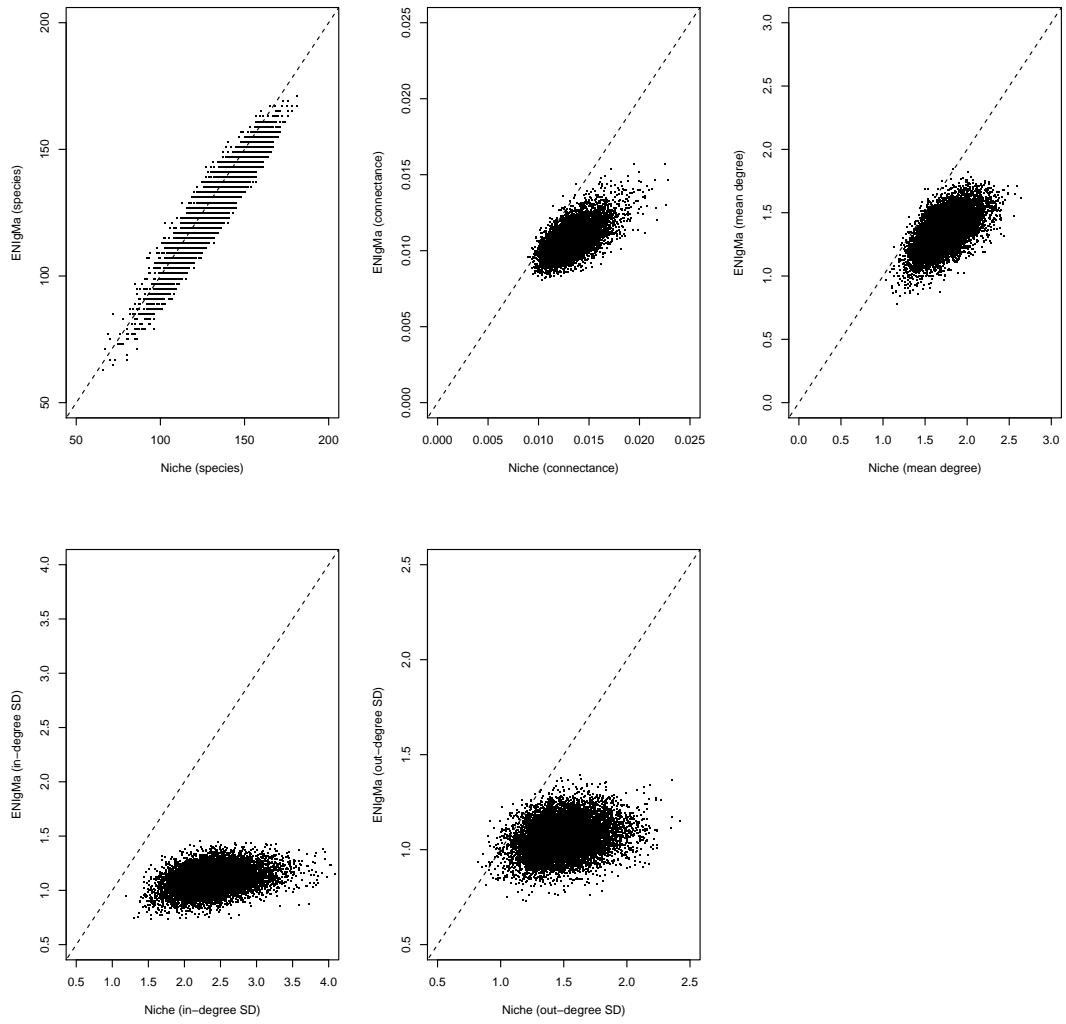


Figure S2. Comparisons of raw structural measures for the assembly (y-axis) and Niche model (x-axis). If the models produce similar structures, metrics will tend to fall on the 1:1 line (drawn). While the values for both models are similar, connectance, mean degree, and the standard deviation of in- and out-degree are all lower for the assembly model relative to those measures for the Niche model.

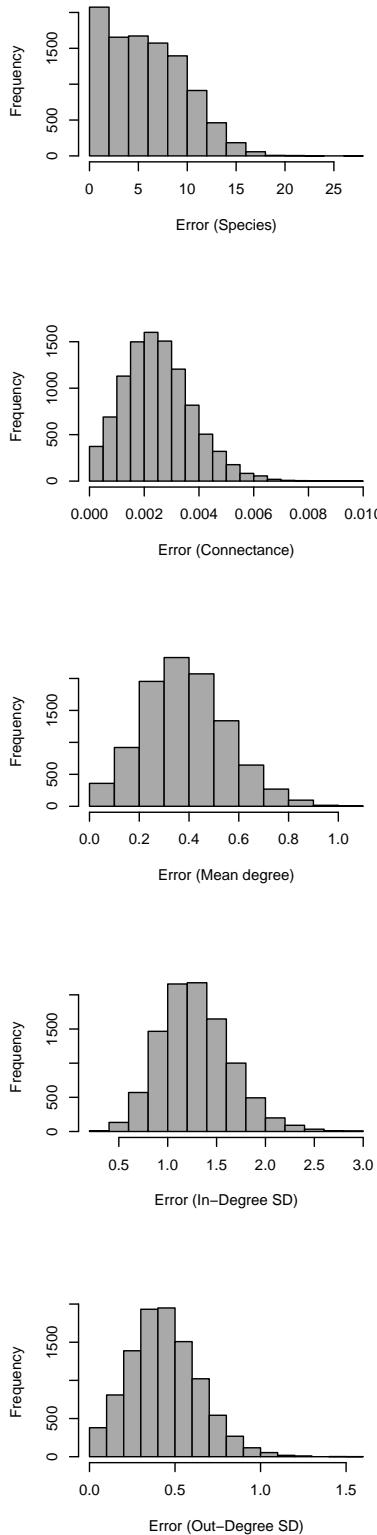


Figure S3. Error between structural measures of the assembly and Niche models. Error is measured as $\sqrt{(m_i - m_j)^2}$, where m_i and m_j are structural metrics for the assembly and Niche model, respectively. Only the trophic network of the assembly model used to assess metrics.

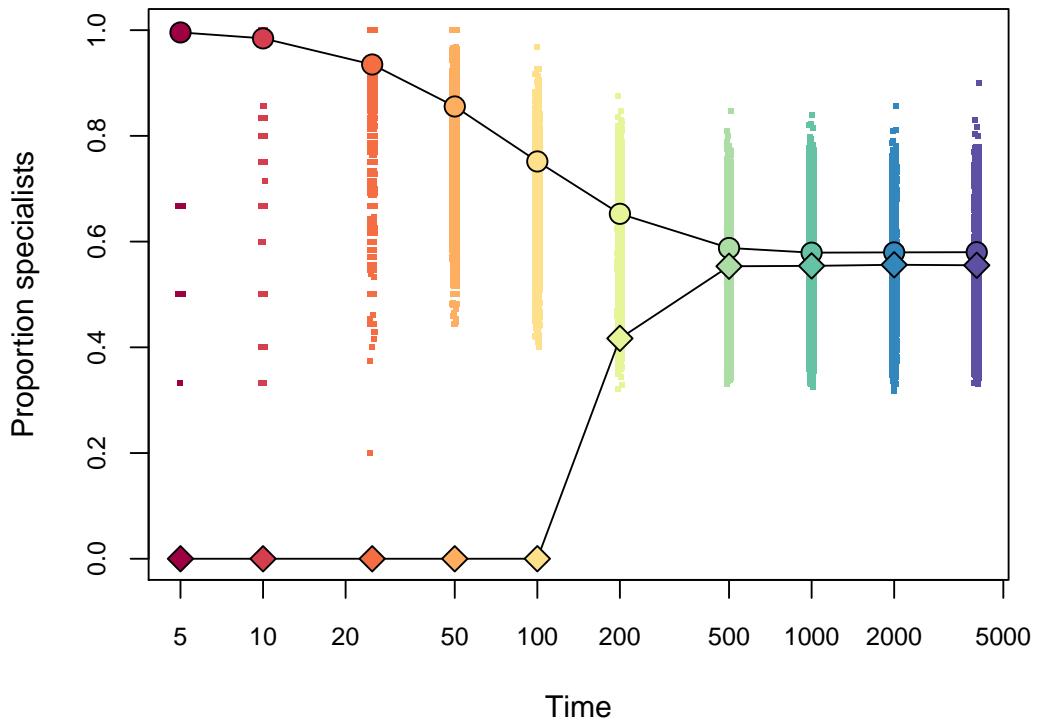


Figure S4. The proportion of specialists as a function of assembly time, where a specialist is defined as a species with a generality index $G_i < 1$. Measures of G_i are shown normalized to different measures of link-density. Circles: G_i^{all} where L accounts for all links in the food web and S accounts 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 (averaged across replicates). This measure is the one presented in the main text and most similar to that used to evaluate assembling mangrove food webs³⁴.

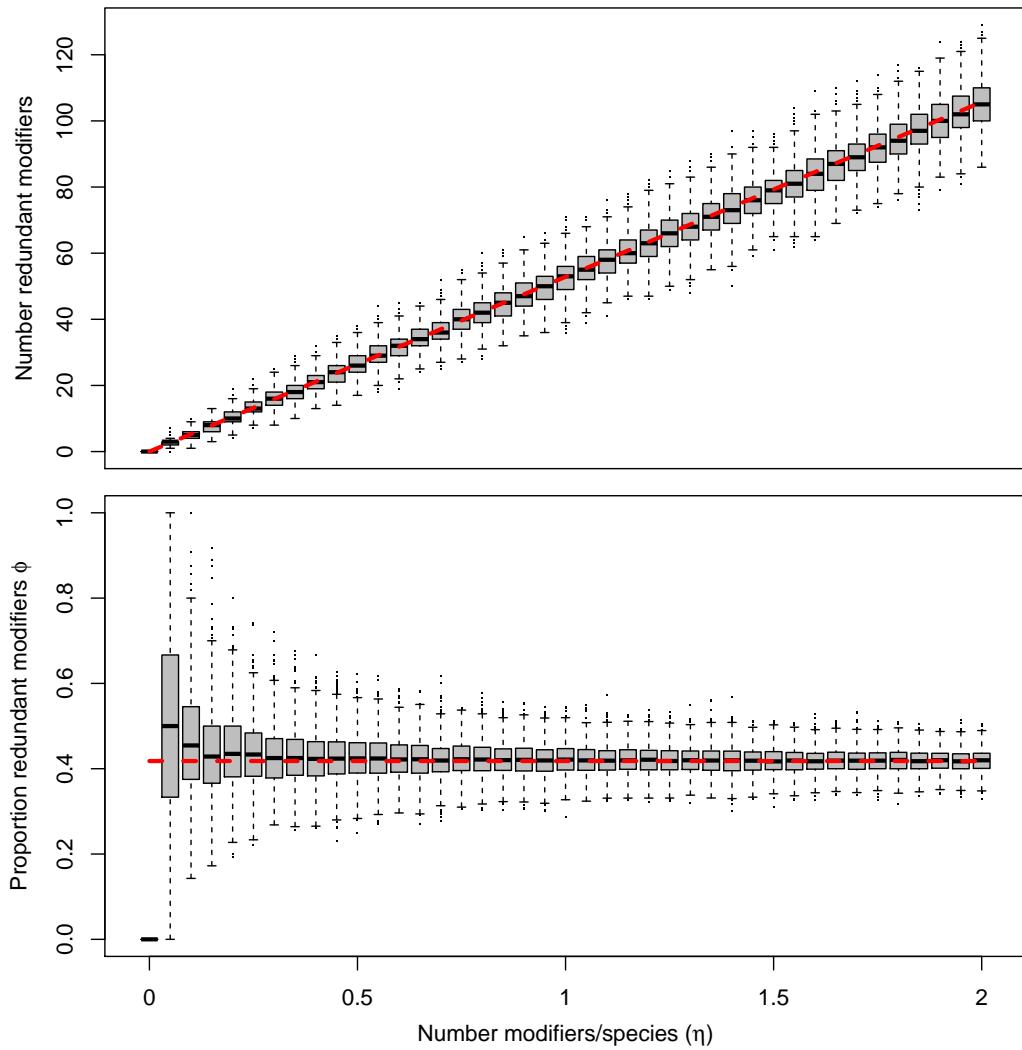


Figure S5. **a**, Number of redundant modifiers in the source pool as a function of the expected number of modifiers made per species η . The red dashed line shows the analytical expectation (Eq. S3). **b**, Proportion of redundant modifiers ϕ versus the total number of modifiers in the source pool as a function of the expected number of modifiers made per species η . The red dashed line shows the analytical expectation of $\phi \approx 0.418$ (Eq. S4).

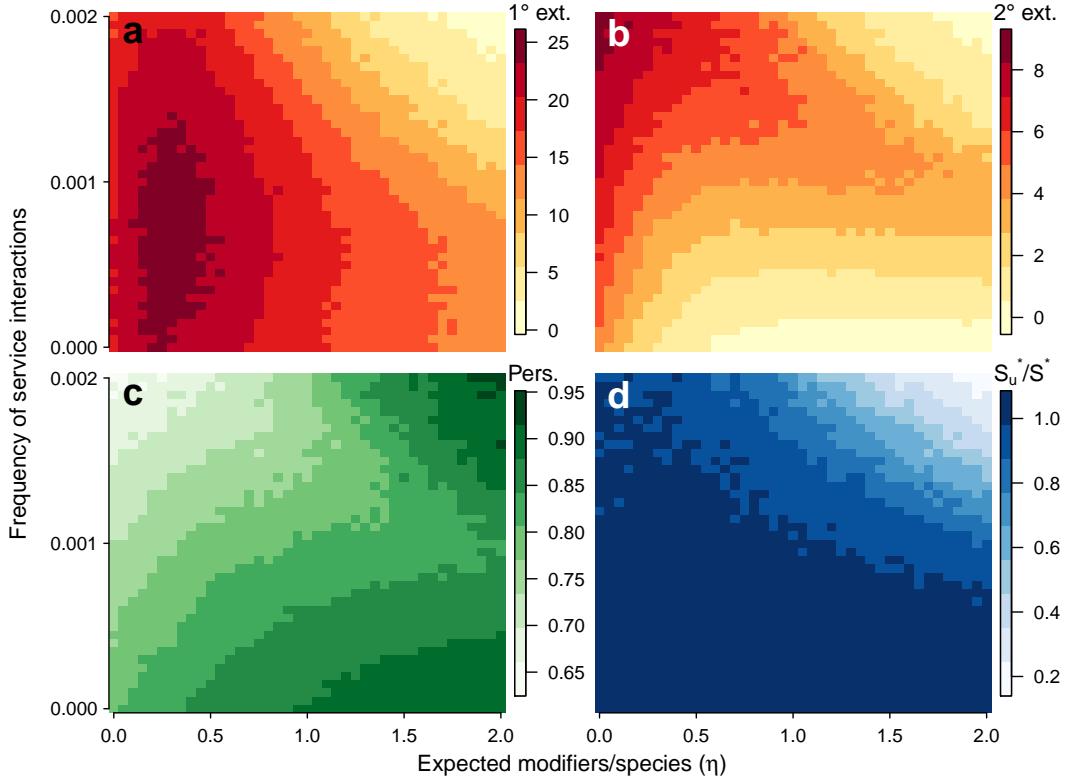


Figure S6. Measures of community robustness as a function of the frequency of service interactions and number of modifiers per species, where each modifier is uniquely made by an engineer. **a**, Mean rates of primary extinction, where primary extinctions occur from competitive exclusion of consumers over shared resources. **b**, Mean rates of secondary extinction, which cascade from primary extinctions. **c**, Mean species persistence, defined as the percent simulation time the community is occupied by a given species, averaged across all species that successfully colonize. **d**, The ratio S_u^*/S^* , where S_u^* denotes steady states for systems where all engineered modifiers are unique to each engineer, and S^* denote steady states for systems with redundant engineering. Lower values of S_u^*/S^* mean that systems with redundant engineers have higher steady states than those without redundancies. Values are averaged over 50 replicates for each parameterization. See Materials and Methods for default parameter values.

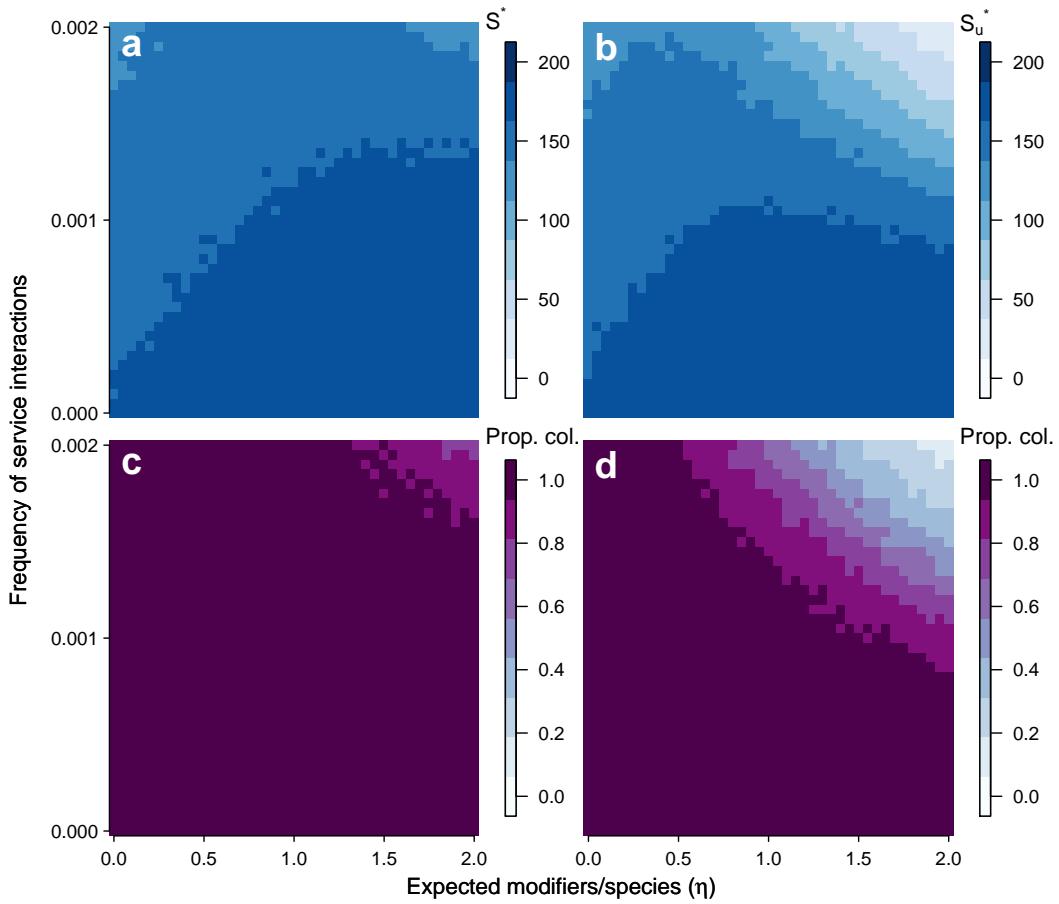


Figure S7. **a**, Steady state community richness with redundant engineering. **b**, Steady state community richness without redundant engineering. **c**, Proportion of species in the source pool that colonize the community at least once throughout the simulation (with redundant engineering). **d**, Proportion of species in the source pool that colonize the community at least once throughout the simulation (without redundant engineering).