



## Supplementary Materials for

### **Emergent phases of ecological diversity and dynamics mapped in microcosms**

Jiliang Hu *et al.*

Corresponding author: Jeff Gore, gore@mit.edu

*Science* **378**, 85 (2022)  
DOI: 10.1126/science.abm7841

#### **The PDF file includes:**

Materials and Methods  
Figs. S1 to S31  
Tables S1 to S3  
References

#### **Other Supplementary Material for this manuscript includes the following:**

MDAR Reproducibility Checklist  
Data S1

## Materials and Methods

### Bacterial isolates, media and culturing conditions

We constructed the library of 48 bacterial species using 24 bacterial isolates from soil samples taken at Middlesex Fells Reservation in Somerville, Massachusetts, and 24 isolates from the *C. elegans* intestine. This library is phylogenetically diverse, with isolates coming from 26 different families among 4 phyla: Proteobacteria, Firmicutes, Bacteroidota and Actinobacteriota (Fig. S13, 14).

In the case of low interaction strength (low nutrients concentration) conditions, experimental communities were cultured in Base Medium (BM):  $1\text{ gL}^{-1}$  yeast extract and  $1\text{ gL}^{-1}$  soytone from Becton Dickinson,  $10\text{ mM}$  sodium phosphate,  $0.1\text{ mM}$   $\text{CaCl}_2$ ,  $2\text{ mM}$   $\text{MgCl}_2$ ,  $4\text{ mgL}^{-1}$   $\text{NiSO}_4$  and  $50\text{ mgL}^{-1}$   $\text{MnCl}_2$ , pH adjusted to 6.5. For intermediate interaction strength (medium nutrients concentration) conditions, we used BM supplemented with  $5\text{ gL}^{-1}$  glucose and  $4\text{ gL}^{-1}$  urea. For the high interaction strength (high nutrients concentration) condition, we used BM supplemented with  $20\text{ gL}^{-1}$  glucose and  $16\text{ gL}^{-1}$  urea. All media were filter sterilized using Bottle Top Filtration Units (VWR). All of the chemicals were purchased from Sigma–Aldrich unless otherwise stated.

Both monocultures and communities of the bacterial isolates were grown in 96-deepwell plates (Deepwell plate 96/1000 $\mu\text{l}$ ; Eppendorf) covered with AeraSeal adhesive sealing films (Excel Scientific). The incubation temperature was  $30\text{ }^\circ\text{C}$  for all communities. The deepwell plates were shaken at 1,200 r.p.m. on Titramax shakers (Heidolph). To minimize evaporation, the plates were incubated inside custom-built acrylic boxes.

### Pre-cultures, daily dilutions, dispersal, and biomass measurements

Before each experiment, pre-cultures were initiated by thawing the bacteria and inoculating individual species into  $600\text{ }\mu\text{L}$  of BM. The resulting monocultures were exposed to 5 daily cycles of growth and (30-fold) dilution into fresh media. At the beginning of each experiment, aliquots of these monocultures were mixed in equal volume proportions to form the synthetic communities. During the experiment, the monocultures were exposed to further dilution cycles and used to apply the daily dispersal into the synthetic communities as described below.

We created 63 different synthetic communities using randomly generated subsets of the library of isolates, each subset constituting the species pool (of size  $S$ ) for each community. After mixing monocultures in equal volumes, each experimental community was initiated by inoculating  $20\text{ }\mu\text{L}$  of its initial mix of isolates into  $600\text{ }\mu\text{L}$  of BM, repeating the process to generate a total of 3 biological replicates per community. To form the communities with  $S \leq 12$ , we created random subsets of species from the soil isolates, and for  $S > 12$  we randomly matched both soil and *C. elegans* isolates. To form synthetic communities with  $3 \leq S \leq 12$ , we distributed the 24 soil isolates into 8 groups (group A-H), where each group included 3 randomly chosen, different isolates. We used these groups to form 3-species communities ( $S=3$ ). To form eight 6-species communities ( $S=6$ ), we combined the eight pairs of 3-species groups (A-H) in the following way: (A+B), (C+D), (E+F), (G+H), (A+H), (B+G), (C+F), (D+E). To form eight 12-species communities ( $S=12$ ), we

combined the eight pairs of 3-species groups (A-H) in the following way: (A+B+C+D), (E+F+G+H), (A+H+B+G), (C+F+D+E), (A+B+E+F), (C+D+G+H), (A+H+C+F), (B+G+D+E).

To form synthetic communities with  $S=24$ , we randomly distributed the 24 soil isolates into 4 groups (group A-D) and the 24 *C. elegans* isolates into another 4 groups (group E-H), each of these groups including 6 different isolates. To form eight 24-species communities ( $S=24$ ), we combined the eight pairs of 6-species groups (A-H) in the following way: (A+B+C+D), (E+F+G+H), (A+H+B+G), (C+F+D+E), (A+B+E+F), (C+D+G+H), (A+H+C+F), (B+G+D+E). We used the whole species pool (24 soil isolates and 24 *C. elegans* isolates) to form the single 48-species community ( $S=48$ ).

The resulting synthetic communities were cultured under serial dilution cycles with dispersal as follows. To apply a  $10^{-6}$  dispersal rate, every 24hr monoculture aliquots of the species in each community pool were mixed at equal volumes, and then diluted by a  $10^4$  factor before inoculating 6 $\mu$ L of this mix into the wells containing the corresponding experimental community matching each species pool. After this, the experimental cultures were thoroughly mixed using a 96-well pipettor (Viaflo 96, Integra Biosciences; settings: pipette/mix program, 5 mixing cycles, mixing volume 300  $\mu$ L, speed 6) before applying a 30-fold dilution by transferring 20  $\mu$ L of the cultures into a new plate with 600  $\mu$ L of fresh media.

Experiments were extended to a total of 10 daily cycles. At the end of every daily cycle, 150 $\mu$ L samples of each culture were used to measure the OD (600nm), a proxy for the total biomass in the cultures, using a Varioskan Flash (Thermo Fisher Scientific) plate reader. The remaining culture volume was stored at -80 °C for subsequent DNA extraction.

We tested the reproducibility of community dynamics under different choices of carbon sources. We replaced 2% glucose by 2% succinate in the media with high nutrients concentration, and still observed biomass fluctuations in some communities (Fig. S27). Among the eight different communities, two communities fluctuate in both glucose and succinate, and four communities reach steady state in both glucose and succinate. There are two communities only that fluctuate in medium with glucose while reaching a steady state in medium with succinate. The results show that emergent fluctuations in communities are reproducible with different carbon sources. Furthermore, we cultured the 12-species communities under high nutrients concentration and diluted every 48 hours. We found that both 24-hours-transfer and 48-hours-transfer regimes yield analogous biomass dynamics for the eight 12-species communities. The communities that reach fluctuating (stable) states with 24-hour-transfers also reach a similar fluctuating (stable) state with 48-hour-transfers (Fig. S27). These results demonstrate that the observed community dynamics are robust to different choices of dilution time regimes. To classify stable and fluctuating communities under succinate and 48-hours-transfer, we calculated the standard deviation of biomass over the last four data points (day 7-10 for communities in succinate, and day 10-16 for communities with 48h transfers). The standard deviation of biomass over time is shown in Fig. S27, where the classification of fluctuating (orange points) and stable (purple points) communities are based on the  $K$ -mean clustering method. The Std of biomass exhibits a sharp decrease from fluctuating to stable communities, and the classification remains the same across the three replicate communities under each condition (Fig. S27).

### DNA extraction, 16S rRNA sequencing and data analysis.

To monitor the dynamics of the microbial communities, we measured community composition via 16S ribosomal RNA (rRNA) amplicon sequencing. DNA extraction was performed with the QIAGEN DNeasy PowerSoil HTP 96 Kit following the protocol provided by the manufacturer. The obtained DNA was used for 16S (V4 region) amplicon sequencing. Library preparation and Illumina MiSeq sequencing were performed by the Environmental Sample Preparation and Sequencing Facility at Argonne National Laboratory. We used the R package DADA2 to obtain the amplicon sequence variants (ASVs) as described by Callahan *et al.*(39). Taxonomic identities were assigned to the ASVs by using SILVA (version 132) as a reference database. For each sample, species richness was calculated as the number of ASVs with a relative abundance  $\geq 0.1\%$ , which corresponds to the 0.1% extinction threshold used in simulation (Fig. S1, S2). The phylogenetic tree (Fig. S14) was constructed using Simple Phylogeny (40) by the EMBL's European Bioinformatics Institute. Taxonomic identities were assigned to ASVs using Randomized Axelerated Maximum Likelihood (RAxML) using default parameters. All plot of relative ASV abundances with stack bars in this paper show the results of one replicate.

### Accuracy and limitations of 16s RNA amplicon sequencing data

Each step in the workflow to assess community composition through amplicon sequencing presents its own biases (41). This includes taxonomical biases in DNA extraction, PCR amplification (e.g., differences in 16S gene copy number), sequencing, and bioinformatics processing. In our study, such biases can significantly compromise the quantitative accuracy of the reported relative abundance of community members, although they are unlikely to significantly compromise our results qualitatively. In particular, these biases could lead to underestimations in community diversity if species fall below the extinction threshold as a result. Similarly, quantitative measures of abundance fluctuations could also be affected. However, these quantitative changes (e.g., under- or overestimation of specific community member) should be comparable across samples, making our qualitative results (e.g., transitions between dynamical modes in a specific order across the phase space) robust to these biases.

In our sequencing dataset, sequencing depth varied from 3579 to 67354 reads, with an average of 21609 reads. This means that we could not effectively resolve any species abundance on the order of .01% or below. Our main observables, diversity and fluctuation fraction, were calculated (Methods) only from species abundances that exceed a threshold of 0.1% (the extinction threshold). On the one hand, we were able to detect abundances for all the members of each species pool in all the data points for Day 0 (Fig. 2 and S20-S22). Considering that community inoculation consisted in mixing monocultures at equal volumes at Day 0, this suggests that species-specific, sequencing-associated errors are relatively modest in our dataset. On the other hand, our communities are composed of members of a defined set of 48-species library. We did not detect any reads from any ASV that does not correspond to a member of the 48-species library, which suggests an absence of significant contamination during experimental data acquisition and processing and is an additional indicator of reliability of the sequencing data for the purposes of this study.

## Pairwise interactions between microbial species

To measure the strength of pairwise interactions in the experiments, we randomly chose six isolates—selected from the different genera Leuconostoc, Pseudomonas, Yersinia, Pantoea, Klebsiella, and Acinetobacter—from the bacterial library. We first measured the carrying capacity  $K_i$  (*i.e.*, the species abundance at equilibrium in the absence of any competitor species) of these isolates through exposing them to 7 cycles of daily dilutions in monoculture, followed by plating and colony counting at the end of the 7<sup>th</sup> cycle. We then co-cultured 15 pairs (all possible combinations) of these isolates over 7 dilution cycles and measured the species abundance at the end point (via sample dilution and colony counting on agar plates). Together with the measured carrying capacities, these species abundances were used to assess the strength of interactions via the relationship  $\alpha_{ij} = \frac{(K_i - N_i)K_j}{N_j K_i}$ , which can be easily derived from the gLV model. Table S1 shows that all pairs of isolates coexist under low nutrient concentrations ( $\alpha_{ij} < 1$  for all the experimentally measured interactions). For higher nutrient concentrations, we considered 2 initial relative abundance for each pair of species (initial species ratios 95:5 and 5:95, measured via culture volume), which allows to identify cases of bistability in which either species can lead its competitor to extinction. For coexisting pairs, the value of  $\alpha_{ij}$  was calculated as stated above. In cases of competitive exclusion (species  $i$  always drives species  $j$  to extinction), we inferred that  $\alpha_{ij} < 1$  and  $\alpha_{ji} > 1$ . For bistability (the high-abundance species drives the low-abundance one to extinction), we inferred that  $\alpha_{ij} > 1$  and  $\alpha_{ji} > 1$ . Tables S2 and S3 show the measured interaction matrices under medium and high nutrients concentrations. We found the interaction matrices measured in the experiment are densely connected matrices (Tables S1-S3), which means that  $\alpha_{ij} \neq 0$  for most (or all) the species interactions. This result is consistent with previously observed microbial community interaction networks (7) and supports the assumption a dense interaction matrix in our theoretical model.

Ecological communities, including microbial communities different from the ones in our experiments, need not to be densely connected. Our model can account for this fact through incorporating the average connectance  $C$  (the fraction of non-zero interactions in the interaction network) in the stability criteria (16). In this way, the main effect of network connectance in community dynamics is equivalent to replacing  $S$  by  $SC$  on the horizontal axis of the phase diagrams in Figs. 1E and 1F (17). Therefore, a finite fraction of zero (or negligibly small) interaction strengths does not qualitatively change the phases of community dynamics and their relative positions on the phase diagram.

## Numerical methods

We modeled the long-term dynamics and diversity of ecological communities using the well-known generalized Lotka-Volterra (gLV) model, modified to include dispersal from a species pool:

$$\frac{dN_i}{dt} = r_i N_i \left( 1 - \sum_{j=1}^s \alpha_{ij} N_j / K_i \right) + D \quad (1)$$

where  $N_i$  is the abundance of species  $i$  (normalized to its carrying capacity),  $\alpha_{ij}$  is the interaction strength that captures how strongly species  $j$  inhibits the growth of species  $i$  (with self-regulation

$\alpha_{ii} = 1$ ), and  $D$  is the dispersal rate from an outside species pool to the focal community. For simplicity and without qualitatively changing our results, we considered the same growth rate  $r_i = 1$  and the same carrying capacity  $K_i = 1$  for all species in the main text. Fig. S6 shows that sampling growth rates from a uniform distribution has little effect on the phase diagram of survival fraction and fluctuation fraction. Fig. S6 shows that sampling carrying capacities from a normal distribution increases the partial coexistence phase while shrinking both the full coexistence phase and fluctuation phase, but does not affect the order of the phases. We further tested the theoretical predictions when considering the existence of positive (facilitative) interspecies interactions (Fig. S9) and varying the symmetry of the interaction matrix (Fig. S8). We also considered different dispersal rates (Fig. S7), and the effects of incorporating daily dilutions (Fig. S9) in these *in silico* communities. These additional results show that our qualitative phase diagrams and conclusions are robust to different choices of ecological network structure and parameters. Although the patterns of ecological diversity and dynamics do not change as the dispersal rate varies from  $D=10^{-7}$  to  $D=10^{-6}$  (Fig. S7), we found that communities with zero dispersal rate exhibit lower fluctuation fraction and survival fraction in the persistent fluctuation phase. Our results show that non-zero dispersal rates can sustain persistent fluctuations.

To show that symmetry and anti-symmetry do not qualitatively change the phase diagram, we simulated communities in two scenarios of non-zero reciprocity,  $\gamma = 0.5$  and  $\gamma = -0.5$ , where the reciprocity of interactions is given by  $\gamma = \text{corr}(\alpha_{ij}, \alpha_{ji})$ . The qualitative patterns and order of transitions in the phase diagram (Fig. S8) are robust to the presence of reciprocity, although  $\gamma$  shifts the stability boundary (solid line in Fig. S8). Positive and negative reciprocity decreases and increases the values of  $S$  and  $\langle \alpha_{ij} \rangle$  at which communities lose stability, respectively. Moreover, positive (negative) reciprocity yields lower (higher) survival fraction and fluctuation fraction of communities. At full symmetry and anti-symmetry ( $\gamma = 1, -1$ ) there is no fluctuating phase (21), but those do not appear to be relevant from the experiment pair-competition results (Table. S1-S3).

To test whether the existence of positive (facilitative) interactions in the ecological network will change our conclusions, we sampled values of  $\alpha_{ij}$  from a uniform distribution  $[-\alpha_0, \alpha_0]$ , where  $\alpha_0$  varies between  $[0, 1.4]$  on the phase diagram. In this simulation, the linear interaction function in the gLV ( $\alpha_{ij}N_j$ ) is replaced with Monod function ( $\alpha_{ij}N_j/(N_j + 1)$ ) to avoid unbounded growth due to positive interactions (21, 42). We observed similar patterns of survival fraction and fluctuation fraction between pure competitive interactions and considering positive interactions (Fig. S9). The first and second moment of the distribution of  $\alpha_{ij}$  should be considered to quantify interaction strength in the stability criteria (17). Here we use  $\text{Std}(\alpha_{ij})$  to quantify the interaction strength because the first moment of  $\alpha_{ij}$  distribution is zero. Our results demonstrate that the existence of three phases (full coexistence, partial coexistence, persistence fluctuation) and the order of transitions are robust to the interaction types in the model.

All simulations used the Runge-Kutta method on Matlab to numerically solve the LV equations (with an integration step of 0.05). A definition of  $100 \times 100$  pixels was used for each phase diagram, linearly segmenting the parameter space in the ranges  $\langle \alpha_{ij} \rangle \in [0, 1.5]$  and  $S \in [1, 100]$ . In each phase diagram, each pixel shows the average result for  $10^3$  simulations. The total

simulation time is  $10^4$  to guarantee the survival fraction and fluctuation fraction have reached steady states as shown in Fig. S5.

To test whether the total biomass fluctuation is consistent with species abundance fluctuation in our simulations, we simulated the time series of community biomass under various conditions (Fig. S3). To quantify the dynamics of biomass in the simulation, we calculated the sum of species abundance ( $\sum_i N_i(t)$ ). Our results demonstrate the fluctuation in species abundance is in agreement with fluctuation in total biomass.

The similar fluctuations between replicates in some experiments (*e.g.*, Fig. 2, medium nutrients,  $S=48$ ) could be explained by the slow divergence that chaotic trajectories can exhibit during moderately long-time windows, or, alternatively, by possible limit cycle oscillations (Fig. S3). We focused on chaotic fluctuations when discussing the model predictions because previous theory shows that all persistent fluctuations will be chaotic as number of species in the pool  $S$  grows (21), though limit cycle oscillations only happen under finite  $S$ .

### Reaching steady state in simulations

We define the steady state of simulated communities as the community state in which neither the survival fraction nor the fluctuation fraction significantly changes as time goes on. In order to consistently analyze the steady state results for all the simulated communities, we analyzed the dependence of the phase diagrams on the simulated time. Fig. S5 shows that neither the survival fraction nor the fluctuation fraction significantly changes after  $t=5\times 10^3$ . Accordingly, the phase diagrams in the paper show the state of communities at  $t=10^4$ , unless otherwise stated.

### Extinction threshold *in silico*

The presence of dispersal from the species pool in Eq. (1) guarantees that all species exhibit strictly positive abundances in Fig 1B and C. Nevertheless, we consider that a species is extinct if its abundance lays below a  $10^{-3}$  threshold. Around this threshold, the dispersal rate becomes the main factor preventing abundance decay (Fig. S1). The species abundance distribution in the partial coexistence phase is bimodal (Fig. S2); the extinction threshold 0.001 clearly separates the high-abundance surviving species from low-abundance species that will go extinct if dispersal ceases (Fig.S1).

### Definition of stable and fluctuating dynamics *in silico*

To differentiate between stable and fluctuating communities, we computed the average coefficient of variation of  $N_i$  between  $t=5\times 10^3$  and  $t=10^4$ . We define communities with this average coefficient of variation higher (lower) than  $10^{-3}$  as fluctuating (stable) communities (Fig. S1).

### Survival fraction *in silico*

To compute the survival fraction, we computed the fraction of species whose abundance exceeded the extinction threshold at any time during the last 100 units of time in the simulation. Our choice of including a time window when measuring diversity is motivated by the fact that, for the case of unstable communities, species abundances fluctuate above and below the extinction threshold over

time. Since we measured diversity and species compositions every 24 hours in the experiment, we consider an analogous window of 100 time units in simulations.

### Analytical curves for boundaries between phases, and sharpness of the transitions

Starting with the pioneering work by Robert May (16), ecologists have sought to predict community behaviors using coarse-grained parameters including the number of species and the first two moments of the distribution of interaction strengths between species. The analytical boundary between the stable phase (II) and the fluctuating phase (III) was derived in Bunin 2017 (21). For equal carrying capacities, it is shown that the boundary lies at the average standing species richness  $S^* = S/2$ , when  $\sigma \equiv \sqrt{S} \text{std}(\alpha_{ij})/(1 - \langle \alpha_{ij} \rangle) = \sqrt{2}$ . There and in (43) it is also shown that the loss of stability of the equilibrium coincides with real parts of some community matrix ( $-\alpha_{ij}$ ) eigenvalues becoming positive. For any distribution of interaction strengths (uniform, exponential, *etc.*),  $\text{std}(\alpha_{ij})$  and  $\langle \alpha_{ij} \rangle$  can be calculated, and this criterion applied. The boundary between the fully-coexisting (I) and stable (II) phases is given by taking the prediction for the average standing species richness  $S^*$  given in (21), and setting  $S^* = S - 1$ , namely the parameters when one species has gone extinct. This is not expected to be exact at large  $S$ , since the prediction in (21) is exact in conditions where  $S^*/S$  is finite at large  $S$ , while here the boundary is at  $S^*/S = 1 - 1/S$  which approaches 1 at large  $S$ . Nonetheless, it gives good results in the range of pool diversities shown in Fig. 1. Other techniques for analyzing the transition are also possible (44).

*Sharpness of transitions* – The term “phase transition” in physics implies that the transition is sharp in systems with many degrees of freedom. The transition between phases (II) and (III) is known to be sharp (21, 22): at high  $S$ , communities that lie above the phase boundary (e.g.,  $\sigma > \sqrt{2}$  with all species exhibiting equal carrying capacities) always (with probability one) exhibit persistent fluctuations, while communities below the phase boundary (e.g., when  $\sigma < \sqrt{2}$  for equal carrying capacities) reach a stable equilibrium .

The transition between phases (I) and (II) is also sharp when  $S$  is large, in the following sense. Fig. S30 shows the probability of full coexistence as a function of  $\langle \alpha_{ij} \rangle$ , for different values of  $S$ . The x-axis is normalized by  $\langle \alpha_{ij} \rangle$  where the analytical boundary is expected. This makes all curves decrease to zero in the same region, but the width of the crossover regime becomes narrower with increasing  $S$ . In other words, the width of the crossover region between the phases is small compared to the width of phase (I), for large  $S$ .

The fact that all curves decrease to zero at the same region, shows that the analytical expression indeed captures the correct dependence of the boundary in  $\langle \alpha_{ij} \rangle$  on  $S$ . The fact that the crossover happens around a value of  $\sim 2.1$  rather than around 1, is due to the inexact theory used, as explained above.

An additional dynamical regime known as a Gardner phase has been theoretically proposed in a model with symmetric interactions (45). The Gardner phase is a regime where there are many deep basins of attraction, and within each one there is a further structure of many close to marginally-stable basins. Altieri et al. discussed what happens when the full symmetry of the interactions is

broken, which seems to be the experimentally relevant situation (Table. S1-S3). In that case, the authors remark that the internal marginally-stable structure is sensitive, and easily washed out by the asymmetry. Although the deep basins might still survive when the interactions are asymmetrical, the existence of these deep basins seems to require a pool size  $S$  that is larger than we have, with multiple stable states only found for  $S >\sim 100$ . Indeed, the simulations in the paper (45) were done for  $S=500, 2000$ . This is an interesting direction for future research, but likely not relevant to our present setting, where  $S \sim 50$ .

### Theoretical alternatives to the Lotka-Volterra model

The Lotka-Volterra model provides a phenomenological representation of bacterial growth and interactions, but our conclusions are not tied to this specific model. It is important to understand whether our central theoretical predictions generalize: how broadly do we expect a similar qualitative map of dynamical phases, where extinctions start to occur before the onset of fluctuations (as we increase either species pool size or interaction strength)? This empirically-observed ordering of phases is not self-evident: it is straightforward to construct few-species models that display fluctuations without extinctions, e.g. predator-prey pairs. Therefore, if our qualitative phase ordering appears across a range of many-species models and experiments, we may be seeing a broad mechanism, one that is presumably collective rather than driven by particular species.

One interesting question is whether these collective dynamics emerge from many independent pairwise species interactions (as in our random Lotka-Volterra model), or whether they are driven by one or a few system-wide factors, such as public goods impacting all species. Ratzke et al. (14) performed bacterial experiments and introduced a different model where all interactions are mediated by modifications of the environmental pH by the bacteria, whose growth is in turn modified by pH. This pH-based model, and a variant, reproduced some experimental results in (14, 33). We slightly amend the model in (14) to represent continuous-time dilution and dispersal from the species pool (we have also simulated discrete daily dilutions and dispersal, with no impact on our conclusions).

$$\frac{dN_i}{dt} = k_{growth} N_i (1 - N_i) f(p - p_{0i}) - c_{dilution} N_i + D \quad (2)$$

$$\text{where } f(x) = \begin{cases} 1 & \text{if } x \in [-p_c, p_c] \\ -1 & \text{otherwise} \end{cases}$$

$$\frac{dp}{dt} = \sum_i c_i N_i + c_{dilution} (7 - p) \quad (3)$$

Equation (2) represents the growth of bacterial abundance  $N_i$ , which follows a logistic equation with parameter  $k_{growth} = 10$  modulated by the pH value  $p$ : growth is maximal when this value is equal to species  $i$ 's pH optimum  $p_{0i}$ , (drawn uniformly over [4.5,9.5]), with a tolerance given by  $p_c = 2$ . In addition, parameter  $c_{dilution} = 3.4$  (continuous-time equivalent to our 1:30 daily dilution) encompasses losses due to dilution, whereas  $D = 10^{-6}$  represents dispersal from the species pool. Equation (3) represents the change of pH induced by the bacteria, and the return toward the neutral pH value of 7 due to dilution. The distribution of  $c_i$ , drawn uniformly over  $[-c,$

$c]$ , thus determines the impact of bacteria on pH, and indirectly, the strength of interactions between bacteria.

We find that this model reproduces some of the main predictions of the Lotka-Volterra model: as seen in Fig. S10, the phase diagram with its three phases is qualitatively preserved in this second model, which suggests that it is highly robust to variations in modelling assumptions. Nevertheless, we believe that the LV model better reproduces our experimental results, whereas the pH-based model was more adequate in (14, 33), plausibly due to different taxa (species that exhibit ecological suicide do not appear in our species pool here) and experimental conditions (lower dilution rate) that lead to stronger impacts of pH. Indeed, there are three points that make the LV model more plausible here:

- 1) In our experiments, pH fluctuations are only moderately correlated (Pearson correlation coefficient: 0.54) with fluctuations in optical density or species composition (whereas the latter two are highly correlated), suggesting that pH is not the sole driving factor of stability here. (Fig S11A)
- 2) The pH-based models in (14, 33) notably aimed to allow for ecological suicide, i.e. species growing then going extinct in monoculture, which does not occur in our experiment (In our initial species pool, we only retained species that survived in all of our growth conditions for the convenience of daily dispersal, Fig. S26). In the high-interaction strength regime and large initial species pool size regime, the pH-based model has two main outcomes: either many-species fluctuations, or total or near-total extinction with zero or few species surviving, and total biomass < 5% of carrying capacity (Fig S11B). Only the former behavior is seen in our experiments; the latter prediction of pH-based model was never observed in our experiments (Fig. 2C and Fig. S26).
- 3) The LV model displays positive correlation between the number of surviving species and the intensity of fluctuations, as in our experiments (Fig. 4 and Fig. S11), whereas the pH-based model displays a negative correlation (we exclude communities falling under the above extinction criterion, which were never observed in our experiment, Fig. S11C). The positive relation between diversity and fluctuations, validated in our experiments, is key to our theoretical argument.

These three reasons lead us to retain the Lotka-Volterra model as our prime example (see Fig. S12 for a best-fit illustration). This cannot rule out a different pH-based model, or a combination of pH-based and direct interactions, as a good description of the biological mechanisms at play in our experimental setting.

We reach two main conclusions:

- i) The random Lotka-Volterra model provides a straightforward prototype of a mechanism that fits our theoretical argument. The dynamical phases are unambiguously *emergent*, in the sense of being driven by species diversity and local pairwise interactions. The picture is as follows: the fluctuating phase must appear after extinctions and turnover in species composition, since fluctuations are driven by a tendency to jump between alternate sets of surviving species, each of them unstable.

ii) We expect this central prediction to hold broadly under different modelling choices. Indeed, we find that the pH-based model displays an analogous set of regimes and transitions (Fig. S10), even though its predictions do not match some of our experimental results as accurately as the LV model (Fig. S11).

In both models, increasing the number of species and the strength of their interactions, whether direct or mediated through an abiotic factor, will typically lead to extinctions before it leads to loss of stability. This results in a phase of partial coexistence preceding a phase of instability, no matter whether species diversity and extinctions are necessary (as in the random LV model) or not (as in the pH-based model) for the mechanism that drives fluctuations. We propose that this ordering, which is robust in many-species models but need not be in few-species models, can be an indicator of emergent collective behavior.

#### Definition of stable and fluctuating experimental communities

To differentiate between stable and fluctuating communities in experiments, we computed the average coefficient of variation (CV) for species abundances from day 7 to day 10. This corresponds to the average value of the standard deviation for the absolute abundance of each species  $N_i$  (over day 7, day 8, day 9 and day 10) scaled by average species abundance. Communities for which the average coefficient of variation in this time range is below (above) a 0.25 threshold are considered stable (fluctuating) communities (Fig. S15A). We also find that the biomass and species compositions fluctuate asynchronously across replicates of fluctuating communities starting from the same initial conditions. The significant differences in relative abundances ( $N_i^* = N_i / \sum N_i$ ) across replicates is an additional indicator of community fluctuations. Fig. S13B shows that the two clusters identifying different community dynamics occur independently of whether differences in absolute abundance or relative abundance are used to assess the (temporal) coefficient of variation. Among the two methods, calculating the temporal variability via relative abundances yields a larger variability across samples, which leads to quantitative, but not qualitative, differences in the results. The cluster of communities' dynamics is consistent between the metric by average coefficient of variation and by final relative composition difference across replicates, as shown in Fig. S15B. To further show the variability of relative abundances over time, we've calculated the variability of relative abundances over time and discussed the results in Fig. S15. The results show that fluctuating communities exhibit larger variability of relative abundances between replicates and over time, which is consistent with the results considering absolute species abundance (Fig. S15B). The average coefficient of variation (Fig. S15, S16) for species abundances was calculated based on only replicate for which we sequenced the whole time series, and the average difference in relative species abundances community across the three replicates for each community (Fig. S15, S16) was calculated based on relative abundances at day 10.

We tested whether the classification of community dynamics as stable or fluctuating is robust to the choice of time window (day 7 to day 10) for calculating the average CV of species abundances. To do so, we calculated the average CV at different time points of the species abundance time series, with a fixed-length time window (4 days). We moved the first day of the time window from day 0 to day 7 and calculated the corresponding average CV of communities in each case. Fig. S16

shows that the average CV for both fluctuating and stable communities reaches steady state before the last time window (from day7 to day10). We chose the 12-species communities under high nutrients concentration for this analysis because they exhibit equal numbers of stable and fluctuating communities ( $n=4$  for both stable and fluctuating communities). These results show that the average CV calculated in the last time window (from day7 to day10) converges rapidly to either small or large values, respectively indicating stability or long-lasting fluctuations in experimental communities.

To demonstrate that our classification of stable and fluctuating communities is robust to other classification methods, we applied a  $K$ -means clustering algorithm considering both average coefficient of variation ( $\langle \text{Std of } N_i(t) \rangle / \langle \text{Mean of } N_i(t) \rangle$ , where  $t$  runs from day7 to day10 to calculate the Std and Mean of  $N_i(t)$  for the replicate with the 16S-sequenced time series; the brackets average across all species in the community) and relative differences in species abundances across replicates (Euclidean distance of  $N_i$  over the 3 replicates at day 10) to classify the dynamics of our communities. Fig. S16 shows the results of this classification method, which are almost identical to those obtained through the stability criteria based on a CV threshold (Fig. S15). There is only one community which is differently classified by each method (open circle in purple in Fig. S16). Classifying this individual community as either stable or fluctuating does not change the three phases in the experimental phase diagram and the order of phase transitions (losing species before losing stability). Therefore, our conclusions are robust to different choices of stability criteria. Since the  $K$ -means clustering algorithm does not require set any threshold of CV, the consistence between results of  $K$ -means clustering and setting stability threshold of CV (0.25) demonstrates the classification of fluctuating and stable communities is robust to different algorithm

Although none of our conclusions on dynamical phases and transitions depends on the classification of the only community that is differently classified by each stability criteria (purple open circle in Fig. S16), we further analyzed both the biomass and sequencing data for the three replicates of this community. Fig. S25 shows that this specific community ( $S=12$ , Community 5) exhibits moderate differences in species abundances at day 10 across the three replicates. The replicate community for which we sequenced the time series also moderate fluctuations in both species' abundances (Fig. S22,  $S=12$ , Community 5) and biomass (Fig. S19, blue curves,  $S=12$ , replicate 1). This places this community at the boundary of fluctuation and stability regardless of the classification method (Figs. 15 and S16). Given that the biomass time series of the other two replicates of this community (Fig. S19, blue curves,  $S=12$ , replicates 2 and 3) exhibit relatively larger fluctuations in biomass than replicate 1, we classified this community as a fluctuating one (in agreement with the CV threshold criteria) in the main text.

#### Quantification and statistical test of species survival fractions

To statistically test differences in species survival fractions between fluctuating communities and stable communities in Fig. 4B, we first calculated the difference of survival fraction between each community (purple or orange points in Fig. 4B) and the corresponding average survival fraction (blue points in Fig. 4B) for each species pool size ( $S$ ). We then performed an analysis of variance (ANOVA) test on the distance to the mean survival fraction for all the fluctuating communities against all the stable communities. This proved the statistical significance on the differences

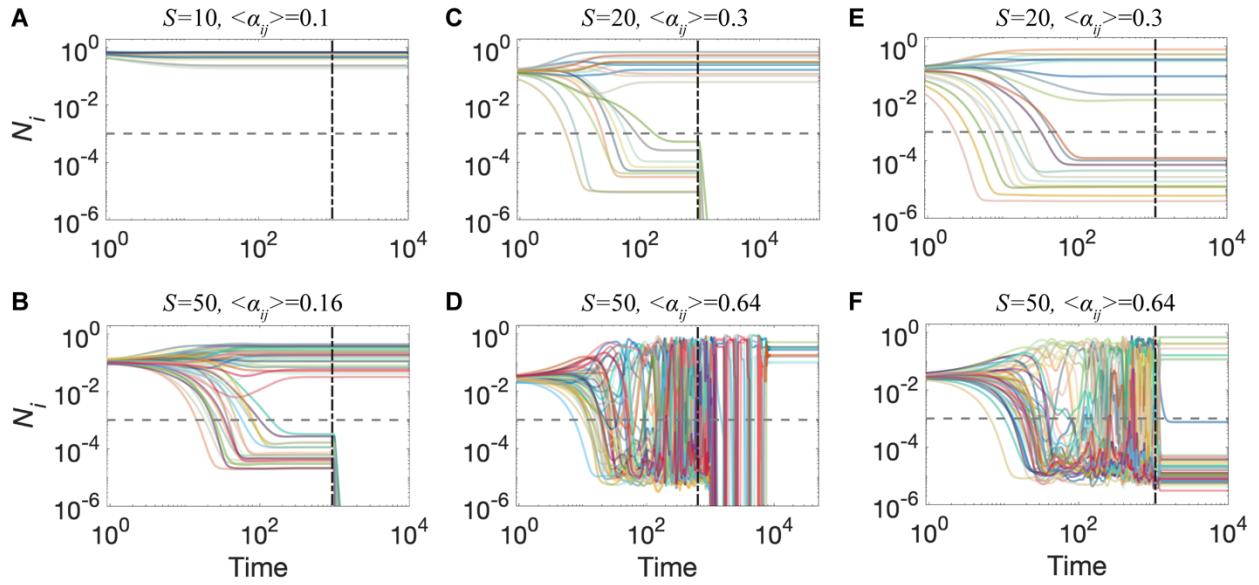
between the two groups, with the probability of observing these differences in a null model being very small ( $p < 0.01$ ).

#### Calculation of error bars

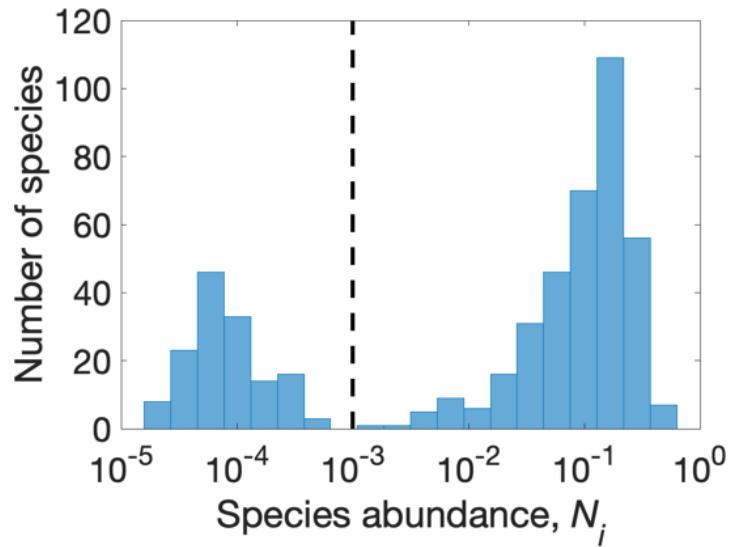
The error bars in the main text figures represent the standard error of the mean (s.e.m.). For the s.e.m. of survival fractions (Fig. 3A and Fig. 4B), we firstly calculated the mean survival fraction over each community's three replicates, and we then calculated the average and s.e.m. of these mean survival fractions across all n communities with the same pool size and nutrient conditions ( $n=30$  when  $S=2$ ;  $n=8$  when  $S=3, 6, 12, 24$ ). For the s.e.m. of fluctuation fractions (Fig. 3B), we could only use one replicate per community for which 16s rDNA sequencing was performed over the entire time series. Outcomes for each community were coded as Boolean values (1 for fluctuating communities, 0 for stable communities). We could then calculate the s.e.m. of this binomial distribution ( $n=30$  when  $S=2$ ;  $n=8$  when  $S=3, 6, 12, 24$ ). The definition of fluctuating and stable communities is given in Fig. S15.

#### Discussion on diversity-stability relationship

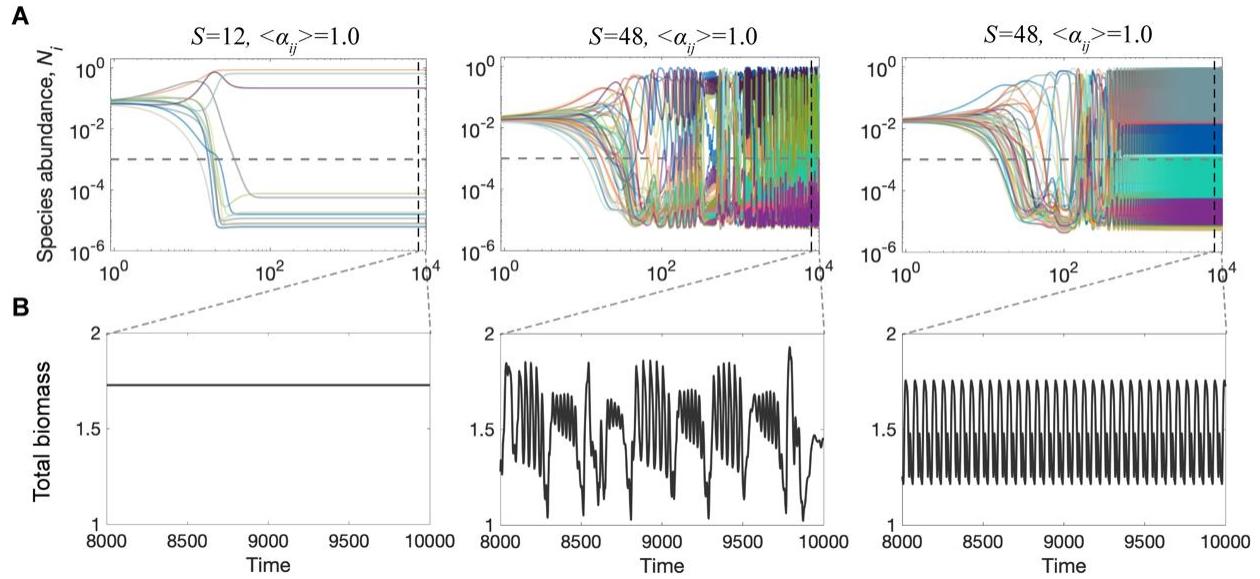
Our model and experiment show that persistent fluctuations and high realized diversity require and allow each other (Fig. 4 and S28), which is consistent with two major ideas in theoretical ecology: in one direction, May's suggestion that complexity leads to instability (16), and in the other direction, Chesson's argument that temporal fluctuations can help maintain diversity (34). The fact that two coarse-grained parameters can independently shape the phase space for community diversity and dynamics argues for caution when interpreting observed links between biodiversity and stability. On the one hand, for any given value of the interaction strength, stability negatively correlates with both size of species pool and realized diversity (number of surviving species): communities with more species are less stable (Fig 1F, 3D, 4A and 4B). On the other hand, for identical species pool sizes, stability positively correlates with diversity: weakly interacting communities exhibit relatively high stability and high realized diversity, while strongly interacting communities are relatively less stable and less diverse (Fig. 1E and 3C). We believe that the interplay between the two parameters that shape the phase diagram could underlie some of the seemingly contradictory results from field experiments (6) addressing the diversity-stability relationship.



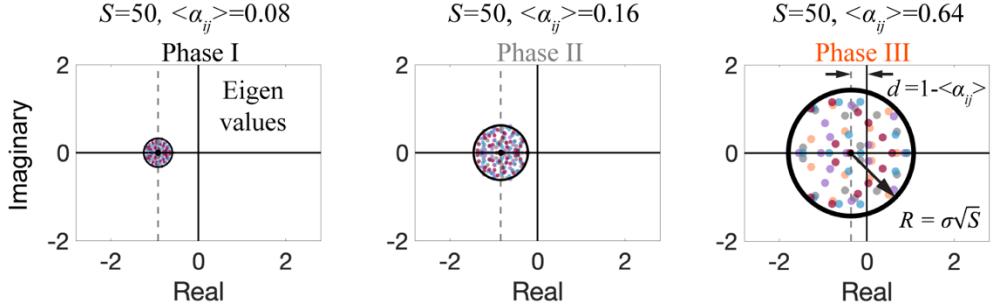
**Fig. S1. High diversity and persistent fluctuations allow and require each other, and are both sustained by dispersal.** (A)-(D) Representative time series for communities in which the dispersal rate is suddenly interrupted. At  $t=10^3$  (vertical dashed line), the dispersal rate changes from  $D=10^{-6}$  to  $D=0.0$  for the rest of the simulation. (A) Before  $t=10^3$ , a community in phase I reaches a stable state with full coexistence. The dynamics after  $t=10^3$  shows that interrupting dispersal does not significantly modify the abundances of the species. (B)-(C) Before  $t=10^3$ , communities in phase II reach an equilibrium in which species coexist at stable abundances, with some species laying below the extinction threshold. After stopping dispersal, only the species that are above the extinction threshold survive at stable abundances, and the rest undergo extinction. (D) A community in phase III exhibits persistent fluctuations while exposed to dispersal. After dispersal is interrupted, extinctions occur as species fall below the extinction threshold due to abundance fluctuations. After some time (approximately  $t=10^4$ ) species extinctions have significantly reduced diversity in the community, and the surviving species reach a stable equilibrium. For the indicated parameter values, and over  $10^3$  simulations, 90% of the simulated communities reached equilibrium after interrupting dispersal. (E-F) Representative time series for communities in which the most abundant species at  $t=10^3$  is pinned (its abundance is artificially kept constant) for the rest of the simulation. (E) For communities that have reached stability, in this case in phase II, pinning the most abundant species has no effect on community dynamics. (F) In phase III, after a fast transient following the species pinning at  $t=10^3$  (vertical dashed line), the community reaches a stable partial coexistence where some of the species lay below the extinction threshold. Out of  $10^3$  simulations, 93% of the communities reached equilibrium after pinning the most abundant species.



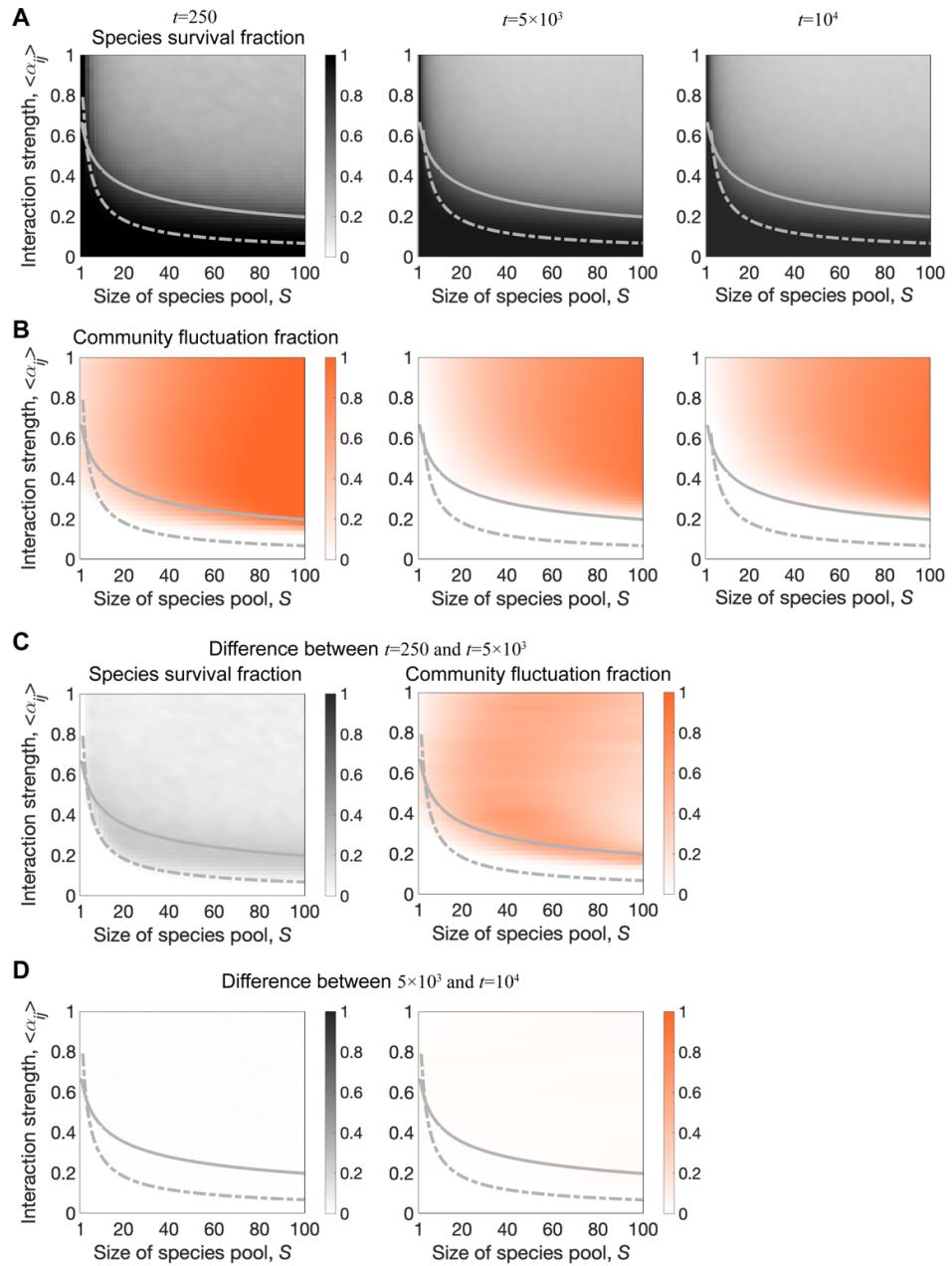
**Fig. S2. At steady state, species abundances exhibit a bimodal distribution in the partial coexistence phase.** The extinction threshold 0.001 (vertical dashed line) clearly separates the high-abundance, surviving species from the low-abundant “extinct” species. Such “extinct” species would reach zero abundance if dispersal is interrupted (Fig. S1). The histogram shows the number of species exhibiting the indicated abundances at steady state. The corresponding dataset was generated from 10 *in silico* communities randomly sampled from the stable partial coexistence phase ( $S=50$ ,  $\langle \alpha_{ij} \rangle=0.2$ ).



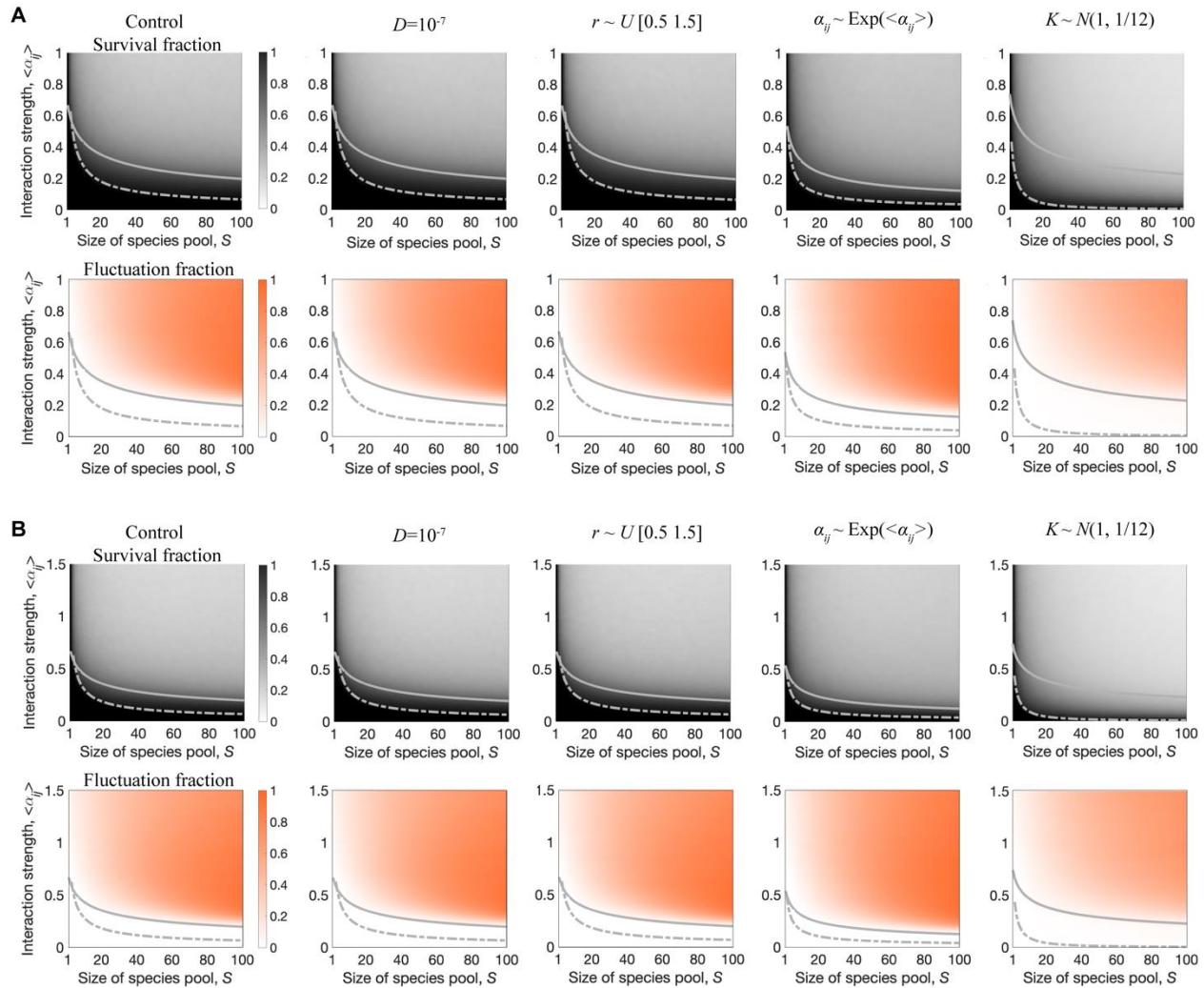
**Fig. S3. Biomass fluctuations (stability) are consistent with fluctuations (stability) of species abundances in simulations.** (A)-(B) As increasing the size of species pool in communities with strong interaction strength ( $\langle \alpha_{ij} \rangle = 1.0$ ), the species abundances and total biomass ( $\sum_i N_i(t)$ ) of communities consistently lose stability and exhibit persistent fluctuations. The species abundances and biomass of communities can also exhibit limit cycle oscillations (right panels) in addition to chaotic fluctuations (middle panels), in the persistent fluctuation phase. The biomass trajectories in (B) show the last 2000 time units in (A) as indicated by the vertical dashed lines.



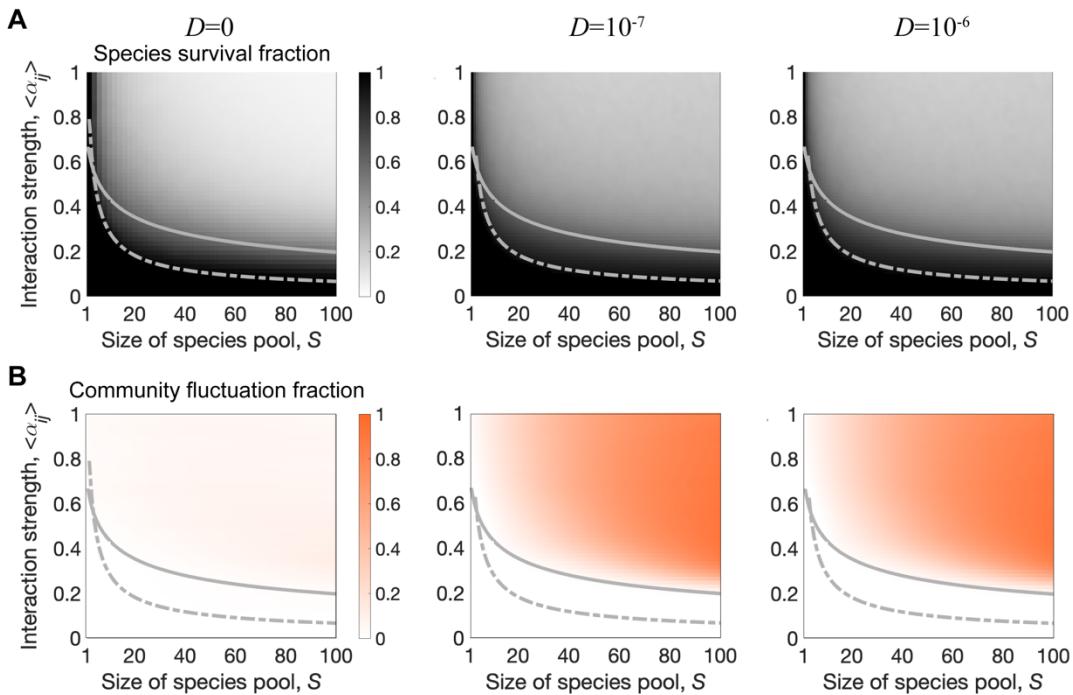
**Fig. S4. Unstable communities have (one or more) eigenvalues of the community matrix ( $-\alpha_{ij}$ ) with positive real parts.** From left to right, the panels show the eigenvalues of representative community matrices for three different values of the average interaction strength. Within each panel, different colors correspond to the eigenvalues of 4 different community matrices. All the eigenvalues lie within a circle with radius  $R$  centered at  $d$  (16, 17). For communities in phase III, where persistent fluctuations occur, some of the community matrix eigenvalues exhibit a positive real part. It was shown that the loss of stability of the equilibrium coincides with real parts of some community matrix ( $-\alpha_{ij}$ ) eigenvalues becoming positive, although it is not the Jacobian matrix (43): the circular distribution of eigenvalues for interaction matrix  $\alpha_{ij}$  is replaced by a “guitar-shaped” distribution for Jacobian matrix (46). Although the shape of eigenvalues distributions is different between interaction matrix and Jacobian matrix, the stability criterion and the signs of eigenvalues are the same for both matrices (46).



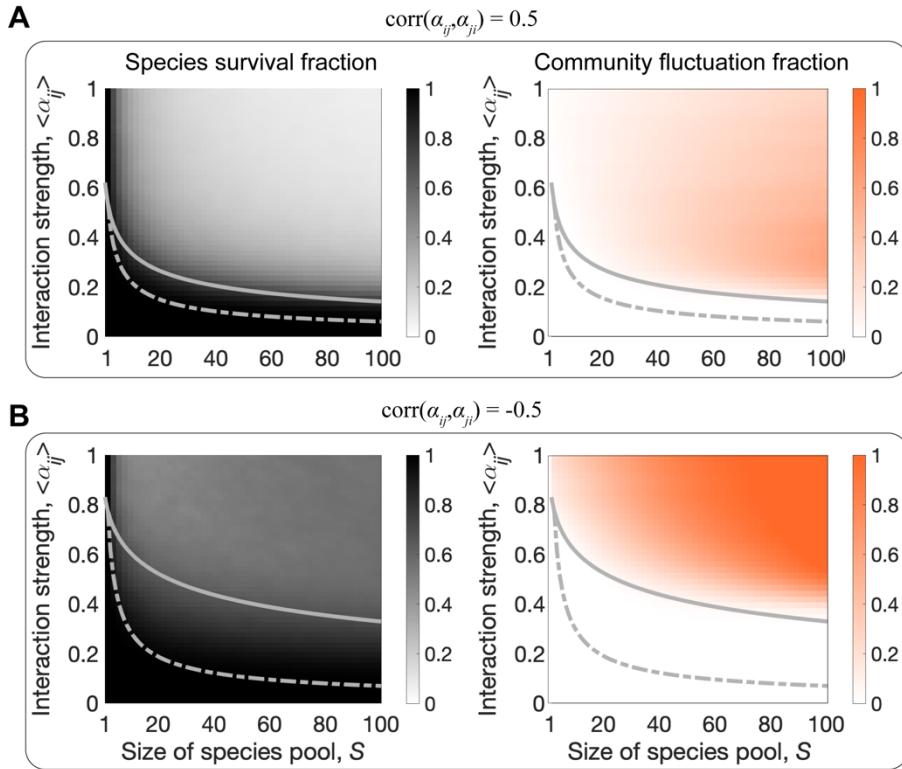
**Fig. S5. After an initial transient, the survival fraction and the fluctuation fraction of simulated communities reach stable values.** From left to right, the panels show phase diagrams of survival fraction (A) and fluctuation fraction (B) for communities at three different simulation times. At  $t=250$  (left), communities have not yet reached steady state, as the phase diagrams quantitatively change as time goes on. The middle panels ( $t=5 \times 10^3$ ) are quantitatively different from the earlier-time phase diagrams ( $t=250$  on the left), but do not significantly differ from phase diagrams computed at later times ( $t=10^4$  on the right). This shows that these two community properties have reached a steady state before  $t=10^4$ . (C) Difference between the survival fractions (left) and the fluctuation fractions (right) computed at  $t=250$  and  $t=5 \times 10^3$ . (D) Difference between the survival fractions (left) and the fluctuation fractions (right) computed at  $t=5 \times 10^3$  and  $t=10^4$ . The dashed line and solid line in the figures represent the survival boundary and stability boundary, respectively.



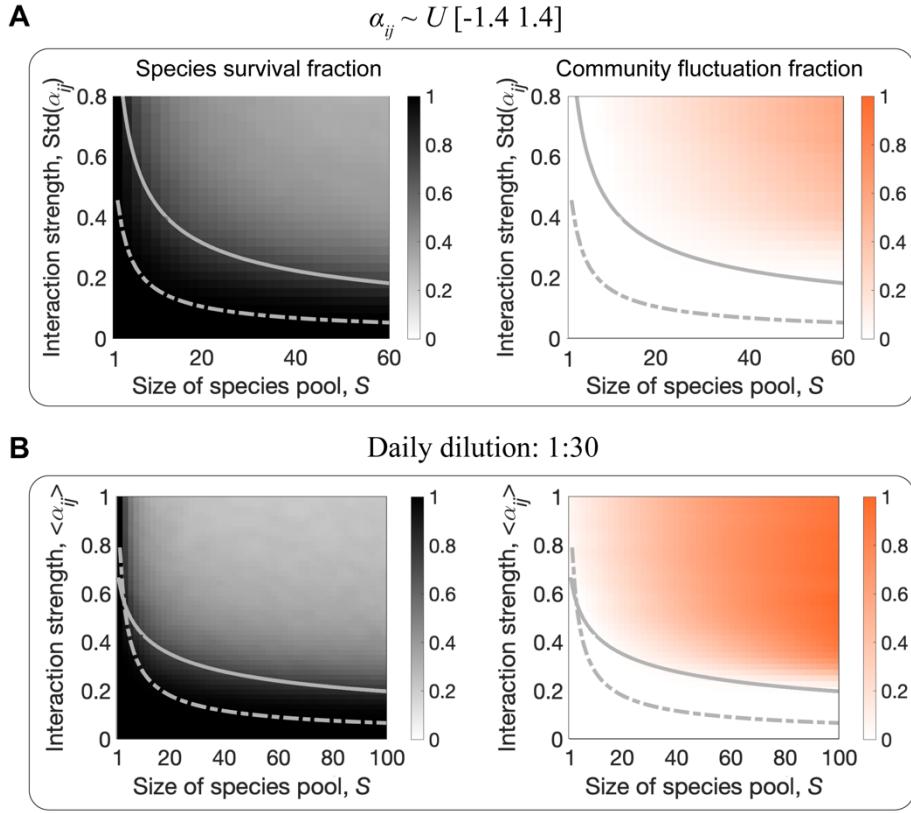
**Fig. S6. The three dynamical phases are robust to modeling choices.** (A) Panels on the left (Control) show the numerical (color map) and analytical (curves) phase diagrams as in Fig. 1E and F. From left to right, the additional phase diagrams show the effects of lowering the dispersal rate to  $D=10^{-7}$ , sampling species growth rates from a uniform distribution, sampling interaction strengths from an exponential distribution, and sampling the carrying capacities from a Gaussian distribution. All non-specified parameter values are identical to the control case (Fig. 1E and F). (B) Phase diagrams analogous to those in (A), but for a higher average interaction strength  $\langle \alpha_{ij} \rangle = 1.5$ . Overall, these phase diagrams show that the three dynamical phases are qualitatively robust to different modeling choices (21).



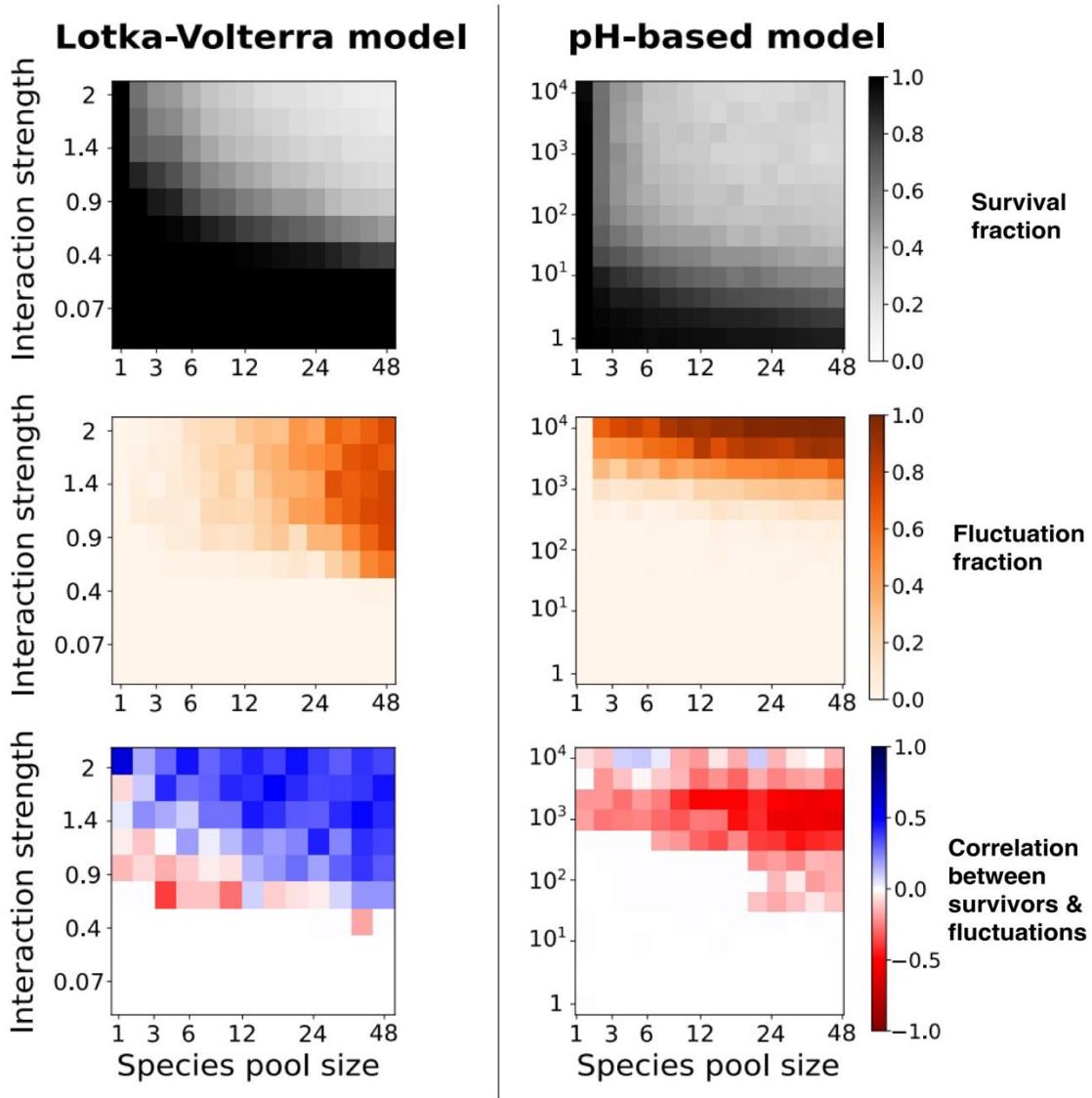
**Fig. S7. Dispersal sustains persistent fluctuations and promotes diversity.** The panels show the theoretical phase diagrams of species survival fraction and community fluctuation fraction under different dispersal rates ( $D=0$ ,  $D=10^{-7}$ ,  $D=10^{-6}$ ). Communities under no dispersal ( $D=0$ , left panels) exhibit lower survival fraction (A) and lower fluctuation fraction (B) in the persistent fluctuation phase. The patterns of ecological diversity and dynamics do not significantly change as the dispersal rate varies from  $D=10^{-7}$  (middle panels) to  $D=10^{-6}$  (right panels). The dashed line and solid line in the figures represent survival boundary and stability boundary, respectively.



**Fig. S8. The three dynamical phases are qualitatively robust to the presence of reciprocity in interspecies interactions.** The panels show the theoretical phase diagrams of species survival fraction (left) and community fluctuation fraction (right) for two cases of non-zero reciprocity,  $\gamma = \text{corr}(\alpha_{ij}, \alpha_{ji}) \neq 0$ . (A) The fluctuating phase (partial coexistence phase) is larger (smaller) in the presence of positive reciprocity ( $\text{corr}(\alpha_{ij}, \alpha_{ji})=0.5$ ) than in the absence of reciprocity ( $\text{corr}(\alpha_{ij}, \alpha_{ji})=0$ , Fig. 1E and F). The fluctuation fraction also increases with positive reciprocity. (B) The fluctuating phase (partial coexistence phase) is smaller (larger) in the presence of negative reciprocity ( $\text{corr}(\alpha_{ij}, \alpha_{ji})=-0.5$ ) than in the absence of reciprocity (Fig. 1E and F). The fluctuation fraction is higher in communities with negative reciprocity than communities with zero reciprocity. The dashed line and solid line in the figures represent survival boundary and stability boundary, respectively. Overall, the same qualitative phases and ordering are found as for communities with zero reciprocity (Fig. 1E and F), with non-zero reciprocity leading to quantitative differences.

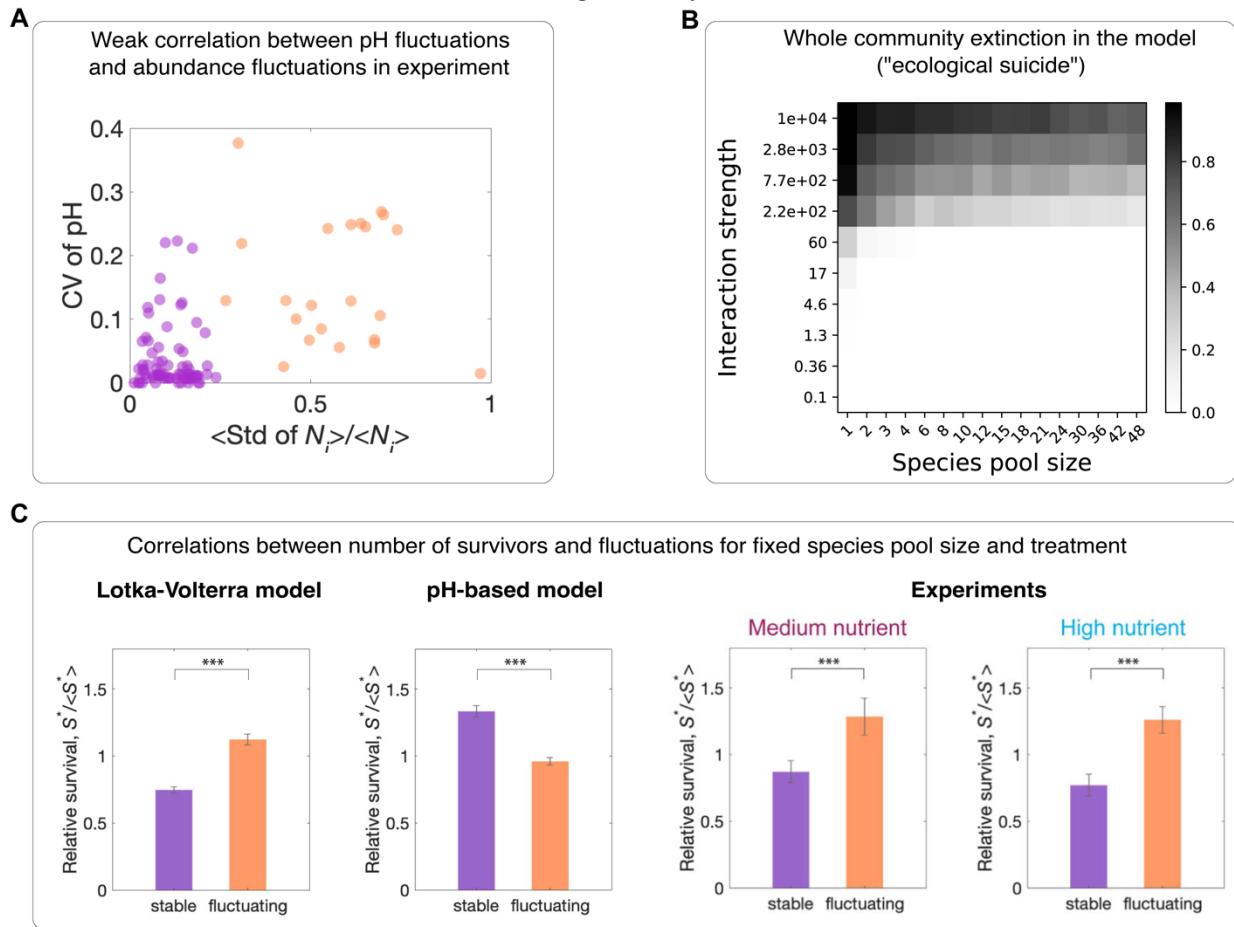


**Fig. S9. The three dynamical phases are qualitatively robust to the presence of positive interactions and serial dilutions in gLV.** (A) To test whether the existence of positive (facilitative) interactions in the ecological network could change our conclusions, we sampled values of  $\alpha_{ij}$  from a uniform distribution  $[-\alpha_0, \alpha_0]$ , where  $\alpha_0$  varies between  $[0, 1.4]$  on the phase diagram. We observed patterns of species survival fraction (left panel) and fluctuation fraction (right panel) analogous to those exhibited by communities with exclusively negative interactions (Fig. 1). The dashed line and solid line in both panels represent survival boundary and stability boundary, respectively. Note that the strength of interactions coincides with  $\text{Std}(\alpha_{ij})$  in this case, since the mean of  $\alpha_{ij}$  is zero (both moments factor into the interaction strength metric  $\text{std}(\alpha_{ij})/(1-\langle \alpha_{ij} \rangle)$  that determines stability (17)). In these simulations, the linear interaction function in the gLV ( $\alpha_{ij}N_j$ ) was replaced with Monod function ( $\alpha_{ij}N_j/(N_j + 1)$ ) to avoid unbounded growth due to positive interactions (21, 42). These phase diagrams results demonstrate that the existence of three phases (full coexistence, partial coexistence, persistence fluctuation) and the order of transitions are robust to varying interaction types in the model (communities lose species before losing stability as the size of species pool  $S$  or interaction strength increases, as in Fig. 1E and F). The dashed line and solid line in the figures represent survival boundary and stability boundary, respectively. (B) *In silico* communities undergoing serial dilutions exhibit the same three dynamical phases (full coexistence, partial coexistence, and fluctuation) as in simulations without dilution. The two phase diagrams show that communities exposed to serial dilutions (1:30 dilution every 24 hours) lose species before losing stability as the size of species pool  $S$  or interaction strength increases, which is consistent with simulations of the continuous (no dilutions) model (Fig. 1E and F). The dashed line and solid line in the figures represent the survival boundary and stability boundary, respectively.

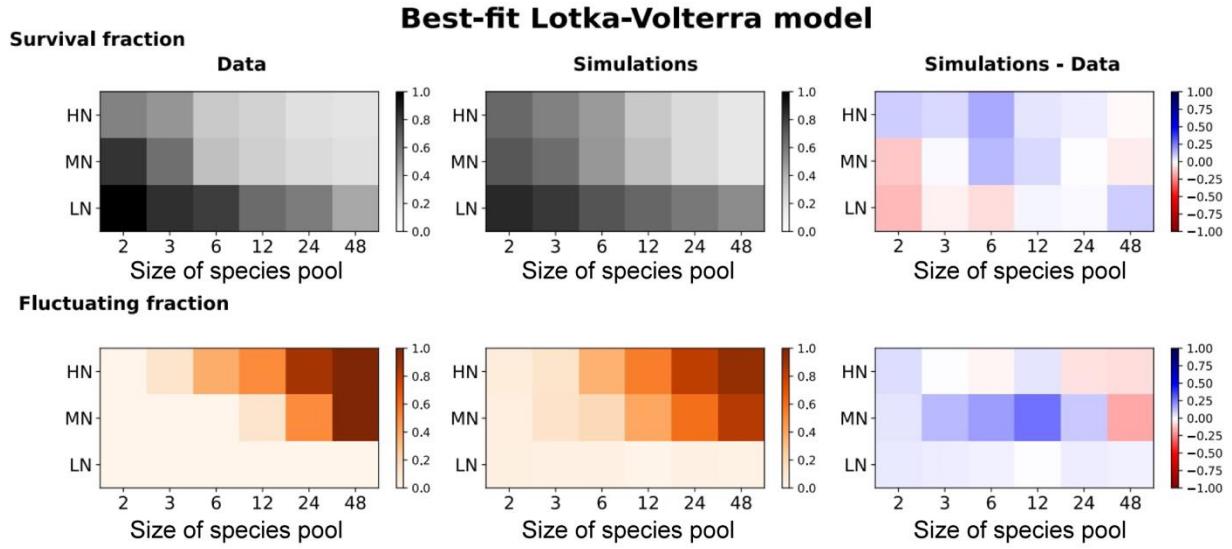


**Fig. S10. Comparison between the Lotka-Volterra and pH-based models.** The topmost two rows display survival and fluctuation fraction, demonstrating that they similarly depend on the parameters of species pool size and interaction strength, in the Lotka-Volterra model (left column, equation 1, with interaction strength given by  $\langle \alpha_{ij} \rangle$ ) and a pH-based model proposed for previous experiments (14) (right column, equations 2 and 3, with interaction strength given by  $\max(c_i)$ ). We conclude that the three dynamical phases are qualitatively robust to different modeling choices. On the other hand, the third row shows that the two models disagree regarding correlations between the number of surviving species and the presence of fluctuations (for different communities with the same parameters) as further discussed in Fig. S11.

### Evidence against the pH model



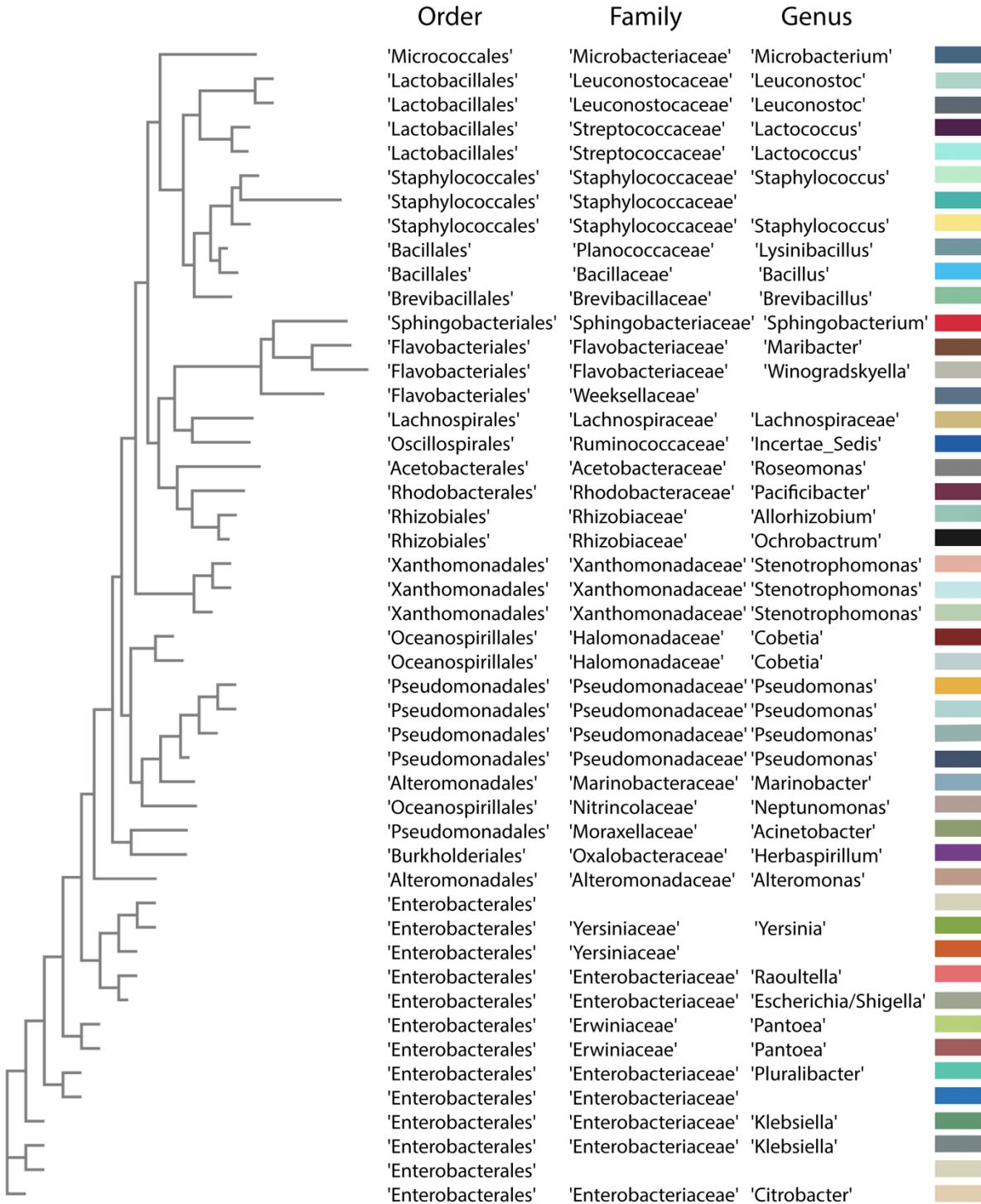
**Fig S11. The LV model reproduces experimental observations better than the pH-based model.** (A) Empirical data shows rather weak correlation (Pearson correlation coefficient: 0.54) between the intensity of fluctuations of pH and of abundances, suggesting that pH is not the sole or main driving factor in species dynamics. (B) The pH-based model proposed in (14) often displays community extinction (total abundance  $< 5\%$  of carrying capacity), which we do not observe in our experiments. (C) In experiments and in the LV model, conditioning on species pool size and nutrients, various communities show positive correlations between fluctuations and diversity, measured here as  $S^*/\langle S^* \rangle$  the number of surviving species relative to the average number of survivors for that same pool size and nutrients. The pH-based model displays negative correlations (we exclude cases of whole community extinction which were never observed in our experiments). The relative survival fraction is statistically higher (lower) in fluctuating communities than stable communities in experiments and LV model (pH-based model,  $p < 0.001$ ). Error bars, s.e.m.,  $n=100$  for simulation data,  $n=51$  (33) for stable communities in high (medium) nutrient,  $n=45$  (15) for fluctuating communities in high (medium) nutrient.



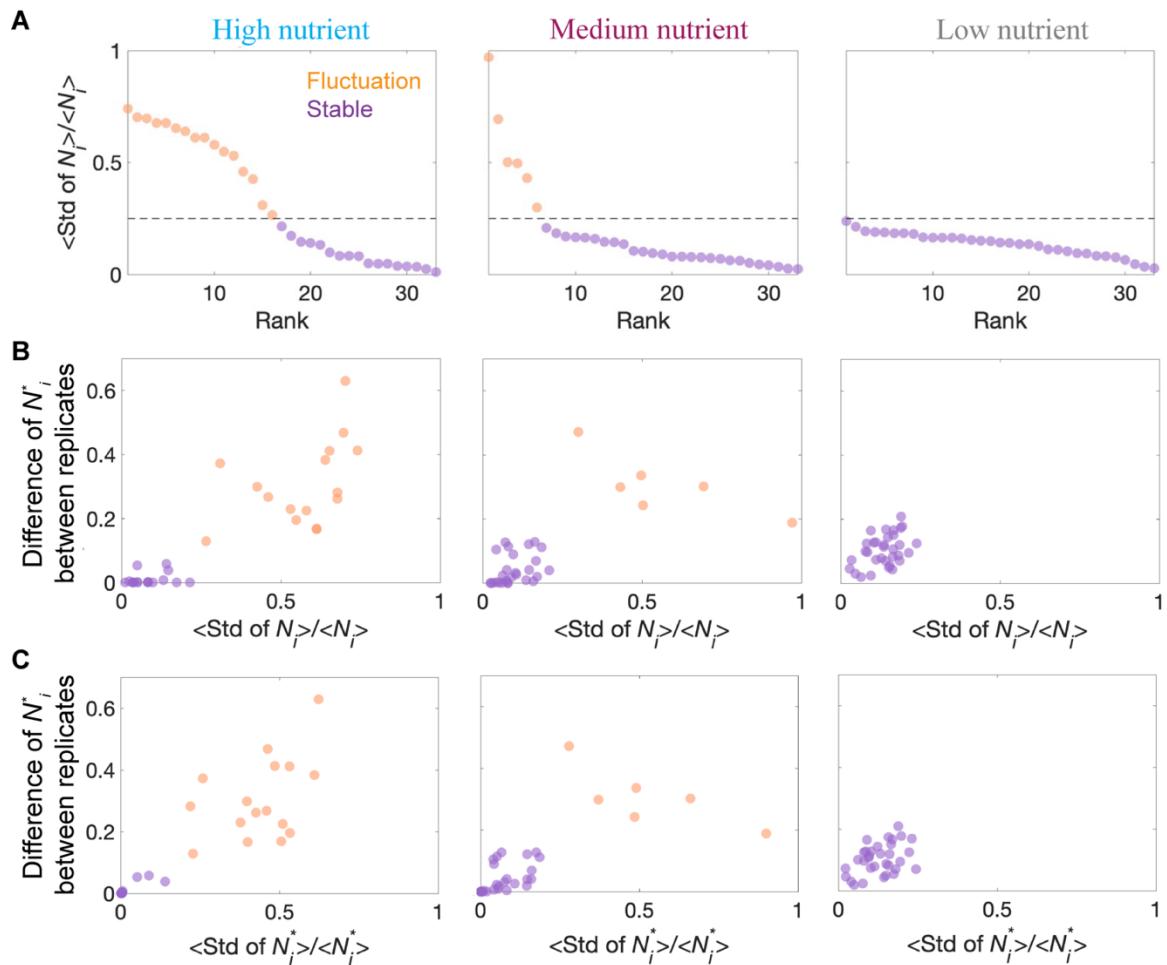
**Fig S12. Best-fit Lotka-Volterra model.** We show the heatmaps of survival fraction (top) and fraction of fluctuating communities (bottom) for data, best-fit simulations, and the difference between the two (left to right). The best-fit simulations are obtained from a Lotka-Volterra with normally-distributed interactions with the following parameters: High nutrient treatment (HN):  $\langle \alpha_{ij} \rangle = 0.87$ , std ( $\alpha_{ij} \rangle = 0.22$ . Medium nutrient treatment (MN):  $\langle \alpha_{ij} \rangle = 0.66$ , std ( $\alpha_{ij} \rangle = 0.41$ . Low nutrient treatment (LN):  $\langle \alpha_{ij} \rangle = 0.14$ , std ( $\alpha_{ij} \rangle = 0.17$ .



**Fig. S13. Taxonomic identity of the 48 bacterial isolates.** The identities have been inferred from the ASV (Methods) of 16S samples taken from monocultures, which allow the classification of the 48 isolates down to the genus level. Colors are consistent with those in the main text and other supplementary figures.

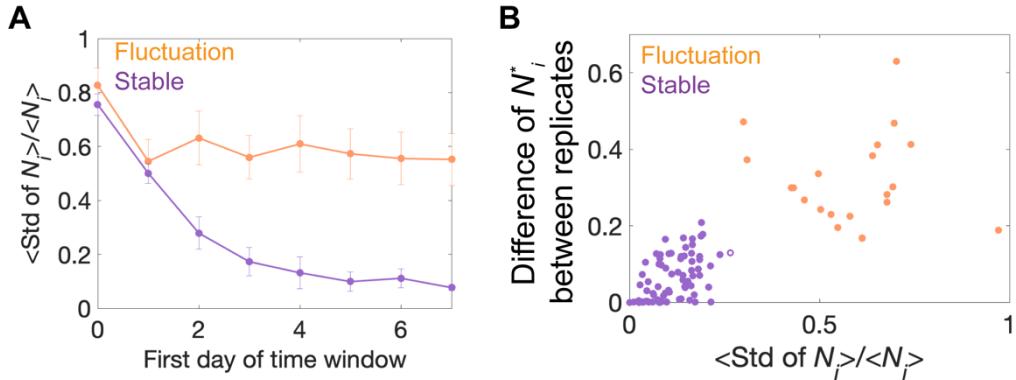


**Fig. S14. Phylogenetic tree of the 48 bacterial isolates.** The tree, generated with the Distance-matrix method from EMBL-EBL (40), shows the relative phylogenetic distance between the 48 bacterial isolates. The library contains bacterial isolates from either soil or *C. elegans* gut samples and spans 19 different orders and 26 different families. The rectangles display the color that is used in the figures (both in the Main Text and the SI) to show the abundance of each species in the different experiments.

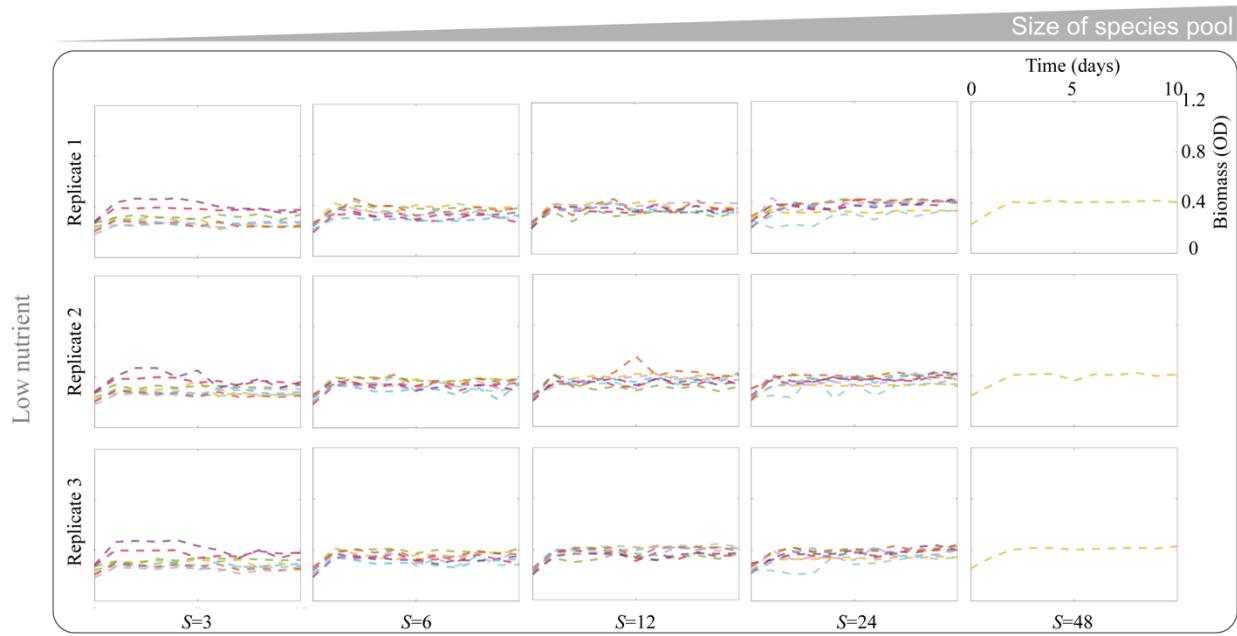


**Fig. S15. 3 different metrics consistently differentiate stable from fluctuating communities.**

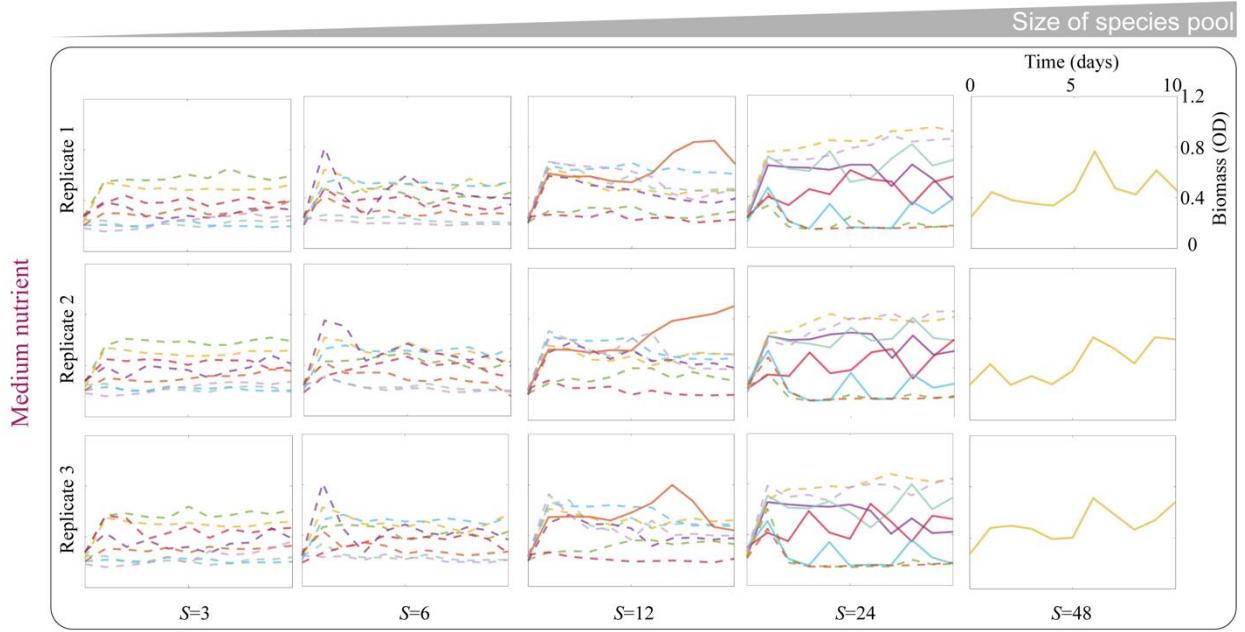
(A) The three panels show the average coefficient of (temporal) variation (see Methods) for absolute species abundances ( $N_i$ , computed as the product of total biomass per species relative abundance) in the experimental communities in the three different nutrients concentrations. We use a stability threshold of 0.25 (dashed line) to classify communities into stable (purple) and fluctuating (orange) ones. The number of fluctuating communities increases with the average interaction strength (nutrients concentration), with all the weakly interacting (low nutrients concentration) communities exhibiting stability. (B) Average difference (Euclidean distance) in the relative abundance of each species ( $N_i^* = N_i / \sum N_i$ ) across replicate communities as a function of the community's average coefficient of (temporal) variation. Stable and fluctuating communities, defined as in (A), span in different regions, with stable communities clustering near the origin. (C) Average difference (Euclidean distance) in the relative abundance of each species ( $N_i^* = N_i / \sum N_i$ ) across replicate communities as a function of the community's average coefficient of (temporal) variation in relative abundance ( $N_i^*$ ). Stable and fluctuating communities, defined as in (A), span in different regions, with stable communities clustering near the origin. The average coefficient of variation (Fig. S15, S16) for species abundances was calculated based on only replicate for which we sequenced the whole time series, and the average difference in relative species abundances community across the three replicates for each community (Fig. S15, S16) was calculated based on relative abundances at day 10.



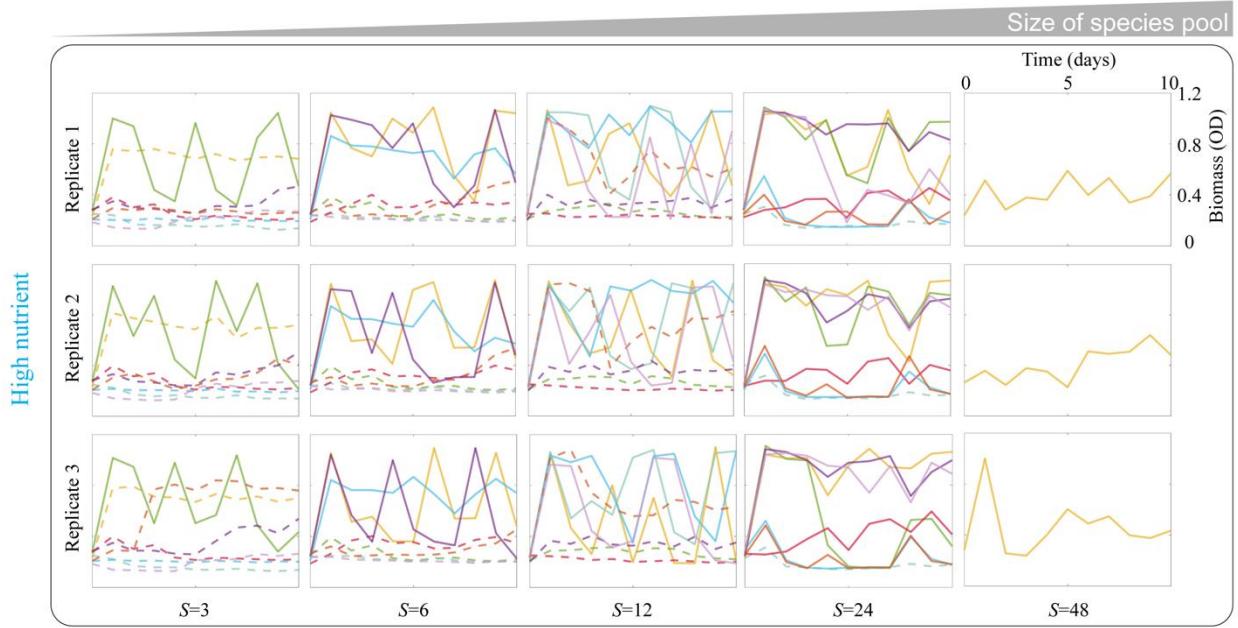
**Fig. S16. The classification of fluctuating communities and stable communities is robust to choices of classification algorithm.** (A) The average coefficient of variation for species abundances reaches a steady state before day 7, enabling the classification of communities into stable and fluctuating ones. For 12-species communities under high nutrient concentrations, the average CV of both fluctuating communities (orange line, n=4) and stable communities (purple line, n=4) reaches a plateau (a constant value) before day 7. The two different plateaus of average CV demonstrate that the dynamics of communities (persistent fluctuations or stability) have reached steady states before the time window (from day 7 to day 10) that we use to calculate the average CV in Fig. S15. Error bars, s.e.m. (B) using a K-means clustering algorithm considering both average CV and differences between species relative abundances across replicates confirms that the classification of fluctuating and stable communities is consistent with the CV threshold (0.25) criteria in Fig. S15. There is only one community (empty circle) that is differently classified by the K-means method. The classification of this single community as either stable or fluctuating changes neither the three phases in the experimental phase diagram nor the order of phase transitions (lose species before losing stability).



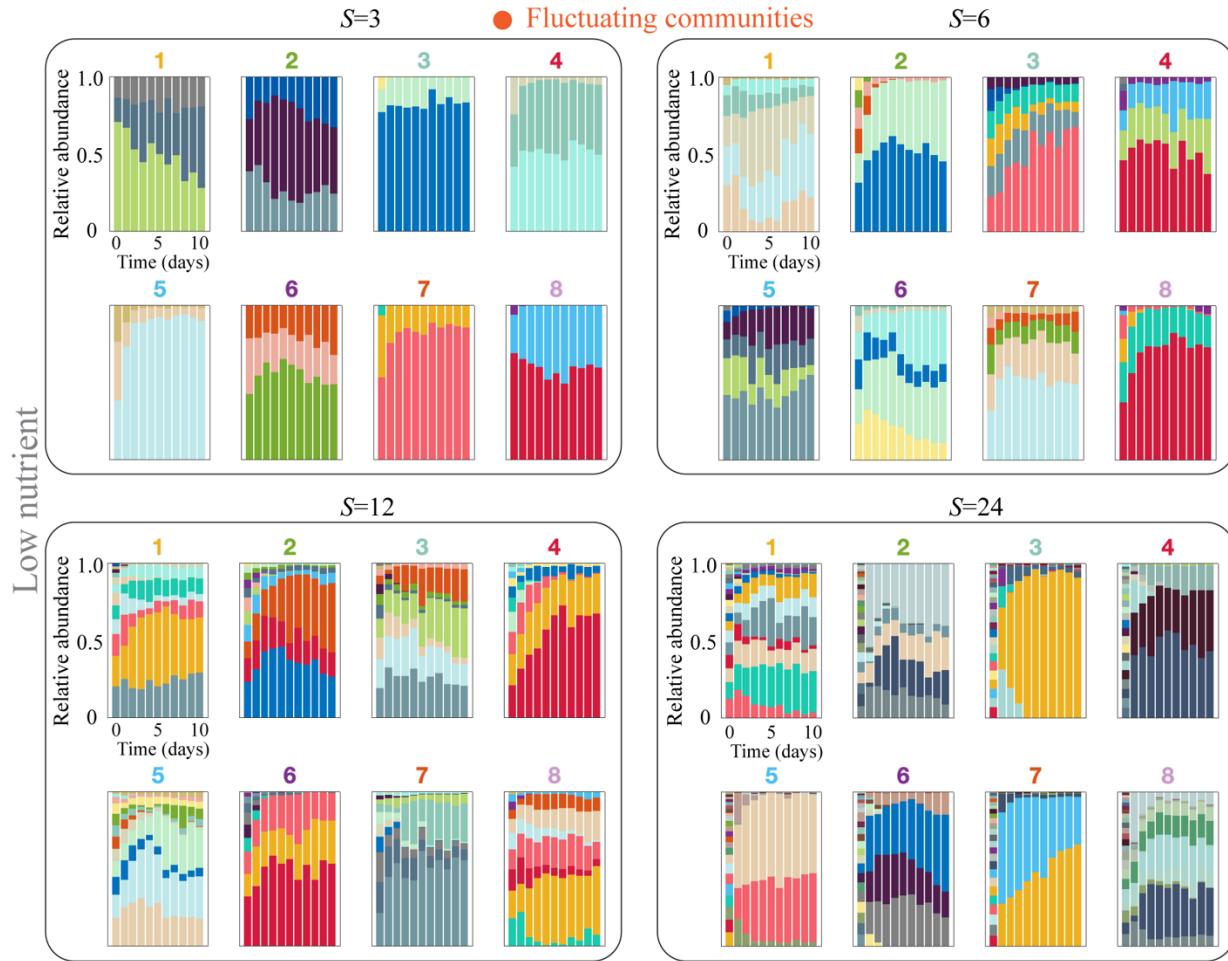
**Fig. S17. Total biomass reaches equilibrium in communities under low nutrients concentration (low interaction strength).** Each panel shows the time series for the OD (600nm) of the eight communities with different species pool composition (depicted by different colors). Each column stands for a different species pool size  $S$  (for the case of  $S=48$ , there is only one community containing the full library of bacterial species). Each row shows the data for a different replicate of the experiment.



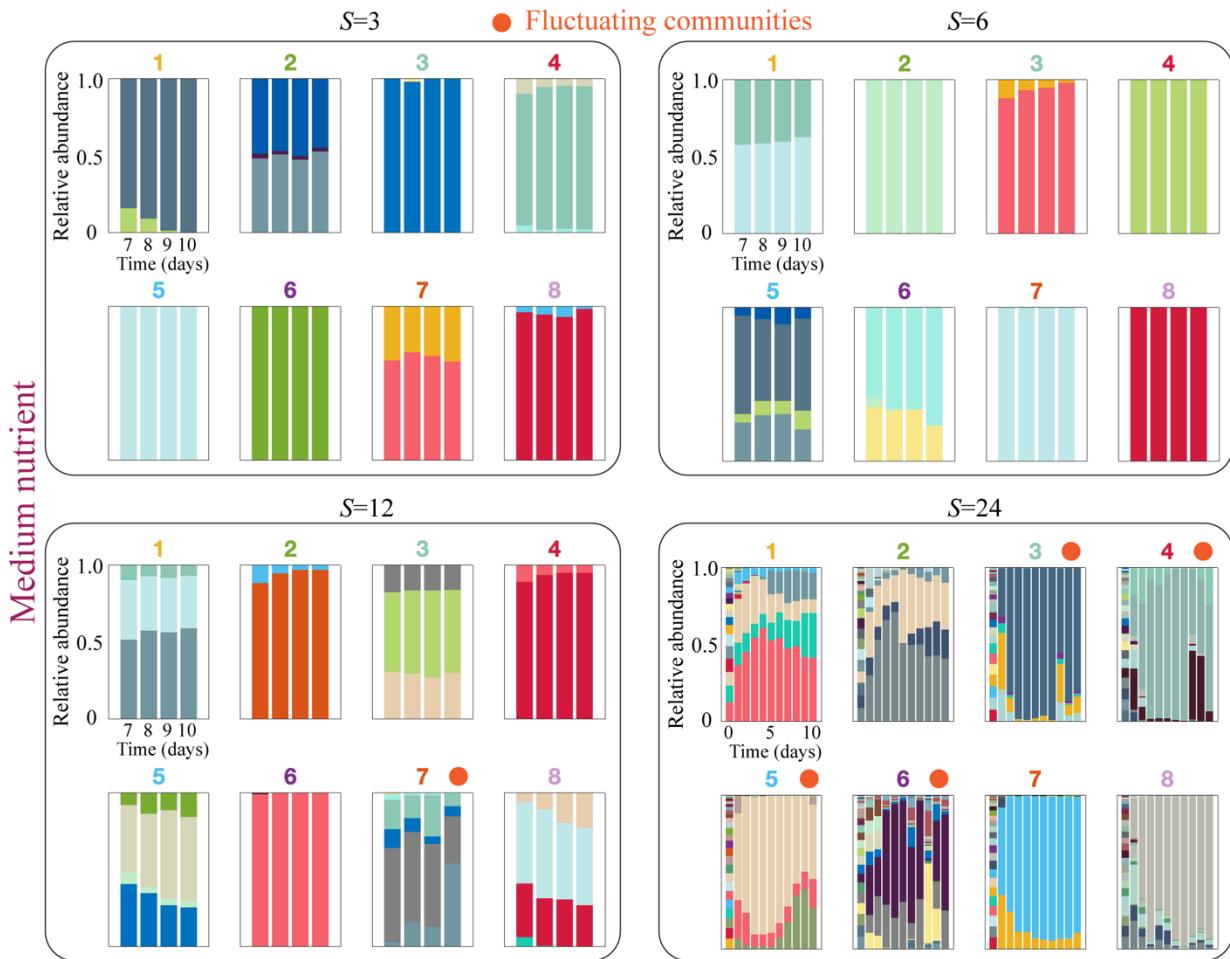
**Fig. S18. Increasing the species pool size leads to persistent fluctuations in total biomass under medium nutrients concentration (medium interaction strength).** Each panel shows the time series for the OD (600nm) of the eight communities with different species pool composition (depicted by different colors). Each column stands for a different species pool size  $S$  (for the case of  $S=48$ , there is only one community containing the full library of bacterial species). Each row shows the data for a different replicate of the experiment. Solid lines (dashed lines) represent fluctuating (stable) communities, the OD fluctuations between day 7 and day 10 were considered to differentiate fluctuating and stable communities as shown in Fig. S15.



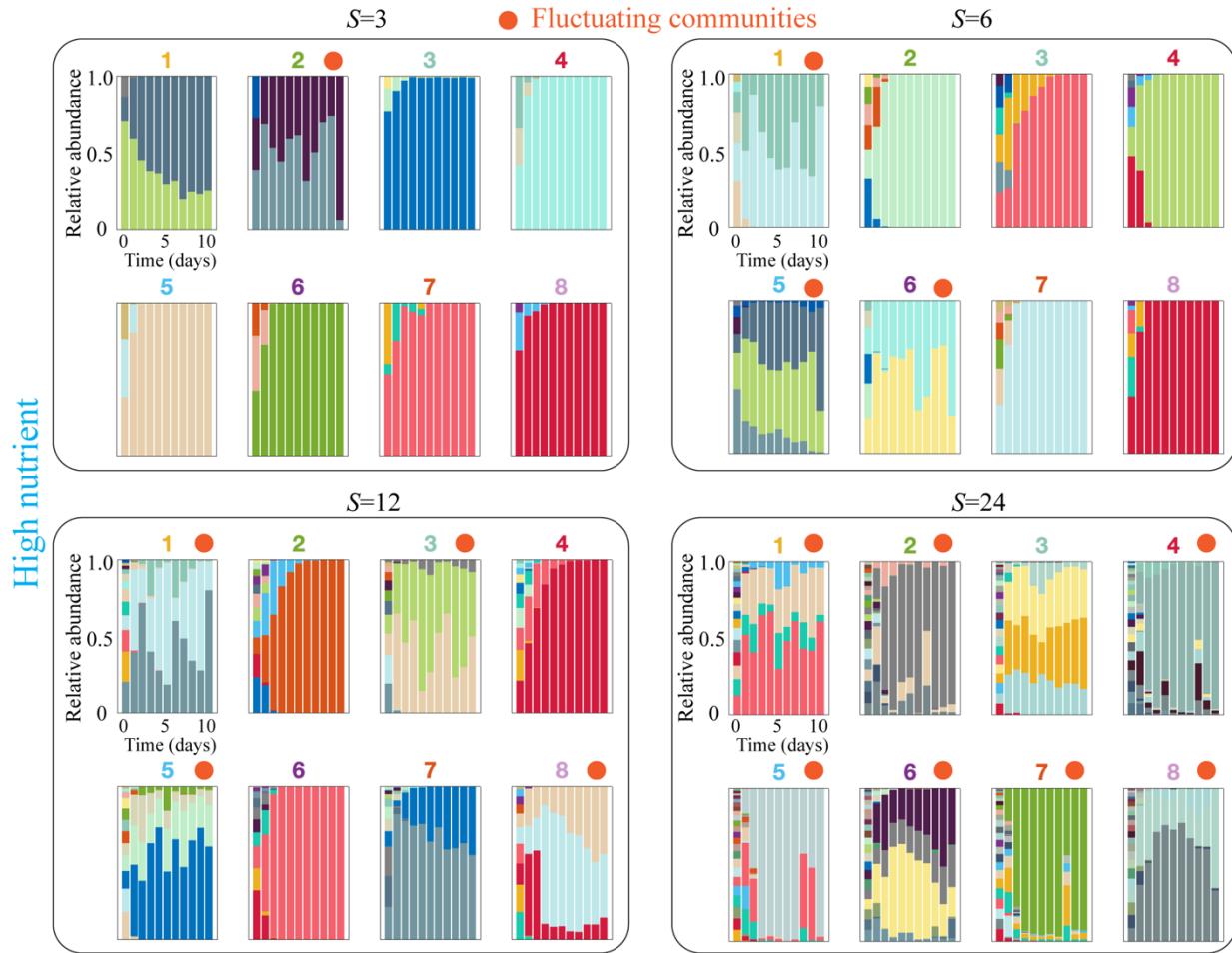
**Fig. S19. Increasing the species pool size leads to persistent fluctuations in total biomass under high nutrients concentration (high interaction strength).** Each panel shows the time series for the OD (600nm) of the eight communities with different species pool composition (depicted by different colors). Each column stands for a different species pool size  $S$  (for the case of  $S=48$ , there is only one community containing the full library of bacterial species). Each row shows the data for a different replicate of the experiment. Solid lines (dashed lines) represent fluctuating (stable) communities, the OD fluctuations between day 7 and day 10 were considered to differentiate fluctuating and stable communities as shown in Fig. S15.



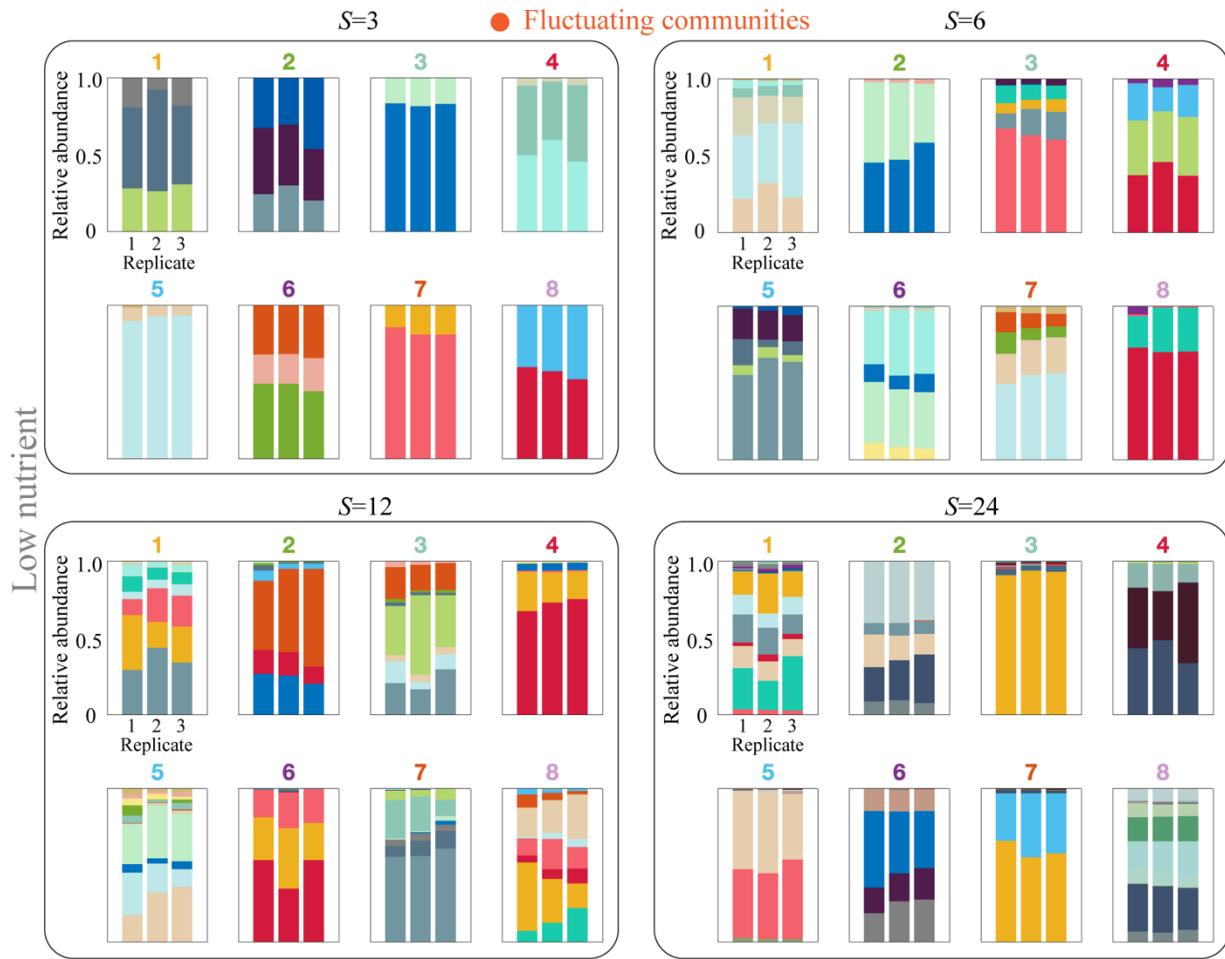
**Fig. S20. Time series for the relative species abundances of the experimental communities with low average interaction strength (low nutrients concentration).** Each panel shows the full time series for each of the 8 communities with the indicated species pool size ( $S=3, 6, 12$  and  $24$ ). Bar colors stand for species identities as in Fig. S13, S14. Under this nutrients condition, all of the communities reached a stable equilibrium (Methods). The color of the number on the top of each panel corresponds to the color assigned to the same community in Fig. S17.

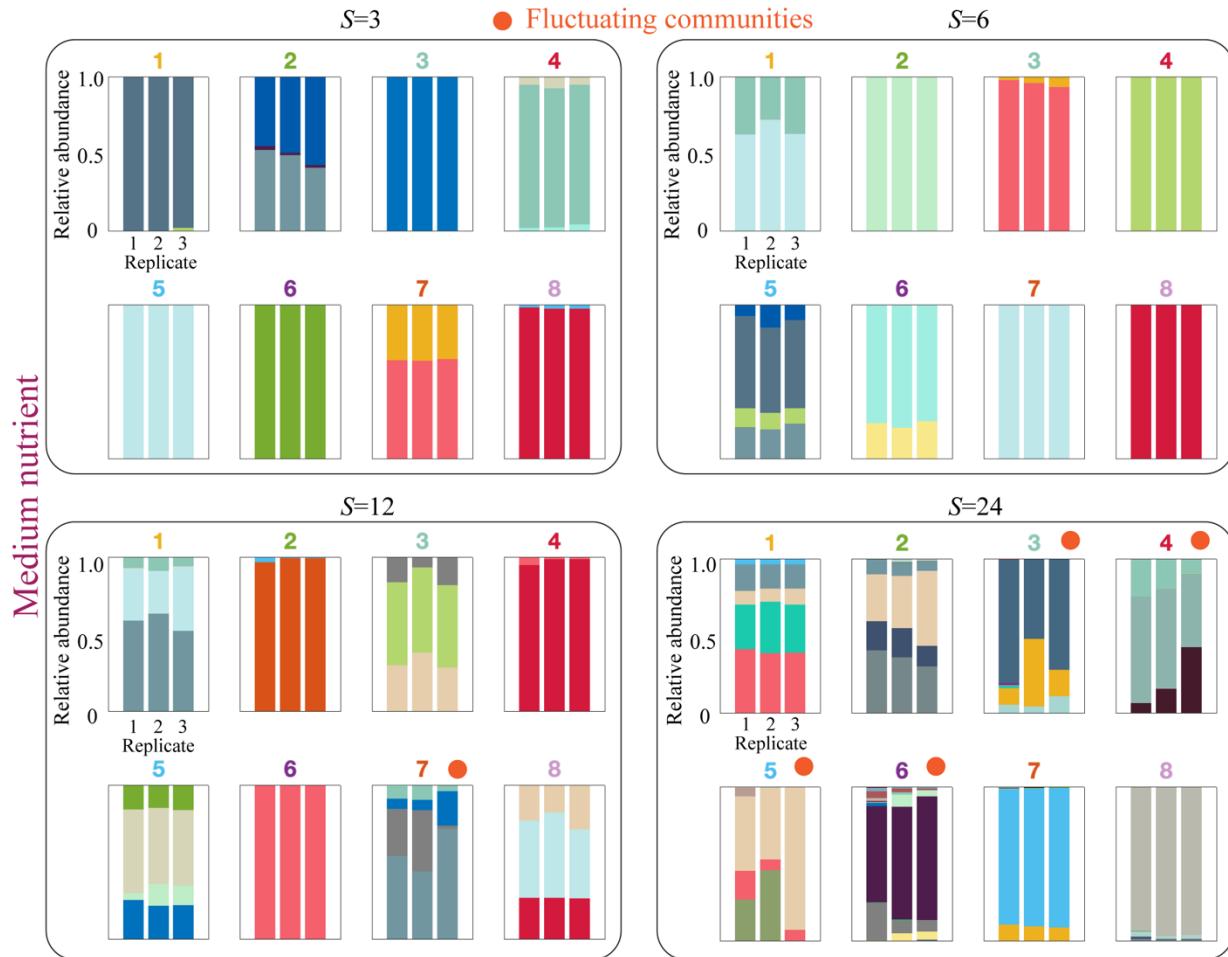


**Fig. S21. Time series for the relative species abundances of the experimental communities with medium average interaction strength (medium nutrients concentration).** Each panel shows the full time series for each of the 8 communities with the indicated species pool size ( $S=3$ ,  $6$ ,  $12$  and  $24$ ). Bar colors stand for species identities. The orange dot on top of some panels indicates that the community exhibits persistent fluctuations (Methods). The color of the number on the top of each panel corresponds to the color assigned to the same community in Fig. S18. For  $S=3$ ,  $6$ , and  $12$ , we only sequenced samples of the last 4 days (7 to 10) of the experiment.

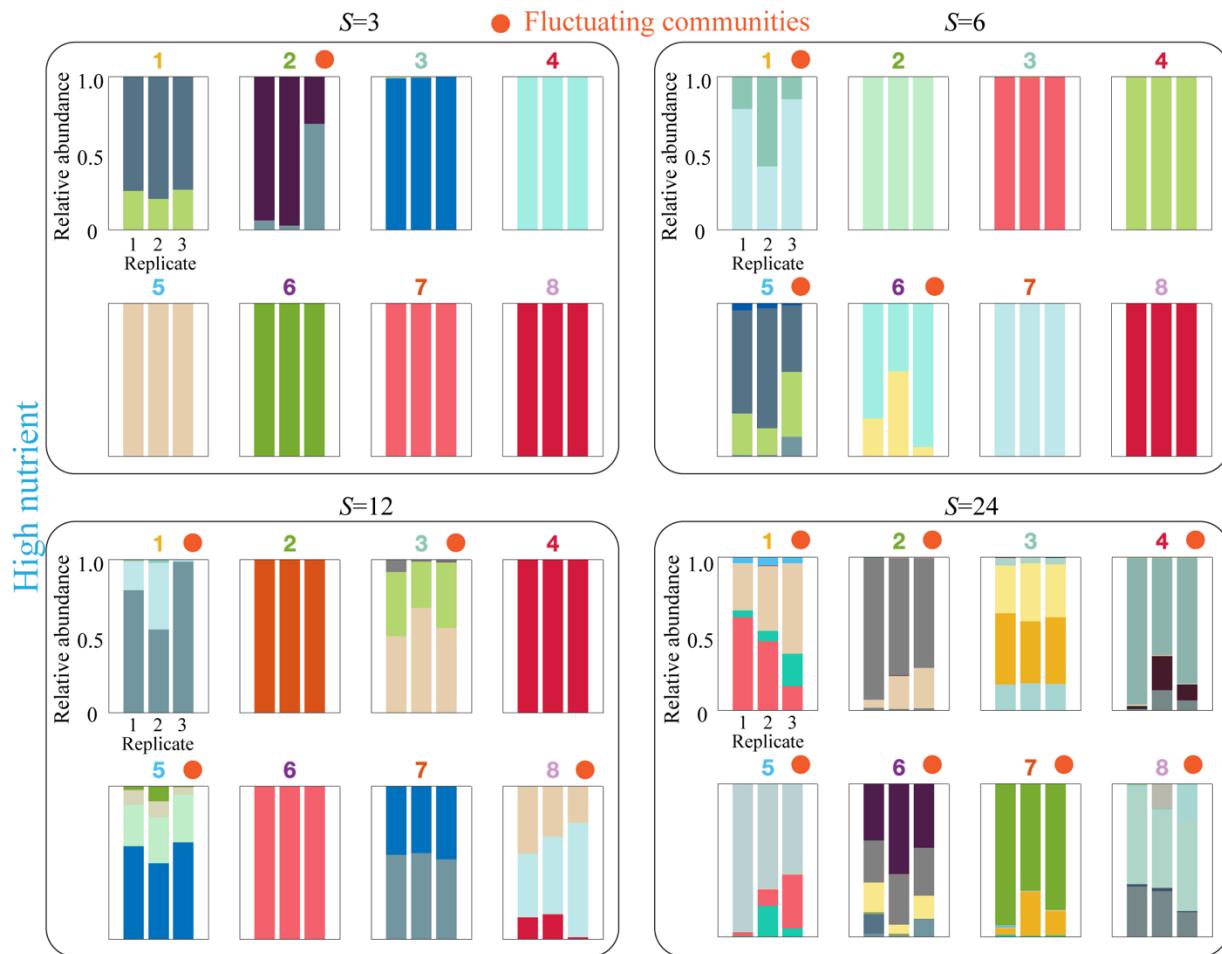


**Fig. S22. Time series for the relative species abundances of the experimental communities with high average interaction strength (high nutrients concentration).** Each panel shows the full time series for each of the 8 communities with the indicated species pool size ( $S=3, 6, 12$  and  $24$ ). Bar colors stand for species identities. The orange dot on top of some panels indicates that the community exhibits persistent fluctuations (Methods). The color of the number on the top of each panel corresponds to the color assigned to the same community in Fig. S19.

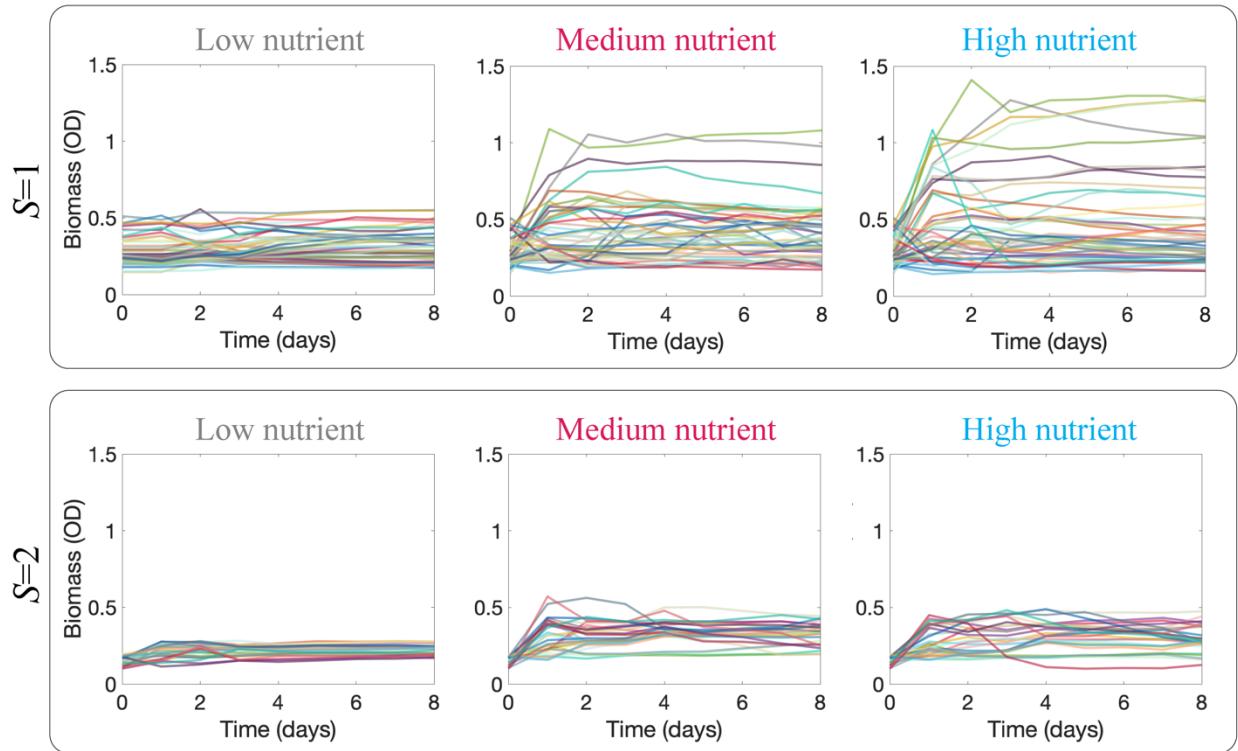




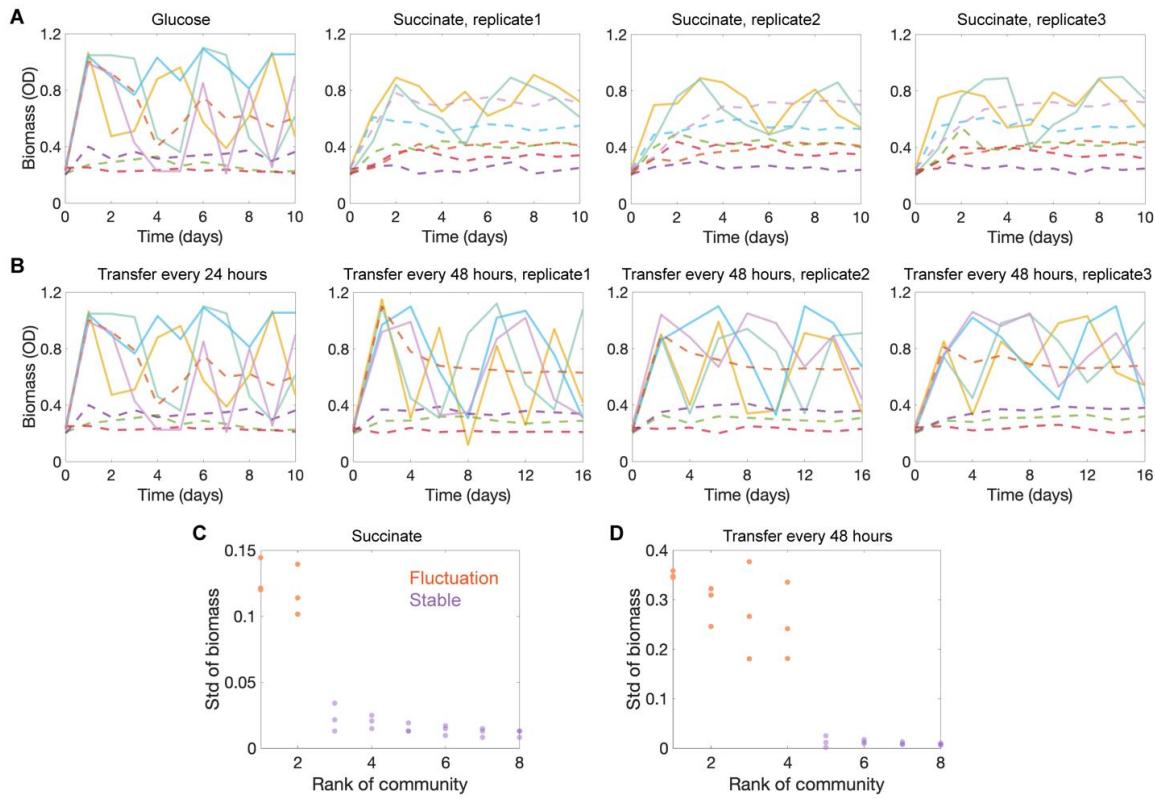
**Fig. S24. Species abundance at the end of the experiment under medium nutrients concentration.** Each panel shows the species relative abundances at the end experiment for each of the 3 replicate communities across 8 different compositions of the species pool for each species pool size ( $S=3, 6, 12$  and  $24$ ). Bar colors stand for different species identities. The orange dot on top of some panels indicates that the community exhibits persistent fluctuations (Methods). The color of the number on the top of each panel corresponds to the color assigned to the same community in Fig. S18.



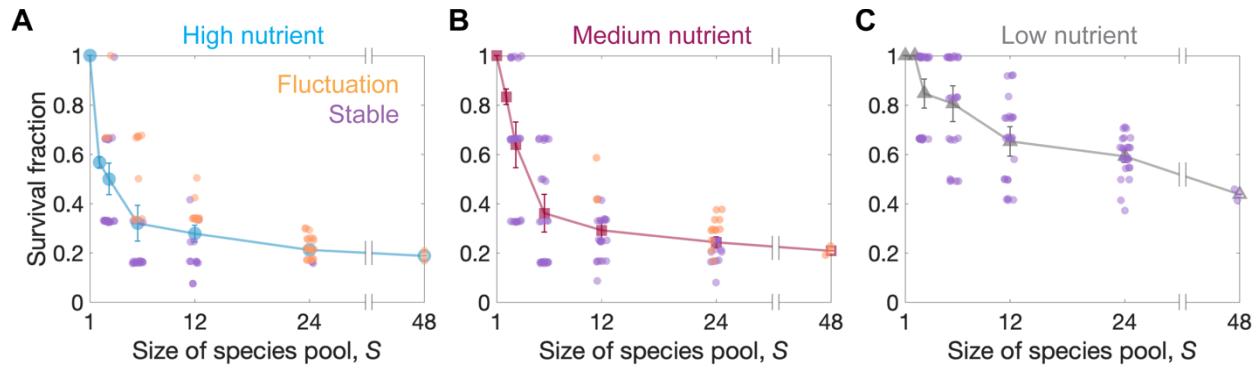
**Fig. S25. Species abundance at the end of the experiment under high nutrients concentration.** Each panel shows the species relative abundances at the end experiment for each of the 3 replicate communities across 8 different compositions of the species pool for each species pool size ( $S=3$ ,  $6$ ,  $12$  and  $24$ ). Bar colors stand for different species identities. The orange dot on top of some panels indicates that the community exhibits persistent fluctuations (Methods). The color of number on the top of each panel corresponds to the color assigned to the same community in Fig. S19.



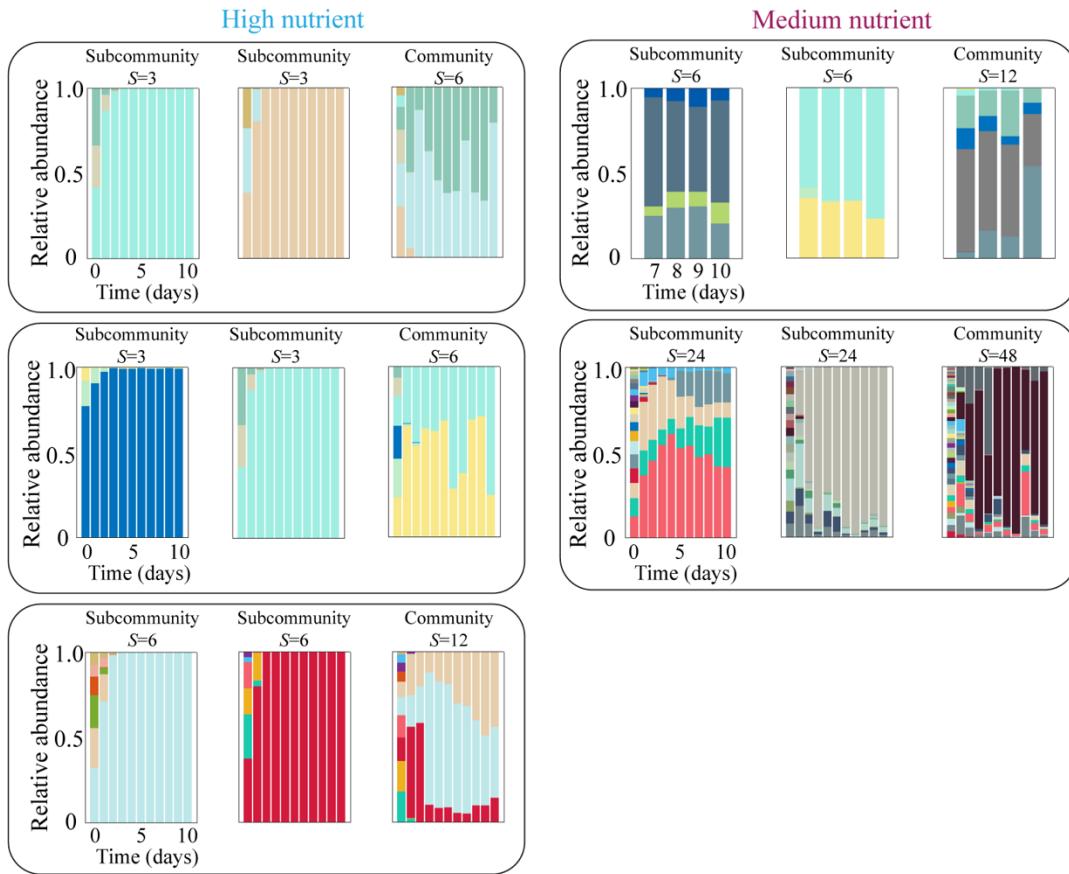
**Fig. S26. Monocultures and 2-species cocultures tend to reach stability in total biomass.** On top, monocultures tend to reach a stable OD (600nm) value at the end of each daily cycle. The width of the observed range of OD values increases with nutrient concentration (low, medium, and high, from left to right). On bottom, time series for the OD (600nm) of 15 different species pair cocultures. To detect bistability, in which the outcome depends on initial species abundances, we considered two initial compositions (5:95 and 95:5 culture volume ratios between species) for each pair of species. Therefore, there are 30 pairwise cocultures tested. The variability of the OD reached in pairwise coculture also increases with nutrient concentration, but to a less extent than it does for monocultures. Different colors stand for different species identities (top) and different species pairs (bottom).



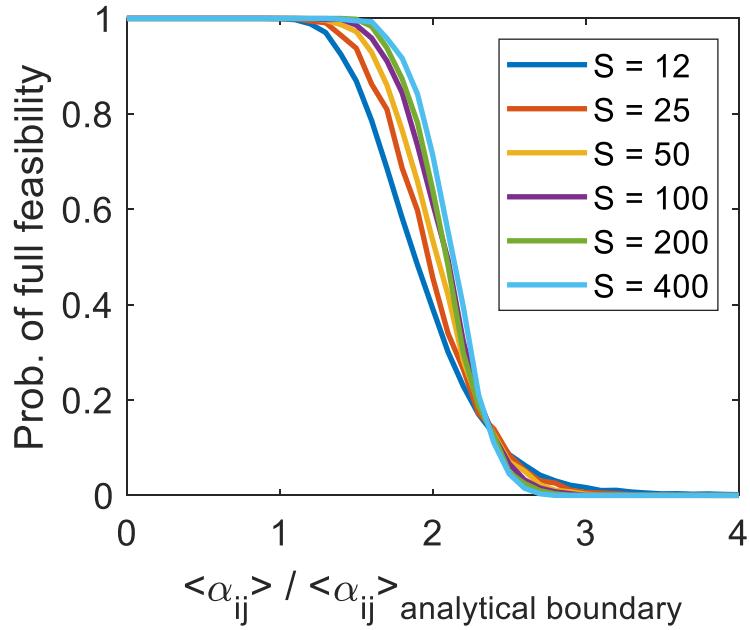
**Fig. S27. The phases of community dynamics are robust to changes in carbon source and dilution frequency.** (A) Persistent fluctuations can occur under different carbon sources. After replacing glucose (leftmost panel) by succinate (right panels) in the media, high nutrients concentration still yields biomass fluctuations in some communities. Two out of eight communities fluctuate in both glucose and succinate, four reach steady state in both glucose and succinate, and the remaining two communities fluctuate in medium with glucose, while reaching a steady state in medium with succinate. Each panel shows the time series for the OD of the eight communities with different species pool composition (depicted by different colors). Solid lines (dashed lines) represent fluctuating (stable) communities. (B) Community dynamics are robust to different choices of dilution frequency. 24-hours transfers (leftmost panel) and 48-hours transfers (right panels) yield analogous biomass dynamics for the eight different 12-species communities under high nutrients concentration. Communities that reach persistent fluctuations (stability) under 24-hours-transfers also reach persistent fluctuations (stability) under 48-hours-transfers. In both (A) and (B), each panel shows the time series for the OD of the eight communities with different species pool composition (depicted by different colors). Each of the three rightmost panels show the results for one of the three experimental replicates performed. (C) Rank plot for the standard deviation of biomass between days 7 and 10 for communities under succinate. The rank of each community was based on the mean value of the standard deviation of the three replicates. A *K*-means clustering algorithm considering standard deviation of biomass over days 7-10 clusters communities into two fluctuating (orange points) ones and six stable (purple points) ones. (D) Rank plot for the standard deviation of biomass between days 10 and 16 for communities under 48-hours transfers. In this case, the *K*-means clustering algorithm considering standard deviation of biomass over day10-16 clusters communities into four fluctuating (orange points) ones and four stable (purple points) ones. The three replicates of each community are consistently classified as either fluctuating or stable ones, which are shown as three data points for each community rank.



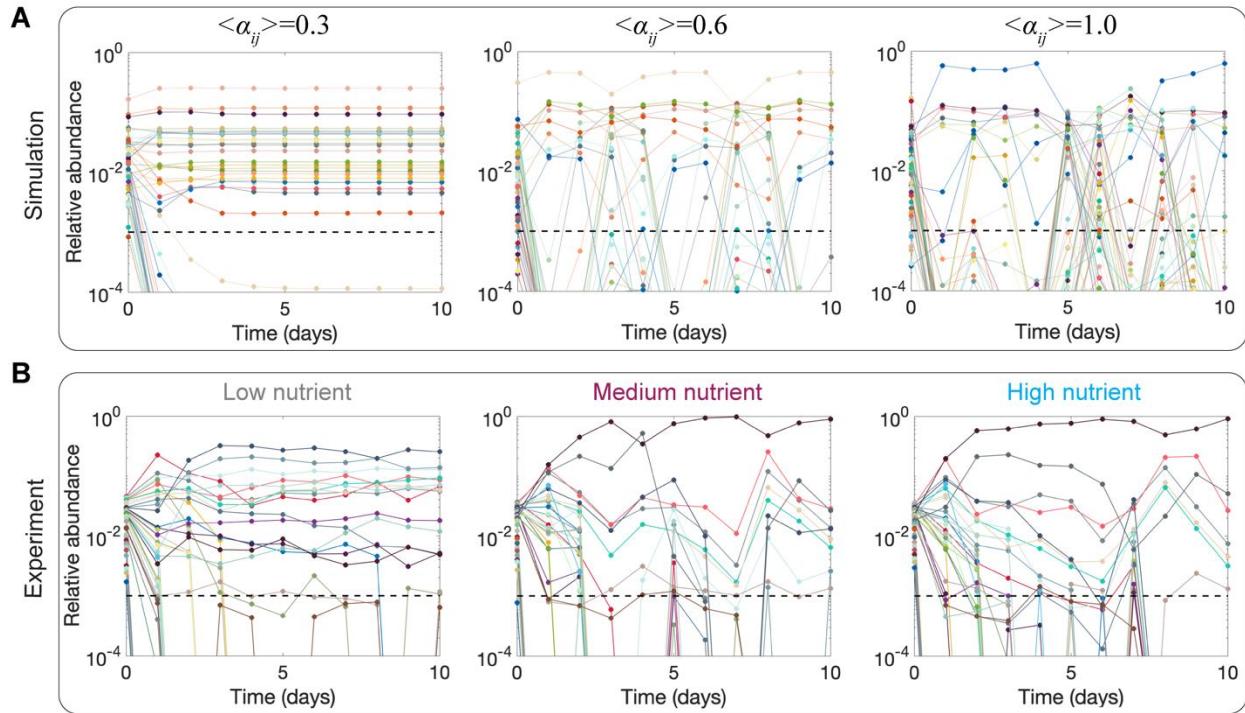
**Fig. S28. Fluctuating communities are more diverse than stable communities under the same conditions.** As the average survival fraction decreases with increasing species pool size  $S$  in high (A), medium (B) and low (C) nutrient concentrations, more communities exhibit fluctuations in species abundances (orange points). For any given  $S$  and nutrients concentration, fluctuating microbial communities exhibit statistically higher survival fractions than stable communities (purple points).



**Fig. S29. Increasing species pool size can lead to emergent fluctuations in species abundances.**  
The panels show representative examples in which a pair of different communities reach stability, while a community with larger species pool, composed by all the species present in that pair, exhibits persistent fluctuations. Each rectangle encloses a different example involving a specific set of communities and experimental condition (nutrients concentration). The top right rectangle shows data for the last 4 days of the experiment (the only days in which these communities were sampled for sequencing), and the rest show the full time series for the 10-day experiment.



**Fig. S30. The probability of full coexistence as a function of the mean interaction strength  $\langle \alpha_{ij} \rangle$  exhibits a sharp phase transition between phases (I) and (II) when  $S$  is large in simulations.** The x-axis is normalized by  $\langle \alpha_{ij} \rangle$  where the analytical survival boundary is expected. While all curves decrease to zero in the same region, the width of the crossover regime becomes narrower with increasing  $S$ . The fact that all curves decrease to zero at the same region, shows that the analytical expression indeed captures the correct dependence of the boundary in  $\langle \alpha_{ij} \rangle$  on  $S$ .



**Fig. S31: Simulated and experimental communities exhibit analogous dynamics of relative species abundances.** (A) Time series of relative species abundances in a representative simulation for  $S=48$ , under low nutrients (low interaction strength, left panel), medium nutrients (medium interaction strength, middle panel), and high nutrients (high interaction strength, right panel) concentrations. We used species abundance data sampled every 24 hours of simulated time in order to match the experimental data sampling frequency. (B) Experimental time series obtained through 16S data in analogous conditions to the panels in (A). Some low abundance species (abundances below the  $10^{-3}$  survival threshold, shown as horizontal dashed lines) exhibit fluctuation in the low nutrient concentration experiment, which can be explained by small numbers effect such as finite 16s sequencing depth.

**Table S1. Experimentally measured interspecies interaction matrix  $\alpha_{ij}$  under low nutrients concentrations.** Each of the 15 pairs resulting from combinations of six randomly chosen isolates from different genera (Leuconostoc, Pseudomonas, Yersinia, Pantoea, Klebsiella, Acinetobacter) in the bacterial library were cocultured for 7 days (with daily dilutions). We measured the equilibrium abundance  $N_i$  via sample dilution and colony counting at the end of the experiment. The value of  $\alpha_{ij}$  was calculated through the expression  $\alpha_{ij} = \frac{(K_i - N_i)K_j}{N_j K_i}$ , where  $K_i$  is the carrying capacity of species (independently measured as the species abundance in monoculture after 7 dilution cycles). The errors indicate the standard deviation of parameter values measured in three replicates.

	<b>Leu</b>	<b>Pse</b>	<b>Yer</b>	<b>Pan</b>	<b>Kle</b>	<b>Aci</b>
<b>Leu</b>	1	0.25±0.03	0.31±0.01	0.29±0.04	0.24±0.01	0.03±0.02
<b>Pse</b>	0.55±0.05	1	0.10±0.05	0.63±0.04	0.67±0.09	0.33±0.01
<b>Yer</b>	0.30±0.04	0.28±0.05	1	0.60±0.04	0.28±0.06	-0.05±0.04
<b>Pan</b>	0.39±0.01	0.33±0.03	0.33±0.05	1	0.04±0.02	0.17±0.08
<b>Kle</b>	0.44±0.03	0.54±0.05	0.06±0.03	0.05±0.01	1	0.69±0.13
<b>Aci</b>	0.40±0.07	0.60±0.03	-0.29±0.08	0.47±0.05	0.64±0.02	1

**Table S2. Experimentally measured interspecies interaction matrix  $\alpha_{ij}$  under medium nutrient concentrations.** Each of the 15 pairs resulting from combinations of six randomly chosen isolates from different genera (Leuconostoc, Pseudomonas, Yersinia, Pantoea, Klebsiella, Acinetobacter) in the bacterial library were cocultured for 7 days (with daily dilutions). We measured the equilibrium abundance  $N_i$  via sample dilution and colony counting at the end of the experiment. The value of  $\alpha_{ij}$  was calculated through the expression  $\alpha_{ij} = \frac{(K_i - N_i)K_j}{N_j K_i}$ , where  $K_i$  is the carrying capacity of species (independently measured as the species abundance in monoculture after 7 dilution cycles). For competitive exclusion (species  $i$  always drive species  $j$  to extinction), it can be inferred that  $\alpha_{ij} < 1$  and  $\alpha_{ji} > 1$ . For bi-stability (the high abundant species drives the low abundant one to extinction), it can be inferred that  $\alpha_{ij} > 1$  and  $\alpha_{ji} > 1$ . The errors indicate the standard deviation of parameter values measured in three replicates.

	<b>Leu</b>	<b>Pse</b>	<b>Yer</b>	<b>Pan</b>	<b>Kle</b>	<b>Aci</b>
<b>Leu</b>	1	0.69±0.05	<1	>1	0.08±0.03	0.11±0.02
<b>Pse</b>	0.31±0.03	1	0.41±0.06	0.99±0.04	0.26±0.01	0.25±0.04
<b>Yer</b>	>1	0.24±0.07	1	>1	0.20±0.05	0.21±0.03
<b>Pan</b>	<1	0.98±0.08	<1	1	>1	<1
<b>Kle</b>	0.21±0.04	0.36±0.01	0.18±0.05	<1	1	0.40±0.02
<b>Aci</b>	0.05±0.02	0.32±0.01	0.13±0.06	>1	0.87±0.11	1

**Table S3. Experimentally measured interspecies interaction matrix  $\alpha_{ij}$  under high nutrient concentrations.** Each of the 15 pairs resulting from combinations of six randomly chosen isolates from different genera (Leuconostoc, Pseudomonas, Yersinia, Pantoea, Klebsiella, Acinetobacter) in the bacterial library were cocultured for 7 days (with daily dilutions). We measured the equilibrium abundance  $N_i$  via sample dilution and colony counting at the end of the experiment. The value of  $\alpha_{ij}$  was calculated through the expression  $\alpha_{ij} = \frac{(K_i - N_i)K_j}{N_j K_i}$ , where  $K_i$  is the carrying capacity of species (independently measured as the species abundance in monoculture after 7 dilution cycles). For competitive exclusion (species  $i$  always drive species  $j$  to extinction), it can be inferred that  $\alpha_{ij} < 1$  and  $\alpha_{ji} > 1$ . For bi-stability (the high abundant species drives the low abundant one to extinction), it can be inferred that  $\alpha_{ij} > 1$  and  $\alpha_{ji} > 1$ . The errors indicate the standard deviation of parameter values measured in three replicates.

	<b>Leu</b>	<b>Pse</b>	<b>Yer</b>	<b>Pan</b>	<b>Kle</b>	<b>Aci</b>
<b>Leu</b>	1	0.09±0.04	<1	>1	0.49	<1
<b>Pse</b>	0.03±0.02	1	<1	>1	<1	<1
<b>Yer</b>	>1	>1	1	>1	>1	>1
<b>Pan</b>	<1	<1	<1	1	>1	<1
<b>Kle</b>	0.94±0.07	>1	<1	>1	1	>1
<b>Aci</b>	>1	>1	<1	>1	>1	1

## References and Notes

1. R. M. May, How many species are there on Earth? *Science* **241**, 1441–1449 (1988).  
[doi:10.1126/science.241.4872.1441](https://doi.org/10.1126/science.241.4872.1441) [Medline](#)
2. J. J. Faith, J. L. Guruge, M. Charbonneau, S. Subramanian, H. Seedorf, A. L. Goodman, J. C. Clemente, R. Knight, A. C. Heath, R. L. Leibler, M. Rosenbaum, J. I. Gordon, The long-term stability of the human gut microbiota. *Science* **341**, 1237439 (2013).  
[doi:10.1126/science.1237439](https://doi.org/10.1126/science.1237439) [Medline](#)
3. E. Benincà, B. Ballantine, S. P. Ellner, J. Huisman, Species fluctuations sustained by a cyclic succession at the edge of chaos. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6389–6394 (2015).  
[doi:10.1073/pnas.1421968112](https://doi.org/10.1073/pnas.1421968112) [Medline](#)
4. G. W. Luck, G. C. Daily, P. R. Ehrlich, Population diversity and ecosystem services. *Trends Ecol. Evol.* **18**, 331–336 (2003). [doi:10.1016/S0169-5347\(03\)00100-9](https://doi.org/10.1016/S0169-5347(03)00100-9)
5. K. S. McCann, The diversity-stability debate. *Nature* **405**, 228–233 (2000).  
[doi:10.1038/35012234](https://doi.org/10.1038/35012234) [Medline](#)
6. A. R. Ives, S. R. Carpenter, Stability and diversity of ecosystems. *Science* **317**, 58–62 (2007).  
[doi:10.1126/science.1133258](https://doi.org/10.1126/science.1133258) [Medline](#)
7. J. Friedman, L. M. Higgins, J. Gore, Community structure follows simple assembly rules in microbial microcosms. *Nat. Ecol. Evol.* **1**, 109 (2017). [doi:10.1038/s41559-017-0109](https://doi.org/10.1038/s41559-017-0109) [Medline](#)
8. O. S. Venturelli, A. C. Carr, G. Fisher, R. H. Hsu, R. Lau, B. P. Bowen, S. Hromada, T. Northen, A. P. Arkin, Deciphering microbial interactions in synthetic human gut microbiome communities. *Mol. Syst. Biol.* **14**, e8157 (2018). [doi:10.15252/msb.20178157](https://doi.org/10.15252/msb.20178157) [Medline](#)
9. F. K. Balagaddé, H. Song, J. Ozaki, C. H. Collins, M. Barnet, F. H. Arnold, S. R. Quake, L. You, A synthetic *Escherichia coli* predator-prey ecosystem. *Mol. Syst. Biol.* **4**, 187 (2008). [doi:10.1038/msb.2008.24](https://doi.org/10.1038/msb.2008.24) [Medline](#)
10. B. Blasius, L. Rudolf, G. Weithoff, U. Gaedke, G. F. Fussmann, Long-term cyclic persistence in an experimental predator-prey system. *Nature* **577**, 226–230 (2020).  
[doi:10.1038/s41586-019-1857-0](https://doi.org/10.1038/s41586-019-1857-0) [Medline](#)
11. G. F. Fussmann, S. P. Ellner, K. W. Shertzer, N. G. Hairston Jr., Crossing the hopf bifurcation in a live predator-prey system. *Science* **290**, 1358–1360 (2000).  
[doi:10.1126/science.290.5495.1358](https://doi.org/10.1126/science.290.5495.1358) [Medline](#)
12. W. Shou, S. Ram, J. M. G. Vilar, Synthetic cooperation in engineered yeast populations. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1877–1882 (2007). [doi:10.1073/pnas.0610575104](https://doi.org/10.1073/pnas.0610575104) [Medline](#)
13. E. Benincà, J. Huisman, R. Heerkloss, K. D. Jöhnk, P. Branco, E. H. Van Nes, M. Scheffer, S. P. Ellner, Chaos in a long-term experiment with a plankton community. *Nature* **451**, 822–825 (2008). [doi:10.1038/nature06512](https://doi.org/10.1038/nature06512) [Medline](#)
14. C. Ratzke, J. Barrere, J. Gore, Strength of species interactions determines biodiversity and stability in microbial communities. *Nat. Ecol. Evol.* **4**, 376–383 (2020).

[doi:10.1038/s41559-020-1099-4](https://doi.org/10.1038/s41559-020-1099-4) [Medline](#)

15. J. E. Goldford, N. Lu, D. Bajić, S. Estrela, M. Tikhonov, A. Sanchez-Gorostiaga, D. Segré, P. Mehta, A. Sanchez, Emergent simplicity in microbial community assembly. *Science* **361**, 469–474 (2018). [doi:10.1126/science.aat1168](https://doi.org/10.1126/science.aat1168) [Medline](#)
16. R. M. May, Will a large complex system be stable? *Nature* **238**, 413–414 (1972).  
[doi:10.1038/238413a0](https://doi.org/10.1038/238413a0) [Medline](#)
17. S. Allesina, S. Tang, Stability criteria for complex ecosystems. *Nature* **483**, 205–208 (2012).  
[doi:10.1038/nature10832](https://doi.org/10.1038/nature10832) [Medline](#)
18. U. Bastolla, M. A. Fortuna, A. Pascual-García, A. Ferrera, B. Luque, J. Bascompte, The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature* **458**, 1018–1020 (2009). [doi:10.1038/nature07950](https://doi.org/10.1038/nature07950) [Medline](#)
19. E. Thébault, C. Fontaine, Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science* **329**, 853–856 (2010).  
[doi:10.1126/science.1188321](https://doi.org/10.1126/science.1188321) [Medline](#)
20. A. Mougi, M. Kondoh, Diversity of interaction types and ecological community stability. *Science* **337**, 349–351 (2012). [doi:10.1126/science.1220529](https://doi.org/10.1126/science.1220529) [Medline](#)
21. G. Bunin, Ecological communities with Lotka-Volterra dynamics. *Phys. Rev. E* **95**, 042414 (2017). [doi:10.1103/PhysRevE.95.042414](https://doi.org/10.1103/PhysRevE.95.042414) [Medline](#)
22. M. Opper, S. Diederich, Phase transition and 1/f noise in a game dynamical model. *Phys. Rev. Lett.* **69**, 1616–1619 (1992). [doi:10.1103/PhysRevLett.69.1616](https://doi.org/10.1103/PhysRevLett.69.1616) [Medline](#)
23. D. A. Kessler, N. M. Shnerb, Generalized model of island biodiversity. *Phys. Rev. E* **91**, 042705 (2015). [doi:10.1103/PhysRevE.91.042705](https://doi.org/10.1103/PhysRevE.91.042705) [Medline](#)
24. J. Huisman, F. J. Weissing, Biodiversity of plankton by species oscillations and chaos. *Nature* **402**, 407–410 (1999). [doi:10.1038/46540](https://doi.org/10.1038/46540)
25. M. T. Pearce, A. Agarwala, D. S. Fisher, Stabilization of extensive fine-scale diversity by ecologically driven spatiotemporal chaos. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 14572–14583 (2020). [doi:10.1073/pnas.1915313117](https://doi.org/10.1073/pnas.1915313117) [Medline](#)
26. F. Roy, M. Barbier, G. Biroli, G. Bunin, Complex interactions can create persistent fluctuations in high-diversity ecosystems. *PLOS Comput. Biol.* **16**, e1007827 (2020).  
[doi:10.1371/journal.pcbi.1007827](https://doi.org/10.1371/journal.pcbi.1007827) [Medline](#)
27. J. C. Allen, W. M. Schaffer, D. Rosko, Chaos reduces species extinction by amplifying local population noise. *Nature* **364**, 229–232 (1993). [doi:10.1038/364229a0](https://doi.org/10.1038/364229a0) [Medline](#)
28. R. V. Solé, D. Alonso, A. McKane, Self-organized instability in complex ecosystems. *Philos. Trans. R. Soc. London Ser. B* **357**, 667–671 (2002). [doi:10.1098/rstb.2001.0992](https://doi.org/10.1098/rstb.2001.0992) [Medline](#)
29. J. D. Touboul, A. C. Staver, S. A. Levin, On the complex dynamics of savanna landscapes. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E1336–E1345 (2018). [doi:10.1073/pnas.1712356115](https://doi.org/10.1073/pnas.1712356115) [Medline](#)
30. A. Hastings, C. L. Hom, S. Ellner, P. Turchin, H. C. J. Godfray, Chaos in ecology: Is mother nature a strange attractor? *Annu. Rev. Ecol. Syst.* **24**, 1–33 (1993).  
[doi:10.1146/annurev.es.24.110193.000245](https://doi.org/10.1146/annurev.es.24.110193.000245)

31. D. S. Maynard, Z. R. Miller, S. Allesina, Predicting coexistence in experimental ecological communities. *Nat. Ecol. Evol.* **4**, 91–100 (2020). [doi:10.1038/s41559-019-1059-z](https://doi.org/10.1038/s41559-019-1059-z)
32. Materials and methods are available as supplementary materials.
33. C. Ratzke, J. Gore, Modifying and reacting to the environmental pH can drive bacterial interactions. *PLOS Biol.* **16**, e2004248 (2018). [doi:10.1371/journal.pbio.2004248](https://doi.org/10.1371/journal.pbio.2004248) [Medline](#)
34. P. Chesson, Multispecies competition in variable environments. *Theor. Popul. Biol.* **45**, 227–276 (1994). [doi:10.1006/tpbi.1994.1013](https://doi.org/10.1006/tpbi.1994.1013)
35. M. Barbier, J. F. Arnoldi, G. Bunin, M. Loreau, Generic assembly patterns in complex ecological communities. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 2156–2161 (2018). [doi:10.1073/pnas.1710352115](https://doi.org/10.1073/pnas.1710352115) [Medline](#)
36. I. Dalmedigos, G. Bunin, Dynamical persistence in high-diversity resource-consumer communities. *PLOS Comput. Biol.* **16**, e1008189 (2020). [doi:10.1371/journal.pcbi.1008189](https://doi.org/10.1371/journal.pcbi.1008189) [Medline](#)
37. J. D. O’Sullivan, J. C. D. Terry, A. G. Rossberg, Intrinsic ecological dynamics drive biodiversity turnover in model metacommunities. *Nat. Commun.* **12**, 3627 (2021). [doi:10.1038/s41467-021-23769-7](https://doi.org/10.1038/s41467-021-23769-7) [Medline](#)
38. Y. Yonatan, G. Amit, J. Friedman, A. Bashan, Complexity-stability trade-off in empirical microbial ecosystems. *Nat. Ecol. Evol.* **6**, 693–700 (2022). [doi:10.1038/s41559-022-01745-8](https://doi.org/10.1038/s41559-022-01745-8) [Medline](#)
39. J. Hu, D. R. Amor, M. Barbier, G. Bunin, J. Gore, Code from the manuscript “Emergent phases of ecological diversity and dynamics mapped in microcosms.” Zenodo (2022); <https://doi.org/10.5281/zenodo.7017202>.
40. B. J. Callahan, K. Sankaran, J. A. Fukuyama, P. J. McMurdie, S. P. Holmes, Bioconductor workflow for microbiome data analysis: From raw reads to community analyses. *F1000 Res.* **5**, 1492 (2016). [doi:10.12688/f1000research.8986.2](https://doi.org/10.12688/f1000research.8986.2) [Medline](#)
41. F. Madeira, Y. M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A. R. N. Tivey, S. C. Potter, R. D. Finn, R. Lopez, The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019). [doi:10.1093/nar/gkz268](https://doi.org/10.1093/nar/gkz268) [Medline](#)
42. L. W. Hugerth, A. F. Andersson, Analysing microbial community composition through amplicon sequencing: From sampling to hypothesis testing. *Front. Microbiol.* **8**, 1561 (2017). [doi:10.3389/fmicb.2017.01561](https://doi.org/10.3389/fmicb.2017.01561) [Medline](#)
43. J. J. Qian, E. Akçay, The balance of interaction types determines the assembly and stability of ecological communities. *Nat. Ecol. Evol.* **4**, 356–365 (2020). [doi:10.1038/s41559-020-1121-x](https://doi.org/10.1038/s41559-020-1121-x) [Medline](#)
44. F. Roy, G. Biroli, G. Bunin, C. Cammarota, Numerical implementation of dynamical mean field theory for disordered systems: Application to the Lotka-Volterra model of ecosystems. *J. Phys. A Math. Theor.* **52**, 484001 (2019). [doi:10.1088/1751-8121/ab1f32](https://doi.org/10.1088/1751-8121/ab1f32)
45. P. Bizeul, J. Najim, Positive solutions for large random linear systems. *Proc. Am. Math. Soc.*

**149**, 2333–2348 (2021). [doi:10.1090/proc/15383](https://doi.org/10.1090/proc/15383)

46. A. Altieri, F. Roy, C. Cammarota, G. Biroli, Properties of equilibria and glassy phases of the random Lotka-Volterra model with demographic noise. *Phys. Rev. Lett.* **126**, 258301 (2021). [doi:10.1103/PhysRevLett.126.258301](https://doi.org/10.1103/PhysRevLett.126.258301) [Medline](#)
47. L. Stone, The feasibility and stability of large complex biological networks: A random matrix approach. *Sci. Rep.* **8**, 8246 (2018). [doi:10.1038/s41598-018-26486-2](https://doi.org/10.1038/s41598-018-26486-2) [Medline](#)