

# 1 Diverse interactions and ecosystem engineering stabilize community assembly

2 Justin D. Yeakel,<sup>1,2</sup> Mathias M. Pires,<sup>3</sup> Marcus A. M. de Aguiar,<sup>3</sup> James L.  
3 O'Donnell,<sup>4</sup> Paulo R. Guimarães Jr.,<sup>5</sup> Dominique Gravel,<sup>6</sup> and Thilo Gross<sup>7</sup>

4 <sup>1</sup>*University of California, Merced, Merced, CA 95340, USA*

5 <sup>2</sup>*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA*

6 <sup>3</sup>*Universidade Estadual de Campinas, Campinas - SP, Brazil*

7 <sup>4</sup>*University of Washington, Seattle, WA 98195, USA*

8 <sup>5</sup>*Universidade de São Paulo, São Paulo, Brazil*

9 <sup>6</sup>*Université de Sherbrooke, Sherbrooke, QCJ1K0A5, Canada*

10 <sup>7</sup>*University 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 stability 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.

11 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<sup>1–3</sup> or pollination networks capturing a specific mutualistic interaction<sup>4</sup>, and continues to lead to significant breakthroughs in our understanding of the dynamical consequences of community structure<sup>5–7</sup>, assembly<sup>8</sup>, and coevolution<sup>9</sup>.

23 Paragraph on the importance of assembly and models of assembly on structure/dynamics. The structure  
24 evolves from a feedback between different types of interactions between species and their environment, which  
25 leads into the next section.

28 Recent interest in ‘multilayer networks’ comprising multiple interaction types (multitype interactions) may provide additional insight into these processes<sup>10,11</sup>. 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<sup>12,13</sup> or more generally niche construction<sup>14,15</sup>, are quite common in nature and exist in almost every ecosystem.

37 Diverse interactions occur not only between species but indirectly through the effects that species have on their environment<sup>13,16,17</sup>. Elephants root out large saplings and small trees, enabling the formation and maintenance of grasslands<sup>18</sup> and creating habitat for smaller vertebrates<sup>19</sup>. Burrowing rodents create shelter and pro-

43 mote primary production by aerating the soil<sup>20</sup>, salmon and aquatic invertebrates create freshwater habitats by changing stream morphology<sup>21</sup>, and leaf-cutter ants alter microclimates, influencing seedling survival and plant growth<sup>22</sup>. These examples illustrate ecosystem engineering, where the engineering organism alters the environment on timescales longer than its own<sup>23</sup>.

50 Ecosystem engineering not only impacts communities on ecological timescales, but has profoundly shaped the evolution of life on Earth<sup>24</sup>. For example, the emergence of multicellular cyanobacteria fundamentally altered the atmosphere during the Great Oxidation Event of the Proterozoic roughly 2.5 Byrs BP<sup>24,25</sup>, 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 plankton<sup>26</sup>, illustrating that engineering clades can have much larger, sometimes global-scale effects.

62 The effect of the environment on species is commonly included in models of ecological dynamics<sup>27–29</sup> 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<sup>17,30</sup>, the absence of engineered effects in network models was addressed by Odling-Smeet et al.<sup>13</sup>, 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

75 interactions.

76 Here we model the assembly of an ecological network  
 77 where nodes represent ecological entities, including engi-  
 neering species, non-engineering species, and the effects  
 79 of the former on the environment, which we call abiotic  
 80 *modifiers*. The links of the network that connect both  
 81 species and modifiers represent trophic (*eat* interactions),  
 82 service (*need* interactions), and engineering depen-  
 83 cies, respectively (Fig. 1; see Materials and Methods for  
 84 a full description). Trophic interactions represent both  
 85 predation as well as parasitism, whereas service interac-  
 86 tions account for all non-trophic interactions such as pol-  
 87 lination or seed dispersal. In our framework a traditional  
 88 mutualism (such as a plant-pollinator interaction) con-  
 89 sists of a service (need) interaction in one direction and  
 90 a trophic (eat) interaction in the other. These multitype  
 91 interactions between species and modifiers thus embed  
 92 multiple dependent ecological sub-systems into a single  
 93 network (Fig. 1). Modifiers in our framework overlap  
 94 conceptually with the ‘abiotic compartments’ described  
 95 in Odling-Smee et al.<sup>13</sup>. Following Pillai et al.<sup>31</sup>, we  
 96 do not track the abundances of biotic or abiotic enti-  
 97 ties but only track their presence or absence. We use  
 98 this framework to explore the dynamics of ecosystem as-  
 99 sembly, where the colonization and extinction of species  
 100 within a community depends on the constraints imposed  
 101 by the trophic, service, and engineering dependencies.

102 We then show how observed network structures emerge  
 103 from the process of assembly, compare their attributes  
 104 with those from empirical systems, and examine the ef-  
 105 fects of ecosystem engineers.

106 Our results offer four key insights into the roles of  
 107 multitype interactions and ecosystem engineering in  
 108 driving community assembly. First, we show that the  
 109 assembly of communities in the absence of engineering  
 110 reproduces many features observed in empirical systems.  
 111 These include changes in the proportion of generalists  
 112 over the course of assembly that accord with measured  
 113 data and trophic diversity similar to empirical observa-  
 114 tions. Second, we show that increasing the frequency  
 115 of mutualistic interactions leads to the assembly of  
 116 ecological networks that are more nested, a common  
 117 feature of diverse mutualistic systems<sup>32</sup>, but are also less  
 118 stable. Our third key result shows that increasing the  
 119 proportion of ecosystem engineers within a community  
 120 has nonlinear effects on observed extinction rates. While  
 121 we find that a low amount of engineering increases  
 122 extinction rates, a high amount of engineering has the  
 123 opposite effect. Finally we show that redundancies  
 124 in engineered effects promote community diversity by  
 125 lowering the barriers to colonization.

126  
 127 **Assembly without ecosystem engineering.** Com-  
 128 munities assemble by random colonization from a source  
 129 pool. A species from the source pool can colonize if it  
 130 finds at least one resource that it can consume (one *eat*  
 131 interaction is satisfied; cf. Ref. 33) and all of its non-  
 132 trophic needs are met (all *need* interactions are satisfied).

133 As such, the service interactions are assumed to be ob-  
 134 ligate, whereas trophic interactions are flexible. While a  
 135 basal resource is always assumed to be present, (Fig. 1A)  
 136 following the establishment of an autotrophic base, the  
 137 arrival of mixotrophs (i.e. mixing auto- and heterotro-  
 138 phy) and lower trophic heterotrophs create opportunities  
 139 for organisms occupying higher trophic levels to invade.  
 140 This expanding niche space initially serves as an acceler-  
 141 ator for community growth.

142 Following the initial colonization phase, extinctions be-  
 143 gin to slow the rate of community growth. Primary ex-  
 144 tinctions occur if a given species is not the strongest com-  
 145 petitor for at least one of its resources (see Methods). A  
 146 species’ competition strength is determined by its inter-  
 147 actions: competition strength is enhanced by the number of its  
 148 of need interactions and penalized by the number of its  
 149 prey (favoring trophic specialists) and predators (favor-  
 150 ing species with fewer predators). Secondary extinctions  
 151 occur when species lose its last trophic or any of its ser-  
 152 vice requirements. See Fig. 1D, E for an illustration  
 153 of the assembly process. As the colonization and ex-  
 154 tinction rates converge, the community reaches a steady  
 155 state around which it oscillates (Fig. 2A). See Methods  
 156 and Supplementary Appendix 1 for a complete descrip-  
 157 tion of the assembly process. Specific model parameter-  
 158 izations for each section are described in Supplementary  
 159 Appendix 2.

160 Assembly of ecological communities in the absence of  
 161 engineering results in interaction networks with struc-  
 162 tures consistent with empirical observations. As the com-  
 163 munity reaches steady state, we find that the connectance  
 164 of trophic interactions ( $C = L/S^2$ , where  $S$  is species  
 165 richness and  $L$  is the number of links) decays to a value  
 166 similar to that of the source pool (Fig. S1). Decay-  
 167 ing connectance has been documented in the assembly of  
 168 mangrove communities<sup>34</sup>, however this decay is a statis-  
 169 tical inevitability, as early in the assembly process small  
 170 food webs will have high link density, from which it can  
 171 only decline. Compared to trophic networks constructed  
 172 using the Niche model<sup>35</sup> given similar species richness  
 173 and connectance, our framework results in networks with  
 174 degree distributions of similar means but with reduced  
 175 variance (Supplementary Appendix 3).

176 Recent empirical work has suggested that general-  
 177 ist species may predominate early in assembly, whereas  
 178 specialists colonize after a diverse resource base has  
 179 accumulated<sup>33,34</sup>. Here the trophic generality of species  
 180  $i$  is defined  $G_i = k_i^{\text{in}}/(L/S)$  where  $k_i^{\text{in}}$  is the in-degree, or  
 181 number of resources consumed, by species  $i$ <sup>35</sup>. A species  
 182 is classified as a generalist if  $G_i > 1$  and a specialist if  
 183  $G_i < 1$ . If generalism is scaled to the steady state link  
 184 density (see Supplementary Appendix, section 4), we ob-  
 185 serve that generalists dominate early in assembly, with an  
 186 increase in specialists as assembly progresses (Fig. 2B).  
 187 This confirms expectations from the trophic theory of is-  
 188 land biogeography<sup>33</sup>, where early communities with lower  
 189 richness are less likely to support specialist consumers  
 190 than late, species-rich communities. At steady state the

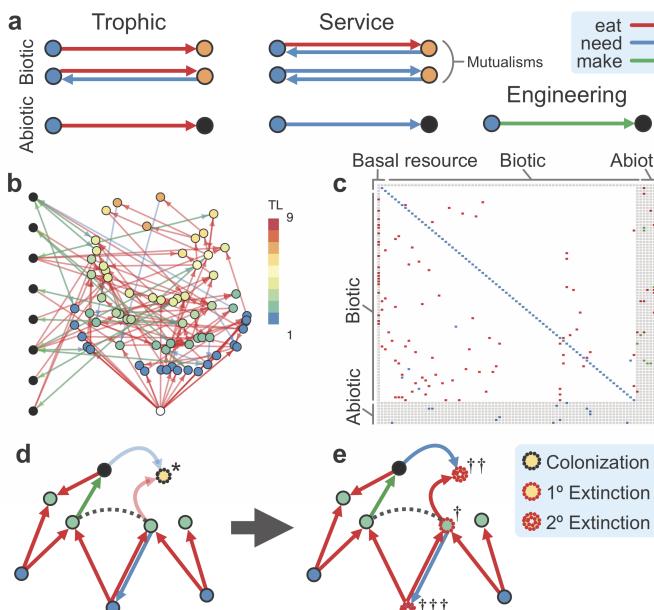


Figure 1. **a**, Multitype interactions between species (colored nodes) and abiotic modifiers (black nodes). **b**, An assembling food web with species (colored nodes; color denotes trophic level, TL) and modifiers (black nodes). The basal resource is the white node rooted at the bottom of the network. **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 ( $\dagger$ ). The extinction of species ( $\dagger$ ) will cascade to affect those connected by trophic ( $\dagger\dagger$ ) and service ( $\dagger\dagger\dagger$ ) dependencies.

proportion of specialists is ca. 56%, similar to empirical observations of assembling food webs<sup>34</sup>.

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 TL = 1.6. Four trophic levels are typically established by  $t = 50$ , where colonization is still dominant, and by  $t = 100$  the time communities reach steady state the interaction networks are characterized by an average  $TL_{max}$  ( $\pm$  standard deviation) =  $11 \pm 2.8$  (Fig. 2C). While the maximum trophic level is higher than that measured in most predator-prey systems<sup>36</sup>, it is not unreasonable if parasitic interactions (which we do not differentiate from other consumers) are included<sup>37</sup>. Overall, the most common trophic level among species at steady state is ca. 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 models<sup>38</sup>.

**Structure and dynamics of mutualisms.** Nested interactions, where specialist interactions are subsets of generalist interactions, are a distinguishing feature of mutualistic networks<sup>32</sup>. A nested structure has been shown to maximize the structural stability of mutualistic networks<sup>39</sup>, emerge naturally via adaptive foraging behaviors<sup>40,41</sup> and neutral processes<sup>42</sup>, and promote the influence of indirect effects in driving coevolutionary dynamics<sup>9</sup>. While models and experiments of trophic networks suggest that compartmentalization confers greater stabilizing properties<sup>43,44</sup>, interaction asymmetry among species may promote nestedness in both trophic<sup>45</sup> and mutualistic systems<sup>46</sup>. Processes that operate on different temporal and spatial scales may have a significant influence on these observations<sup>47</sup>. For example, over evolutionary time, coevolution and speciation may degrade nested structures in favor of modularity<sup>48</sup>, and there is some evidence from Pleistocene food webs that geographic insularity may reinforce this process<sup>49</sup>.

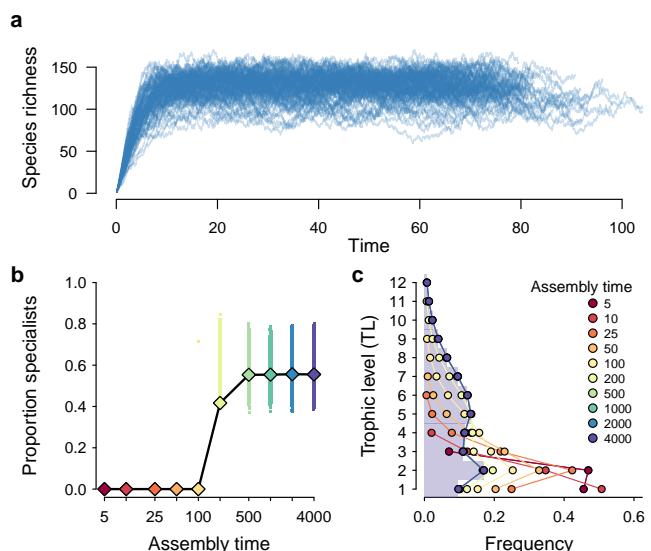


Figure 2. **a**, Assembling communities over time from a pool of 200 non-engineering species. Steady state species richness is reached by  $t = 250$ . **b**, The proportion of specialists as a function of assembly time (iterations), where a specialist is defined as a species with a generality index  $G_i < 1$ . 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 TL = 1. Measures were evaluated across  $10^4$  replicates; see Materials and Methods for parameter values.

Does the assembly of ecological networks favor nested-

ness when mutualistic interactions are frequent? Increasing the number of need interactions (Fig. 1) increases the number of both service-resource and service-service dependencies. Consider how species with more service interactions compare to those with fewer. More service interactions *i)* increase a species' competition strength, lowering its primary extinction risk while also *ii)* increasing inter-species dependencies and its secondary extinction risk. While the occurrence of a mutualism assumes a fitness advantage for the species receiving the service, the latter highlights the potential risk associated with losing mutualistic partners<sup>50,51</sup>. Indeed, the balance that mutualists must maintain with their partners may have large implications for the future of global biodiversity<sup>52</sup>.

We find that as we increase the frequency of service interactions (holding constant trophic interactions; see Appendix 2), the assembled community at steady state becomes more nested (Fig. 3). In this case, nestedness emerges from the assembly process and provides structural stability. This stability 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. S5). As such, 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 (Fig. 3a). 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 state<sup>8</sup>.

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<sup>8,40,41,54,55</sup>. We expect that the increased rate of secondary extinctions attributable to the loss of

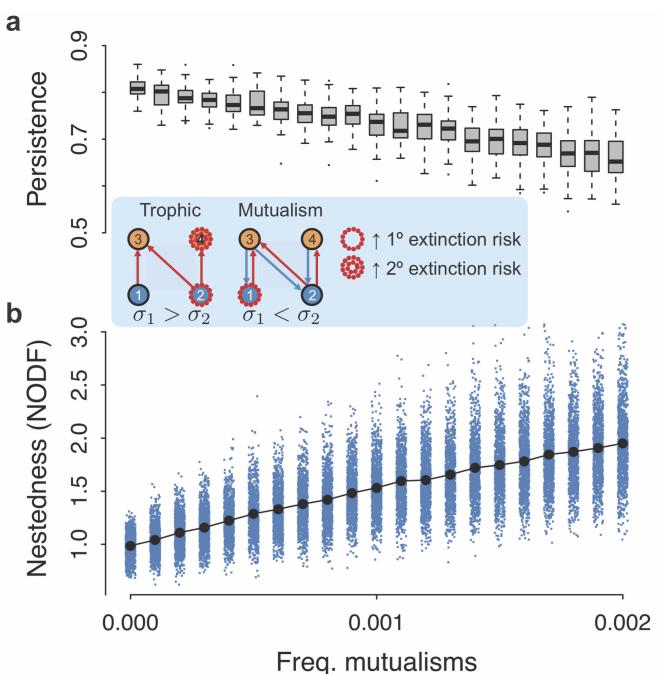


Figure 3. **a**, Species persistence with increasing frequencies of mutualistic (service) interactions without ecosystem engineers. **b**, Structural nestedness of communities, measured as NODF (Nestedness based on Overlap and Decreasing Fill)<sup>53</sup>. Measures were evaluated across  $10^4$  replicates; 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  $\sigma$ , 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.

**298** obligate mutualistic partners to have greater impact on  
**299** system stability than the potential loss of non-obligate  
**300** mutualistic partners. As such, we do not expect inclu-  
**301** sion of non-obligate mutualisms to alter the qualitative  
**302** nature of our findings.

**304** **Assembly with ecosystem engineering.** The con-  
**305** cept of ecosystem engineering, or more generally niche  
**306** construction, has both encouraged an extended evo-  
**307** lutionary synthesis<sup>56</sup> while also garnering considerable  
**308** controversy<sup>57,58</sup>. Models that explore the effects of  
**309** ecosystem engineering are relatively few, but have cov-  
**310** ered important ground<sup>13,23</sup>. For example, engineering  
**311** has been shown to promote invasion<sup>59</sup>, alter primary  
**312** productivity<sup>60</sup>, and change the selective environment  
**313** over eco-evolutionary timescales<sup>61,62</sup> which can lead to  
**314** unexpected outcomes such as the fixation of deleteri-  
**315** ous alleles<sup>63</sup>. On smaller scales, microbiota construct  
**316** shared metabolic resources that have a significant influ-  
**317** ence on microbial communities<sup>64</sup>, the dynamics of which

318 may even serve as the missing ingredient stabilizing some 376 is the habitat provided to invertebrates by the recently  
 319 complex ecological systems<sup>65</sup>. 377 discovered rock-boring teredinid shipworm (*Lithoredo*  
 320 We next explore the effects of ecosystem engineering 378 *abatanica*)<sup>66</sup>. Here, freshwater invertebrates are serviced  
 321 by allowing species to produce abiotic modifiers as addi- 379 by the habitat modifications engineered by the shipworm,  
 322 tional nodes in the ecological network (Fig. 1). These 380 linking species indirectly via an abiotic effect (in our  
 323 modifier nodes produced by engineers can serve to fulfill 381 framework via a modifier node). As the frequency of  
 324 resource or service requirements for other species. The 382 service interactions increases, the negative effects asso-  
 325 parameter  $\eta$  defines the mean number of modifiers pro- 383 ciated with rare engineers is diminished (Fig. 4A). In-  
 326 duced per species in the pool, drawn from a Poisson dis- 384 creasing service interactions both elevates the competi-  
 327 tribution (see Methods and Supplementary Appendix 1 385 tive strength of species receiving services (from species  
 328 for details). If a species makes  $\leq 1$  modifier, we label it 386 and/or modifiers), while creating more interdependen-  
 329 an engineer. As the mean number of modifiers/species 387 cies between and among species. As trophic interactions  
 330  $\eta$  increases, both the number of engineers in the pool as 388 are replaced by service interactions, previously vulnera-  
 331 well as the number of modifiers each engineer makes in- 389 ble species gain a competitive foothold and persist (Fig.  
 332 creases. As detailed in Supplementary Appendix 1, mul- 390 3, inset), lowering rates of primary extinctions (Fig. 4A).  
 333 tiple engineers can make the same modifier, such that 391 The costs of these added services to the community are  
 334 engineering redundancies are introduced when  $\eta$  is large. 392 an increased rate of secondary extinctions (Fig. 4B) and  
 335 When an engineer colonizes the community, so do its 393 higher species turnover (Fig. 4C). Low rates of primary  
 336 modifiers, which other species in the system may interact 394 extinction coupled with high rates of secondary extinc-  
 337 with. When engineers are lost, their modifiers will also be 395 tion mean that extinctions are less common but lead to  
 338 lost, though can linger in the community for a period of 396 larger cascades.

341 Increasing engineering has significant consequences for 397 community stability, but these effects also are sensitive  
 342 to the frequency of service interactions within the com- 398 munity. We measure community stability by *i*) rates of  
 343 primary versus secondary extinctions, *ii*) species persis- 399 tence, and *iii*) steady state community diversity. All  
 344 measures were averaged over each species within the 400 community across assembly time.

349 As the number of engineers increase, mean rates of pri- 401 mary extinction are first elevated and then decline (Fig.  
 350 4a). At the same time, the mean rates of secondary ex- 402 tinction systematically decline and persistence systemat-  
 351 ically increases (Fig. 4b-c). Higher rates of primary ex- 403 tinction coupled with lower rates of secondary extinction  
 352 that occur when modifiers are rare ( $0 < \eta \leq 0.5$ ) mean 404 that extinctions are common, but of limited magnitude  
 353 such that disturbances are compartmentalized. As mod- 405 ifiers become more common both primary and secondary  
 354 extinction rates decline, which corresponds to increased 406 persistence. We suggest two mechanisms that may pro-  
 355 duce the observed results. First, when engineers and 407 361 modifiers are present but rare, they provide additional  
 356 resources for consumers. This stabilization of consumers 408 362 ultimately results in increased vulnerability of prey, such  
 357 that the cumulative effect is increased competitive ex- 409 363 clusion of prey and higher rates of primary extinction  
 358 410 (Fig. 4a). Second, when engineers and their modifiers  
 359 411 are common ( $\eta > 0.5$ ) the available niche space expands,  
 360 412 lowering competitive overlap and suppressing both pri-  
 361 413 mary and secondary extinctions. Notably the presence  
 362 414 of even a small number of engineers serves to limit the  
 363 415 magnitude of secondary extinction cascades.

364 Increasing the frequency of service interactions pro- 416 While the importance of engineering timescales has  
 365 motes service interactions between species and engi- 417 been emphasized previously<sup>23</sup>, redundant engineering  
 366 neered modifiers (Fig. 1). A topical example of the latter 418 has been assumed to be unimportant<sup>12</sup>. We argue that  
 367 419 redundancy may be an important component of highly  
 368 420 engineered systems, and particularly relevant when the  
 369 421 effects of engineers increase their own fitness<sup>59</sup> as is gen-

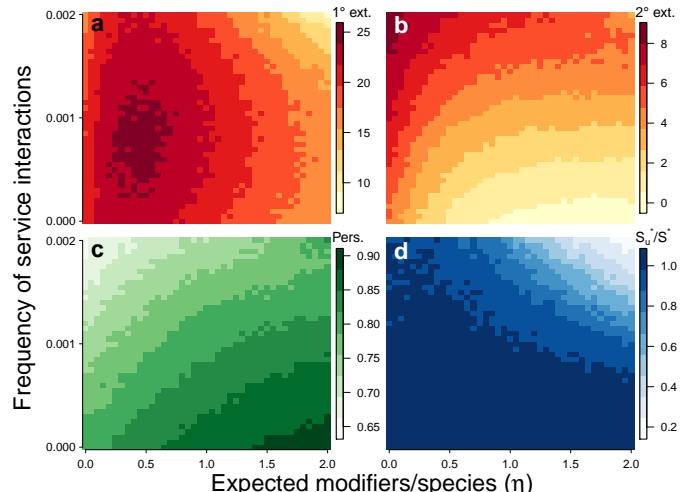


Figure 4. Community stability as a function of the frequency of service interactions and modifiers per species. **a**, Mean rates of primary extinction, where primary extinctions occur from competitive exclusion of consumers over shared resources. **b**, Mean rates of secondary extinction, which cascade 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.

397 While the importance of engineering timescales has  
 398 been emphasized previously<sup>23</sup>, redundant engineering  
 399 has been assumed to be unimportant<sup>12</sup>. We argue that  
 400 redundancy may be an important component of highly  
 401 engineered systems, and particularly relevant when the  
 402 effects of engineers increase their own fitness<sup>59</sup> as is gen-

403 really assumed to be the case with niche construction<sup>62</sup>. 461 multitype interactions can produce model communities  
 404 The vast majority of contemporary ecosystem engineer- 462 with realistic structures and dynamics. Moreover, the  
 405 ing case studies focus on single taxa, such that redundant 463 inclusion of ecosystem engineering by way of modifier  
 406 engineers appear rare<sup>12</sup>. However if we consider longer 464 nodes reveals that low levels of engineering may be  
 407 timescales, increasing diversity of engineering clades may 465 expected to produce higher rates of extinction while  
 408 promote redundancy, and in some cases this may feed 466 limiting the size of extinction cascades, and that engi-  
 409 back to accelerate diversification<sup>14</sup>. Such positive feed- 467 neering redundancy – whether it is common or rare  
 410 back mechanisms likely facilitated the global changes in- 468 – may have considerable dynamical implications. We  
 411 duced by cyanobacteria in the Proterozoic<sup>24,25</sup> among 469 suggest that including the effects of engineers, either  
 412 other large-scale engineering events in the history of 470 explicitly as we have done here, or otherwise, is vital  
 413 life<sup>24</sup>. Engineering redundancies are likely important on 471 for understanding the inter-dependencies that define  
 414 shorter timescales as well. For example, diverse sessile 472 ecological systems. As past ecosystems have funda-  
 415 epifauna on shelled gravels in shallow marine environ- 473 mentally altered the landscape on which contemporary  
 416 ments are facilitated by the engineering of their ances- 474 communities interact, future ecosystems will be defined  
 417 tors, such that the engineered effects of the clade de- 475 by the influence of engineering today. Given the rate  
 418 termine the future fitness of descendants<sup>67</sup>. In the mi- 476 and magnitude that humans are currently engineering  
 419 crobiome, redundant engineering may be very common 477 environments<sup>69</sup>, understanding the role of ecosystem  
 420 due to the influence of horizontal gene transfer in struc- 478 engineers is tantamount to understanding our own.

421 turing metabolite production<sup>68</sup>. In these systems, re-  
 422 dundancy in the production of shared metabolicic re-  
 423 sources may play a key role in community structure and  
 424 dynamics<sup>64,65</sup>.

425 When there are few engineers, each modifier in the  
 426 community tends to be unique to a particular engineer-  
 427 ing species. Engineering redundancies increase linearly  
 428 with  $\eta$  (Supplementary Appendix 1; Fig. S6), such that  
 429 the loss of an engineer will not necessarily lead to the  
 430 loss of engineered modifiers. We examine the effects of  
 431 this redundancy by comparing our results to those pro-  
 432 duced by the same model, but where each modifier is  
 433 uniquely produced by a single species. Surprisingly, the  
 434 lack of engineering redundancies does not alter the gen-  
 435 eral relationship between engineering and measures of  
 436 community stability (Fig. S7). However we find that  
 437 redundancies play a central role in maintaining species  
 438 diversity. When engineering redundancies are allowed,  
 439 steady state community richness  $S^*$  does not vary con-  
 440 siderably with increasing service interactions and engi-  
 441 neering (Fig. S8A). In contrast, when redundant engi-  
 442 neering is not allowed, steady state community richness  
 443  $S_u^*$  declines sharply (Figs. 4D, S8B).

444 Communities lacking redundancy have lower species  
 445 richness because species' trophic and service dependen-  
 446 cies are unlikely to be fulfilled, precluding colonization  
 447 (Fig. S8C,D). Colonization occurs only when trophic and  
 448 service dependencies are fulfilled. A species requiring  
 449 multiple engineered modifiers, each uniquely produced,  
 450 means that each required entity must precede coloniza-  
 451 tion. This magnifies the role of priority effects in con-  
 452 straining assembly order<sup>15</sup>, precluding many species from  
 453 colonizing. In contrast, redundancy increases the tempo-  
 454 ral stability of species' niches while minimizing priority  
 455 effects by allowing multiple engineers to fulfill dependen-  
 456 cies. Our results thus suggest that redundant engineers  
 457 may play important roles in assembling ecosystems by  
 458 lowering the barriers to colonization thereby promoting  
 459 community diversity.

460 We have shown that the dynamics of assembly with

461 multitype interactions can produce model communities  
 462 with realistic structures and dynamics. Moreover, the  
 463 inclusion of ecosystem engineering by way of modifier  
 464 nodes reveals that low levels of engineering may be  
 465 expected to produce higher rates of extinction while  
 466 limiting the size of extinction cascades, and that engi-  
 467 neering redundancy – whether it is common or rare  
 468 – may have considerable dynamical implications. We  
 469 suggest that including the effects of engineers, either  
 470 explicitly as we have done here, or otherwise, is vital  
 471 for understanding the inter-dependencies that define  
 472 ecological systems. As past ecosystems have funda-  
 473 mentally altered the landscape on which contemporary  
 474 communities interact, future ecosystems will be defined  
 475 by the influence of engineering today. Given the rate  
 476 and magnitude that humans are currently engineering  
 477 environments<sup>69</sup>, understanding the role of ecosystem  
 478 engineers is tantamount to understanding our own.

## 479 Methods

480 We model an ecological system with a network where nodes rep-  
 481 resent *ecological entities* such as populations of species and/or the  
 482 presence of abiotic modifiers affecting species such as (examples).  
 483 Following Pilai et al.<sup>31</sup>, we do not track the abundances of entities  
 484 but only track their presence or absence. The links of the network  
 485 represent interactions between pairs of entities (x,y). We distin-  
 486 guish three types of such interactions: x eats y, x needs y to be  
 487 present, x makes modifier y.

488 The assembly process entails two steps: first a source pool of  
 489 species is created, followed by colonization/extinction into/from a  
 490 local community. The model is initialized by creating  $S$  species  
 491 and  $M = \eta S$  modifiers, such that  $N = S + M$  is the average total  
 492 number of entities and  $\eta$  is the average number of modifiers per  
 493 species in the system. For each pair of species (x,y) there is a  
 494 probability  $p_e$  that x eats y and probability  $p_n$  that x needs y. For  
 495 each pair of species x and modifier m, there is a probability  $q_e$   
 496 that x eats m and a probability  $q_n$  that species x needs modifier  
 497 m. Additionally, each species makes a number of modifiers that is  
 498 drawn from a Poisson distribution with mean  $\mu = \eta e / (e - 1)$  where  
 499 e is Euler's number. Once the number of modifiers per species is  
 500 determined, each modifier is assigned to a species independently.  
 501 This means that multiple species may make the same modifier, and  
 502 that there may be some modifiers that are not made by any species,  
 503 which are eliminated from the pool.

504 In addition to interactions with ecosystem entities, there can be  
 505 interactions with a basal resource, which is always present. The  
 506 first species always eats this resource, such that there is always a  
 507 primary producer in the pool. Other species eat the basal resource  
 508 with probability  $p_e$ . Species with zero assigned trophic interac-  
 509 tions are assumed to be primary producers. See Supplementary  
 510 Appendix 1 for additional details on defining the source pool.

511 We then consider the assembly of a community which at any  
 512 time will contain a subset of entities in the pool and always the  
 513 basal resource. In time, the entities in the community are updated  
 514 following a set of rules. A species from the pool can colonize the  
 515 community if the following conditions are met: 1) all entities that  
 516 a species needs are present in the community, and 2) at least one  
 517 entity that a species eats is present in the community. If a colo-  
 518 nization event is possible, it occurs stochastically in time with rate  
 519  $r_c$ .

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

$$\sigma_i = c_{nn} n_i - c_{ee} e_i - c_{vv} v_i, \quad (1)$$

523 where  $n_i$  is the number of entities that species  $i$  needs,  $e_i$  is the  
 524 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

#### 558 Data availability

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

#### 561 Code availability

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

- 
- [1] Paine, R. T. Food web complexity and species diversity. *Am. Nat.* **100**, 65–75 (1966).
- [2] Dunne, J. A., Williams, R. J. & Martinez, N. D. Food-web structure and network theory: the role of connectance and size. *Proc Natl Acad Sci USA* **99**, 12917–12922 (2002).
- [3] Pascual, M. & Dunne, J. *Ecological Networks: Linking Structure to Dynamics in Food Webs* (Oxford University Press, Oxford, UK, 2006). URL <http://www.google.com/search?client=safari&rls=en-us&q=Ecological+Networks&ie=UTF-8&oe=UTF-8>.
- [4] Bascompte, J. & Jordano, P. *Mutualistic Networks*. Monographs in Population Biology (Princeton University Press, Princeton, NJ, 2013). URL <https://books.google.com/books?id=xVWzAAAAQBAJ>.
- [5] May, R. M. Will a large complex system be stable? *Nature* **238**, 413–414 (1972).
- [6] Gross, T., Levin, S. A. & Dieckmann, U. Generalized models reveal stabilizing factors in food webs. *Science* **325**, 747–750 (2009).
- [7] Allesina, S. & Tang, S. Stability criteria for complex ecosystems. *Nature* **483**, 205–208 (2012).
- [8] Ponisio, L. C., Gaiarsa, M. P. & Kremen, C. Opportunistic attachment assembles plant–pollinator networks. *Ecol Lett* **20**, 1261–1272 (2017).
- [9] Guimarães Jr, P. R., Pires, M. M., Jordano, P., Bascompte, J. & Thompson, J. N. Indirect effects drive co-evolution in mutualistic networks. *Nature* **18**, 586 (2017).
- [10] Kéfi, S., Miele, V., Wieters, E. A., Navarrete, S. A. & Berlow, E. L. How structured is the entangled bank? the surprisingly simple organization of multiplex ecological networks leads to increased persistence and resilience. *PLoS Biol* **14**, e1002527 (2016).
- [11] Pilosof, S., Porter, M. A., Pascual, M. & Kéfi, S. The multilayer nature of ecological networks. *Nature Ecol Evol* **1**, 0101 (2017).
- [12] Lawton, J. H. What do species do in ecosystems? *Oikos* **71**, 367–374 (1994).
- [13] Odling-Smeel, J., Erwin, D. H., Palkovacs, E. P., Feldman, M. W. & Laland, K. N. Niche construction theory: a practical guide for ecologists. *Q Rev Biol* **88**, 4–28 (2013).
- [14] Odling-Smeel, F., Laland, K. & Feldman, M. *Niche Construction: The Neglected Process in Evolution (MPB-37)*. Monographs in Population Biology (Princeton University Press, Princeton, NJ, 2013). URL <https://books.google.com/books?id=0k-NujtWMZOC>.
- [15] Fukami, T. Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. *Annu Rev Ecol Evol Syst* **46**, 1–23 (2015).
- [16] Jones, C. G., Lawton, J. H. & Shachak, M. Organisms as ecosystem engineers. *Oikos* **69**, 373–386 (1994).
- [17] Olff, H. et al. Parallel ecological networks in ecosystems. *Philos T Roy Soc B* **364**, 1755 (2009).
- [18] Haynes, G. Elephants (and extinct relatives) as earth-movers and ecosystem engineers. *Geomorphology* **157–158**, 99 – 107 (2012).
- [19] Pringle, R. M. Elephants as agents of habitat creation for small vertebrates at the patch scale. *Ecology* **89**, 26–33 (2008).
- [20] Reichman, O. & Seabloom, E. W. The role of pocket gophers as subterranean ecosystem engineers. *Trends Ecol Evol* **17**, 44 – 49 (2002).
- [21] Moore, J. W. Animal Ecosystem Engineers in Streams. *BioScience* **56**, 237–246 (2006). arXiv:<http://oup.prod.sis.lan/bioscience/article-pdf/56/3/237/26897844/56-3-237.pdf>.
- [22] Meyer, S. T., Leal, I. R., Tabarelli, M. & Wirth, R. Ecosystem engineering by leaf-cutting ants: nests of atta cephalotes drastically alter forest structure and microclimate. *Ecol Entomol* **36**, 14–24 (2011).
- [23] Hastings, A. et al. Ecosystem engineering in space and time. *Ecol Lett* **10**, 153–164 (2007).
- [24] Erwin, D. H. Macroevolution of ecosystem engineering, niche construction and diversity. *Trends Ecol Evol* **23**, 304 – 310 (2008).
- [25] Schirrmeyer, B. E., de Vos, J. M., Antonelli, A. & Bagheri, H. C. Evolution of multicellularity coincided with increased diversification of cyanobacteria and the great oxidation event. *Proc Natl Acad Sci USA* **110**, 1791–1796 (2013). arXiv:<https://www.pnas.org/content/110/5/1791.full.pdf>.

- [26] Loladze, I. & Elser, J. J. The origins of the Red-field nitrogen-to-phosphorus ratio are in a homoeostatic protein-to-rRNA ratio. *Ecol Lett* **14**, 244–250 (2011).
- [27] Woodward, G., Perkins, D. M. & Brown, L. E. Climate change and freshwater ecosystems: impacts across multiple levels of organization. *Philos T Roy Soc B* **365**, 2093–2106 (2010).
- [28] Brose, U. *et al.* Climate change in size-structured ecosystems. *Philos T Roy Soc B* **367**, 2903–2912 (2012).
- [29] Gibert, J. P. Temperature directly and indirectly influences food web structure. *Sci Rep-UK* **9**, 5312 (2019).
- [30] Getz, W. M. Biomass transformation webs provide a unified approach to consumer-resource modelling. *Ecol Lett* **14**, 113–124 (2011).
- [31] Pillai, P., Gonzalez, A. & Loreau, M. Metacommunity theory explains the emergence of food web complexity. *Proc Natl Acad Sci USA* **108**, 19293–19298 (2011).
- [32] Bascompte, J., Jordano, P., Melián, C. J. & Olesen, J. M. The nested assembly of plant-animal mutualistic networks. *Proc Natl Acad Sci USA* **100**, 9383–9387 (2003).
- [33] Gravel, D., Massol, F., Canard, E., Mouillot, D. & Mouquet, N. Trophic theory of island biogeography. *Ecol Lett* **14**, 1010–1016 (2011).
- [34] Piechnik, D. A., Lawler, S. P. & Martinez, N. D. Food-web assembly during a classic biogeographic study: species “trophic breadth” corresponds to colonization order. *Oikos* **117** (2008).
- [35] Williams, R. J. & Martinez, N. D. Simple rules yield complex food webs. *Nature* **404**, 180–183 (2000).
- [36] Williams, R. & Martinez, N. Limits to trophic levels and omnivory in complex food webs: Theory and data. *Am Nat* **163**, 458–468 (2004). arXiv:<https://doi.org/10.1086/381964>.
- [37] Lafferty, K. D., Dobson, A. P. & Kuris, A. M. Parasites dominate food web links. *Proc Natl Acad Sci USA* **103**, 11211–11216 (2006).
- [38] Turney, S. & Buddle, C. M. Pyramids of species richness: the determinants and distribution of species diversity across trophic levels. *Oikos* **125**, 1224–1232 (2016).
- [39] Rohr, R. P., Saavedra, S. & Bascompte, J. On the structural stability of mutualistic systems. *Science* **345**, 1253497–1253497 (2014).
- [40] Valdovinos, F. S. *et al.* Niche partitioning due to adaptive foraging reverses effects of nestedness and connectance on pollination network stability. *Ecol Lett* **19**, 1277–1286 (2016).
- [41] Valdovinos, F. S. Mutualistic networks: moving closer to a predictive theory. *Ecol Lett* **0** (2019).
- [42] Krishna, A., Guimarães Jr, P. R., Jordano, P. & Bascompte, J. A neutral-niche theory of nestedness in mutualistic networks. *Oikos* **117**, 1609–1618 (2008).
- [43] Stouffer, D. B. Compartmentalization increases food-web persistence. *Proc Natl Acad Sci USA* **108**, 3648–3652 (2011).
- [44] Gilarranz, L. J., Rayfield, B., Liñán-Cembrano, G., Bascompte, J. & Gonzalez, A. Effects of network modularity on the spread of perturbation impact in experimental metapopulations. *Science* **357**, 199–201 (2017).
- [45] Araújo, M. S. *et al.* Nested diets: a novel pattern of individual-level resource use. *Oikos* **119**, 81–88 (2010).
- [46] Pires, M. M., Prado, P. I. & Guimarães Jr, P. R. Do food web models reproduce the structure of mutualistic networks? *PLoS ONE* **6**, e27280 (2011).
- [47] Massol, F. *et al.* Linking community and ecosystem dynamics through spatial ecology. *Ecol Lett* **14**, 313–323 (2011).
- [48] Ponisio, L. C. *et al.* A network perspective for community assembly. *Front Ecol Evol* **7**, 103 (2019).
- [49] Yeakel, J. D., Guimarães Jr, P. R., Bocherens, H. & Koch, P. L. The impact of climate change on the structure of Pleistocene food webs across the mammoth steppe. *Proc Roy Soc B* **280**, 20130239–20130239 (2013).
- [50] Bond, W. J., Lawton, J. H. & May, R. M. Do mutualisms matter? Assessing the impact of pollinator and disperser disruption on plant extinction. *Phil Trans Roy Soc B* **344**, 83–90 (1994).
- [51] Colwell, R. K., Dunn, R. R. & Harris, N. C. Coextinction and persistence of dependent species in a changing world. *Ann Rev Ecol Evol Sys* **43**, 183–203 (2012).
- [52] Dunn, R. R., Harris, N. C., Colwell, R. K., Koh, L. P. & Sodhi, N. S. The sixth mass coextinction: are most endangered species parasites and mutualists? *Proc Roy Soc B* **276**, 3037–3045 (2009).
- [53] Almeida-Neto, M., Guimarães, P., Guimarães Jr, P. R., Loyola, R. D. & Ulrich, W. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos* **117**, 1227–1239 (2008).
- [54] Ramos-Jiliberto, R., Valdovinos, F. S., Moisset de Españoñas, P. & Flores, J. D. Topological plasticity increases robustness of mutualistic networks. *J Anim Ecol* **81**, 896–904 (2012).
- [55] Vieira, M. C. & Almeida Neto, M. A simple stochastic model for complex coextinctions in mutualistic networks: robustness decreases with connectance. *Ecol Lett* **18**, 144–152 (2015).
- [56] Laland, K. N. *et al.* The extended evolutionary synthesis: its structure, assumptions and predictions. *Proc Roy Soc B* **282**, 20151019 (2015).
- [57] Gupta, M., Prasad, N., Dey, S., Joshi, A. & Vidya, T. Niche construction in evolutionary theory: the construction of an academic niche? *J Gen* **96**, 491–504 (2017).
- [58] Feldman, M. W., Odling-Smee, J. & Laland, K. N. Why Gupta et al.’s critique of niche construction theory is off target. *J Gen* **96**, 505–508 (2017).
- [59] Cuddington, K. Invasive engineers. *Ecol Model* **178**, 335–347 (2004).
- [60] Wright, J. P. & Jones, C. G. Predicting effects of ecosystem engineers on patch-scale species richness from primary productivity. *Ecology* **85**, 2071–2081 (2004).
- [61] Kylafis, G. & Loreau, M. Ecological and evolutionary consequences of niche construction for its agent. *Ecol Lett* **11**, 1072–1081 (2008).
- [62] Krakauer, D. C., Page, K. M. & Erwin, D. H. Diversity, dilemmas, and monopolies of niche construction. *Am Nat* **173**, 26–40 (2009).
- [63] Laland, K. N., Odling-Smee, F. J. & Feldman, M. W. Evolutionary consequences of niche construction and their implications for ecology. *Proc Natl Acad Sci USA* **96**, 10242–10247 (1999).
- [64] Kallus, Y., Miller, J. H. & Libby, E. Paradoxes in leaky microbial trade. *Nat Commun* **8**, 1361 (2017).
- [65] Muscarella, M. E. & O’Dwyer, J. P. Ecological insights from the evolutionary history of microbial innovations. *bioRxiv* 220939 (2017).
- [66] Shipway, J. R. *et al.* A rock-boring and rock-ingesting freshwater bivalve (shipworm) from the philippines. *Proc Roy Soc B* **286**, 20190434 (2019).

- 775 [67] Kidwell, S. M. Taphonomic feedback in Miocene assem-  
 776 blages: testing the role of dead hardparts in benthic com-  
 777 munities. *Palaios* **1**, 239–255 (1986).
- 778 [68] Polz, M. F., Alm, E. J. & Hanage, W. P. Horizontal  
 779 gene transfer and the evolution of bacterial and archaeal  
 780 population structure. *Trends Genet* **29**, 170 – 175 (2013).
- 781 [69] Corlett, R. T. The anthropocene concept in ecology and  
 782 conservation. *Trends in Ecology & Evolution* **30**, 36 – 41  
 783 (2015).
- 784 [70] Gillespie, D. T. Exact stochastic simulation of coupled  
 785 chemical reactions. *J Phys Chem* **81**, 2340–2361 (1977).
- 786 [71] Williams, R. J. & Purves, D. W. The probabilistic niche  
 787 model reveals substantial variation in the niche structure  
 788 of empirical food webs. *Ecology* **92**, 1849–1857 (2011).
- 789 [72] Warren, C. P., Pascual, M., Lafferty, K. D. & Kuris,  
 790 A. M. The inverse niche model for food webs with para-  
 791 sites. *Theor Ecol* **3**, 285–294 (2010).

792 **Acknowledgements**

793 We would like to thank Uttam Bhat, Irina Birskis Barros, Emmet  
 794 Brickowski, Ashkaan Fahimipour, Jean Philippe Gibert, Chris P  
 795 Kempes, Taran Rallings, Samuel Scarpino, Megha Suswaram, and  
 796 Ritwika VPS for insightful discussions and comments throughout  
 797 the lengthy gestation of this manuscript. The original idea  
 798 was conceived at the Networks on Networks Working Group in  
 799 Göttingen, Germany (2014) and the Santa Fe Institute (2015).  
 800 This work was formerly prepared as a part of the Ecological  
 801 Network Dynamics Working Group at the National Institute  
 802 for Mathematical and Biological Synthesis (2015–2019), spon-  
 803 sored by the National Science Foundation through NSF Award  
 804 DBI-1300426, with additional support from The University of  
 805 Tennessee, Knoxville. Infinite revisions were conducted at the  
 806 Santa Fe Institute made possible by travel awards to JDY and TG.  
 807 Additional support came from UC Merced startup funds to JDY,  
 808 the International Centre for Theoretical Physics ICTP-SAIFR,  
 809 FAPESP (2016/01343-7) and CNPq (302049/2015-0) to MAMA,  
 810 CNPq and FAPESP (2018/14809-0) to PRG, and DFG research  
 811 unit 1748 and EPSRC (EP/N034384/1) to TG.

812

813

814 **Author contributions**

815 JDY and TG conceived of the model framework. JDY, MMP,  
 816 MAMA, and TG designed the analyses. JDY, MMP, MAMA,  
 817 JLOD, PRG, DG, and TG analyzed the results and contributed  
 818 to multiple versions of the manuscript.

819

820 **Competing interests**

821 The authors declare no competing interests.

822

## SUPPLEMENTARY METHODS

876

## Appendix 2: Model parameterizations

823

## Appendix 1: Building the source pool

824 Here and henceforth, we refer to the assembly model presented  
 825 in the main text as the ENIgMa model (E:eat, N:need, Ig:ignore,  
 826 Ma:make). To initiate the ENIgMa assembly model, we must first  
 827 construct the source pool, where each ecological entity (species and  
 828 modifiers) is defined by its potential interactions with each other.  
 829 The source pool interaction matrix  $\mathbf{P}$  is generated by first setting  
 830 the number of species in the pool  $S_P$  and determining the number  
 831 of modifiers  $M_P$  that are made by ecosystem engineers. The result-  
 832 ing matrix is  $N_P \times N_P$  where  $N_P = S_P + M_P$ , and is subdivided  
 833 into four quadrants, only two of which play a role here: species-  
 834 species interactions and species-modifier interactions (see Fig. 1).

835 In these two quadrants, the expected frequency of eat interactions  
 836  $E\{p_e\}$  and the expected frequency of need interactions  $E\{p_n\}$  are  
 837 free parameters, as is the expected number of modifiers made per  
 838 species  $E\{M_i\} = \eta$ . Here and throughout, we simplify this pa-  
 839 rameter space by assuming that the frequency of eat and need in-  
 840 teractions for species-species (SS) interactions and species-modifier  
 841 (SM) interactions are equivalent, such that  $E_{SS}\{p_e\} = E_{SM}\{p_e\}$   
 842 and  $E_{SS}\{p_n\} = E_{SM}\{p_n\}$ . For each species, a set number of modi-  
 843 fiers is drawn from  $Pois(\eta)$ , such that the expected proportion of  
 844 species that are engineers (species that make modifiers) is  $1 - e^{-\eta}$ .  
 845 If a particular modifier is randomly and independently drawn for  
 846 a given engineer from a complete list of all possible modifiers, such  
 847 that multiple species – with some probability – can make the same  
 848 modifier, the expected number of modifiers is

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

849 where  $e$  is Euler's number. The frequency of engineering (make)  
 850 interactions is then calculated as

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

851 Finally the frequency of the non-interaction is calculated as  
 852  $E_{SS}\{p_\emptyset\} = 1 - E_{SS}\{p_e\} + E_{SS}\{p_n\}$  and  $E_{SM}\{p_\emptyset\} = 1 - E_{SM}\{p_e\} +$   
 853  $E_{SM}\{p_n\} + E_{SM}\{p_m\}$  for species-species and species-modifier in-  
 854 teractions, respectively. Pairwise interactions are assigned ran-  
 855 domly from these probabilities between species-species and species-  
 856 modifiers independently in both quadrants, such that the source  
 857 pool matrix has no imbued structure apart from the number of  
 858 species, the number of modifiers, and the frequency of each direc-  
 859 tional interaction type. Each source pool is provided a *basal re-*  
 860 *source* (the first row/column). A species with a trophic interaction  
 861 to this resource is identified as an autotroph (or mixotroph depend-  
 862 ing on its other trophic interactions). If they do not have service  
 863 dependencies with other species/modifiers, it is these species that  
 864 are uniquely able to initiate assembly.

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

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

870 such that the proportion of redundant modifiers  $\phi$  is

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

871 Accordingly, we find that the number of redundant modifiers in-  
 872 creases linearly with  $\eta$ , while the proportion of modifiers that are  
 873 redundant is fixed. Figure S6A,B shows both analytical expecta-  
 874 tions and numerically-derived measures for  $E\{M_P\}_{\text{redundant}}$  and  
 875  $\phi$ , respectively.

877 Simulations described in the main text have default parameter-  
 878 izations of  $S = 200$ ,  $p_e = 0.01$ ,  $c_n = \pi$ ,  $c_e = \sqrt{2}$ ,  $c_v = 1$ , and 4000  
 879 iterations (time-steps).

880 **Assembly without ecosystem engineering** Here we set the  
 881 average number of modifiers made per species  $\eta = 0$  and the prob-  
 882 ability of need interactions in the species pool  $p_n = 0.002$ .

883 **Structure and dynamics of mutualisms** Again we used the de-  
 884 fault parameterizations but set  $\eta = 0$ , while varying  $p_n \in [0, 0.002]$ .

885 **Assembly with ecosystem engineering** Here we used the de-  
 886 fault parameterizations but varied  $\eta \in [0, 2]$  and  $p_n \in [0, 0.002]$ .

## 887 Appendix 3: Comparison to Niche Model

888 We compared certain structural features of ENIgMa at steady  
 889 state to those of the Niche Model<sup>35</sup>. Comparisons were restricted  
 890 to networks constructed in the absence of engineering because engi-  
 891 neers introduce indirect effects that are not considered in static food  
 892 web models, and may make such comparisons irrelevant. While  
 893 there are many similarities, there are also some important differ-  
 894 ences, some of which are highlighted in the main text. While we  
 895 consider a comparison of our framework with other food web mod-  
 896 els such as the Niche Model relevant, we emphasize that the mo-  
 897 tivations underlying both are distinct. Our approach is intended  
 898 to provide a deeper understanding into how multitype depen-  
 899 dencies between species and the environment impact the dynamics of  
 900 community assembly. While capturing general qualitative features  
 901 of empirical systems demonstrates that the dynamics we consider  
 902 are ecologically relevant, the goal of our approach is distinct from  
 903 that of static food web models, which aim to maximize structural  
 904 similarities between model and empirical systems<sup>35,71</sup>.

905 We compared steady state ecological networks that emerge  
 906 from ENIgMa (described in Materials and Methods, main text)  
 907 with food webs constructed from the Niche Model<sup>35</sup> with simi-  
 908 lar species richness and connectance. Because species richness and  
 909 connectance of the Niche Model are often altered by eliminating dis-  
 910 connected species, we compared *i*) species richness, *ii*) connectance,  
 911 *iii*) mean species degree, *iv*) standard deviation of out-degree dis-  
 912 tributions, and *v*) standard deviation of in-degree distributions av-  
 913 eraged across 1000 replicates for each model.

914 We found that all measures resulted in fairly similar values be-  
 915 tween ENIgMa and the Niche Model food webs with a some im-  
 916 portant differences (Figs. S2,S3). While similar, ENIgMa pro-  
 917 duces consistently lower values of connectance, mean species de-  
 918 gree, as well as standard deviations of the in- and out-degree dis-  
 919 tributions. This means that the food webs produced by ENIgMa  
 920 are more sparsely connected with less variance between species.  
 921 These results were expected, as the Niche Model assumes system-  
 922 atically increasing dietary ranges with higher niche values, whereas  
 923 the trophic interactions assigned to species in the source pool of  
 924 ENIgMa are drawn independently. An important difference be-  
 925 tween the Niche Model and ENIgMa is that we do not distinguish  
 926 between predators and parasites. A different framework known  
 927 as the Inverse Niche Model<sup>72</sup> has been proposed to address par-  
 928 asitic interactions. The Inverse Niche Model assumes increasing  
 929 specialization with feeding hierarchies, which would serve to lower  
 930 the average generality of species (lower degree). In addition, the  
 931 Inverse Niche model outputs lower standard deviations of in- and  
 932 out-degree distributions. Together these trends suggest that the  
 933 qualitative structural differences that we observe for the assembly  
 934 and Niche model may reflect an important structural distinction  
 935 between food webs that do and do not include parasitic species.

936

#### Appendix 4: Measures of generality

937 The trophic breadth of potential colonizers is thought to play an  
 938 important role in community assembly. The definition of a specialist  
 939 or generalist to some degree depends on the size and connectance  
 940 of the larger food web. Trophic generality for a species  $i$  is defined  
 941  $G_i = k_i^{\text{in}}/(L/S)$ , where  $k_i^{\text{in}}$  is the in-degree, or number of resources  
 942 consumed by species  $i^{35}$ . A species is classified as a generalist if  
 943 the number of its trophic interactions is greater than the average  
 944 number of links per species, or  $G_i > 1$ , and a specialist if  $G_i < 1$ ,  
 945 where a community can be described by the proportion of special-  
 946 ists found therein. For interaction networks that are assembling  
 947 over time, generality can be scaled by a number of different mea-  
 948 sures of  $L/S$ , and this has a large effect on our interpretation of the  
 949 role of generality in community assembly. For instance,  $L/S$  may  
 950 be quantified by either including all autotrophic species or only au-  
 951 totrophic functional groups. Furthermore, the scaling of generality  
 952 may be made with respect to the current state of the community  
 953 at each point in time, or with respect to the community at steady  
 954 state. For instance, in their investigation of assembling mangrove  
 955 food webs, Piechnik et al.<sup>34</sup> scaled trophic breadth to a standard  
 956 steady state value of  $L^*/S^* = 0.2$  averaged across 102 food webs.

957 To examine how our assessment of the role of generalism over  
 958 the course of assembly changes based on the application of differ-  
 959 ent scalings, we employ three different measures of  $L/S$  to calculate  
 960  $G_i$ : 1)  $G_i^{\text{all}}$ , where  $L$  accounts for all links in the food web and  $S$   
 961 accounts for all species relative to each time interval in the assem-  
 962 bly process (circles; Fig. S4b); 2)  $G_i^{\text{hetero}}$ , where we consider only  
 963 the links and species richness of heterotrophs, excluding autotrophs

964 (points; Fig. S4b); 3)  $G_i^*$ , where  $L$  and  $S$  are measured with re-  
 965 spect to the communities at steady state, which is most similar to  
 966 the measure used to evaluate assembling mangrove food webs (dia-  
 967 monds; Fig. S4b). Whether trophic breadth is scaled to the current  
 968 state of  $L/S$  or the steady state value of  $L^*/S^*$  has a large influence  
 969 on the estimated proportion of generalists in the community, par-  
 970 ticularly when the size of the system is small. We observe that for  
 971  $G_i^{\text{all}}$ , the system is initially assembled by specialist species, though  
 972 over the course of assembly the proportion of specialists relative  
 973 to generalists declines to intermediate values (circles represent-  
 974 ing the average over replicates in Fig. S4). If only the trophic links  
 975 between non-autotrophs are considered as in  $G_i^{\text{hetero}}$ , specialists  
 976 still dominate early in assembly, but there is a greater range, such  
 977 that some systems can be described by a mixed proportion of spe-  
 978 cialists and generalists (individual points representing independent  
 979 replicates in Fig. S4).

980 The different normalizations by which generality is measured  
 981 will impact the interpretation of both empirical and model systems  
 982 alike. In our framework, species colonizing early in the assem-  
 983 bly process are generalists compared to how the term is defined  
 984 at the steady state, but they are functionally specialists with re-  
 985 spect to the assembling community. For example, a species that is  
 986 trophically connected to 10 resource species in the source pool may  
 987 colonize a community where it is consuming a small subset of its  
 988 potential range. As the community grows, that species may real-  
 989 ize more of its trophic niche if those resource species subsequently  
 990 colonize the system. To what end we label this species a generalist  
 991 or specialist relative to the assembling community is thus subject  
 992 to multiple interpretations.

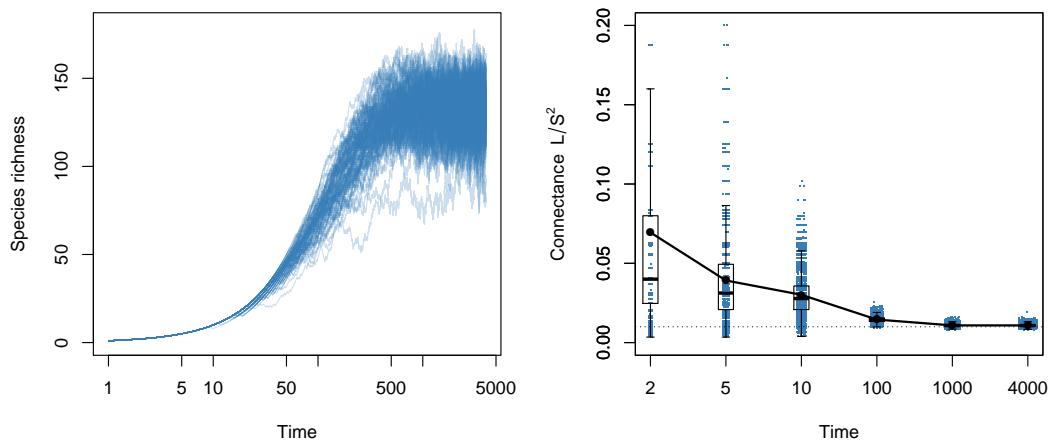


Figure S1. Left: Assembly of communities over time results in steady state species richness by ca. time-step 250. Right: Trophic connectance early in assembly is high because few species are tightly connected. Over time, connectance decays as species richness increases, and the density of trophic interactions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

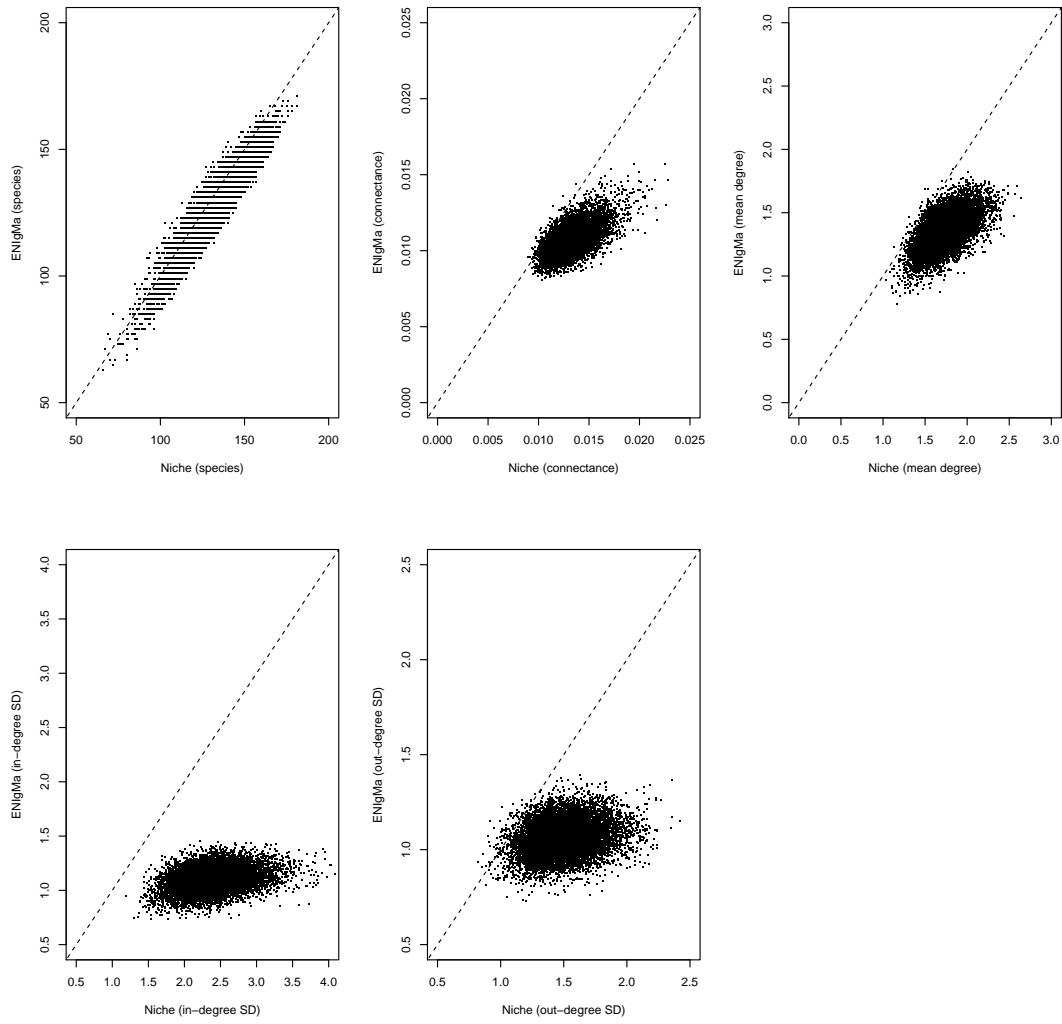


Figure S2. Comparisons of raw structural measures for the assembly (y-axis) and Niche model (x-axis). If the models produce similar structures, metrics will tend to fall on the 1:1 line (drawn). While the values for both models are similar, connectance, mean degree, and the standard deviation of in-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

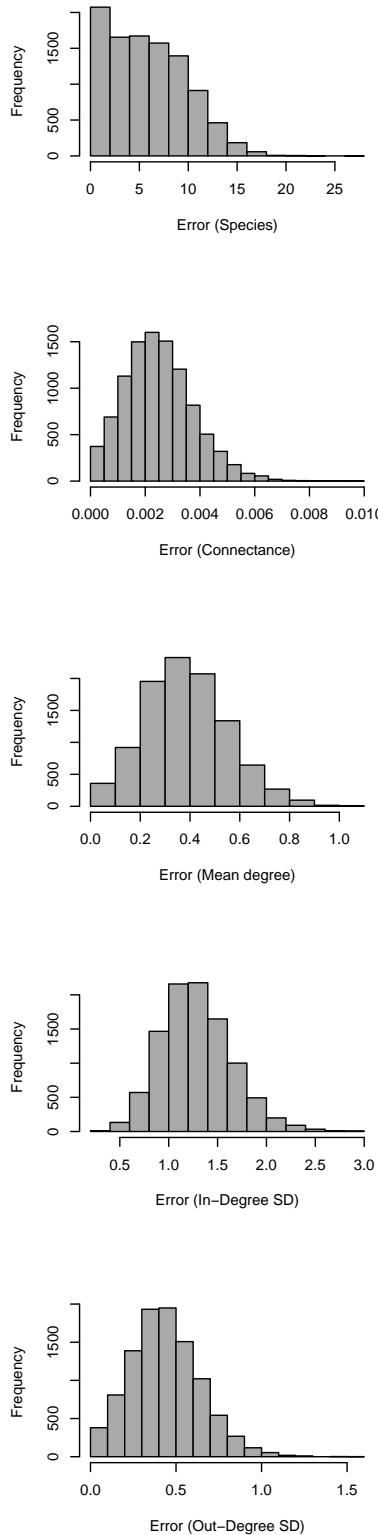


Figure S3. Error between structural measures of the assembly and Niche models. Error is measured as  $\sqrt{(m_i - m_j)^2}$ , where  $m_i$  and  $m_j$  are structural metrics for the assembly and Niche model, respectively. Only the trophic network of the assembly model used to assess metrics.

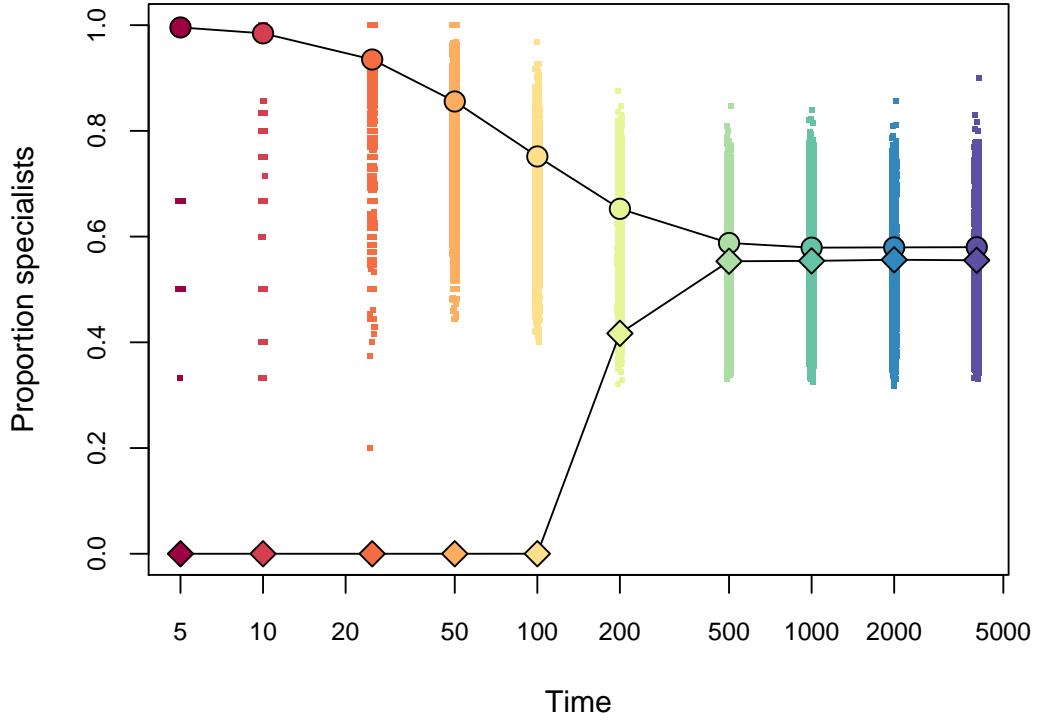


Figure S4. The proportion of specialists as a function of assembly time, where a specialist is defined as a species with a generality index  $G_i < 1$ . Measures of  $G_i$  are shown normalized to different measures of link-density. Circles:  $G_i^{\text{all}}$  where  $L$  accounts for all links in the food web and  $S$  accounts for all species relative to each time interval in the assembly process (averaged across replicates). Points:  $G_i^{\text{hetero}}$ , 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 webs<sup>34</sup>.

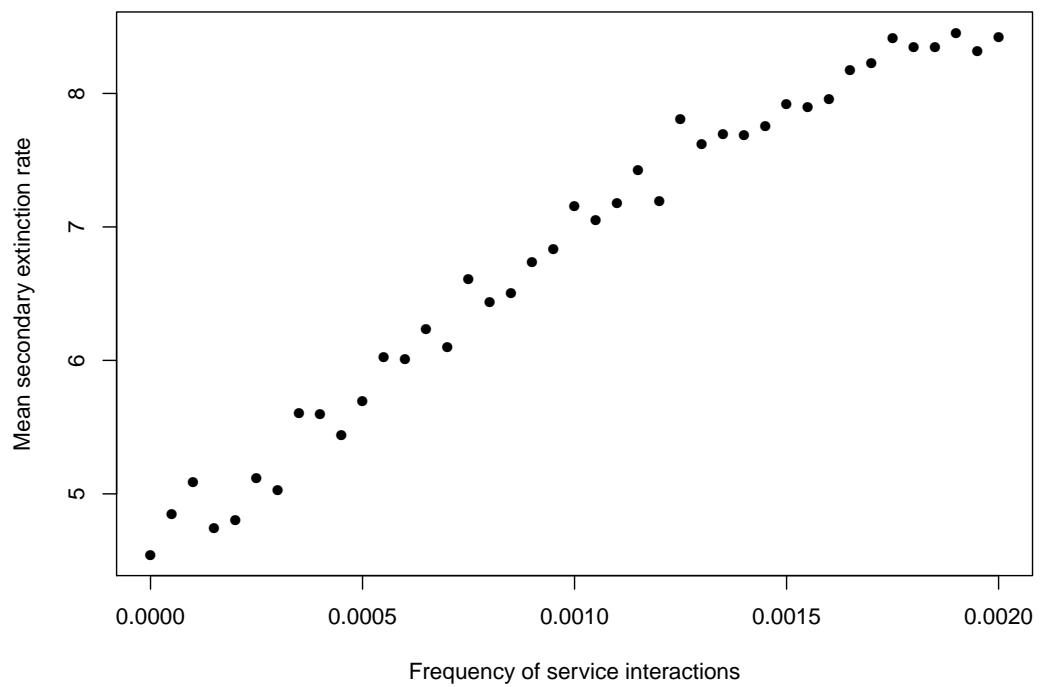


Figure S5. Mean rates of secondary extinction with increasing service interactions.

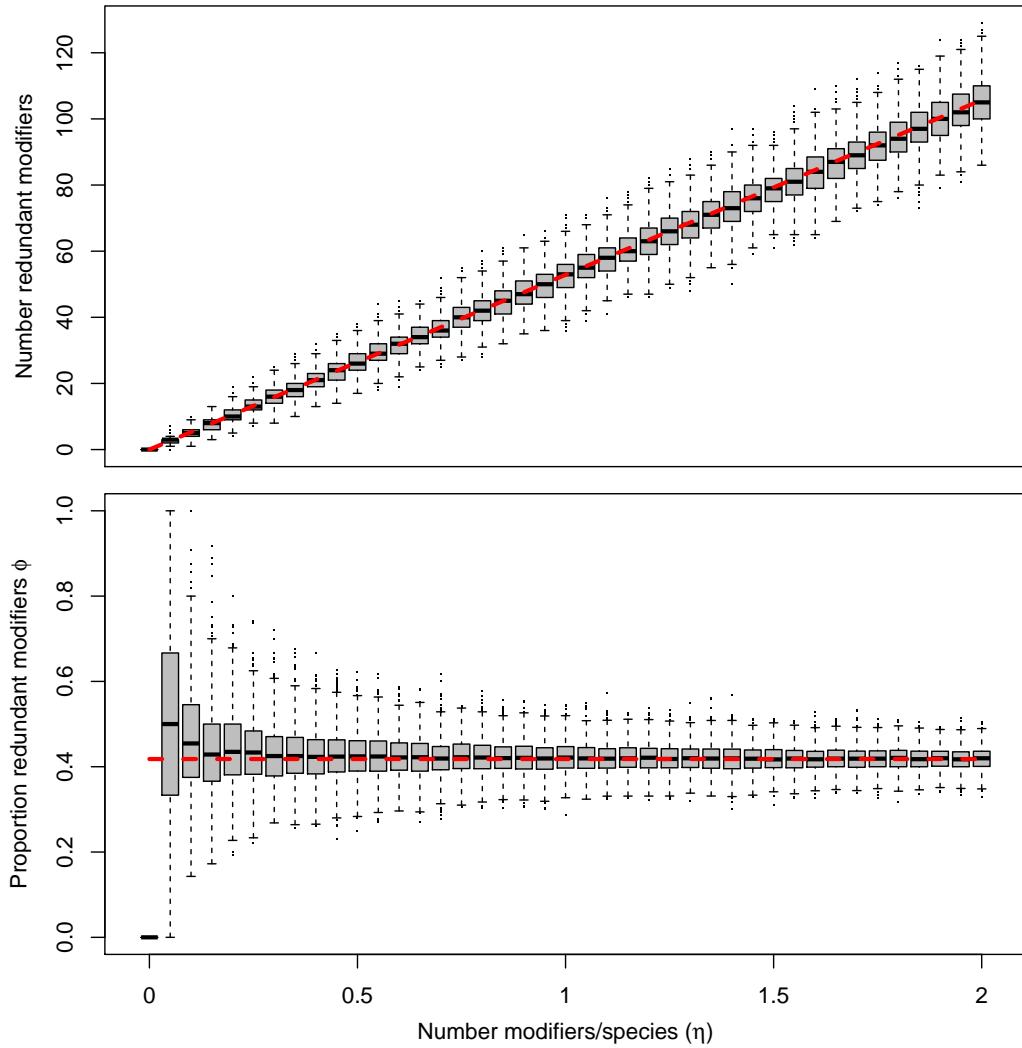


Figure S6. **a**, Number of redundant modifiers in the source pool as a function of the expected number of modifiers made per species  $\eta$ . The red dashed line shows the analytical expectation (Eq. S3). **b**, Proportion of redundant modifiers  $\phi$  versus the total number of modifiers in the source pool as a function of the expected number of modifiers made per species  $\eta$ . The red dashed line shows the analytical expectation of  $\phi \approx 0.418$  (Eq. S4).

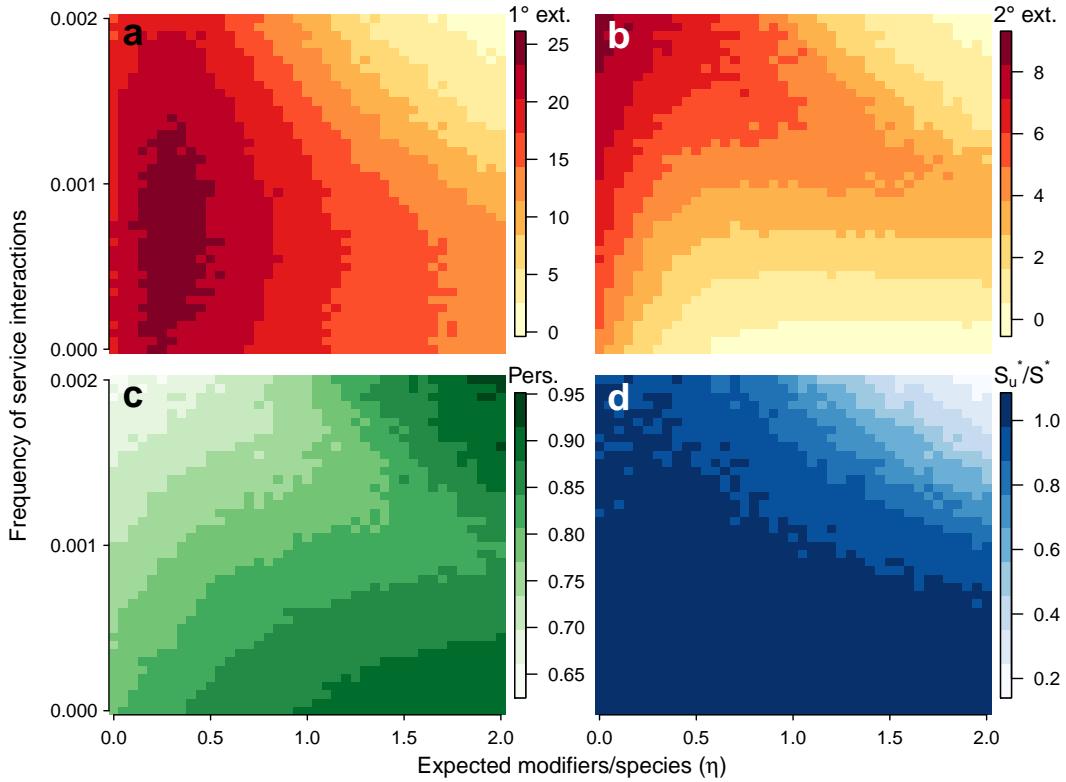


Figure S7. Measures of community stability as a function of the frequency of service interactions and number of modifiers per species, where each modifier is uniquely made by an engineer. **a**, Mean rates of primary extinction, where primary extinctions occur from competitive exclusion of consumers 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.

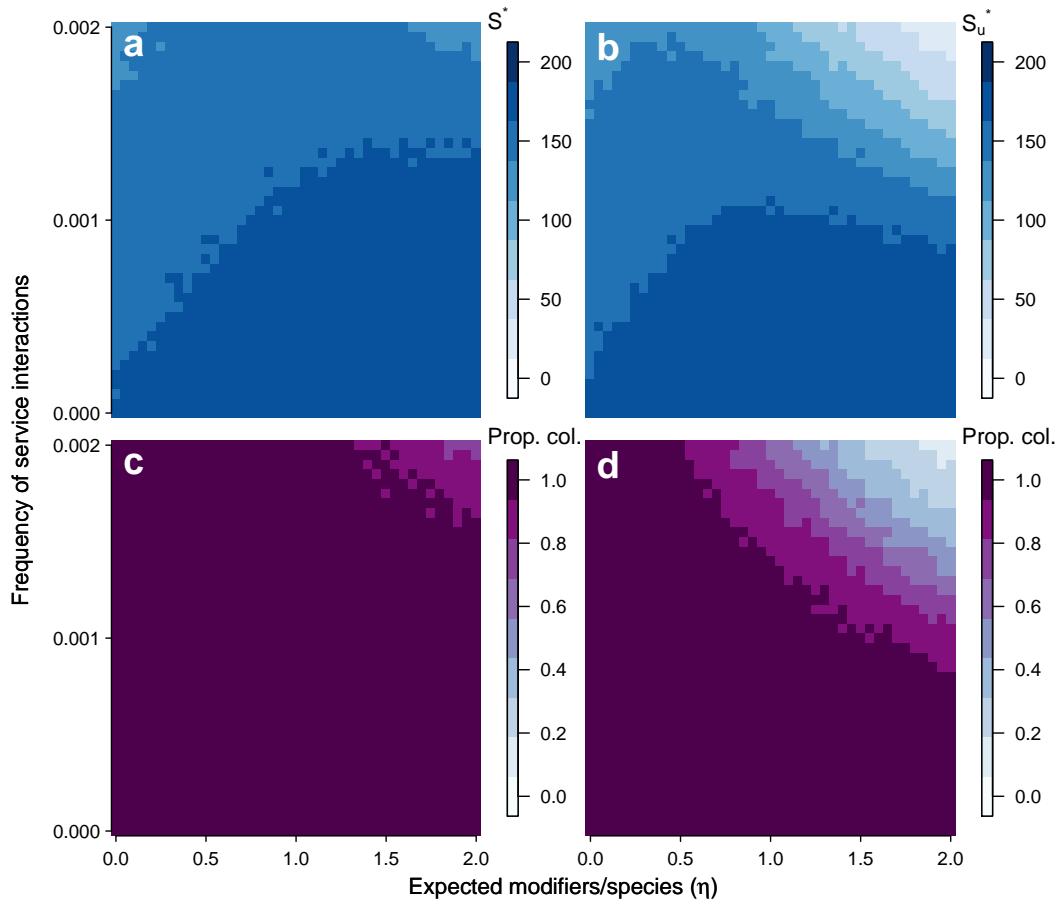


Figure S8. **a**, Steady state community richness with redundant engineering. **b**, Steady state community richness without redundant engineering. **c**, Proportion of species in the source pool that colonize the community at least once throughout the simulation (with redundant engineering). **d**, Proportion of species in the source pool that colonize the community at least once throughout the simulation (without redundant engineering).