

MIT Open Access Articles

Community structure follows simple assembly rules in microbial microcosms

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Friedman, Jonathan, et al. "Community Structure Follows Simple Assembly Rules in Microbial Microcosms." *Nature Ecology & Evolution* 1, 5 (March 2017): 0109 © 2017 Macmillan Publishers Limited, part of Springer Nature

As Published: <http://dx.doi.org/10.1038/s41559-017-0109>

Publisher: Springer Science and Business Media LLC

Persistent URL: <https://hdl.handle.net/1721.1/123084>

Version: Original manuscript: author's manuscript prior to formal peer review

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike



1

1 **Community structure follows simple assembly rules in microbial microcosms**

2

3 Jonathan Friedman¹, Logan M. Higgins^{1,2} & Jeff Gore¹

4 ¹Physics of Living Systems, Department of Physics, Massachusetts Institute of Technology,
5 Cambridge, MA 02139, USA

6 ²Microbiology Graduate Program, Massachusetts Institute of Technology, Cambridge,
7 Massachusetts, 02139, USA

8

9

10 **Introduction**

11 **Microbes typically form diverse communities of interacting species, whose activities have**
12 **tremendous impact on the plants, animals, and humans they associate with¹⁻³, as well as on the**
13 **biogeochemistry of the entire planet⁴. The ability to predict the structure of these complex**
14 **communities is crucial to understanding, managing, and utilizing them⁵. Here, we propose a**
15 **simple, qualitative assembly rule that predicts community structure from the outcomes of**
16 **competitions between small sets of species, and experimentally assess its predictive power**
17 **using synthetic microbial communities. The rule's accuracy was evaluated by competing**
18 **combinations of up to eight soil bacterial species, and comparing the experimentally observed**
19 **outcomes to the predicted ones. Nearly all competitions resulted in a unique, stable**
20 **community, whose composition was independent of the initial species fractions. Survival in**
21 **three-species competitions was predicted by the pairwise outcomes with an accuracy of ~90%.**
22 **Obtaining a similar level of accuracy in competitions between sets of seven or all eight species**
23 **required incorporating additional information regarding the outcomes of the three-species**
24 **competitions. Our results demonstrate experimentally the ability of a simple bottom-up**
25 **approach to predict community structure. Such an approach is key for anticipating the**
26 **response of communities to changing environments, designing interventions to steer existing**
27 **communities to more desirable states, and, ultimately, rationally designing communities *de***
28 **novo^{6,7}.**

29

30 **Main text**

31 Modeling and predicting microbial community structure is often pursued using bottom-up
32 approaches that assume that species interact in a pairwise manner⁸⁻¹¹. However, pair interactions
33 may be modulated by the presence of additional species^{12,13}, an effect that can significantly alter
34 community structure¹⁴ and may be common in microbial communities¹⁵. While it has been shown
35 that such models can provide a reasonable fit to sequencing data of intestinal microbiomes^{16,17}, their
36 predictive power remains uncertain, as it has rarely been directly tested experimentally (^{18,19} are
37 notable exceptions).

38 Current approaches to modeling microbial communities commonly employ a specific parametric
39 model, such as the generalized Lotka-Volterra (gLV) model²⁰⁻²². Generating predictions from such
40 models requires fitting a large number of parameter values from empirical data, which is often
41 challenging and prone to over-fitting. In addition, the exact form of the interactions needs to be
42 assumed, and a failure of the model can reflect a misspecification of the type of pairwise
43 interaction, rather than the presence of higher-order interactions²³.

44 Here we take an alternative approach in which qualitative information regarding the survival of
45 species in competitions between small sets of species (e.g., pairwise competitions) is used to predict
46 survival in more diverse multispecies competitions (**Fig. 1**). While this approach forgoes the ability
47 to predict exact species abundances, it does not require specifying and parameterizing the exact

48 form of interactions. Therefore, it is robust to model misspecification, and requires only survival
49 data, which can be more readily obtained than exact parameter values.

50 Intuitively, competitions typically result in the survival of a set of coexisting species, which cannot
51 be invaded by any of the species that went extinct during the competition. To identify sets of species
52 that are expected to coexist and exclude additional species, we first use the outcomes of pairwise
53 competitions. We propose the following assembly rule: in a multispecies competition, species that
54 all coexist with each other in pairs will survive, whereas species which are excluded by any of the
55 surviving species will go extinct. This rule provides a formalizes out intuition, and can be used to
56 systemically predict community structure from pairwise outcomes (Methods, **Fig. S1**).

57 To directly assess the predictive power of this approach, we used a set of eight heterotrophic soil-
58 dwelling bacterial species as a model system (**Fig. 2a, Methods**). Competition experiments were
59 performed by co-inoculating species at varying initial fractions, and propagating them through five
60 growth-dilution cycles (**Fig. S2**). During each cycle, cells were cultured for 48 hours and then
61 diluted by a factor of 1500 into fresh media, which corresponds to ~ 10.6 cellular divisions per
62 growth cycle, and ~ 53 cellular divisions over the entire competition period. The overall competition
63 time was chosen such that species extinctions would have sufficient time to occur, while new
64 mutants would typically not have time to arise and spread. Community compositions were assessed
65 by measuring the culture optical density (OD), as well as by plating on solid agar media and
66 counting colonies, which are distinct for each species²⁵. These two measurements quantify the
67 overall abundance of microbes in the community, and the relative abundances of individual species,
68 respectively. All experiments were done in duplicate.

69 Pairwise competitions resulted in stable coexistence or competitive exclusion of one of the species.
70 We performed competitions between all species pairs and found that in the majority of the pairs
71 ($19/28 = 68\%$, **Fig. 2b**) both species could invade each other, and thus stably coexisted. In the
72 remaining pairs ($9/28 = 32\%$) competitive exclusion occurred, where only one species could invade
73 the other (Time trajectories from one coexisting pair and one pair where exclusion occurs are shown
74 in **Fig. 2c**. Outcomes for all pairs are shown in **Fig. 2d**). Species' growth rate in monoculture was
75 correlated with their average competitive ability, but, in line with previous reports²⁶, it could not
76 predict well the outcome of specific pair competitions (**Fig. S3**).

77 Next, we measured the outcome of competition between all 56 three-species combinations. These
78 competitions typically resulted in a stable community whose composition was independent of the
79 starting fractions (**Table S1**). However, 2 of the 56 trios displayed inconsistent results with high
80 variability between replicates. This variability likely resulted from rapid evolutionary changes that
81 occurred during the competition (**Fig. S4**). All but one of the other trio competitions resulted in
82 stable communities with a single outcome, independent of starting conditions. This raises the
83 question of whether this unique outcome could be predicted based upon the experimentally
84 observed outcomes of the pairwise competitions.

85 Trios were grouped by the topology of their pairwise outcome network, which was used to predict
86 their competitive outcomes. The most common topology involved two coexisting pairs, and a pair
87 where competitive exclusion occurs ($30/56 = 54\%$). To illustrate this scenario, consider a set of
88 three species, labeled A, B, and C, where species A and C coexist with B in pairwise competitions,
89 whereas C is excluded when competing with A. In this case, our proposed assembly rule predicts
90 that the trio competition will result in the survival of species A and B, and exclusion of C (**Fig. 3a**).
91 This predicted outcome occurred for a majority of the experimentally observed trios (**Fig. 3b**), but
92 some trio competitions resulted in less intuitive outcomes (**Fig. 3c**). For example, 1 of the 30 trios
93 with this topology led to the extinction of A and the coexistence of B and C (**Fig. 3c**). The
94 experimentally observed outcomes of competition in this trio topology highlights that our simple
95 assembly rule typically works, and the failures provide a sense of alternative outcomes that are
96 possible given the same underlying topology of pairwise outcomes. Unpredicted outcomes may
97 occur due to several mechanisms, which are discussed at the last part of the manuscript.

98 Another frequent topology was coexistence between all three species pairs ($15/56 = 27\%$), in which
99 case none of the species is predicted to be excluded in the trio competition (**Fig. 3d**). Such trio
100 competitions resulted either in coexistence of all three species, as predicted by our assembly rule
101 (**Fig. 3e**), or in the exclusion of one of the species (**Fig. 3f**). Overall, 5 different trio layouts, and 11
102 competitive outcomes have been observed (**Fig. 3g-k**). Notably, all observed trio outcomes across
103 all topologies can be generated from simple pairwise interactions, including the outcomes which
104 were not correctly predicted by our assembly rule²⁴. An incorrect prediction of our simple assembly
105 rule is therefore not necessarily caused by higher-order interactions.

106 Overall, survival in three-species competitions was well predicted by pairwise outcomes. The
107 assembly rule predicted species survival across all the three-way competitions with an 89.5%
108 accuracy (**Fig. 4a**), where accuracy is defined as the fraction of species whose survival was
109 correctly predicted. To get a sense of how the observed accuracy compares to the accuracy
110 attainable when pairwise outcomes are not known, as a null model, we considered the case where
111 the only information available is the average probability that a species will survive in a trio
112 competition (note that this probability is not assumed to be available in our simple assembly rule).
113 Using this information, trio outcomes could only be predicted with a 72% accuracy (**Fig. 4a**,
114 **Methods**). We further compared the observed accuracy to the accuracy expected when species
115 interact solely in a pairwise manner, according to the gLV equations with a random interaction
116 matrix (**Methods**). We found that the observed accuracy is consistent with the accuracy obtained in
117 simulations of competitions that parallel our experimental setup ($p=0.29$, **Fig. 4b**). Survival of
118 species in pairwise competition is therefore surprisingly effective in predicting survival when
119 species undergo trio competition.

120 Nonetheless, there are exceptional cases where qualitative pairwise outcomes are not sufficient to
121 predict competitive outcomes of trio competitions. Accounting for such unexpected trio outcomes
122 may improve prediction accuracy for competitions involving a larger set of species. We encode
123 unexpected trio outcomes by creating effective modified pairwise outcomes, which replace the
124 original outcomes in the presence of an additional species. For example, competitive exclusion will
125 be modified to an effective coexistence when two species coexist in the presence of a third species
126 despite one of them being excluded from the pair competition. The effective, modified outcomes
127 can be used to make predictions using the assembly rule as before (**Methods, Fig. S1**). By
128 accounting for unexpected trio outcomes, the assembly rule extends our intuition, and predicts
129 community structure in the presence of potentially complex interactions.

130 The ability of the assembly rule to predict the outcomes of more diverse competitions was assessed
131 by measuring survival in competitions between all seven-species combinations, as well as the full
132 set of eight species (**Fig. 5a**). Using only the pairwise outcomes, survival in these competitions
133 could only be predicted with an accuracy of 62.5%, which is barely higher than the 61% accuracy
134 obtained when using only the average probability that a species will survive these competitions
135 (**Fig. 5b**). A considerably improved prediction accuracy of 86% was achieved by incorporating
136 information regarding the trio outcomes (**Fig. 5b**). As in the trio competitions, the observed
137 accuracies are consistent with those obtained in gLV simulations that parallel the experimental
138 setup, both when predicting using pairwise outcomes alone ($p=0.53$), or in combination with trio
139 outcomes ($p=0.21$, **Fig. 5c**).

140 Our assembly rule makes predictions that match our intuition, but there are several conditions under
141 which these predictions may be inaccurate. First, community structure can be influenced by initial
142 species abundances²⁷, as has recently been demonstrated in pairwise competitions between bacteria
143 of the genus *Streptomyces*²⁸. Our assembly rule may be able to correctly predict the existence of
144 multiple stable states, as it identifies all putative sets of coexisting, non-invasible species in a given
145 species combination. However, we did not have sufficient data to evaluate the rule's accuracy in
146 such cases, as multistability was observed in only one of all our competition experiments.

147 Complex ecological dynamics, such as oscillations and chaos, can also have a significant impact on
148 species survival^{29,30}, making it difficult to predict the community structure. These dynamics can
149 occur even in simple communities containing only a few interacting species. For example,
150 oscillatory dynamics occur in gLV models of competition between as few as three species²⁴, and
151 have been experimentally observed in a cross-protection mutualism between a pair of bacterial
152 strains³¹. In contrast, our competitions predominantly resulted in a unique and stable final
153 community. This occurred despite the fact that we observed complex inter-species interactions
154 involving interference competition and facilitation (**Fig S4**). These results indicate that complex
155 ecological dynamics may in fact be rare, though it remains to be seen whether they become more
156 prevalent in more diverse assemblages. Relatedly, prediction is challenging in the presence of
157 competitive cycles (e.g. “Rock-Paper-Scissors” interactions), which often lead to oscillatory
158 dynamics, and are thought to increase species survival and community diversity^{32,33}. Such non-
159 hierarchical relationships are absent from our competitive network, and thus their effect cannot be
160 evaluated here.

161 In the absence of multistability or complex dynamics, our approach may still fail when competitive
162 outcomes do not provide sufficient information regarding the interspecies interactions. This could
163 be due to higher-order interactions, which only manifest in the presence of additional species, or
164 because only qualitative information regarding survival is utilized. The observed accuracy of the
165 assembly rule was consistent with the one found in gLV simulations, but this does not necessarily
166 indicate that our species interact in a linear, pairwise fashion. In fact, fitting the gLV model directly
167 to our pairwise data does not improve predictability (**Fig. S6**). Determining whether, in any
168 particular competition, predictions fail due to insufficient information regarding the strength of
169 linear interactions, non-linear interactions, or higher-order interactions will require more detailed
170 measurements.

171

172 Controlling and designing microbial communities has numerous important application areas ranging
173 from probiotic therapeutics, to bioremediation and biomanufacturing⁵. The ability to predict what
174 community will be formed by a given set of species is crucial for determining how extinctions and
175 invasions will affect existing communities, and for engineering desired communities. Our results
176 suggest that, when measured in the same environment, community structure can be predicted from
177 the outcomes of competitions between small sets of species, demonstrating the feasibility of a
178 bottom-up approach to understanding and predicting community structure. While these results are
179 encouraging, they were obtained using a small set of closely related species in well-controlled
180 laboratory settings. It remains to be seen to what extent these results hold in other systems and in
181 more natural settings, involving more diverse assemblages which contain additional trophic levels,
182 in the presence of spatial structure, and over evolutionary time scales.

183

184

185 **Methods**

186 **Species and media.** The eight soil bacterial species used in this study are *Enterobacter aerogenes*
187 (Ea, ATCC#13048), *Pseudomonas aurantiaca* (Pa, ATCC#33663), *Pseudomonas chlororaphis* (Pch,
188 ATCC#9446), *Pseudomonas citronellolis* (Pci, ATCC#13674), *Pseudomonas fluorescens*
189 (ATCC#13525), *Pseudomonas putida* (ATCC#12633), *Pseudomonas veronii* (ATCC#700474), and
190 *Serratia marcescens* (Sm, ATCC#13880). All species were obtained from ATCC. The base growth
191 media was M9 minimal media²⁵, which contained 1X M9 salts (Sigma Aldrich, M6030), 2mM
192 MgSO₄, 0.1mM CaCl₂, 1X trace metals (Teknova, T1001). For the final growth media, the base
193 media was supplemented with 1.6mM galacturonic acid and 3.3mM serine as carbon sources, which
194 correspond to 10mM of carbon for each of these substrates. These carbon sources were chosen from
195 a set of carbon sources commonly used to characterize soil microbes (Biolog, EcoPlate) to ensure

196 that each of the eight species survives in monoculture. Nutrient broth (0.3% yeast extract, 0.5%
197 peptone) was used for initial inoculation and growth prior to experiment. Plating was done on 10cm
198 Petri dishes containing 25mL nutrient agar (nutrient broth with 1.5% agar added).

199 **Competition experiments.** Frozen stocks of individual species were streaked out on nutrient agar
200 Petri plates, grown at room temperature for 48hr, and then stored at 4°C for up to 2 weeks. Prior to
201 competition experiments, single colonies were picked and each species was grown separately in
202 50mL Falcon tubes, first in 5ml nutrient broth for 24hr and next in 5ml of the experimental M9
203 media for 48hr. During the competition experiments, cultures were grown in Falcon flat-bottom 96-
204 well plates (BD Biosciences), with each well containing a 150µl culture. Plates were incubated at
205 25°C without shaking, and were covered with a lid and wrapped in Parafilm. For each growth-
206 dilution cycle, the cultures were incubated for 48hr and then serially diluted into fresh growth media
207 by a factor of 1500.

208 Initial species mixtures were performed by diluting each species separately to an optical density
209 (OD) of 3×10^{-4} . Different species were then mixed by volume to the desired composition. This
210 mixture was further diluted to an OD of 10^{-4} , from which all competitions were initialized. For each
211 set of competing species, competitions were conducted from all the initial conditions in which each
212 species was present at 5%, except for one more abundant species. For example, for each species
213 pair there were 2 initial conditions with one species at 95% and the other at 5%, whereas for the 8
214 species competition there were 8 initial conditions each with a different species at 65% and the rest
215 at 5%. For a few species pairs (Fig. 2a-b), we conducted additional competitions starting at more
216 initial conditions. All experiments were done in duplicate.

217 **Measurement of cell density and species fractions.** Cell densities were assessed by measuring
218 optical density at 600nm using a Varioskan Flash plate reader. Relative abundances were measured
219 by plating on nutrient agar plates. Each culture was diluted by a factor between 10^5 and 10^6 in
220 phosphate-buffered saline, depending on the culture's OD. For each diluted culture, 75µl were
221 plated onto an agar plate. Colonies were counted after 48h incubation in room temperature. A
222 median number of 85 colonies per plate were counted. To determine species extinction in
223 competition between a given set of species, we combined all replicates and initial conditions from
224 that competition, and classified as extinct any species whose median abundance was less than 1%,
225 which is just above our limit of detection.

226 **Assembly rule predictions and accuracy.** For any group of competing species, predictions were
227 made by considering all possible competitive outcomes (e.g. survival of any single species, any
228 species pair, etc.). Outcomes that were consistent with our assembly rule were those that were
229 predicted to be a possible outcome of the competition (**Fig. S1**). For any given competition, there
230 may be several such feasible outcomes, however a unique outcome was predicted for all our
231 competition experiments.

232 Pairwise outcomes were modified using trio outcomes as following: Exclusion was replaced with
233 coexistence for pairs that coexisted in the presence of any additional species. Coexistence was
234 replaced with exclusion whenever a species went extinct in a trio competition with two species with
235 which it coexisted when competed in isolation. Only modifications cause by the surviving species,
236 or an invading species were considered. Therefore, a new set of modified pairwise outcomes was
237 generated for each putative set of surviving species being evaluated.

238 The prediction accuracy was defined as the fraction of species whose survival was correctly
239 predicted. When the assembly rule identified multiple possible outcomes, which occurred only in
240 the gLV simulations, accuracy was averaged over all such feasible outcomes. Additionally, when the
241 competitive outcome depended on the initial condition, accuracy was averaged across all initial
242 conditions.

243 For reference, we computed the accuracy of predictions made based on the probability that a species
244 will survive a competition between the same number of species. For example, for predicting trio

6

245 outcomes, we used the proportion of species that survived, averaged across all trio competitions.
246 Using this information, the highest accuracy would be achieved by predicting that all species
247 survive in all competitions, if the average survival probability is > 0.5 , and predicting that all
248 species go extinct otherwise.

249 **Simulated competitions.** To assess the assembly rule's expected accuracy in a simple case in which
250 species interact in a purely pairwise manner, we simulated competitions using the generalized
251 Lotka-Volterra (gLV) dynamics:

$$\dot{x}_i = r_i x_i \left(1 - x_i + \sum_{i \neq j} \alpha_{ij} x_j \right),$$

252

253 where x_i is the density of species i (normalized to its carrying capacity), r_i is the species' intrinsic
254 growth rate, and α_{ij} is the interaction strength between species i and j . For each simulation, we
255 created a set of species with random interactions where the α_{ij} parameters were independently
256 drawn from normal distribution with a mean of 0.6 and a standard deviation of 0.46. Results were
257 insensitive to variations in growth rates, thus they were all set to 1 for simplicity. These parameters
258 recapitulate the proportions of coexistence and competitive exclusion observed in our experiments,
259 and yield a distribution of trio layouts similar to the one (**Fig. S7**). The probability of generating
260 bistable pairs in these simulations is low ($\sim 3.7\%$, corresponding to one bistable pair in a set of eight
261 species), and we further excluded the bistable pairs that were occasionally generated by chance,
262 since we had not observed any such pairs in the experiments.

263 The accuracy of the assembly rules in gLV systems was estimated by running simulations that
264 parallel our experimental setup: A set of 8 species with random interaction coefficients was
265 generated, and the pairwise outcomes were determined according to their interaction strengths.
266 These outcomes were used to generate predictions for the trio competitions using our assembly rule.
267 Next, all 3-species competitions were simulated with the same set of initial conditions used in the
268 experiments. Finally, the predicted trio outcomes were compared to the simulation outcomes across
269 all trios to determine the prediction accuracy. Thus, a single accuracy value was recorded for each
270 set of 8 simulated species. Similarly, for each simulated 8-species set, the pair and trio outcomes
271 were used to generate predictions for the 7-species and 8-species competitions, and their accuracy
272 was assessed by comparing them to the outcomes of simulated competitions. Prediction accuracy
273 distributions were estimated using Gaussian kernel density estimation from the accuracy values of
274 100 simulated sets of 8 species.

275 One-sided P-values evaluating the consistency of the experimentally observed accuracies with the
276 simulation results were defined as the probability that a simulation would yield an accuracy which
277 is at least as high as the experimentally observed one.

278 **Code availability.** An implementation of the assembly rule and the gLV simulations, as well as
279 routines for evaluating the rule's accuracy are freely available online at:
280 <https://bitbucket.org/yonatanf/assembly-rule>.

281

1. Berendsen, R. L., Pieterse, C. M. & Bakker, P. A. The rhizosphere microbiome and plant health.
Trends Plant Sci. **17**, 478–486 (2012).

2. Ezenwa, V. O., Gerardo, N. M., Inouye, D. W., Medina, M. & Xavier, J. B. Animal Behavior and the Microbiome. *Science* **338**, 198–199 (2012).
3. Flint, H. J., Scott, K. P., Louis, P. & Duncan, S. H. The role of the gut microbiota in nutrition and health. *Nat. Rev. Gastroenterol. Hepatol.* **9**, 577–589 (2012).
4. Falkowski, P., Fenchel, T. & Delong, E. The microbial engines that drive Earth's biogeochemical cycles. *Science* **320**, 1034–9 (2008).
5. Widder, S. *et al.* Challenges in microbial ecology: building predictive understanding of community function and dynamics. *ISME J.* (2016). doi:10.1038/ismej.2016.45
6. Großkopf, T. & Soyer, O. S. Synthetic microbial communities. *Curr. Opin. Microbiol.* **18**, 72–77 (2014).
7. Fredrickson, J. K. Ecological communities by design. *Science* **348**, 1425–1427 (2015).
8. Faust, K. & Raes, J. Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* **10**, 538–550 (2012).
9. Bucci, V. & Xavier, J. B. Towards Predictive Models of the Human Gut Microbiome. *J. Mol. Biol.* **426**, 3907–3916 (2014).
10. Berry, D. & Widder, S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front Microbiol* **5**, 10–3389 (2014).
11. Carrara, F., Giometto, A., Seymour, M., Rinaldo, A. & Altermatt, F. Inferring species interactions in ecological communities: a comparison of methods at different levels of complexity. *Methods Ecol. Evol.* n/a-n/a (2015). doi:10.1111/2041-210X.12363
12. Billick, I. & Case, T. J. Higher Order Interactions in Ecological Communities: What Are They and How Can They be Detected? *Ecology* **75**, 1530–1543 (1994).
13. Momeni, B. & Shou, W. The validity of pairwise models in predicting community dynamics. *bioRxiv* 60988 (2016). doi:10.1101/060988
14. Wootton, J. T. The nature and consequences of indirect effects in ecological communities. *Annu. Rev. Ecol. Syst.* 443–466 (1994).
15. Kelsic, E. D., Zhao, J., Vetsigian, K. & Kishony, R. Counteraction of antibiotic production and degradation stabilizes microbial communities. *Nature* **521**, 516–519 (2015).
16. Stein, R. R. *et al.* Ecological Modeling from Time-Series Inference: Insight into Dynamics and Stability of Intestinal Microbiota. *PLOS Comput Biol* **9**, e1003388 (2013).

17. Bucci, V. *et al.* MDSINE: Microbial Dynamical Systems INference Engine for microbiome time-series analyses. *Genome Biol.* **17**, 121 (2016).
18. Vandermeer, J. H. The competitive structure of communities: an experimental approach with protozoa. *Ecology* 362–371 (1969).
19. Dormann, C. F. & Roxburgh, S. H. Experimental evidence rejects pairwise modelling approach to coexistence in plant communities. *Proc. R. Soc. Lond. B Biol. Sci.* **272**, 1279–1285 (2005).
20. Mounier, J. *et al.* Microbial Interactions within a Cheese Microbial Community. *Appl. Environ. Microbiol.* **74**, 172–181 (2008).
21. Fisher, C. K. & Mehta, P. Identifying Keystone Species in the Human Gut Microbiome from Metagenomic Timeseries Using Sparse Linear Regression. *PLOS ONE* **9**, e102451 (2014).
22. Marino, S., Baxter, N. T., Huffnagle, G. B., Petrosino, J. F. & Schloss, P. D. Mathematical modeling of primary succession of murine intestinal microbiota. *Proc. Natl. Acad. Sci.* **111**, 439–444 (2014).
23. Case, T. J. & Bender, E. A. Testing for Higher Order Interactions. *Am. Nat.* **118**, 920–929 (1981).
24. Zeeman, M. L. Hopf bifurcations in competitive three-dimensional Lotka–Volterra systems. *Dyn. Stab. Syst.* **8**, 189–216 (1993).
25. Celiker, H. & Gore, J. Clustering in community structure across replicate ecosystems following a long-term bacterial evolution experiment. *Nat. Commun.* **5**, (2014).
26. Concepción-Acevedo, J., Weiss, H. N., Chaudhry, W. N. & Levin, B. R. Malthusian Parameters as Estimators of the Fitness of Microbes: A Cautionary Tale about the Low Side of High Throughput. *PLOS ONE* **10**, e0126915 (2015).
27. Fukami, T. Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. *Annu. Rev. Ecol. Evol. Syst.* **46**, 1–23 (2015).
28. Wright, E. S. & Vetsigian, K. H. Inhibitory interactions promote frequent bistability among competing bacteria. *Nat. Commun.* **7**, 11274 (2016).
29. Armstrong, R. A. & McGehee, R. Competitive Exclusion. *Am. Nat.* **115**, 151–170 (1980).
30. Huisman, J. & Weissing, F. Biodiversity of plankton by species oscillations and chaos. *Nature* **402**, 407 (1999).

31. Yurtsev, E. A., Conwill, A. & Gore, J. Oscillatory dynamics in a bacterial cross-protection mutualism. *Proc. Natl. Acad. Sci.* **113**, 6236–6241 (2016).
32. Kerr, B., Riley, M. A., Feldman, M. W. & Bohannan, B. J. Local dispersal promotes biodiversity in a real-life game of rock–paper–scissors. *Nature* **418**, 171–174 (2002).
33. Allesina, S. & Levine, J. A competitive network theory of species diversity. *Proc. Natl. Acad. Sci.* **108**, 5638 (2011).

282

283

284

285

286

287

288 **Acknowledgements** We would like to thank A. Perez-Escudero, N. Vega, E. Yurtsev and members
289 of the Gore laboratory for for critical discussions and comments on the manuscript. This work was
290 supported by the DARPA BRICS program, an NIH New Innovator Award (NIH DP2), an NSF
291 CAREER Award, a Sloan Research Fellowship, the Pew Scholars Program and the Allen
292 Investigator Program.

293

294 **Author Contributions** J.F. and J.G. designed the study. J.F. and L.H. performed the experiments
295 and analysis. J.F., L.H. and J.G. wrote the manuscript.

296

297 **Author Information** The authors declare no competing financial interests. Correspondence and
298 requests for materials should be addressed to J.F. (yonatanf@mit.edu) or J.G. (gore@mit.edu).

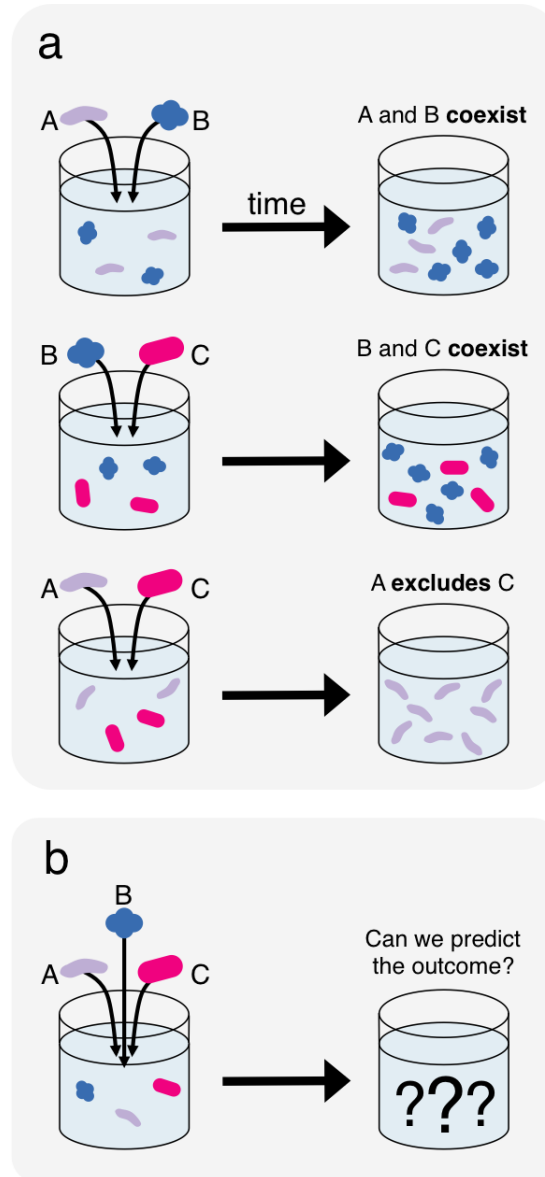
10

299 **Figure 1. A bottom-up approach to predicting community composition from qualitative**
300 **competitive outcomes.** Qualitative information regarding the survival of species in competitions
301 between small sets of species, such as pairwise competitions (a) is used to predict survival in more
302 diverse multispecies competitions, such as trio competitions (b). The particular pairwise outcomes
303 illustrated here reflect the true outcomes observed experimentally in one set of three species (**Fig.**
304 **3b**).

305

306

307

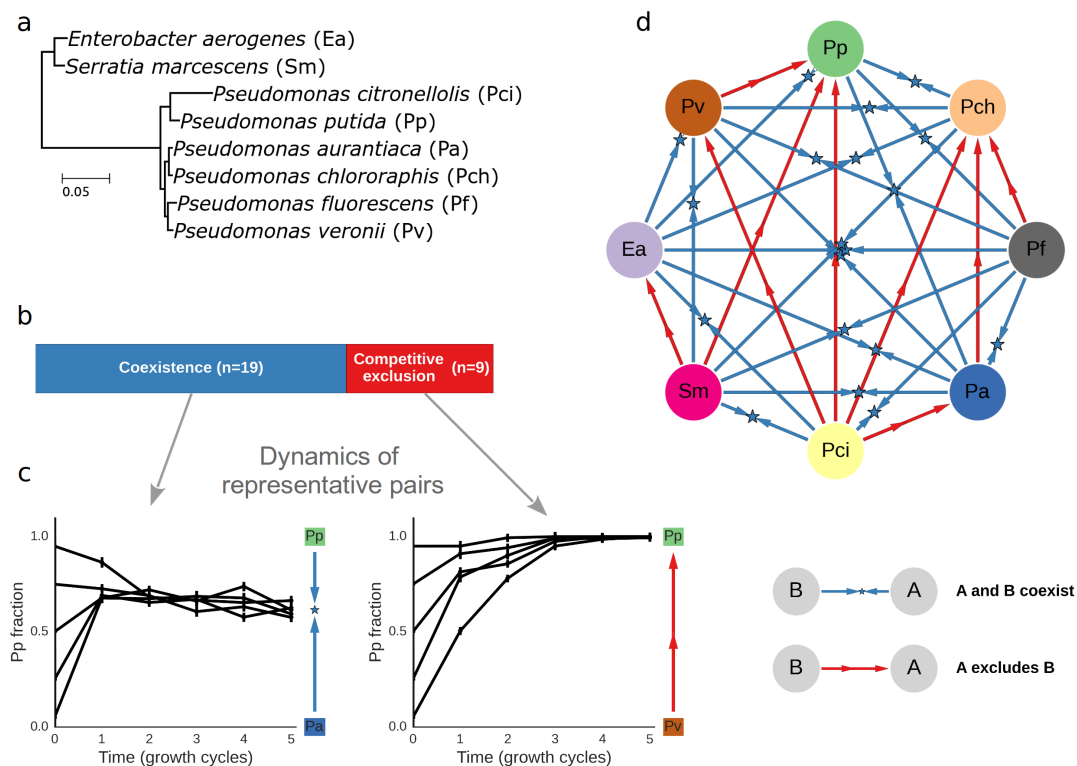


11

308 **Figure 2. Pairwise competitions resulted in stable coexistence or competitive exclusion.** (a)
309 Phylogenetic tree of the set of eight species used in this study. The tree is based on the full 16S
310 gene, and the branch lengths indicate the number of substitutions per base pair. (b) Coexistence was
311 observed for 19 of the 28 pairs, whereas competitive exclusion was observed for 9 of the 28 pairs.
312 (c) Changes in relative abundance over time in one pair where competitive exclusion occurred, and
313 one coexisting pair. The y-axis indicates the fraction of one of the competing species. In the
314 exclusion example, the species fraction increased for all initial conditions, resulting in the exclusion
315 of the competitor. In contrast, in the coexistence case, fractions converged to an intermediate value
316 and both species were found at the end of the competition. Blue and red arrows to the right indicate
317 the qualitative competitive outcome, with the star marking the final fraction in the case of
318 coexistence. Error bars represent the standard deviation of the posterior Beta distribution of the
319 fractions, based on colony counts averaged across replicates. (d) Network diagram of the outcomes
320 of all pairwise competitions.

321

322

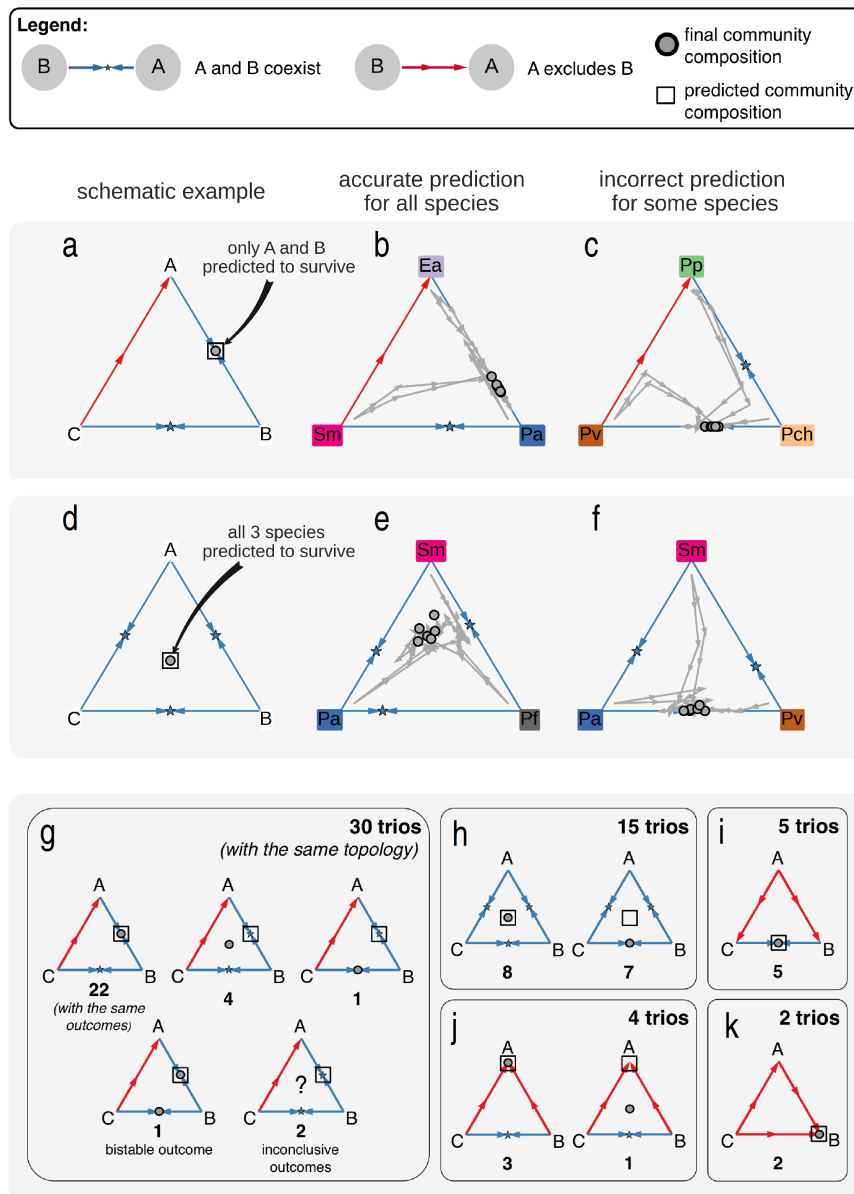


12

323 **Figure 3. Trio competitions typically resulted in a unique outcome.** Changes in species fraction
 324 were measured over time for several trio competitions. (a-c) Trios involving 2 coexisting pairs and 1
 325 pair where competitive exclusion occurs. In these plots, each triangle is a simplex denoting the
 326 fractions of the three competing species. The simplex vertices correspond to a community
 327 composed solely of a single species, whereas edges correspond to a 2-species mixture. The edges
 328 thus denote the outcomes of pair competitions, which were performed separately. Trajectories begin
 329 at different initial compositions, and connect the species fractions measured at the end of each
 330 growth cycle. Dots mark the final community compositions. (a) schematic example, showing that
 331 only species A and B are predicted to coexist for this pattern of pairwise outcomes. (b) example of a
 332 trio competition which resulted in the predicted outcome. (c) An example of an unpredicted
 333 outcome. (d-f) Similar to a-c, but for trios where all species coexist in pairs. (g-k) All trio layouts
 334 and outcomes, grouped by the topology of the pairwise outcomes network. Dots denote the final
 335 community composition (not exact species fractions, but rather species survivals). One trio
 336 displayed bistability, which is indicated by two dots representing the two possible outcomes. Two
 337 trios displayed inconsistent results with high variability between replicates, which is indicated by a
 338 question mark.

339

340



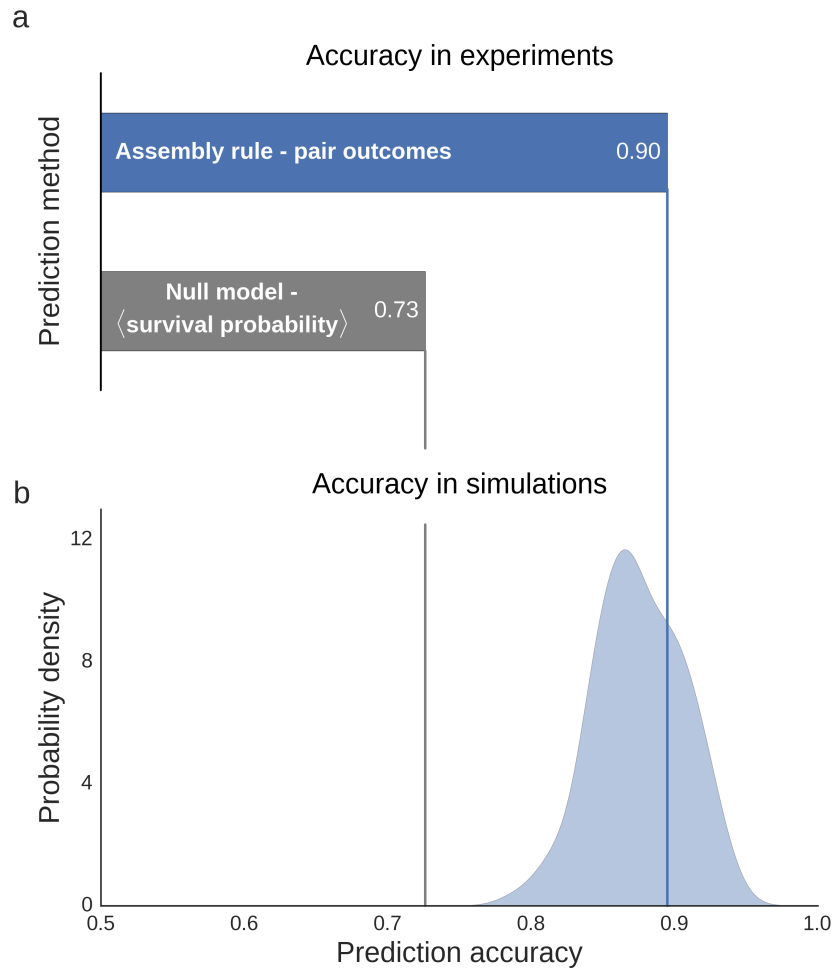
13

341 **Figure 4. Survival in trio competitions is well predicted by pairwise outcomes.** (a) Prediction
342 accuracy of the assembly rule and the null model, where predictions are made solely based on the
343 average probability that species survive in trio competitions. (b) The distribution of accuracies of
344 prediction made using the assembly rule from gLV simulations which mirror our experimental
345 design. The experimentally observed accuracy is consistent with those found in the simulations.

346

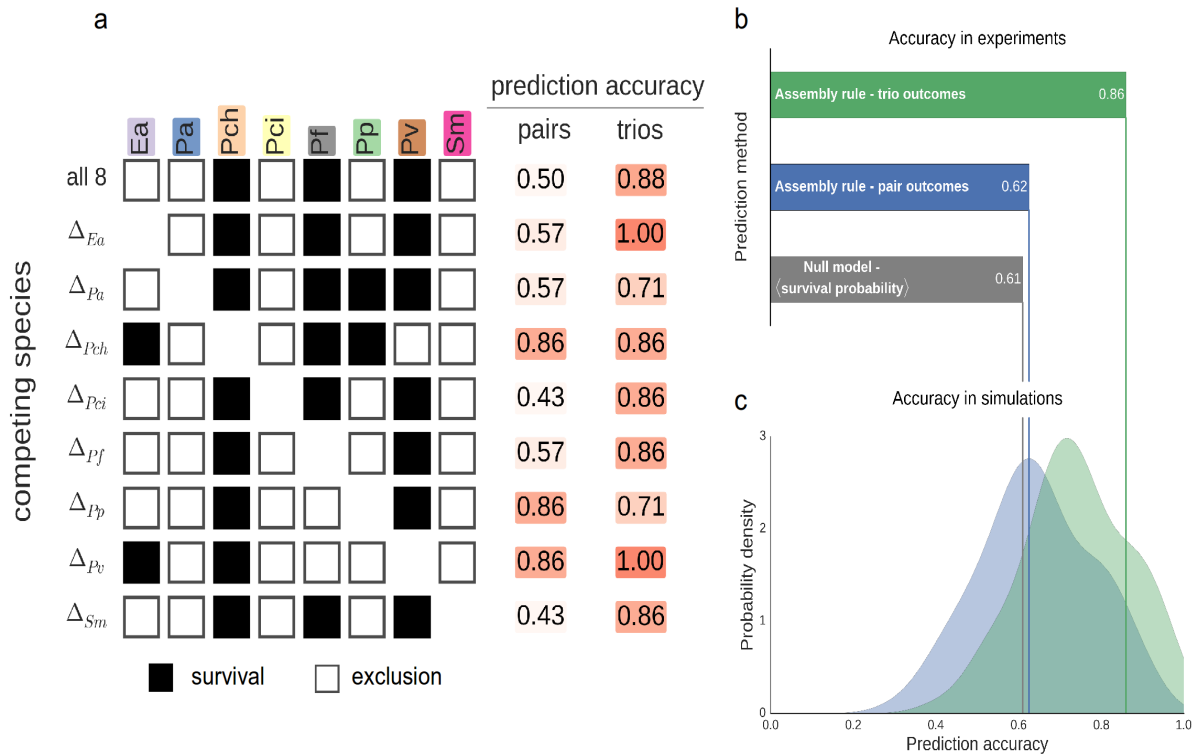
347

348



14

349 **Figure 5. Predicting survival in more diverse competition required incorporating the**
 350 **outcomes of the trio competitions.** (a) Species survival when competing all 8 species, and all sets
 351 of 7 species. Bright and pale colors indicate survival and extinction, respectively. Survival is
 352 predicted either using only pair outcomes, or using both pair and trio outcomes. (b) Prediction
 353 accuracy of the null model and the assembly rule, using either pair outcomes only, or pair and trio
 354 outcomes (c) The distribution of accuracies of prediction made using the assembly rule from gLV
 355 simulations which mirror our experimental design. In these simulations, predictions were made
 356 using either pair outcomes only, or pair and trio outcomes. In both cases, the experimentally
 357 observed accuracies are consistent with those found in the simulations.
 358



360

361

362