

1 **Title:**

2 “Neutral and niche dynamics in a synthetic microbial community”

3 **Authors:**

4 NJ Cira<sup>1</sup>, MT Pearce<sup>2</sup>, SR Quake<sup>1\*</sup>

5 **Affiliations:**

6 <sup>1</sup> Department of Bioengineering, Stanford University, Stanford, CA 94305.

7 <sup>2</sup> Department of Physics, Stanford University, Stanford, CA 94305.

8 \*Corresponding author. Email: quake@stanford.edu

9 **Abstract**

10 Ecologists debate the relative importance of niche versus neutral processes in understanding  
11 biodiversity<sup>1,2</sup>. This debate is especially pertinent to microbial communities, which play crucial roles in  
12 biogeochemical cycling<sup>3,4</sup>, food production<sup>5</sup>, industrial processes<sup>6,7</sup>, and human health and disease<sup>8</sup>.  
13 Here we created a synthetic microbial community using heritable genetic barcodes and tracked  
14 community composition over time across a range of experimental conditions. We show that a transition  
15 exists between the neutral and niche regimes, and, consistent with theory, the crossover point depends  
16 on factors including immigration, fitness, and population size. We find that diversity declined most  
17 rapidly at intermediate population sizes, which can be explained by a tradeoff between replacement by  
18 migration and duration of growth. We then ran an experiment where the community underwent abrupt  
19 or gradual changes in size, the outcome of which highlights that selecting the correct model is essential  
20 to managing diversity. Taken together these results emphasize the importance of including niche effects  
21 to obtain realistic models across a wide range of parameters, even in simple systems.

22 **Main text:**

23           Microorganisms are essential players in areas such as health, disease, industry, and the  
24 environment. Furthermore, it is often the community that gives rise to the output or property of  
25 interest<sup>9</sup>, rather than any individual organism. Understanding microbial communities is therefore  
26 important in a wide variety of systems for predicting responses to anthropogenic<sup>4</sup> and natural  
27 perturbations, engineering desired outputs<sup>10,11</sup>, and understanding native functions. The study of  
28 microbial communities has been aided by increasing quantities of data as sequencing technologies have  
29 rapidly advanced, but in order to move from taxonomic descriptions to deeper understanding, there  
30 have been calls for placing these results in the context of a theoretical framework<sup>12–18</sup>.

31           Borrowed from macroscopic ecology, theoretical frameworks of microbial ecology can be  
32 divided into two types: models based on niche theory and neutral models. Models based on niche  
33 theory take a wide variety of forms, but critically differ from neutral theories in that they specify explicit  
34 differences between community members. For instance, one species may grow faster in certain abiotic  
35 conditions or might be killed as prey to feed another species<sup>19</sup>. These models can make detailed  
36 predictions, but often necessitate measurements or estimates of many parameters<sup>20,21</sup>. Conversely,  
37 neutral models take into account only random mechanisms. In ecology, they draw no distinction  
38 between individuals even across different species, which act in competition for a single limiting  
39 resource. Each species has the same fitness, and each species' relative abundance changes only due to  
40 random processes such as immigration and drift from random sampling<sup>1</sup>. Despite obvious species  
41 differences documented by decades of observation, neutral models can, perhaps surprisingly, predict  
42 frequently observed patterns of natural communities, such as lognormal-like species relative abundance  
43 distributions<sup>22</sup> and power law-like species area relationships<sup>23</sup>. Neutral models are thus a potentially  
44 enticing way to understand communities by abstracting away complicated differences between species,

45 but a natural question arises about their applicability under varying balances between niche and random  
46 processes.

47         As in macroscopic ecology, both niche and neutral theories have been discussed in microbial  
48 ecology, and studies have reached a wide range of conclusions about the relative contributions of niche  
49 and neutral processes to community assembly and structure<sup>24–35</sup>. As a result, researchers have  
50 emphasized the need for controlled time course experiments to explore how the balance between  
51 different mechanisms such as selection, drift, and immigration compels the application of different  
52 models<sup>14,36</sup>. Toward this end, we created a synthetic microbial “community” to emulate an ecological  
53 community. In this community “species” are represented by unique heritable DNA barcodes<sup>37,38</sup> that  
54 distinguish otherwise clonal *E. coli* bacterial cells. We created and validated a library of 456 different  
55 strains with Sanger sequencing to become the different species in our community. Starting with all  
56 species present, we grew this community to saturation then passaged it through a bottleneck once per  
57 day in 2 mL of shaken rich media using a wide range of bottleneck sizes ( $\sim 10^0$  to  $10^7$  cells, referred to as  
58 “small” and “large” bottlenecks, respectively). After each bottleneck we immigrated a controlled  
59 number of cells from the naïve barcoded “metacommunity” (average 55 cells/round). We took samples  
60 from the saturated growth at each time point, and used high throughput amplicon sequencing of the  
61 inserted barcodes to measure the abundance of each species present in each experimental community  
62 every round for 25 days. (Figure 1). Our approach allowed detection of species down to abundances of  
63 1/1,000, creating a detection threshold analogous to Preston’s veil<sup>39</sup>. Details of the experimental  
64 methods can be found in the Supplementary methods section.

65         We note, of course, that this simplified system far from captures the complexity present in  
66 natural microbial communities. The species exist in a well-mixed environment and, since they started as  
67 clones, they share similar fitness, and the same nutrient requirements. Therefore, interactions between

68 species are essentially a zero-sum competition for a limiting resource and do not include other types of  
69 interactions such as mutualism or predation. Recall, however, that neutral models only account for  
70 competitive interactions between species, and do not account for fitness differences. In fact, deviations  
71 from these conditions might be expected to drive the system away from neutral dynamics, so the  
72 experiments here might be expected to be even more neutral than higher complexity natural  
73 communities.

74 Figure 2A shows the number of species present over time in nine different experiments as the  
75 bottleneck size is varied. We find that the number of species present in each condition declines from the  
76 initial state with a variety of dynamics depending on bottleneck size. For each experiment we can also  
77 visualize the relative abundance of all 456 species over time with Muller plots (Figure 2 C-E,  
78 Supplementary figure 12) which show differences in dynamics between different species within a single  
79 experiment and different patterns of dynamics across bottleneck sizes. In analogy to classic species area  
80 curves<sup>40,41</sup>, where area is often assumed to be proportional to the number of individuals<sup>41</sup>, we also  
81 create log-log plots of the number of species present as a function of number of individuals passing  
82 through the bottleneck. These plots change with time and do not appear to reach an equilibrium (Figure  
83 2B). We also visualize the data with residence time histograms, rank abundance plots, and species  
84 relative abundance histograms. (Supplementary figures 15, 18, 21). For a simple case with no  
85 immigration see Supplementary figure 8.

86 We constructed and simulated a simple neutral model in an attempt to capture the system  
87 dynamics. The simulation has 25 rounds for each time point of the experiment. Each round, the new  
88 community is chosen from the old community by Poisson sampling to account for the bottleneck size,  $N$ .  
89 After the bottlenecking event, a mean number of cells,  $M$ , are immigrated from the original naïve  
90 population to the community, also by a Poisson process. When there is no immigration, this neutral

91 model predicts that the community will eventually contain only one species (Supplementary figure 8).  
92 When immigration is included the number of species present eventually reaches a stable equilibrium  
93 (Figure 3A, Supplementary figure 10) independent of starting conditions (Supplementary figure 9). Like  
94 other neutral models<sup>23</sup>, this model predicts a power law relationship with exponent near  $\frac{1}{4}$  when the  
95 number of species present is plotted against the bottleneck size for simulations that reach equilibrium  
96 (Figure 3I). Characteristic species relative abundance plots are also predicted. (Supplementary figure 22)  
97 Details of the model and expanded results can be found in Supplementary sections 2 and 3. A neutral  
98 model with additional stochasticity such as a very large variance in growth could lead to a fast loss of  
99 diversity as seen in the experiments, however our experimentally extracted estimates of the growth  
100 variance in individual lineages show that this variance is far too small to explain the loss of diversity at  
101 large bottleneck sizes (Supplementary section 4).

102           Community dynamics in many experiments begin to differ drastically from the predictions of the  
103 neutral model (Figure 3A). During the early time points (rounds 1-5) experiments with larger bottlenecks  
104 lose diversity slower relative to those with smaller bottlenecks as predicted by the neutral model, but  
105 experiments with medium and large bottlenecks lose diversity at much faster absolute rates than  
106 predicted. At later time points, the experiments with the two smallest bottlenecks continue to match  
107 the neutral predictions well, but experiments with medium sized bottlenecks have the lowest diversity,  
108 and experiments with large bottlenecks have lost an intermediate amount of diversity. This results in an  
109 experimental species vs area plot which does not follow a monotonic trend or stabilize over the duration  
110 of the experiment and rapidly develops a pronounced minimum at medium bottleneck sizes contrasting  
111 sharply with the neutral prediction. (Figure 3I)

112           A comparison of the Muller plots between the experiment and the neutral model illuminates  
113 the cause of the discrepancy. Muller plots from the neutral model match the experiments with the

114 smallest two bottlenecks reasonably well (Figure 3C vs 3D). However, in all experiments with larger  
115 bottlenecks the neutral model predicts relatively uniform and consistent relative abundances between  
116 species through the simulated time, compared to the experiments where one or more species begin to  
117 take over the population. (Figure 3F vs 3G, Supplementary figures 12&13) As these species become  
118 dominant, the community rapidly loses diversity. The dominant species seem to rise in prevalence  
119 exponentially over time (Figure 3G), typical of a fitness advantage instead of a random process. For  
120 species that appeared to grow adaptively, we extracted their relative fitnesses from the relative  
121 abundance vs time data, correcting for immigration. We obtained maximum per round Malthusian  
122 relative fitnesses,  $R$ , of 100-180%, translating to maximum per replication Malthusian relative fitnesses,  
123  $r$ , of 5-15%,  $R = xr$ , where  $x$  is the number of replication cycles required to grow to  $N_f$ , the final  
124 population size, given by  $x = \log_2 \frac{N_f}{N+M}$ .  $R$  decreased approximately linearly with  $\log N$ , consistent with  
125 constant  $r$  across experiments due to an advantage in the exponential phase of growth. We noticed  
126 overlap in the identities of the fit species across experiments, suggesting the presence of preexisting  
127 fitness differences. Regardless of the underlying source of these fitness differences (For a discussion of  
128 these see Supplementary section 5, Supplementary table 1) they appear to cause deviations from the  
129 neutral predictions for larger bottleneck sizes.

130 A more complex model that captures the dynamics over a larger range of bottleneck sizes might  
131 then depart from neutrality and include fitness differences between species. We changed the model to  
132 include preexisting fitness differences by assigning each species a per replication relative fitness,  
133 constant across experiments, selected randomly from an exponential distribution<sup>42,43</sup>. We then scaled  
134 this fitness by the number of replication rounds per experiment to obtain a per round fitness for each  
135 species, consistent with advantages in the exponential growth phase (Supplementary section 5). This  
136 modification captured many more features over a larger range of the parameter space (Figure 3B)

137 including species relative abundance trajectories where one or more species come to dominate (Figure  
138 3H) and non-monotonic species area curves which do not stabilize over the experiment (Figure 3I).  
139 Further additions to the model such as including the chance for mutations to arise during the course of  
140 the experiment may lead to a more complete picture, especially at timescales beyond those investigated  
141 here. Noting the success of the neutral model at small bottlenecks, we next assess when additional  
142 complexities departing from neutrality become necessary.

143         The fact that small fitness differences can lead to non-neutral dynamics has been understood in  
144 the population genetics literature<sup>44</sup> for some time and has more recently been studied in the context of  
145 neutral ecology models<sup>24,45–47</sup>. Transitions between niche and neutral have been proposed along  
146 speciation<sup>48</sup> and immigration<sup>24</sup> gradients and with different interplays between species interactions and  
147 stochasticity<sup>36,49–51</sup>. In our experiment, the different bottleneck sizes have different proportions of  
148 immigrants,  $m \equiv \frac{M}{M+N}$ , allowing us to explore the transition from neutral to niche along an immigration  
149 gradient in a well-controlled experiment. We can compare our results to theoretical predictions for the  
150 simple scenario of a single species with a fitness advantage,  $R$ . Our discussion of the predictions follows  
151 the derivation by Sloan *et al.*<sup>24</sup> and can be found in Supplementary section 2c.

152         For the neutral case, any given species' mean frequency, is simply equal to that species'  
153 frequency in the incoming immigrant pool,  $f_m$ . When there is a selective advantage, the probability  
154 distribution of the fit species,  $p(f)$ , is shifted toward higher frequencies. This effect is most noticeable  
155 when selection is stronger than stochastic effects, i.e  $R \gg \frac{1}{N_e}$  where  $N_e$  is the effective population size  
156 which is on the order of the total population after immigration,  $N + M$ . Even if selection is strong, the  
157 distribution can appear neutral if immigration is strong enough to compensate. For strong selection and  
158 strong immigration, the equilibrium frequency,  $f_{eq}$ , of the fit species given by the deterministic  
159 dynamics is a good measure for determining whether the distribution appears neutral or not. Figure 4A

160 shows  $f_{eq}$  as a function of the selective advantage,  $R$ , and immigration proportion,  $m$ , with the  
161 experimentally investigated points noted. The equilibrium frequency transitions from  $f_m$  when neutral  
162 up to 1 (indicating near-fixation) when non-neutral.

163           The transition happens when selection and immigration are the same scale. It can be  
164 understood heuristically by considering the effective growth rate of cells already in the population.  
165 Immigration exerts an effective negative fitness effect because cells are replaced by new immigrants.  
166 The frequency would drop to  $\tilde{f} = (1 - m)f_1 = \frac{N}{N+M}f_1$  because the fraction  $m$  are replaced. It is  
167 convenient to define a negative fitness  $\delta$  such that  $\frac{N}{N+M} = e^{-\delta}$ . After immigration, the population grows  
168 again until the end of the cycle. The fit species' frequency is then  $f_2 = e^R \tilde{f} = e^{R-\delta} f_1$ , so  $R - \delta$  acts as  
169 an effective fitness. If  $R - \delta \gtrsim 0$  then the frequency of cells in the population increases despite  
170 replacement by immigration. If we assume that the fitness advantage is an increase in the bacterial  
171 doubling rate by  $r$ , such that the fitness per cycle,  $R$ , scales with the number of doublings, then we can  
172 predict when the transition occurs for our experiments by solving  $R = \delta$ . For  $r \approx 8.5\%$  (and using  $M =$   
173  $55$  and a final cell count of  $6.5 \times 10^9$ ) The transition is predicted when  $m \approx 0.89$ . This compares well with  
174 the results that the smallest bottleneck ( $m \approx 0.93$ ) appears neutral while the second smallest  
175 bottleneck ( $m \approx 0.63$ ) and all larger bottlenecks appear non-neutral.

176           An interesting feature of the experiments is that the fastest exponential takeover and loss of  
177 diversity in the population happens at an intermediate bottleneck size, leading to, at least transiently,  
178 non-monotonic species area curves. If bottlenecking events are understood as a disturbance our results  
179 over this range of bottleneck sizes stand in contrast to the intermediate disturbance hypothesis, which  
180 suggests that diversity is maximized at intermediate levels of disturbance<sup>52,53</sup>. The concept of an  
181 effective fitness is useful in explaining this feature of the data; as the bottleneck size increases, the  
182 chance of being replaced by an immigrant (negative fitness effect) decreases, but the growth phase

183 advantage (positive fitness effect) also decreases. This results in a tradeoff where the effective fitness is  
184 maximized when  $N = \frac{M}{r} \ln 2$ , as found in Supplementary section 2c. With the same  $M$  and  $r$  as before,  
185 this gives  $N = 450$ , in agreement with the observation that a species in the  $N = 325$  bottleneck had the  
186 largest effective fitness extracted from the experiments.

187           In our final experiment we addressed whether knowing which model to apply has practical  
188 implications for understanding and managing how the community recovers from a disturbance.  
189 Communities that have been restricted by severe bottlenecks lose diversity, for example the human gut  
190 microbiome after antibiotic treatment<sup>54</sup>. A neutral model can make predictions about the best way to  
191 recover diversity. In certain parameter regimes, diversity is actually predicted to recover faster from a  
192 slow, rather than abrupt, increase in the bottleneck size. For example, in our system, neutral models  
193 predict that a community maintained with a bottleneck size of 32 recovering to a bottleneck size of 3250  
194 with immigration of 55 individuals per round gains diversity faster when the bottleneck is increased over  
195 several rounds rather than in a single round, prescribing a management strategy where the population  
196 gradually expands in size to maximize the rate at which diversity is recovered. (Figure 4C) Drastically  
197 different from neutral predictions, a model including fitness differences between species predicts that  
198 diversity would only continue to decay as the bottleneck size was increased, regardless of the dynamics  
199 of the increase, since increasing the bottleneck size decreases the immigration fraction, thereby  
200 increasing the effective fitness and only helping the fit species to outcompete. The experiment reveals  
201 that as the bottleneck is relaxed, regardless of relaxation dynamics, diversity is lost, in far better  
202 agreement with the model incorporating fitness differences, since the system transitioned well out of  
203 the regime where a neutral model was appropriate.

204           These results show a transition between niche and neutral regimes, providing an experimental  
205 case where a general guideline using fitness differences and immigration proportion successfully

206 predicts whether the system can be treated as neutral. If these conditions are not met, then non-neutral  
207 explanations are required to understand the community. These results also show that using the correct  
208 model is essential when predicting community response to change and can impact management  
209 strategies. Finally, we note that though these results were obtained using a synthetic microbial  
210 community, the framework, models, and analytical results may be useful in other ecological systems  
211 involving fitness differences and immigration.

212

213

214

215 **Acknowledgements:**

216 We gratefully acknowledge S. Norviel for assistance in validating the barcoded library, DS Fisher, JR  
217 Blundell, A Agarwala, F. Zanini, GR Dick, J Russel, and JT Morton for helpful comments, and C Vollmers,  
218 GR Mantalas, NF Neff, J Okamoto, and B Passarelli for help with sequencing and informatics. NJC and  
219 MTP are supported by NSF GRFP Fellowships, NJC is supported by a Siebel Scholar Fellowship, and MTP  
220 is supported by a Sara Hart Kimball Fellowship as part of Stanford's Graduate Fellowship Program. This  
221 research was supported by NASA Exobiology grant EXO-NNX11AR78G and U.S. National Science  
222 Foundation grants MCB 0546865, OISE 0968421, and DGE-114747.

223

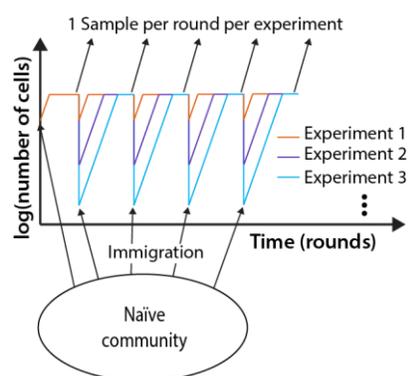
224

225

226

227

228 **Figures with captions:**



229

230 **Figure 1.** Experimental setup. 456 genetically barcoded *E. coli* strains were serially propagated at a

231 variety of dilutions with an influx of immigrant cells after each dilution.

232

233

234

235

236

237

238

239

240

241

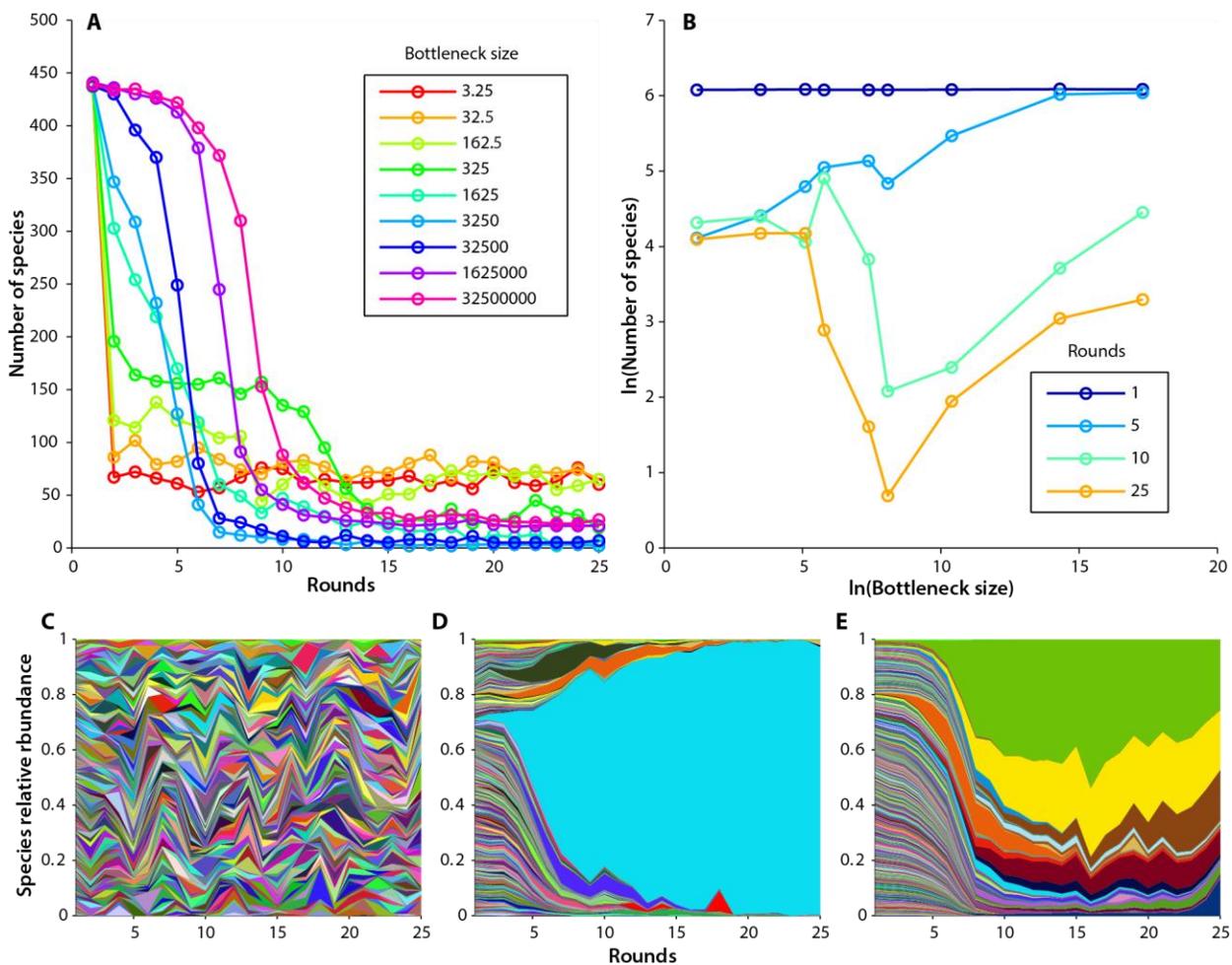
242

243

244

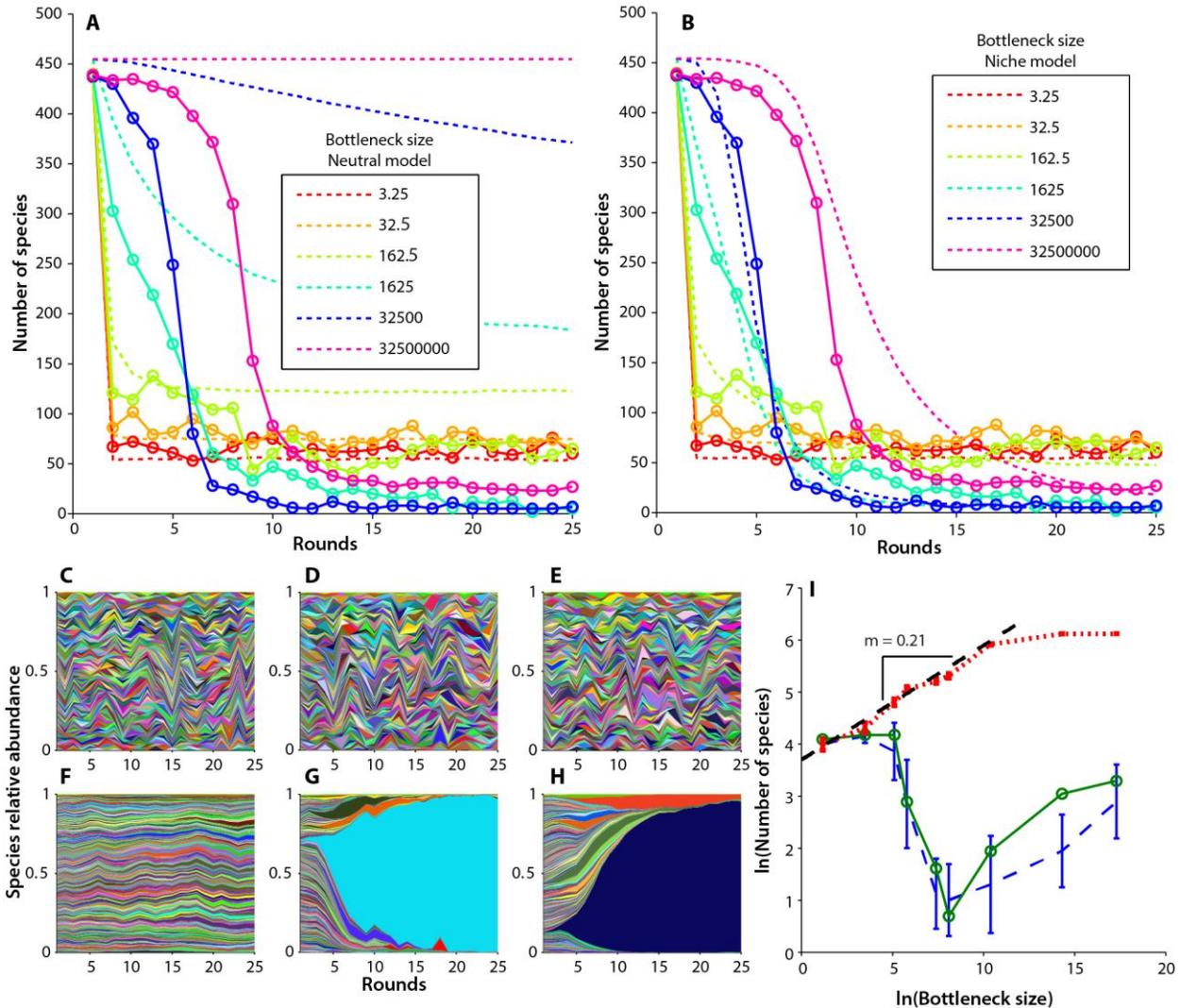
245

246



247  
248 **Figure 2.** Experimental time dynamics. Immigration is  $\sim 55$  cells/round for each experiment. A) Number  
249 of species detected over time for a range of dilutions. B) Species area curves over time. C-E) Muller plots  
250 showing the relative abundance of all 456 strains over time in three different experiments. Each species  
251 is represented by a different color where the proportion of each color represents the relative  
252 abundance of that species. Color/species pairings are consistent from plot to plot. C) bottleneck = 3.2  
253 cells/round. D) bottleneck = 1625 cells/round. E) bottleneck = 1,625,000 cells/round.

254  
255  
256  
257



258

259

260

261

262

263

264

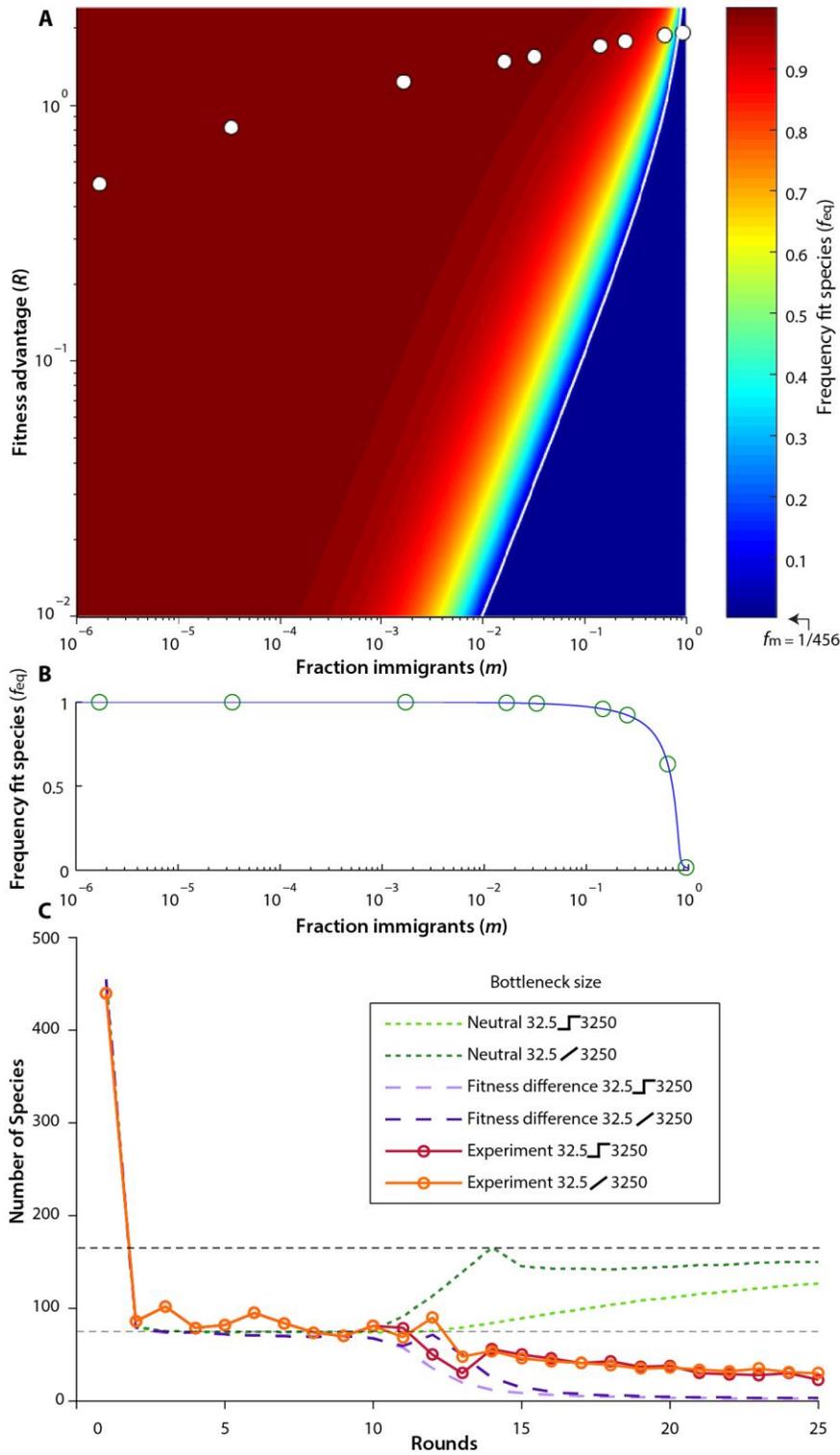
265

266

**Figure 3.** Comparison of neutral and niche models to experimental data. A-B) Number of species detected over time for a range of bottleneck sizes. Solid lines are experiment, dashed lines are model (means of 100 simulations). Note, for ease of visualization, not all bottleneck sizes are displayed. For the remaining bottleneck sizes, see Supplementary section 3 A) Neutral model. B) Niche model. C-H) Muller plots showing the relative abundance of all 456 strains over time for two different bottleneck sizes (representative trials picked for simulations) C-E) bottleneck = 3.25 cells/round F-H) bottleneck = 1625 cells/round for neutral (C & F) and niche (E & H) models compared to experiment (D & G). I) End point (round 25) species area curves for neutral (red dashed line), and niche (blue dashed line) models

267 compared to experiment (solid green line). Note the neutral model predicts a power law relationship for  
268 the smaller bottleneck sizes. Error bars denote one standard deviation from 100 simulated trials.

269



270  
271

272 **Figure 4.** Transition from niche to neutral. A) Phase space of equilibrium frequency ( $f_{eq}$ ) as a function of  
273 per round relative fitness ( $R$ ) vs fraction of immigrants ( $m$ ). Points indicate the experimentally  
274 investigated region where bottleneck size decreases from left to right (assuming a maximum per  
275 replication relative fitness of 8.5%). The immigration fraction of any species is,  $f_m \approx 1/456$ . The white  
276 line indicates the  $R = \delta$  threshold. B) A slice through the phase space along the experimentally tested  
277 conditions. A transition is predicted: at high immigrant fraction fit clones do not rise to high abundance  
278 and the system is neutral, and at the low immigration fraction the fit species dominates the population  
279 and cause departure from neutrality. Each circle indicates the theoretical prediction for  $f_{eq}$ . C)  
280 Community recovery. Here a community is maintained at a bottleneck size of 32.5 for 10 rounds then  
281 the bottleneck is allowed to expand to 3250. The recovery took the form of either a step function, or a  
282 gradual expansion. Though both models predict a similar number of species to the experimental  
283 community before recovery (lower horizontal dashed line), the neutral model (green lines) makes  
284 drastically different predictions than a model incorporating fitness differences (purple lines) after  
285 recovery. The neutral model predicts that the number of species in the community will increase to the  
286 new equilibrium level (upper horizontal dashed line), with the recovery happening much slower for the  
287 step function (light green dashed line) than the gradual increase (dark green dashed line). The model  
288 incorporating fitness differences predicts the community will lose diversity independent of step  
289 (lavender dashed line) or gradual recovery (purple dashed line). In both the step (solid red line) and  
290 gradual recovery (solid orange line) the experimental communities lost diversity.

291

292

293 **References:**

- 294 1. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography*. **17**, (Princeton  
295 University Press, 2001).
- 296 2. Chave, J., Muller-Landau, H. C. & Levin, S. A. Comparing classical community models: theoretical  
297 consequences for patterns of diversity. *Am. Nat.* **159**, 1–23 (2002).
- 298 3. Francis, C. A., Beman, J. M. & Kuypers, M. M. New processes and players in the nitrogen cycle:  
299 the microbial ecology of anaerobic and archaeal ammonia oxidation. *ISME J* **1**, 19–27 (2007).
- 300 4. Bardgett, R. D., Freeman, C. & Ostle, N. J. Microbial contributions to climate change through  
301 carbon cycle feedbacks. *ISME J* **2**, 805–814 (2008).
- 302 5. Ardhana, M. M. & Fleet, G. H. The microbial ecology of cocoa bean fermentations in Indonesia.  
303 *Int. J. Food Microbiol.* **86**, 87–99 (2003).
- 304 6. Wagner, M. *et al.* Microbial community composition and function in wastewater treatment  
305 plants. *Antonie Van Leeuwenhoek* **81**, 665–680 (2002).
- 306 7. Saunders, A. M., Albertsen, M., Vollertsen, J. & Nielsen, P. H. The activated sludge ecosystem  
307 contains a core community of abundant organisms. *ISME J.* **10**, 11–20 (2016).
- 308 8. Tremaroli, V. & Bäckhed, F. Functional interactions between the gut microbiota and host  
309 metabolism. *Nature* **489**, 242–249 (2012).
- 310 9. Lepp, P. W. *et al.* Methanogenic Archaea and human periodontal disease. *Proc. Natl. Acad. Sci. U.*  
311 *S. A.* **101**, 6176–81 (2004).
- 312 10. Shou, W., Ram, S. & Vilar, J. M. G. Synthetic cooperation in engineered yeast populations. *Proc.*  
313 *Natl. Acad. Sci. U. S. A.* **104**, 1877–82 (2007).
- 314 11. Fredrickson, J. K. Ecological communities by design. *Science (80-. )*. **348**, 1425–1427 (2015).
- 315 12. Prosser, J. I. *et al.* The role of ecological theory in microbial ecology. *Nat. Rev. Microbiol.* **5**, 384–  
316 392 (2007).
- 317 13. Faust, K. & Raes, J. Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* **10**,  
318 538–550 (2012).
- 319 14. Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C. & Martiny, J. B. H. Beyond biogeographic  
320 patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.* **10**, 1–10 (2012).
- 321 15. Robinson, C. J., Bohannan, B. J. M. & Young, V. B. From structure to function: the ecology of host-  
322 associated microbial communities. *Microbiol. Mol. Biol. Rev.* **74**, 453–76 (2010).
- 323 16. Widder, S., Allen, R., Pfeiffer, T., Curtis, T. & Wiuf, C. Challenges in microbial ecology: building  
324 predictive understanding of community function and dynamics. *ISME* **10**, 2557–2568 (2016).
- 325 17. Dethlefsen, L., McFall-Ngai, M. & Relman, D. A. An ecological and evolutionary perspective on  
326 humang-microbe mutualism and disease. *Nature* **449**, 811–818 (2007).
- 327 18. Lennon, J. T. & Locey, K. J. Macroecology for microbiology. *Environ. Microbiol. Rep.* (2017).  
328 doi:10.1111/1758-2229.12512
- 329 19. Alfred J. Lotka. *Elements of Physical Biology*. (Williams and Wilkins Company, 1925).  
330 doi:10.2105/AJPH.15.9.812-b
- 331 20. Follows, M. J., Dutkiewicz, S., Grant, S. & Chisholm, S. W. Emergent biogeography of microbial  
332 communities in a model ocean. *Science (80-. )*. **315**, 1843–1846 (2007).
- 333 21. Allison, S. D. A trait-based approach for modelling microbial litter decomposition. *Ecology Letters*  
334 **15**, 1058–1070 (2012).
- 335 22. Hubbell, S. P. Tree dispersion, abundance, and diversity in a tropical dry forest. *Science (80-. )*.  
336 **203**, 1299–1309 (1979).
- 337 23. Bell, G. The distribution of abundance in neutral communities. *Am. Nat.* **155**, 606–617 (2000).
- 338 24. Sloan, W. T. *et al.* Quantifying the roles of immigration and chance in shaping prokaryote  
339 community structure. *Environ. Microbiol.* **8**, 732–740 (2006).

- 340 25. Woodcock, S. *et al.* Neutral assembly of bacterial communities. *FEMS Microbiol. Ecol.* **62**, 171–  
341 180 (2007).
- 342 26. Horner-Devine, M., Silver, J. & Leibold, M. A Comparison of taxon co-occurrence patterns for  
343 macro-and microorganisms. *Ecology* **88**, 1345–1353 (2007).
- 344 27. Ayarza, J. M. & Erijman, L. Balance of Neutral and Deterministic Components in the Dynamics of  
345 Activated Sludge Floc Assembly. *Microb. Ecol.* **61**, 486–495 (2011).
- 346 28. Langenheder, S. & Székely, A. Species sorting and neutral processes are both important during  
347 the initial assembly of bacterial communities. *ISME J.* **5**, 1086–1094 (2011).
- 348 29. Zhang, Q. G., Buckling, A. & Godfray, H. C. J. Quantifying the relative importance of niches and  
349 neutrality for coexistence in a model microbial system. *Funct. Ecol.* **23**, 1139–1147 (2009).
- 350 30. Ofiteru, I. D. *et al.* Combined niche and neutral effects in a microbial wastewater treatment  
351 community. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15345–15350 (2010).
- 352 31. Jeraldo, P. *et al.* Quantification of the relative roles of niche and neutral processes in structuring  
353 gastrointestinal microbiomes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 9692–8 (2012).
- 354 32. Dumbrell, A. J., Nelson, M., Helgason, T., Dytham, C. & Fitter, A. H. Relative roles of niche and  
355 neutral processes in structuring a soil microbial community. *ISME J.* **4**, 337–345 (2010).
- 356 33. Fukami, T., Beaumont, H. J. E., Zhang, X.-X. & Rainey, P. B. Immigration history controls  
357 diversification in experimental adaptive radiation. *Nature* **446**, 436–9 (2007).
- 358 34. Manefield, M., Whiteley, A., Curtis, T. & Watanabe, K. Influence of sustainability and immigration  
359 in assembling bacterial populations of known size and function. *Microb. Ecol.* **53**, 348–354 (2007).
- 360 35. Venkataraman, A. *et al.* Application of a neutral community model to assess structuring of the  
361 human lung microbiome. *MBio* **6**, (2015).
- 362 36. Fisher, C. & Mehta, P. The transition between the niche and neutral regimes in ecology. *Proc.*  
363 *Natl. Acad. Sci. U. S. A.* **111**, 13111–13116 (2014).
- 364 37. Winzeler, E. A. *et al.* Functional characterization of the *S. cerevisiae* genome by gene deletion and  
365 parallel analysis. *Science (80-. )*. **285**, 901–906 (1999).
- 366 38. Levy, S. F. *et al.* Quantitative evolutionary dynamics using high-resolution lineage tracking.  
367 *Nature* **519**, 181–6 (2015).
- 368 39. Preston, F. W. The Commonness, And Rarity, of Species. *Ecology* **29**, 254–283 (1948).
- 369 40. MacArthur, R. H. & Wilson, E. O. *The theory of island biogeography*. **1**, (Princeton University  
370 Press, 1967).
- 371 41. Preston, F. W. The canonical distribution of commonness and rarity: Part I. *Ecology* **43**, 182–215  
372 (1962).
- 373 42. Orr, H. A. The Distribution of Fitness Effects Among Beneficial Mutations. *Genetics* **163**, 1519–  
374 1526 (2003).
- 375 43. Kassen, R. & Bataillon, T. Distribution of fitness effects among beneficial mutations before  
376 selection in experimental populations of bacteria. *Nat. Genet.* **38**, 484–8 (2006).
- 377 44. Fisher, R. A. The distribution of gene ratios for rare mutations. *Proc. R. Soc. Edinb* (1930).
- 378 45. Zhang, D. & Lin, K. The effects of competitive asymmetry on the rate of competitive  
379 displacement: how robust is Hubbell’s community drift model? *J. Theor. Biol.* **188**, 361–367  
380 (1997).
- 381 46. Yu, D., Terborgh, J. & Potts, M. D. Can high tree species richness be explained by Hubbell’s null  
382 model? *Ecol. Lett.* **1**, 193–199 (1998).
- 383 47. He, F., Zhang, D. Y. & Lin, K. Coexistence of nearly neutral species. *J. Plant Ecol.* **5**, 72–81 (2012).
- 384 48. Chisholm, R. A. & Pacala, S. W. Theory predicts a rapid transition from niche-structured to neutral  
385 biodiversity patterns across a speciation-rate gradient. *Theor. Ecol.* **4**, 195–200 (2011).
- 386 49. Haegeman, B. & Loreau, M. A mathematical synthesis of niche and neutral theories in community  
387 ecology. *J. Theor. Biol.* **269**, 150–165 (2011).

- 388 50. Pigolotti, S. & Cencini, M. Species abundances and lifetimes: From neutral to niche-stabilized  
389 communities. *J. Theor. Biol.* **338**, 1–8 (2013).
- 390 51. Gravel, D., Canham, C. D., Beaudet, M. & Messier, C. Reconciling niche and neutrality: The  
391 continuum hypothesis. *Ecol. Lett.* **9**, 399–409 (2006).
- 392 52. Connell, J. H. Diversity in Tropical Rain Forests and Coral Reefs. *Science (80-. )*. **199**, 1302–1310  
393 (1978).
- 394 53. Lubchenco, J. Plant Species Diversity in a Marine Intertidal Community: Importance of Herbivore  
395 Food Preference and Algal Competitive Abilities. *Am. Nat.* **112**, 23–39 (1978).
- 396 54. Dethlefsen, L., Huse, S., Sogin, M. L. & Relman, D. A. The pervasive effects of an antibiotic on the  
397 human gut microbiota, as revealed by deep 16s rRNA sequencing. *PLoS Biol.* **6**, 2383–2400  
398 (2008).
- 399
- 400
- 401
- 402
- 403
- 404
- 405
- 406
- 407
- 408
- 409
- 410
- 411
- 412
- 413
- 414
- 415
- 416
- 417
- 418