Neutral and selective dynamics in a synthetic microbial community

Nate J. Cira[a,1], Michael T. Pearce[b], and Stephen R. Quake[a,b,c,2]

[a]Department of Bioengineering, Stanford University, Stanford, CA 94305; [b]Department of Physics, Stanford University, Stanford, CA 94305; and [c]Chan Zuckerberg Biohub, San Francisco, CA 94158

Contributed by Stephen R. Quake, August 20, 2018 (sent for review May 11, 2018; reviewed by Michael M. Desai and Jeff Gore)

Ecologists debate the relative importance of selective vs. neutral processes in understanding biodiversity. This debate is especially pertinent to microbial communities, which play crucial roles in areas such as health, disease, industry, and the environment. Here, we created a synthetic microbial community using heritable genetic barcodes and tracked community composition over repeated rounds of subculture with immigration. Consistent with theory, we find a transition exists between neutral and selective regimes, and the crossover point depends on the fraction of immigrants and the magnitude of fitness differences. Neutral models predict an increase in diversity with increased carrying capacity, while our selective model predicts a decrease in diversity. The community here lost diversity with an increase in carrying capacity, highlighting that using the correct model is essential for predicting community response to change. Together, these results emphasize the importance of including selection to obtain realistic models of even simple systems.

ecology | synthetic biology | microbiology

Microorganisms are essential players in biogeochemical cycling (1, 2), food production (3), industrial processes (4, 5), and human health and disease (6). Furthermore, it is often the community that gives rise to the output or property of interest (7), rather than any individual organism. Understanding microbial communities is therefore important in a wide variety of systems for predicting responses to anthropogenic (2) and natural perturbations, engineering desired outputs (8, 9), and understanding native functions. The study of microbial communities has been aided by increasing quantities of data as sequencing technologies have rapidly advanced, but to move from taxonomic descriptions to deeper understanding, there have been calls for placing these results in the context of a theoretical framework (10–16).

There exist myriad ecological theories and numerous ways of classifying them. In the framework of Vellend (17), theories can be classified by their inclusion of four classes of process: selection, drift, dispersal, and speciation. Theories that do not include selective effects are considered neutral, and much debate in ecology has focused on the importance of including selection vs. considering only the neutral processes in understanding biodiversity (18–20). Models incorporating selection take a wide variety of forms but critically differ from neutral theories in that they specify explicit differences between community members. For instance, one species may grow faster in certain abiotic conditions or might be killed as prey to feed another species (21). These models can make detailed predictions but often necessitate measurements or estimates of many parameters (22, 23). Conversely, neutral models take into account only random mechanisms. In ecology, they draw no distinction between individuals, even across different species that act in competition for a single limiting resource. Each species has the same fitness, and each species' relative abundance changes only due to random processes such as immigration and drift from random sampling (18). Despite obvious species differences documented by decades of observation, neutral models under certain assumptions can, perhaps surprisingly, recapture frequently observed patterns of natural communities, such as lognormal-like species relative abundance distributions (24) and power law-like species area relationships (25). Neutral models are thus a potentially enticing way to understand communities by abstracting away complicated differences between species, but a natural question arises about their applicability under varying balances between selective and random processes.

As in macroscopic ecology, both selective and neutral theories have been discussed in microbial ecology, and studies have reached a wide range of conclusions about the relative contributions of selective and neutral processes to community assembly and structure (26–37). As a result, researchers have emphasized the need for controlled time-course experiments to explore how the balance among different processes such as selection, drift, and immigration compels the application of different models (12, 38). Toward this end, we created a synthetic microbial "community" to emulate an ecological community. In this community, "species" are represented by unique heritable DNA barcodes (39, 40) that distinguish otherwise clonal *Escherichia coli* bacterial cells. We created and validated a library of 456 different strains with Sanger sequencing to become the different species in our community. Starting with all species present, we grew this community to saturation then passaged it through a bottleneck once per day in 2 mL of shaken rich media using a wide range of bottleneck sizes ($\sim 10^0$ to 10^7 cells from the smallest to largest bottlenecks, respectively). After each bottleneck event, we immigrated a controlled number of cells from the naïve barcoded "metacommunity" (average of 55 cells per round). We took samples from the saturated growth at each time point and used high-throughput amplicon sequencing of the inserted barcodes to measure the abundance of each species present in each experimental community every round for 25 days (Fig. 1). Our approach had a detection limit for species down to abundances of 1/1,000. Details of the experimental methods can be found in *SI Appendix*, *Methods*.

We note, of course, that this simplified system far from captures the complexity present in natural microbial communities.

Significance

We created a synthetic microbial community to help understand how evolution and selection pressure change the species diversity of an ecosystem. Our results show that there is a clear transition between neutral and selective regimes that depends on the rate of immigration as well as the fitness differences.

Author contributions: N.J.C. and S.R.Q. designed research; N.J.C. and M.T.P. performed research; N.J.C., M.T.P., and S.R.Q. analyzed data; and N.J.C., M.T.P., and S.R.Q. wrote the paper.

Reviewers: M.M.D., Harvard University; and J.G., Massachusetts Institute of Technology.

The authors declare no conflict of interest.

[1]Present address: Rowland Institute at Harvard, Harvard University, Cambridge, MA 02142.

[2]To whom correspondence should be addressed. Email: quake@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1808118115/-/DCSupplemental.

Published online September 28, 2018.

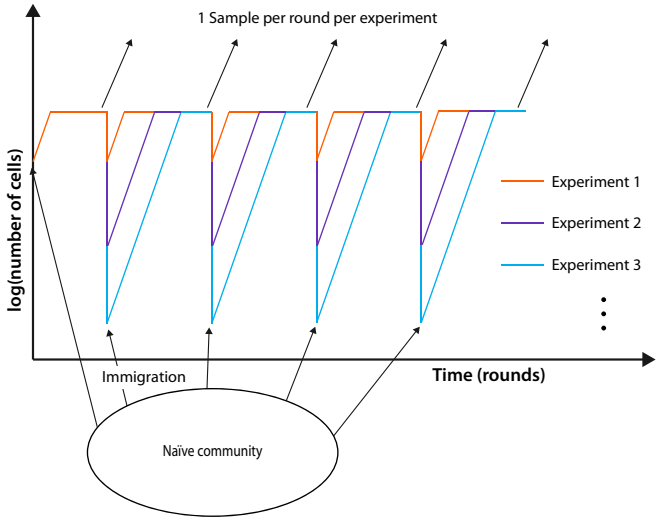

Fig. 1. Experimental setup. A total of 456 genetically barcoded *E. coli* strains were serially propagated at a variety of dilutions with an influx of immigrant cells after each dilution.

The species exist in a well-mixed environment and, since they started as clones, they share similar fitness and the same nutrient requirements. Therefore, interactions between species are essentially a zero-sum competition for a limiting resource and do not include other types of interactions such as mutualism or predation. Recall, however, that neutral models only account for competitive interactions between species and do not account for fitness differences. In fact, deviations from these conditions might be expected to drive the system away from neutral dynamics, so the experiments here might be expected to be even more neutral than higher-complexity natural communities.

Fig. 2*A* shows the number of species present over time in nine different experiments as the bottleneck size was varied. We found that the number of species present in each condition declined from the initial state, with a variety of dynamics depending on bottleneck size. For each experiment, we also visualized the relative abundance of all 456 species over time with Muller plots (Fig. 2 *C–E* and *SI Appendix*, Fig. S12), which show differences in dynamics between different species within a single experiment and different patterns of dynamics across bottleneck sizes. In analogy to classic species area curves (41, 42) in which area is often assumed to be proportional to the number of individuals (42), we also created log–log plots of the number of species present as a function of number of individuals passing through the bottleneck. These plots change with time and do not appear to reach an equilibrium (Fig. 2*B*). We also visualized the data with residence time histograms, rank abundance plots, and species relative abundance histograms (*SI Appendix*, Figs. S15, S18, and S21). For a simple case with no immigration, see *SI Appendix*, Fig. S8.

We constructed and simulated a simple neutral model in an attempt to capture the system dynamics. The simulation has 25 rounds for each time point of the experiment. Each round, the new community is chosen from the old community by Poisson sampling to account for the bottleneck size, *N*. After the bottlenecking event, a mean number of cells, *M*, are immigrated from the original naïve population to the community, also by a Poisson process. When there is no immigration, this neutral model predicts that the community will eventually contain only one species (*SI Appendix*, Fig. S8). When immigration is included, the number of species present eventually reaches a stable equilibrium (Fig. 3*A* and *SI Appendix*, Fig. S10), independent of starting conditions (*SI Appendix*, Fig. S9). Like other neutral models (25), this model predicts a power law relationship with an

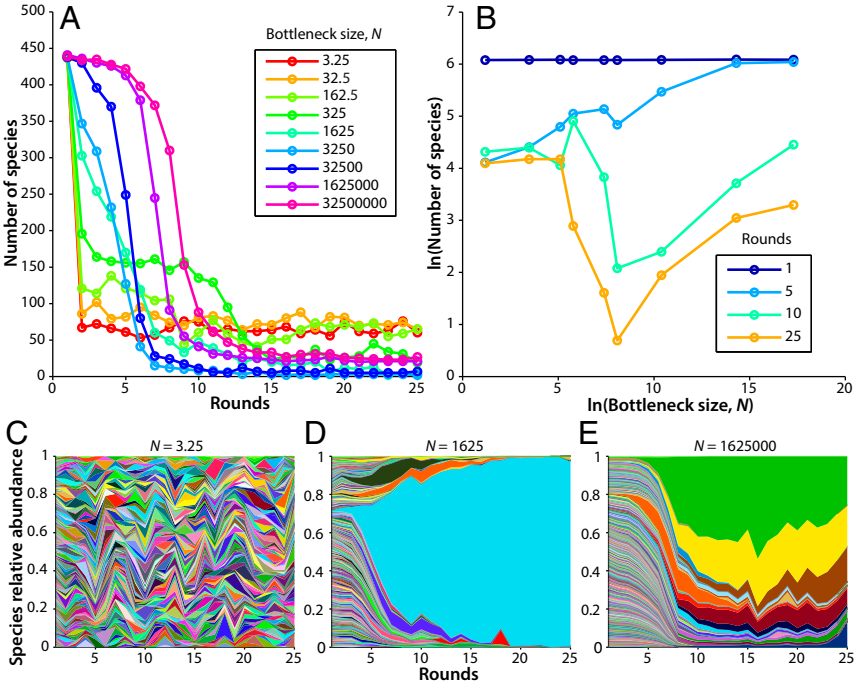

Fig. 2. Experimental time dynamics. Immigration is ∼55 cells per round for each experiment. (*A*) Number of species detected over time for a range of dilutions. (*B*) Species area curves over time. (*C–E*) Muller plots showing the relative abundance of all 456 strains over time in three different experiments. Each species is represented by a different color, whereby the proportion of each color represents the relative abundance of that species. Color/species pairings are consistent from plot to plot. Bottleneck was 3.25 cells per round in *C*; 1,625 cells per round in *D*; and 1,625,000 cells per round in *E*.

exponent near ¼ when the number of species present is plotted against the bottleneck size for simulations that reach equilibrium (Fig. 3*I*). Characteristic species relative abundance plots are also predicted (*SI Appendix*, Fig. S22). Details of the model and expanded results can be found in *SI Appendix*, *Models* and *Data and Comparisons*.

Community dynamics in many experiments begin to differ drastically from the predictions of the neutral model (Fig. 3*A*). During the early time points (rounds 1 to 5) experiments with larger bottlenecks lost diversity slower relative to those with smaller bottlenecks as predicted by the neutral model, but experiments with medium and large bottlenecks lost diversity at much faster absolute rates than predicted. At later time points, the experiments with the two smallest bottlenecks continued to match the neutral predictions well, but experiments with medium-sized bottlenecks had the lowest diversity and those with large bottlenecks lost an intermediate amount of diversity. This resulted in an experimental species-vs.-area plot that does not follow a monotonic trend or stabilize over the duration of the experiment and rapidly develops a pronounced minimum at medium bottleneck sizes, contrasting sharply with the neutral prediction (Fig. 3*I*). Although a neutral model with additional stochasticity (such as a very large variance in growth) could lead to a fast loss of diversity as seen in the experiments, our experimentally extracted estimates of the growth variance in individual lineages show that this variance is far too small to explain the loss of diversity at large bottleneck sizes (*SI Appendix*, *Stochasticity*).

A comparison of the Muller plots between the experiment and the neutral model illuminates the cause of the discrepancy. Muller plots from the neutral model matched the experiments

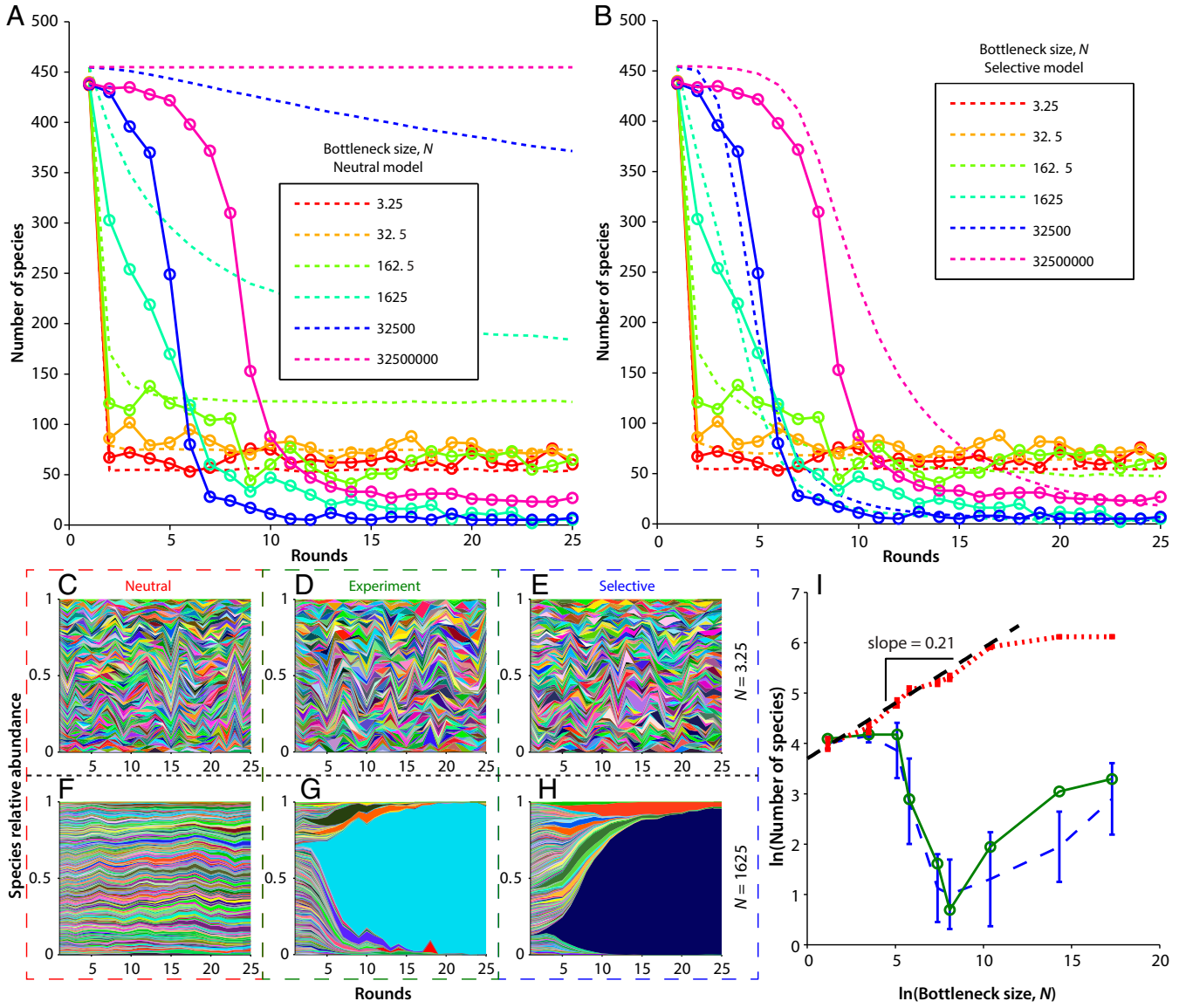

Fig. 3. Comparison of neutral (*A*) and selective (*B*) models to the experimental data. (*A* and *B*) Number of species detected over time for a range of bottleneck sizes. Solid lines are experimental data, dashed lines are model data (means of 100 simulations). Note that for ease of visualization, not all bottleneck sizes are displayed. For the remaining bottleneck sizes, see *SI Appendix*, *Data and Comparisons*. (*C*–*H*) Muller plots showing the relative abundance of all 456 strains over time for two different bottleneck sizes (representative trials picked for simulations). Bottleneck was 3.25 cells per round in *C*–*E* and 1,625 cells per round in *F*–*H* for neutral (*C* and *F*) and selective (*E* and *H*) models compared with experiment (*D* and *G*). (*I*) End point (round 25) species area curves for neutral (red dashed line) and selective (blue dashed line) models compared with experiment (solid green line). Note that the neutral model predicts a power law relationship for the smaller bottleneck sizes. Error bars denote 1 SD from 100 simulated trials and are negligible for larger bottleneck neutral simulations.

with the smallest two bottlenecks reasonably well (Fig. 3*C* vs. Fig. 3*D*). However, in all experiments with larger bottlenecks, the neutral model predicted relatively uniform and consistent relative abundances between species through the simulated time, compared with the experiments in which one or more species began to take over the population (Fig. 3*F* vs. Fig. 3*G* and *SI Appendix*, Figs. S12 and S13). As these species became dominant, the community rapidly lost diversity. The dominant species seemed to rise in prevalence exponentially over time (Fig. 3*G*), typical of a fitness advantage instead of a random process. For species that appeared to grow adaptively, we extracted their relative fitnesses from the relative abundance-vs.-time data, correcting for immigration. We obtained maximum per-round Malthusian relative fitnesses, R, of 100 to 180%, translating to maximum per-replication Malthusian relative fitnesses, r, of 5 to 15%, a similar order of magnitude to per-replication fitness differences measured in experimental evolution experiments starting with clonal microbes (40, 43). Here, $R = xr$, where x is the number of replication cycles required to grow to N_f, the final population size, given by $x = \log_2 \frac{N_f}{N+M}$. R decreased approximately linearly with $\log N$, consistent with constant r across experiments due to an advantage in the exponential phase of growth. We noticed overlap in the identities of the fit species across experiments, suggesting the presence of preexisting fitness differences. Regardless of the underlying source of these fitness differences (for a discussion of these, see *SI Appendix, Distribution and Nature of Selective Advantage* and Table S1), they appear to cause deviations from the neutral predictions for larger bottleneck sizes.

A more complex model that captures the dynamics over a larger range of bottleneck sizes might then depart from neutrality and include selection in the form of fitness differences between species. We changed the model to include preexisting fitness differences by assigning each species a per-replication relative fitness, constant across experiments, selected randomly from an exponential distribution (44, 45). We then scaled this fitness by the number of replication rounds per experiment to obtain a per-round fitness for each species, consistent with advantages in the exponential growth phase (*SI Appendix, Distribution and Nature of Selective Advantage*). Simulations including this modification captured many more features over a larger range of the parameter space (Fig. 3*B*), including species relative abundance trajectories in which one or more species come to dominate (Fig. 3*H*), as well as nonmonotonic species area curves that do not stabilize over the experiment (Fig. 3*I*). Further additions to the model, such as including the chance for mutations to arise during the course of the experiment, may lead to a more complete picture, especially at timescales beyond those investigated here. Noting the success of the neutral model at small bottlenecks, we next assessed under what conditions additional complexities departing from neutrality become necessary.

The fact that small fitness differences can lead to nonneutral dynamics has been understood in the population genetics literature (46) for some time and has more recently been studied in the context of neutral ecology models (26, 47–49). Transitions from neutrality have been proposed along speciation (50) and immigration (26) gradients and with different interplays between species interactions and stochasticity (38, 51–53). In our experiment, the different bottleneck sizes have different proportions of immigrants, $m \equiv \frac{M}{M+N}$, allowing us to explore the transition from neutral to selective along an immigration gradient in a well-controlled experiment. In addition to the simulated models, below we discuss and compare our results to the theoretical predictions for the simple scenario of a single species with a fitness advantage, R, over a neutral background, following the derivation by Sloan et al. (26), which can be found in *SI Appendix, Models*.

For the neutral case, any given species' mean frequency is simply equal to that species' frequency in the incoming immigrant pool, f_m. When there is a selective advantage, the probability distribution of the fit species, $P(f)$, is shifted toward higher frequencies. This effect is most noticeable when selection is stronger than stochastic effects—that is, $R \gg \frac{1}{N_e}$, where N_e is the effective population size, which is on the order of the harmonic mean of the population size as it grows, starting at the total population size after immigration, $N + M$, and ending at the population size after saturated growth. Even if selection is strong, the distribution can appear neutral if immigration is strong enough to compensate. For strong selection and strong immigration, the equilibrium frequency, f_{eq}, of the fit species given by the deterministic dynamics is a good measure for determining whether the distribution appears neutral or not. Fig. 4*A* shows f_{eq} as a function of the selective advantage, R, and immigration proportion, m, with the experimentally investigated points noted. The f_{eq} transitions from f_m when neutral, up to 1 (indicating near-fixation) when nonneutral.

The transition from neutral to selective happens when selection and immigration are the same scale. It can be understood heuristically by considering the effective growth rate of cells already in the population. Immigration exerts an effective negative fitness effect because cells are replaced by new immigrants. Starting at initial frequency, f_1, the frequency would drop to $\tilde{f} = (1-m)f_1 = \frac{N}{N+M}f_1$ because the fraction m is replaced. It is convenient to define a negative fitness, δ, such that $\frac{N}{N+M} = e^{-\delta}$. After immigration, the population grows 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 an effective fitness. If $R - \delta \gtrsim 0$, then the frequency of cells in the population increases despite replacement by immigration, defining the conditions for the transition from a neutral to a selective regime. If we assume that the fitness advantage is an increase in the bacterial doubling rate by r such that the fitness per cycle, R, scales with the number of doublings, then we can predict when the transition occurs for our experiments by solving $R = \delta$. For $r \approx 8.5\%$ (and using $M = 55$ and a final cell count of 6.5×10^9), the transition is predicted when $m \approx 0.89$. This matches well the results in Fig. 3*I* (comparing number of species between the experiment and each simulated model), whereby the smallest bottleneck ($m = 0.94$) appears neutral while the second-smallest bottleneck ($m = 0.63$) departs from neutrality and all larger bottlenecks appear nonneutral ($m \leq 0.25$). Note that the maintenance of diversity by migration is reminiscent of the rescue effect (54) or mass effects (55) from classical ecology.

An interesting feature of the experiments is that the fastest exponential takeover and loss of diversity in the population happens at an intermediate bottleneck size, leading to, at least transiently, nonmonotonic species area curves. This feature is not directly predicted by existing theory and, if bottlenecking events are understood as a disturbance, our results over this range of bottleneck sizes stand in contrast to the intermediate disturbance hypothesis, which suggests that diversity is maximized at intermediate levels of disturbance (56, 57), likely due to lack of a competition-colonization trade-off in this system. The concept of an effective fitness, $R - \delta$, is useful in explaining this feature of the data; as the bottleneck size increases, the chance of being replaced by an immigrant (negative fitness effect, δ) decreases, but the growth phase advantage (positive fitness effect, R) also decreases. This results in a trade-off in which the effective fitness is maximized when $N = \frac{M}{r}\ln 2$, as found in *SI Appendix, Models*. With the same M and r as before, this gives $N = 450$, in agreement with the observation that a species in the $N = 325$ bottleneck had the largest effective fitness extracted from the experiments.

In our final experiment, we addressed whether knowing which model to apply has practical implications for understanding and

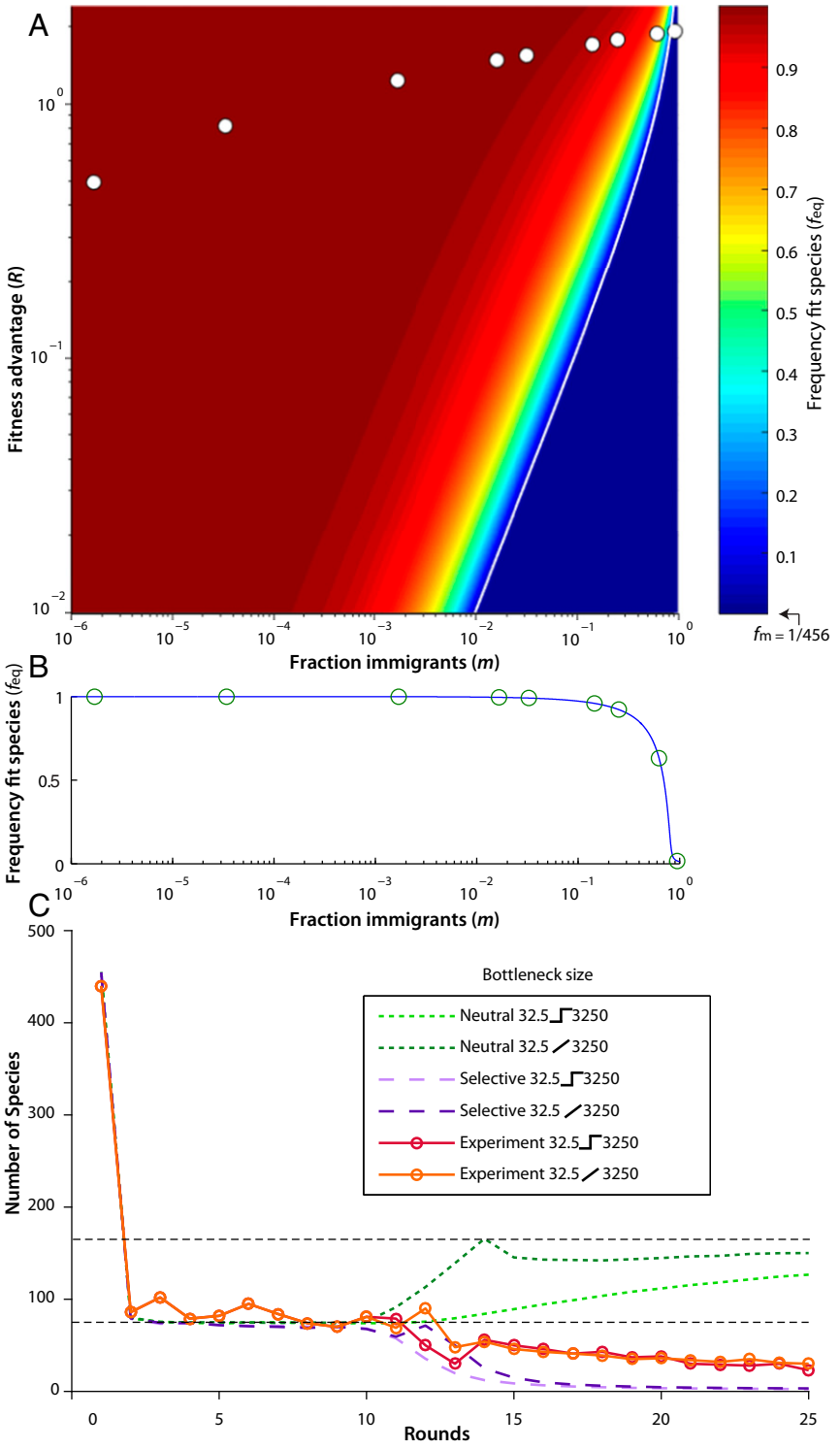

Fig. 4. Transition from selective to neutral. (*A*) Phase space of equilibrium frequency (f_{eq}) of a single-fit clone in a neutral background as a function of per-round relative fitness (R) vs. fraction of immigrants (m). Points indicate the experimentally investigated region, and bottleneck size decreases from left to right (assuming a maximum per-replication relative fitness of 8.5%). The immigration fraction of any species is, $f_m \approx 1/456$. The white line indicates the $R = \delta$ threshold. (*B*) A slice through the phase space along the experimentally tested conditions. A transition is predicted: At high immigrant fraction, fit clones do not rise to high abundance and the system is neutral; and at the low immigration fraction, the fit species dominates the population and causes departure from neutrality. Each circle indicates the theoretical prediction for f_{eq}. (*C*) Community recovery. Here, a community is maintained at a bottleneck size of 32.5 for 10 rounds and then the bottleneck is allowed to expand to 3,250. The recovery took the form of either a step function or a gradual expansion. Although both models predict a similar number of species to the experimental community before recovery (lower horizontal dashed line), the neutral model (green lines) makes drastically different predictions than a selective model (purple lines) after recovery. The neutral model predicts that the number of species in the community will increase to the new equilibrium level (upper horizontal dashed line), with the recovery happening much slower for the step function (light-green dashed line) than the gradual increase (dark-green dashed line). The selective model predicts that the community will lose diversity independent of step (lavender dashed line) or gradual recovery (purple dashed line). In both the step (solid red line) and gradual recovery (solid orange line) the experimental communities lost diversity.

managing how a community responds to change. Communities that have been restricted by severe bottlenecks lose diversity—for example, the human gut microbiome after antibiotic treatment (58). To investigate recovery from severe bottleneck, we maintained a community with a bottleneck size of 32 for 10 rounds and then increased the bottleneck size to 3,250 with immigration of 55 individuals per round. Neutral models predict that as the bottleneck size is increased, the diversity of the community will also increase. Furthermore, under these conditions, a neutral model predicts that diversity is actually predicted to recover faster from a slow, rather than an abrupt, increase in the bottleneck size, prescribing a strategy in which the bottleneck size is gradually increased to maximize the rate at which diversity is recovered (Fig. 4*C*). To test this in the experiment, we carried out both gradual and step increases to the bottleneck size. Opposite the neutral predictions, a selective model predicts that diversity would only continue to decay as bottleneck size is increased, regardless of the dynamics of the increase, since increasing the bottleneck size decreases the immigration fraction, thereby increasing the effective fitness and only helping the fit species to outcompete. Both models captured the experimental results at the initial smaller bottleneck size, but as the bottleneck was increased, regardless of increase dynamics, diversity was lost, which was in far better agreement with the model incorporating fitness differences because the system transitioned well out of the regime in which a neutral model was appropriate. (Fig. 4*C*) This illustrates that using the correct model is important to predicting the outcomes of change and that management strategies based on incorrect models can have unintended consequences.

Perhaps the most surprising aspect of this work is that despite the construction of initial conditions in our system with minimal differences between species and a relatively homogenous environment, a neutral model failed at all but the highest immigration fractions to capture the community dynamics. However, as shown in Fig. 4*A*, if fitness differences had been an order of magnitude or more smaller, a neutral model would have captured communities with much more reasonable immigration fractions. Classic population genetics predicts that if we were to continue running the experiment, as the population approached optimal fitness, the fitness differences between individuals would decrease (59). Could this effect redeem a neutral model for capturing longer-term communities? On the surface, systems such as the long-term evolution experiment show diminishing population-level fitness gains over time (60), suggesting smaller fitness differences between individuals. However, beneath the surface of the population, clear niche differentiation (61–63) and definitively selective changes in allele frequency are occurring, even after 60,000 generations (63), making increased duration unlikely to redeem a purely neutral framework. Similar emergent niche effects have also been observed in other simple systems (64, 65), athough we did not observe them in this experiment, potentially due to the duration and the decreased ability to resolve community members as species were lost and barcode pools became more homogenous. Given the poor performance of a neutral framework in simple systems, we suspect that selective effects cannot typically be neglected in microbial communities, even after long amounts of time.

These results show a transition between selective and neutral regimes, providing an experimental case in which the general $R - \delta$ guideline balancing fitness differences and immigration proportion successfully predicts whether the system can be treated as neutral (only if $R - \delta < 0$). If these conditions are not met, then nonneutral explanations are required to understand the community. These results also show that using the correct model is essential when predicting community response to change and can impact management strategies. Lastly, we note that athough these results were obtained using a synthetic microbial community, the framework, models, and analytical results may be useful in other ecological systems involving fitness differences and immigration.

Materials and Methods

Random DNA barcodes were cloned into plasmids and transformed into otherwise clonal *E. coli* bacteria. A total of 456 strains that passed validation by Sanger sequencing were grown to saturation and combined in an equal volume ratio to create a strain library that was manipulated as a mock ecological community in these experiments. This community was propagated in 2 mL of shaken LB Lennox media at 37 °C via serial passaging once per day across a broad range of dilutions such that different numbers of cells passed from one time point to the next across different experiments. After each dilution, a mean of ∼55 cells from the original (complete) community was added to each experimental condition.

Samples from the saturated growth after each time point were extracted by Miniprep Kit (Qiagen), and amplicon libraries were generated and sequenced on an Illumina MiSeq using a two-stage PCR that introduced unique molecular indices. After demultiplexing, reads were assigned to particular strains from the initial library, and the fraction of reads above a noise threshold of 1/1,000 was taken to be proportional to the fraction of the community made up by each labeled strain. Additional experiments were run in which the dilution factor was changed across rounds to explore the effect of changing population size on community dynamics.

Stochastic simulations were set up and run in MATLAB to model community dynamics, using Poisson sampling to mimic the dilution and immigration steps of each experimental condition. The neutral versions of these simulations did not contain fitness differences between community members, whereas the selective versions incorporated fitness differences. Full details on materials and methods appear in *SI Appendix*.

ACKNOWLEDGMENTS. We thank S. Norviel for assistance in validating the barcoded library; D. S. Fisher, J. R. Blundell, A. Agarwala, F. Zanini, G. R. Dick, B. Chen, A. I. Lee-Richerson, J. Russel, and J. T. Morton for helpful comments; and C. Vollmers, G. R. Mantalas, N. F. Neff, J. Okamoto, and B. Passarelli for help with sequencing and informatics. N.J.C. and M.T.P. are supported by the National Science Foundation Graduate Research Fellowship Program, N.J.C. is supported by a Siebel Scholar Fellowship, and M.T.P. is supported by a Sara Hart Kimball Fellowship as part of Stanford's Graduate Fellowship Program. This research was supported by National Aeronautics and Space Administration Exobiology Grant EXO-NNX11AR78G and US National Science Foundation Grants MCB 0546865, OISE 0968421, and DGE-114747.

1. Francis CA, Beman JM, Kuypers MM (2007) New processes and players in the nitrogen cycle: The microbial ecology of anaerobic and archaeal ammonia oxidation. *ISME J* 1: 19–27.
2. Bardgett RD, Freeman C, Ostle NJ (2008) Microbial contributions to climate change through carbon cycle feedbacks. *ISME J* 2:805–814.
3. Ardhana MM, Fleet GH (2003) The microbial ecology of cocoa bean fermentations in Indonesia. *Int J Food Microbiol* 86:87–99.
4. Wagner M, et al. (2002) Microbial community composition and function in wastewater treatment plants. *Antonie vVan Leeuwenhoek* 81:665–680.
5. Saunders AM, Albertsen M, Vollertsen J, Nielsen PH (2016) The activated sludge ecosystem contains a core community of abundant organisms. *ISME J* 10:11–20.
6. Tremaroli V, Bäckhed F (2012) Functional interactions between the gut microbiota and host metabolism. *Nature* 489:242–249.
7. Lepp PW, et al. (2004) Methanogenic Aarchaea and human periodontal disease. *Proc Natl Acad Sci USA* 101:6176–6181.
8. Shou W, Ram S, Vilar JMG (2007) Synthetic cooperation in engineered yeast populations. *Proc Natl Acad Sci USA* 104:1877–1882.
9. Fredrickson JK (2015) Ecological communities by design. *Science* 348:1425–1427.
10. Prosser JI, et al. (2007) The role of ecological theory in microbial ecology. *Nat Rev Microbiol* 5:384–392.
11. Faust K, Raes J (2012) Microbial interactions: From networks to models. *Nat Rev Microbiol* 10:538–550.
12. Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH (2012) Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nat Rev Microbiol* 10:497–506.
13. Robinson CJ, Bohannan BJM, Young VB (2010) From structure to function: The ecology of host-associated microbial communities. *Microbiol Mol Biol Rev* 74:453–476.
14. Widder S, et al.; Isaac Newton Institute Fellows (2016) Challenges in microbial ecology: Building predictive understanding of community function and dynamics. *ISME J* 10:2557–2568.
15. Dethlefsen L, McFall-Ngai M, Relman DA (2007) An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* 449:811–818.
16. Lennon JT, Locey KJ (2017) Macroecology for microbiology. *Environ Microbiol Rep* 9: 38–40.
17. Vellend M (2010) Conceptual synthesis in community ecology. *Q Rev Biol* 85:183–206.

18. Hubbell SP (2001) *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ Press, Princeton), Vol 17.
19. Chave J, Muller-Landau HC, Levin SA (2002) Comparing classical community models: Theoretical consequences for patterns of diversity. *Am Nat* 159:1–23.
20. Adler PB, Hillerislambers J, Levine JM (2007) A niche for neutrality. *Ecol Lett* 10: 95–104.
21. Lotka AJ (1925) *Elements of Physical Biology* (Williams and Wilkins, Baltimore).
22. Follows MJ, Dutkiewicz S, Grant S, Chisholm SW (2007) Emergent biography of microbial communities in a model ocean. *Science* 315:1843–1846.
23. Allison SD (2012) A trait-based approach for modelling microbial litter decomposition. *Ecol Lett* 15:1058–1070.
24. Hubbell SP (1979) Tree dispersion, abundance, and diversity in a tropical dry forest. *Science* 203:1299–1309.
25. Bell G (2000) The distribution of abundance in neutral communities. *Am Nat* 155: 606–617.
26. Sloan WT, et al. (2006) Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ Microbiol* 8:732–740.
27. Woodcock S, et al. (2007) Neutral assembly of bacterial communities. *FEMS Microbiol Ecol* 62:171–180.
28. Horner-Devine MC, et al. (2007) A comparison of taxon co-occurrence patterns for macro- and microorganisms. *Ecology* 88:1345–1353.
29. Ayarza JM, Erijman L (2011) Balance of neutral and deterministic components in the dynamics of activated sludge floc assembly. *Microb Ecol* 61:486–495.
30. Langenheder S, Székely AJ (2011) Species sorting and neutral processes are both important during the initial assembly of bacterial communities. *ISME J* 5:1086–1094.
31. Zhang QG, Buckling A, Godfray HCJ (2009) Quantifying the relative importance of niches and neutrality for coexistence in a model microbial system. *Funct Ecol* 23: 1139–1147.
32. Ofiteru ID, et al. (2010) Combined niche and neutral effects in a microbial wastewater treatment community. *Proc Natl Acad Sci USA* 107:15345–15350.
33. Jeraldo P, et al. (2012) Quantification of the relative roles of niche and neutral processes in structuring gastrointestinal microbiomes. *Proc Natl Acad Sci USA* 109: 9692–9698.
34. Dumbrell AJ, Nelson M, Helgason T, Dytham C, Fitter AH (2010) Relative roles of niche and neutral processes in structuring a soil microbial community. *ISME J* 4:337–345.
35. Fukami T, Beaumont HJE, Zhang X-X, Rainey PB (2007) Immigration history controls diversification in experimental adaptive radiation. *Nature* 446:436–439.
36. Manefield M, Whiteley A, Curtis T, Watanabe K (2007) Influence of sustainability and immigration in assembling bacterial populations of known size and function. *Microb Ecol* 53:348–354.
37. Venkataraman A, et al. (2015) Application of a neutral community model to assess structuring of the human lung microbiome. *MBio* 6:e02284-14.
38. Fisher CK, Mehta P (2014) The transition between the niche and neutral regimes in ecology. *Proc Natl Acad Sci USA* 111:13111–13116.
39. Winzeler EA, et al. (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285:901–906.
40. Levy SF, et al. (2015) Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature* 519:181–186.
41. MacArthur RH, Wilson EO (1967) *The Theory of Island Biogeography* (Princeton Univ Press, Princeton), Vol 1.
42. Preston FW (1962) The canonical distibution of commonness and rarity: Part I. *Ecology* 43:182–215.
43. Imhof M, Schlotterer C (2001) Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. *Proc Natl Acad Sci USA* 98:1113–1117.
44. Orr HA (2003) The distribution of fitness effects among beneficial mutations. *Genetics* 163:1519–1526.
45. Kassen R, Bataillon T (2006) Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. *Nat Genet* 38:484–488.
46. Fisher RA (1930) The distribution of gene ratios for rare mutations. *Proc R Soc Edinburgh* 50:204–219.
47. Zhang D, Lin K (1997) The effects of competitive asymmetry on the rate of competitive displacement: How robust is Hubbell's community drift model? *J Theor Biol* 188: 361–367.
48. Yu D, Terborgh J, Potts MD (1998) Can high tree species richness be explained by Hubbell's null model? *Ecol Lett* 1:193–199.
49. He F, Zhang DY, Lin K (2012) Coexistence of nearly neutral species. *J Plant Ecol* 5: 72–81.
50. Chisholm RA, Pacala SW (2011) Theory predicts a rapid transition from niche-structured to neutral biodiversity patterns across a speciation-rate gradient. *Theor Ecol* 4:195–200.
51. Haegeman B, Loreau M (2011) A mathematical synthesis of niche and neutral theories in community ecology. *J Theor Biol* 269:150–165.
52. Pigolotti S, Cencini M (2013) Species abundances and lifetimes: From neutral to niche-stabilized communities. *J Theor Biol* 338:1–8.
53. Gravel D, Canham CD, Beaudet M, Messier C (2006) Reconciling niche and neutrality: The continuum hypothesis. *Ecol Lett* 9:399–409.
54. Brown JH, Kodric-Brown A (1977) Turnover rates in insular biogeography: Effect of immigration on extinction. *Ecology* 58:445–449.
55. Shmida A, Wilson MV (1985) Biological determinants of species diversity. *J Biogeogr* 12:1–20.
56. Connell JH (1978) Diversity in tropical rain forests and coral reefs. *Science* 199: 1302–1310.
57. Lubchenco J (1978) Plant species diversity in a marine intertidal community: Importance of herbivore food preference and algal competitive abilities. *Am Nat* 112:23–39.
58. Dethlefsen L, Huse S, Sogin ML, Relman DA (2008) The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol* 6:e280.
59. Fisher RF (1930) *The Genetical Theory of Natural Selection* (Oxford Univ Press, Clarendon, UK).
60. Wiser MJ, Ribeck N, Lenski RE (2013) Long-term dynamics of adaptation in asexual populations. *Science* 342:1364–1367.
61. Rozen DE, Lenski RE (2000) Long-term experimental evolution in Escherichia coli. VIII. Dynamics of a balanced polymorphism. *Am Nat* 155:24–35.
62. Plucain J, et al. (2014) Epistasis and allele specificity in the emergence of a stable polymorphism in *Escherichia coli*. *Science* 343:1366–1369.
63. Good BH, McDonald MJ, Barrick JE, Lenski RE, Desai MM (2017) The dynamics of molecular evolution over 60,000 generations. *Nature* 551:45–50.
64. Rainey PB, Travisano M (1998) Adaptive radiation in a heterogeneous environment. *Nature* 394:69–72.
65. Helling RB, Vargas CN, Adams J (1987) Evolution of *Escherichia coli* during growth in a constant environment. *Genetics* 116:349–358.