Stabilizing role of seed banks and the maintenance of bacterial diversity: Supplementary Information

Nathan I. Wisnoski and Jay T. Lennon

SUPPLEMENTAL METHODS

Sequencing and bioinformatics: After extracting nucleic acids, we used DNase (Invitrogen) to remove DNA from the RNA extractions and then synthesized cDNA with SuperScript III First Strand Synthesis kit and random hexamer primers (Invitrogen). To amplify the 16S rRNA gene (DNA) and transcripts (cDNA), we used barcoded V4 primers (515F and 806R) designed for the Illumina MiSeq platform (Caporaso *et al.* 2012). We then purified the PCR products with AMPure XP, quantified DNA concentrations using PicoGreen, and pooled samples at 10 ng per sample. The resulting libraries were sequenced on an Illumina MiSeq at the Indiana University Center for Genomic and Bioinformatics Sequencing Facility using 250 × 250 bp paired-end reads (Reagent Kit v2). Sequences were subsequently processed using the software package mothur (version 1.41.1) (Schloss *et al.* 2009). We assembled contigs, removed low quality sequences (minimum score of 35), aligned sequences to the SILVA Database (version 132) (Quast *et al.* 2013), removed chimeras using the VSEARCH algorithm (Rognes *et al.* 2016), and created 97% similar operational taxonomic units (OTUs) using the OptiClust algorithm (Westcott & Schloss 2017), and classified sequences with the RDP taxonomy (Cole *et al.* 2009).

Estimating cell abundance: To support inferences about biotic interactions made from relative abundances and negative frequency dependence, we estimated the annual variability in overall community density using flow cytometry for the first year 60 weeks of the time series. We

filtered lake water samples using a 5 µm syringe filter to remove large particles. We stained 1 ml of the sample with 1 µl of eFluor660 (eBioscience, UK), a fixed viability dye that penetrates ruptured cell wells and stains dead cells, at room temperature for 30 minutes. After 30 minutes, cells were fixed with 13.5 µl of 37% formalin. Samples were frozen at -80 ºC. To enumerate overall cell density, we thawed samples on ice in the dark (to preserve eFluor660 staining), then transferred each sample to a 15x75 mm clamp cap tube. In the tube, we added two drops of cell permeable Hoechst 33342 (Chazotte 2011), which stains DNA, 5 µl of 1:1000 cell permeable Pyronin-y stain, which stains RNA, and 1 µ of a bead standard (final concentration of 10^6 beads/ml) for cell counting.

We collected 50,000 bead events on the LSR II flow cytometer using the BD FACSDiva Software (v. 6.1.3) in the Indiana University Flow Cytometry Core Center (samples run by director of the facility, C. Hassel). Bead events were determined using a size-based threshold based on SSC (side scatter) and FSC (forward scatter). We analyzed data using R v.4.0.5 (R Core Team 2020), using the packages "flowCore" (Ellis *et al.* 2020) and "flowStats" (Hahne *et al.* 2020). To estimate total community density, we performed a hyperbolic arc-sine transformation of the channels reading the Hoechst DNA stain and the eFluor660 viability stain. We then created gates, such that cells which stained positive with Hoechst (fluorescence > 9), but not for eFluor660 (8 < eFluor fluorescence < 10) were considered potentially live bacteria. We set the lower thresholds based on the background fluorescence calculated for cell-free controls. The distribution of Hoechst fluorescence was bimodal because of residual background fluorescence for Hoechst. We then used the 'rangeGate' procedure from the 'flowStats' package to select only the population of events with high fluorescence values for the Hoechst stain (to further remove low-fluorescence background noise) using an algorithm that split the histogram of fluorescence

intensity at its lowest point density (typically around fluorescence of 9), then filtered the data to keep only the higher fluorescence events. We then counted the events in this category as true cells.

We used bead data to estimate cell density in the community. Because the flow cytometer stopped after reading 50,000 bead counts, the data is standardized per 50,000 beads. Using the known concentration of beads (10^6 beads/ml) and assuming a homogeneous distribution in the tube (tubes were shaken well before reading), we estimated the number of cells per ml. Cell density was estimated as observed cell counts per 50,000 beads, multiplied by 10^6 beads per ml, to obtain cells per ml. We then visualized estimated cell counts for roughly the first year of sampling (Fig. S1).

Differential response to environment: Because temperature was associated with the major axis of community variation in the RDA and is known to place important constraints on bacterial metabolism, nutrient uptake, and reproduction, we analyzed whether variation in temperature may have facilitated temporal niche partitioning. For each of the persistent OTUs, we compared its relative abundance in the community with the current temperature using linear regression with a quadratic term to accommodate nonlinear responses to temperature. We then compared differences among taxa in how their commonness or rarity varied along the observed temperature gradient (Fig. S5).

REFERENCES

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., *et al.* (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, 6, 1621–1624.

Chazotte, B. (2011). Labeling Nuclear DNA with Hoechst 33342. *Cold Spring Harb Protoc*, 2011, pdb.prot5557.

Cole, J.R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R.J., *et al.* (2009). The Ribosomal Database Project: Improved alignments and new tools for rRNA analysis. *Nucleic Acids Research*, 37, 141–145.

Ellis, B., Haaland, P., Hahne, F., Le Meur, N., Gopalakrishnan, N., Spidlen, J., *et al.* (2020). *flowCore: flowCore: Basic structures for flow cytometry data. R package version 2.2.0.*

Hahne, F., Gopalakrishnan, N., Khodabakhshi, A.H., Wong, C.-J. & Lee, K. (2020). *flowStats: Statistical methods for the analysis of flow cytometry data. R package version 4.2.0. http://www.github.com/RGLab/flowStats*.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., *et al.* (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 41, 590–596.

R Core Team. (2020). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.*

Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, 4, e2584–e2584.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., *et al.* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, 75, 7537–7541.

Westcott, S.L. & Schloss, P.D. (2017). OptiClust, an improved method for assigning amplicon-based sequence data to operational taxonomic units. *mSphere*, 2, e00073-17.

SUPPLEMENTAL TABLES AND FIGURES

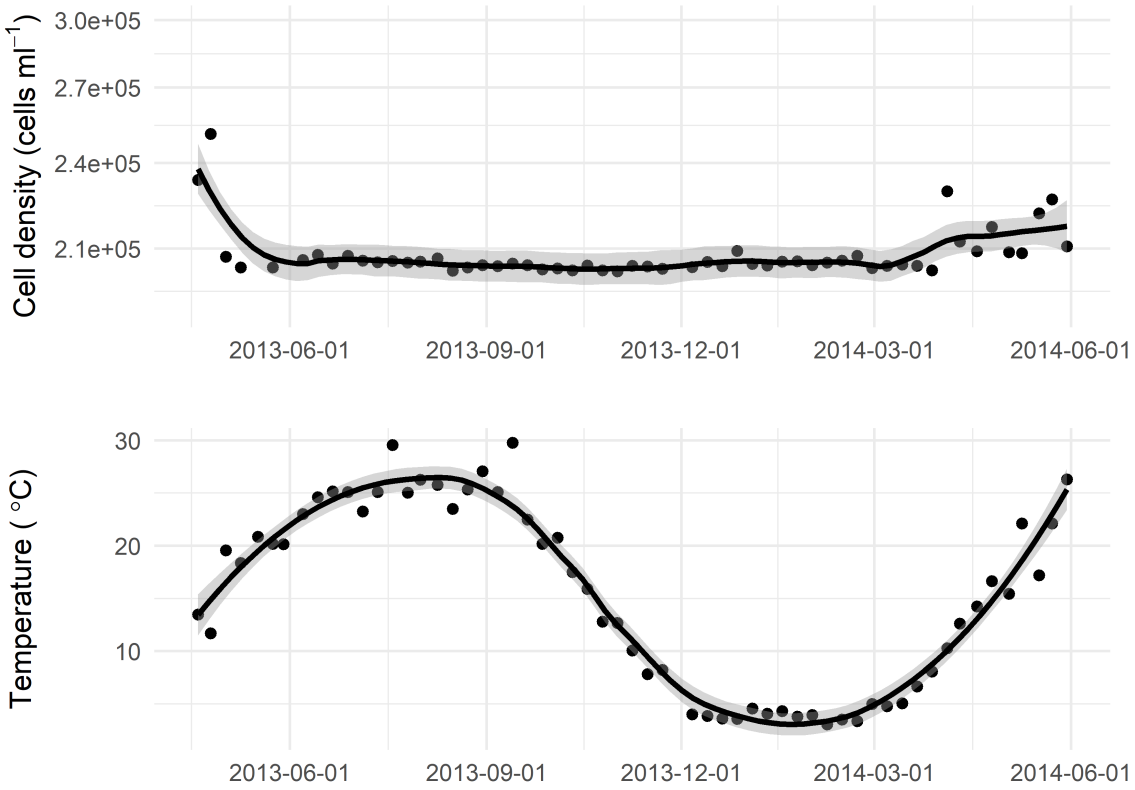
Table S1. Operational taxonomic units (OTUs) that were classified as persistent in the bacterioplankton community based on being detected in ≥80% of the total (i.e., DNA) community samples. The table is sorted by Julian date of max growth.

OTU	Class	Max growth rate (d^{-1})	Date of max growth
Otu00045	Betaproteobacteria	0.622	2014-01-03
Otu00039	Betaproteobacteria	0.486	2014-02-14
Otu00067	Betaproteobacteria	0.314	2014-02-14
Otu00102	Betaproteobacteria	0.442	2015-02-28
Otu00105	Alphaproteobacteria	0.636	2014-02-28
Otu00183	unclassified	0.4	2014-02-21
Otu00065	Sphingobacteriia	0.626	2014-03-21
Otu00129	Alphaproteobacteria	0.329	2014-03-28
Otu00012	Betaproteobacteria	0.892	2014-04-18
Otu00016	Actinobacteria	0.553	2014-04-18
Otu00017	Actinobacteria	0.79	2015-04-04
Otu00021	Gammaproteobacteria	0.815	2014-04-18
Otu00024	Bacteroidetes unclassified	0.644	2015-04-11
Otu00048	Verrucomicrobiae	0.756	2015-04-11
Otu00055	Flavobacteriia	0.564	2014-04-18
Otu00064	Alphaproteobacteria	0.527	2013-04-25
Otu00148	unclassified	0.534	2013-04-25
Otu00172	Gammaproteobacteria	0.442	2015-04-11
Otu00219	Betaproteobacteria	0.52	2014-04-18

Otu00250	Actinobacteria	0.329	2014-04-25
Otu00002	Actinobacteria	0.319	2015-05-03
Otu00008	Actinobacteria	0.365	2013-05-09
Otu00014	Actinobacteria	0.41	2015-05-03
Otu00031	Cytophagia	0.428	2014-05-09
Otu00049	Actinobacteria	0.445	2014-05-17
Otu00051	Flavobacteriia	0.689	2013-05-09
Otu00062	Flavobacteriia	0.665	2013-05-09
Otu00113	Bacteroidetes unclassified	0.564	2013-05-09
Otu00116	Betaproteobacteria	0.607	2014-05-09
Otu00151	Betaproteobacteria	0.495	2013-05-17
Otu00200	unclassified	0.454	2015-05-23
Otu00208	Betaproteobacteria	0.564	2014-05-09
Otu00022	Opitutae	0.771	2013-06-14
Otu00058	Armatimonadia	0.588	2013-06-21
Otu00066	Betaproteobacteria	0.623	2013-06-07
Otu00083	Flavobacteriia	0.773	2015-06-06
Otu00095	Betaproteobacteria	0.413	2015-06-06
Otu00098	Betaproteobacteria	0.58	2013-06-14
Otu00123	Sphingobacteriia	0.495	2014-06-20
Otu00194	Deltaproteobacteria	0.698	2014-06-13
Otu00196	Actinobacteria	0.442	2013-06-07
Otu00294	Alphaproteobacteria	0.465	2013-06-21
Otu00004	Actinobacteria	0.35	2015-07-11
Otu00009	Gammaproteobacteria	1.103	2013-07-26

Otu00010	Proteobacteria unclassified	0.405	2015-07-11
Otu00011	Betaproteobacteria	0.754	2015-07-18
Otu00038	Actinobacteria	0.508	2015-07-11
Otu00195	Actinobacteria	0.396	2014-07-18
Otu00292	Alphaproteobacteria	0.495	2015-07-26
Otu00019	Cytophagia	0.527	2013-08-01
Otu00020	Betaproteobacteria	0.428	2013-08-01
Otu00029	Actinobacteria	0.428	2013-08-23
Otu00036	Alphaproteobacteria	0.527	2013-08-16
Otu00037	Actinobacteria	0.534	2014-08-29
Otu00052	Alphaproteobacteria	0.413	2013-08-09
Otu00073	Betaproteobacteria	0.396	2013-08-01
Otu00076	Actinobacteria	0.413	2014-08-08
Otu00087	Betaproteobacteria	0.504	2014-08-23
Otu00112	Alphaproteobacteria	0.355	2014-08-23
Otu00226	Opitutae	0.442	2015-08-02
Otu00026	Betaproteobacteria	0.41	2015-09-02
Otu00033	Alphaproteobacteria	0.349	2013-09-13
Otu00005	Sphingobacteriia	0.534	2014-10-04
Otu00015	Actinobacteria	0.396	2014-10-17
Otu00034	Alphaproteobacteria	0.343	2014-10-04
Otu00060	Betaproteobacteria	0.773	2013-10-25
Otu00082	Bacteroidetes unclassified	0.57	2014-10-04
Otu00154	Alphaproteobacteria	0.476	2014-10-04
Otu00158	Gammaproteobacteria	0.684	2013-10-04

Otu00192	unclassified	0.23	2014-10-17
Otu00001	Betaproteobacteria	0.286	2013-11-15
Otu00007	Betaproteobacteria	0.471	2013-11-15
Otu00018	Gammaproteobacteria	0.8	2013-11-15
Otu00047	Betaproteobacteria	0.442	2013-11-15
Otu00056	Cytophagia	0.587	2013-11-15
Otu00077	Flavobacteriia	0.585	2014-11-21
Otu00109	Actinobacteria	0.442	2013-11-15
Otu00118	Actinobacteria	0.377	2013-11-22
Otu00177	Proteobacteria unclassified	0.46	2013-11-15
Otu00198	Betaproteobacteria	0.691	2013-11-15
Otu00217	Proteobacteria unclassified	0.486	2013-11-22



102

103 **Figure S1**. Community density remains relatively stable throughout the year. Cell density was

104 estimated by flow cytometry. Density peaks slightly during the spring warm up, but overall, total

105 density remains relatively stable.

106

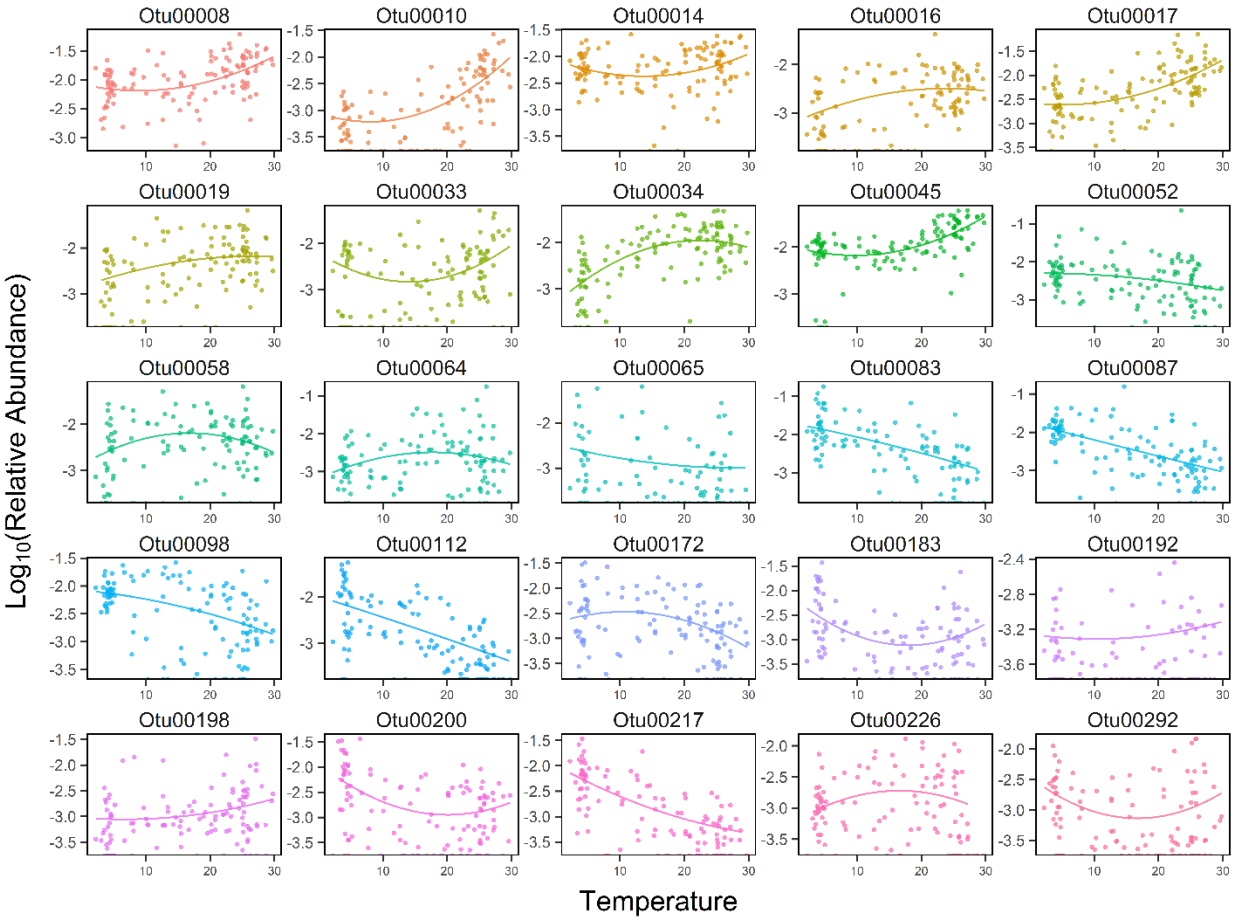
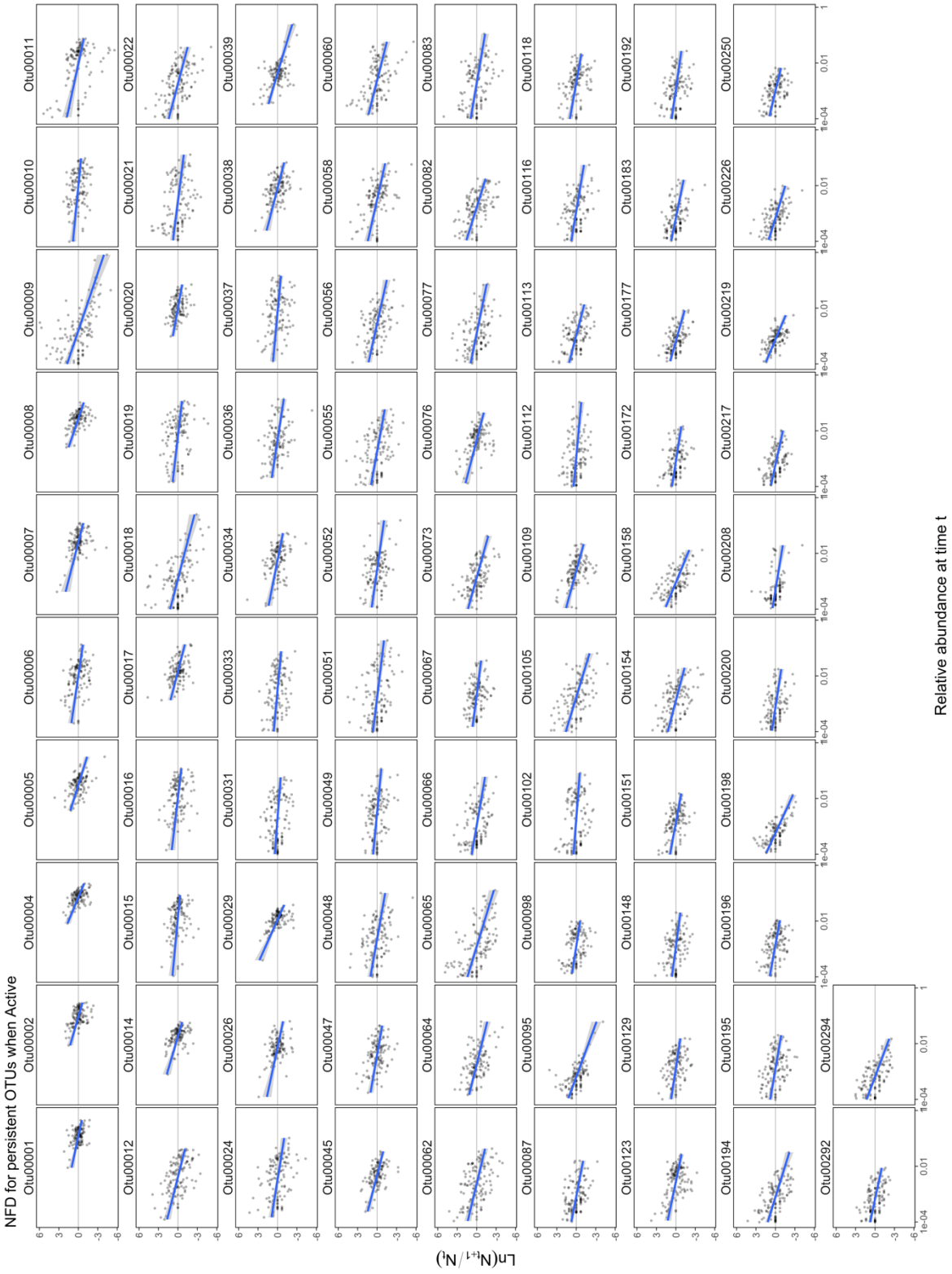


Figure S2. Differential responses of persistent taxa along a temperature gradient. Points indicate the relative abundances of a random sample of 25 of the 82 persistent OTUs in the active portion of the community. Fits are linear regression models with quadratic terms to capture nonlinearities along the temperature gradient. Note that some taxa increase in relative abundance with higher temperatures, while other taxa increase in relative abundance at lower temperatures. Others display unimodal responses to temperature.



115

116 **Figure S3**. Negative frequency dependence for the 82 persistent taxa in the active community.

117 The linear regression lines in this figure are the same as depicted in Fig. S3A.

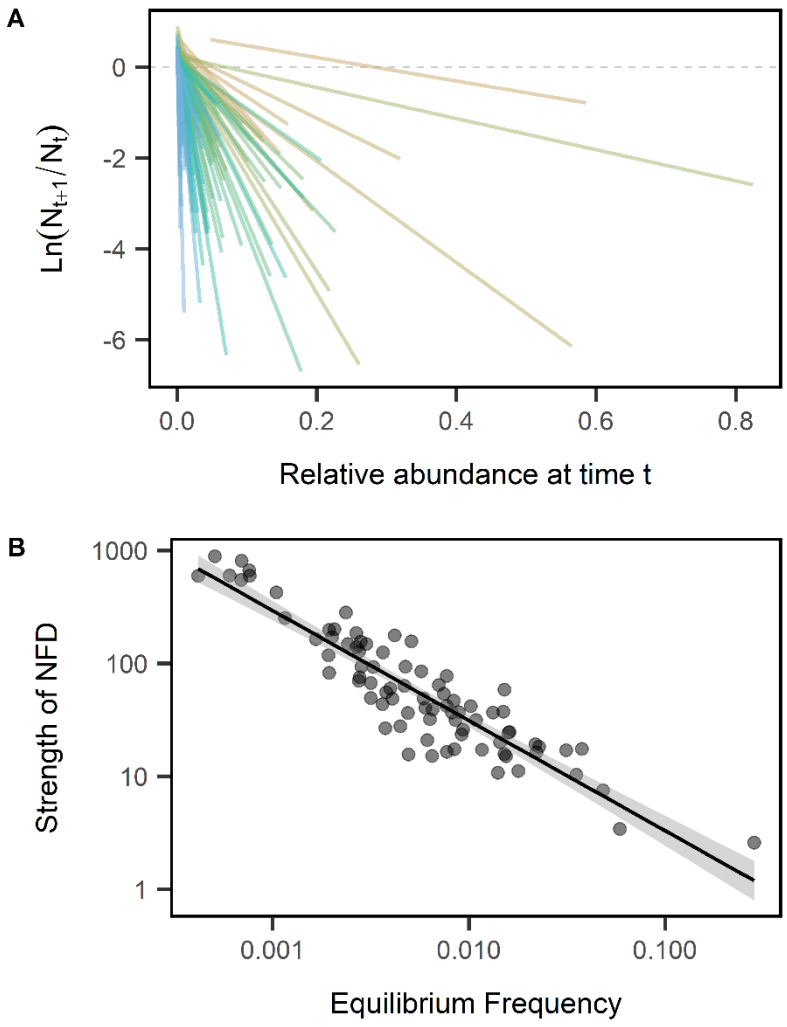
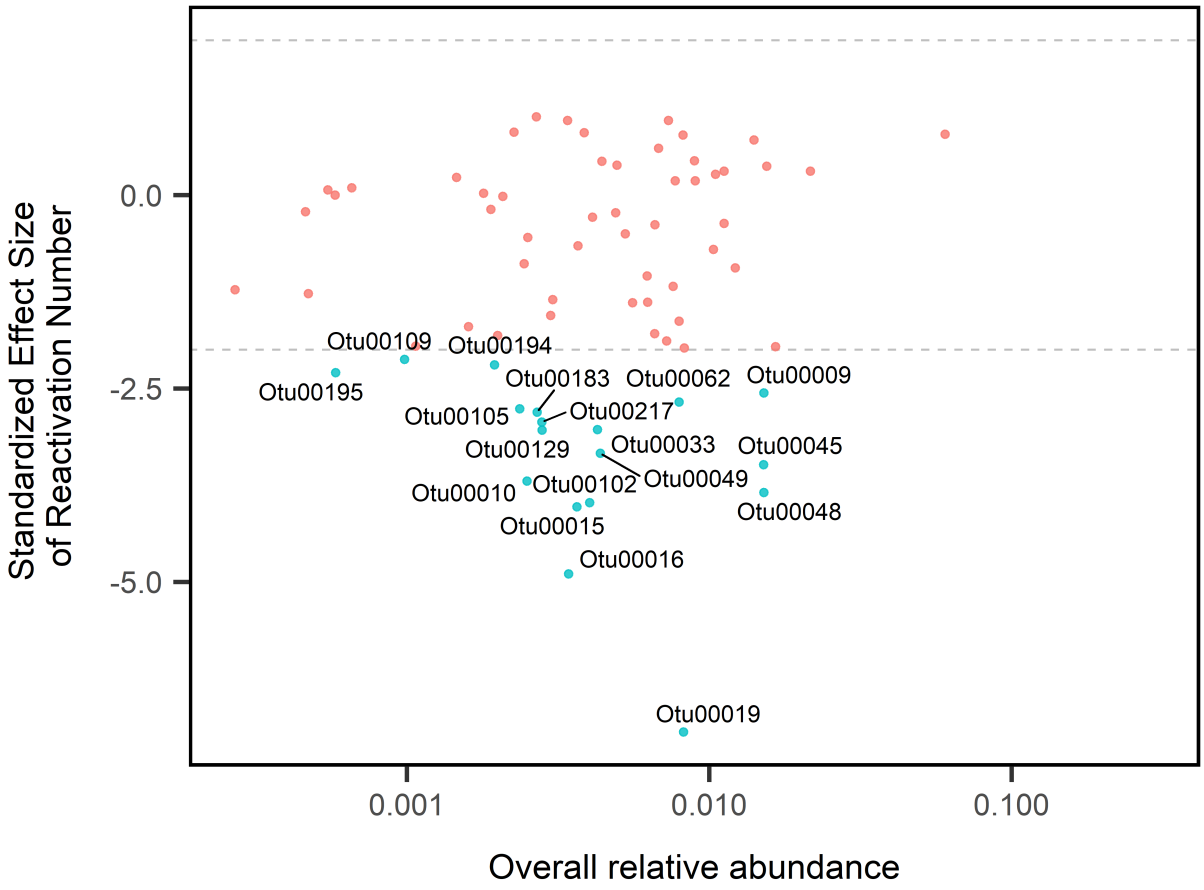


Figure S4. Negative frequency dependence (NFD) in the active portion of the community for the 82 persistent bacterial taxa. (A) Relationship between the rate of change of an OTU and its relative abundance. Depicted in this graph are simple linear-regression fits for the 82 taxa individually (data points not shown to reduce clutter). Negative relationships indicate NFD growth and variation in slopes indicates variation in the strength of NFD. (B) Rare taxa (lower equilibrium frequencies) exhibit stronger NFD, while common taxa (higher equilibrium frequency) have weaker NFD.



127

128 **Figure S5**. Comparison of observed number of reactivations to the number of reactivations
129 expected for the stochastic null model simulations (n = 1000). The deviation between
130 observations and null distributions were quantitatively compared by calculating a standardized
131 effect size for each OTU. We plotted standardized effect sizes for each persistent OTU in
132 relation to its overall relative abundance in the community, and labeled the OTUs with
133 significantly fewer observed reactivations than expected by chance (i.e., two standard deviations
134 below the range of expected values for that taxon).

135