UNIVERSITY OF CALIFORNIA,
IRVINE

Microbiomes: biogeographic patterns and the influence of dispersal

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Kendra Emily Walters

Dissertation Committee:
Professor Jennifer B.H. Martiny, Chair
Assistant Professor Cascade J.B. Sorte
Professor Steven D. Allison

2021

DEDICATION

To

my parents and sister

in recognition of your love and kindness

I will honor all life
wherever and in whatever form it may dwell.

(Diane Ackerman, “School Prayer”)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Thank you to my advisor, Jennifer Martiny, for being an amazing mentor and scientist. I feel incredibly privileged for the opportunity to work with you for these last five years – for the support and encouragement you provided, for our brainstorming and problem-solving sessions, and for sharing the joy of science with you. I cannot imagine a better graduate school experience. Thank you to my two committee members, Cascade Sorte and Steve Allison, for your support and guidance. I always looked forward to committee meetings and appreciate the time you took to help improve this research.

To the fellow members of Jen's lab, past and present, thank you for creating a wonderful community to work in. We had the perfect balance of serious science time and goofy fun time. Of course, I want to give an additional and huge thank you to Claudia Weihe. You were not only immensely helpful to all my lab endeavors and helped me develop all the lab skills that I needed, but you were also an immediate friend. For that, I am incredibly grateful.

Thank you to all the people in the department and university who helped make graduate school enjoyable. To the grad students and post-docs in microbial group, you are all incredible scientists, and I learned a lot from you. Thank you to Phoot Phungus, the best trivia-winning, escape-room-completing, secret-snowflake-gift-giving friends around. You are always there to celebrate the wins and commiserate the losses. You listened to all my presentations and have provided feedback on way too many drafts. I feel so lucky to have found such a wonderful group of people.

To John Powers, my amazing partner, thank you for providing both emotional and scientific support during graduate school. I can always count on you to lend an ear to listen to frustrations or help debug my code. You also make sure I take time away from research to get outside and explore nature, which brings much needed balance into my life.

Last, but certainly not least, thank you to my family. Not only did your unending support, love, and encouragement help me through graduate school, but your passion for the outdoors and thoughtfulness towards balance in the world inspired my love for ecology in the first place.

This work was funded by the National Science Foundation Graduate Research Fellowship Program (DGE-1321846) to KEW and research support through grants from the US Department of Energy (DE-SC0016410; DE-SC0020382) to JBHM. Support was also provided by the University of California, Irvine and the Ridge to Reef National Science Foundation Research Traineeship, award DGE-1735040.

The text of Chapter 1 of this dissertation is a reprint of the material as it appears in PLOS One. This material is distributed under the Creative Commons Attribution (CC BY) license, which "permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited." Jennifer B.H. Martiny, a co-author listed in this publication, directed and supervised research, which forms the basis for this dissertation.

> Walters, K. E., & Martiny, J. B. (2020). Alpha-, beta-, and gamma-diversity of bacteria varies across habitats. PLOS One, 15(9), e0233872.

VITA

Kendra Emily Walters

EDUCATION

2021	Ph.D., Biological Sciences, University of California – Irvine Advisor: Jennifer Martiny
2019	M.S., Biological Sciences, University of California – Irvine
2016	B.S., Biology and Geology, University of Oregon

PUBLICATIONS

K.E. Walters and J.B.H. Martiny. 2020. Alpha-, beta-, and gamma-diversity of bacteria varies across habitats. PLOS One. doi: 10.1371/journal.pone.0233872

J.B.H. Martiny and **K.E. Walters**. 2018. Towards a natural history of soil bacterial communities. Trends in Microbiology. doi: 10.1016/j.tim.2018.02.010

NON-ACADEMIC SCIENCE WRITING

2020	Raising Our Voices for Diversity in the Geosciences. Lucila Houttuijn Bloemendaal, Katarena Matos, **Kendra Walters**, and Aditi Sengupta. Eos.
2019-2020	Fourteen radio stories each covering a recent scientific paper across all disciplines. **Kendra Walters**. Loh Down on Science.

SELECTED FELLOWSHIPS AND AWARDS

2021	UCI School of Biological Sciences Graduate Fellowship Award
2019	Loh Down on Science Fellowship
	AGU Voices for Science Advocate
2018	NSF Graduate Research Fellowship
2016	UCI Graduate Dean's Recruitment Fellowship
	UCI Graduate Fellowship
	UCI Provost Ph.D. Fellowship

SELECTED TEACHING EXPERIENCES

2020	Introduction to Ecology, UCI, Instructor of Record
2019	Microbial Community: Data Analysis Workshop, UCI Microbiome Initiative
2019	Characterizing Microbiome Diversity Workshops, UCI Microbiome Initiative
2017-2018	From Organisms to Ecosystems, UCI

ABSTRACT OF THE DISSERTATION

Microbiomes: biogeographic patterns and the influence of dispersal

by

Kendra Emily Walters

Doctor of Philosophy in Biological Sciences

University of California, Irvine, 2021

Professor Jennifer B.H. Martiny, Chair

Bacteria are essential parts of ecosystems and are the most diverse organisms on the planet. Yet, our understanding of their biogeographic patterns and the processes that drive those patterns is underdeveloped. My dissertation provides a comprehensive analysis of bacterial biogeographic patterns and investigates how dispersal into soil impacts microbiome assembly. To establish global biogeographic patterns, I analyzed alpha-, beta-, and gamma-diversity of bacterial assemblages using 11,680 samples compiled by the Earth Microbiome Project (Chapter 1). This study compared bacterial diversity trends across habitats and provided a big-picture analysis that linked multiple smaller-scale studies. From this study, I found that soils contained the highest bacterial richness within a single sample, prompting a follow-up field experiment to study how dispersal may drive the high diversity found in soils.

Recent evidence suggests that, similar to larger organisms, dispersal is a key driver of microbiome assembly and diversity. Dispersal may be particularly important to leaf litter, the top layer of soil, which is exposed to microbial immigration from multiple sources and plays a key role in nutrient cycling in soils. However, our understanding of the rates and taxonomic composition of microbial dispersal in natural environments is limited. To fill this gap, I

characterized the rate and composition of bacteria dispersing into leaf litter via three dispersal routes (through the air, from nearby vegetation, and up from the bulk soil) and simultaneously quantified the impact of those routes on soil microbiomes (Chapter 2). I found that dispersal from vegetation changed the microbiome composition and functioning, driving the decomposition of leaf litter. That result led to another field experiment to characterize immigration rates and composition of microorganisms dispersing from the plant communities of two ecosystems: a grassland and a shrubland (Chapter 3). Immigrating microorganisms were strongly correlated with the local vegetation, although only within 1 m – a much shorter dispersal distance than previously assumed. Overall, my dissertation highlights the high microbial diversity found in soils and suggests that dispersal into leaf litter from local vegetation plays a central role in the assembly of soil microbiomes.

INTRODUCTION

One of the most long-standing goals of ecology is to understand biogeography, or the distribution of biodiversity over space and time. In the 18th century, Johann Forster and Comte de Buffon were among the first to connect species distributions to climate and geographic distance. They laid the groundwork for scholars, such as Alfred Wallace, Charles Darwin, and Alexander von Humboldt, to establish the basis of modern biogeography in the 19th century. Through extensive fieldwork, they put forth ideas of natural selection, zoogeographic regions, and dispersal limitation into the world. Since then, biogeographic patterns have formed a strong foundation in the field of ecology and sparked curiosity about the drivers of biogeography.

Centuries later, microbiologists followed a similar path. Microbial biogeography may have begun between the 19th and 20th centuries when microbiologists began selecting for specific bacteria through enrichment techniques (Baas-Becking, 1934). However, it took another century for ecologists to organize individual studies and observations into formal hypotheses of microbial biogeography (Costello et al., 2012; Lindström & Langenheder, 2012; Jennifer B. Hughes Martiny et al., 2006; Nemergut et al., 2013; Ramette & Tiedje, 2007). These hypotheses have led to a wide swath of studies that aim to understand the processes that influence microbial communities.

One of the first objectives in the field of microbial biogeography has been to characterize patterns of biodiversity (Torsvik et al., 2002). However, establishing the biogeography of microbes is challenging. Most studies focus on a single geographic region or habitat, such as soil, sediment, or water (Fierer et al., 2012; Fierer & Jackson, 2006; Neufeld & Mohn, 2005), and arbitrary species definitions and methodological biases make comparing independent studies near impossible. This challenge motivated projects that aimed to amass enough data using one

method so that a vast number of samples are directly comparable (e.g., the Earth Microbiome Project and the Tara Oceans project). These data, along with smaller-scale studies, allow for the characterization of the biogeographic patterns of microorganisms like those of plants and animals (Gimmler et al., 2016; Thompson et al., 2017).

Biogeography provides a launch pad for understanding how diversity is generated and maintained. Current research highlights the importance of both stochastic and deterministic processes for shaping biogeographic patterns. Selection by the environment, a deterministic process, was known to be important since Beijerink and Winogradsky developed enrichment techniques around the turn of the 20th century. Recent research has identified environmental factors such as pH, salinity, and C:N ratios as key for explaining variation in microbial community composition (Auguet et al., 2010; Bates et al., 2011; Fierer & Jackson, 2006; Lozupone & Knight, 2007). However, selection is not the whole story. Stochastic processes, including dispersal and drift (Vellend, 2010), are also important drivers of microbial biogeography but have only been recently acknowledged. Toting microbes' small size and high abundance, Baas-Becking claimed in 1934 that "everything is everywhere" (Baas-Becking, 1934). That view persisted for the next seventy years (Finlay, 2002). However, we now understand that microbes are dispersal limited (Foissner, 2006; Jennifer B. Hughes Martiny et al., 2006) and that dispersal, and its influence on drift, impact community composition and biogeographic patterns.

However, the connection between dispersal and composition is not always clear because many studies use correlations between geographic distance and community composition as a proxy for stochastic processes (Hanson et al., 2012). Recent studies that manipulated microbial dispersal demonstrate the potential for dispersal to alter microbiome composition and

functioning (Albright & Martiny, 2018; S. E. Evans et al., 2019; Svoboda et al., 2018). That said, our knowledge of how microorganisms disperse in the field is limited as their small size and high abundance makes quantifying their dispersal logistically challenging (Nemergut et al., 2013). Specifically, microorganisms likely disperse through different routes (combinations of their dispersal source and the vector of transport) which are hypothesized to differentially influence resident communities (Cevallos-Cevallos et al., 2012; Lindström & Langenheder, 2012; Rime et al., 2016). However, no studies have yet studied microbial dispersal through different routes in the field.

My dissertation seeks to identify global biogeographic patterns of bacterial diversity and understand how dispersal drives those biogeographic patterns. To do so, I first used data compiled by the Earth Microbiome Project to analyze global patterns of alpha-, beta-, and gamma-diversity for bacterial communities across multiple habitats (Chapter 1; Walters and Martiny, 2020). From this work, I identified that soil contains the greatest number of bacterial species in a single sample. To understand the drivers of high diversity in soil, I designed a field experiment to characterize dispersal into soil through three routes (through the air, from nearby vegetation, and up from the bulk soil) in a grassland in Southern California (Chapter 2). For each route, I measured the dispersal rate, taxonomic composition, and influence on the soil microbiome. In a follow-up experiment, I focused on the spatial heterogeneity of dispersal into soil from local vegetation by measuring the distance and taxonomic composition of microorganisms dispersing from plant individuals across a landscape. Altogether, my dissertation improves our understanding of the patterns and drivers of microbial biogeography.

CHAPTER 1

Alpha-, beta-, and gamma-diversity of bacteria varies across habitats

ABSTRACT

Bacteria are essential parts of ecosystems and are the most diverse organisms on the planet. Yet, we still do not know which habitats support the highest diversity of bacteria across multiple scales. We analyzed alpha-, beta-, and gamma-diversity of bacterial assemblages using 11,680 samples compiled by the Earth Microbiome Project. We found that soils contained the highest bacterial richness within a single sample (alpha-diversity), but sediment assemblages displayed the highest gamma-diversity. Sediment, biofilms/mats, and inland water exhibited the most variation in community composition among geographic locations (beta-diversity). Within soils, agricultural lands, hot deserts, grasslands, and shrublands contained the highest richness, while forests, cold deserts, and tundra biomes consistently harbored fewer bacterial species. Surprisingly, agricultural soils encompassed similar levels of beta-diversity as other soil biomes. These patterns were robust to the alpha- and beta- diversity metrics used and the taxonomic binning approach. Overall, the results support the idea that spatial environmental heterogeneity is an important driver of bacterial diversity.

INTRODUCTION

Bacteria are the most diverse organisms on the planet (Whitman et al., 1998). Bacterial richness and composition influences ecosystem functioning, whether in host-associated communities, soils, or oceans (Barberán & Casamayor, 2010; T. Bell et al., 2005; Jennifer B. H. Martiny et al., 2017; Philippot et al., 2013; Strickland et al., 2009; Wagg et al., 2014). Nevertheless, we have yet to answer a number of basic questions about bacterial diversity, including "Which habitats contain the highest diversity of bacteria?" More broadly, evaluating geographic patterns in biodiversity across habitats and spatial scales can illuminate the processes influencing and consequences of biodiversity (Bryant et al., 2008; Carson et al., 2010; Ibarbalz et al., 2019; Logares et al., 2020; Mittelbach et al., 2007).

While many studies document spatial patterns of bacterial diversity, most are restricted to a particular geographic region or habitat, such as soil, sediment, or water (Fierer et al., 2012; Fierer & Jackson, 2006; Neufeld & Mohn, 2005). To understand global trends, however, studies that analyze diversity across habitats and geographic regions are needed. Combining data from independent projects is oftentimes infeasible because community variation can be caused simply by differences in methodology. The Earth Microbiome Project (EMP) comprises 27,751 samples from 97 studies from a wide range of habitats and geographic regions that are processed in the exact same way (Thompson et al., 2017). Although there are limitations to PCR-based sequencing surveys (Apprill et al., 2015), this dataset is unique for its size in using standardize methods for all samples. Thus, this dataset provides an opportunity for a rigorous comparison of bacterial diversity across many parts of the globe.

A recent overview from the EMP noted, as has previously been observed, that communities of free-living bacteria are more diverse than host-associated bacteria (Fierer &

Lennon, 2011; Lozupone & Knight, 2007; Torsvik et al., 2002). For example, soil and sediment samples have higher alpha-diversity than animal gut or skin microbiomes. We expand on this initial alpha-diversity analysis by additionally evaluating beta- and gamma-diversity across spatial scales while mitigating for unevenly spaced samples (Fig. 1.1). We also take the opportunity to use the large size of the EMP dataset to test whether different diversity metrics and taxa definitions influence our understanding of microbial diversity patterns.

We specifically ask: which habitats support the highest levels of bacterial diversity? We consider three interrelated aspects of biodiversity: alpha-, beta-, and gamma-diversity. We measure alpha-diversity as the observed richness (number of taxa) or evenness (the relative abundances of those taxa) of an average sample within a habitat type. We quantify beta-diversity as the variability in community composition (the identity of taxa observed) among samples within a habitat (Marti J. Anderson et al., 2006). Finally, we calculate gamma-diversity as the total observed richness of all samples within in a habitat.

We test several predictions about relative, not absolute, diversity patterns because, even in this large dataset, bacterial diversity remains undersampled. First, we predict that sediment and soil support the highest alpha-diversity within a single sample. These habitats are known to have relatively high bacterial diversity, although their relative rankings have not yet reached a consensus (Fierer & Lennon, 2011; Lozupone & Knight, 2007; Thompson et al., 2017; Torsvik et al., 2002). Second, we expect that soil, sediment, inland water, and biofilm/mat habitats will exhibit high beta-diversity. These habitats are spatially separated with less dispersal or mixing than air or marine water. Finally, we predict that soils and sediments will exhibit high gamma-diversity as they are expected to have both high alpha- and beta-diversity.

Within the soil habitat, we hypothesize that soils from biomes higher in plant diversity and productivity (e.g., forests and grasslands) support higher alpha-diversity than soils from biomes with low diversity and productivity (e.g., tundra and deserts) (Kier et al., 2005; Lamb et al., 2011; Zak et al., 2003). Of course, these biomes do not directly influence diversity, but they are defined based on abiotic factors (Holdridge, 1967), such as temperature or precipitation, that do influence diversity. Further, we expect that agricultural soils will exhibit lower beta-diversity than other biomes as common practices (pesticides, tilling, and fertilizer use) and the low diversity of crop plants influences community composition (Jangid et al., 2008; Upchurch et al., 2008). We also compare the relationship between diversity and biomes to those between diversity, and pH or temperature to assess whether plant diversity or abiotic conditions more strongly influence bacterial diversity. We expect bacterial richness to peak at neutral pH and moderate temperatures (Bahram et al., 2018; Fierer & Jackson, 2006; Griffiths et al., 2011), and abiotic factors to be a stronger influence on diversity than plant biomes. Overall, the aim of this study was to compare bacterial diversity trends across habitats. We show that the most diverse habitat depends on the type of diversity (alpha-, beta-, or gamma-diversity).

MATERIALS AND METHODS

Bacterial 16S rDNA (V4 region) sequence data and associated metadata (e.g., sample location, sample type, date of sampling) were downloaded from the Earth Microbiome Project (EMP) on September 1, 2016. Sample processing, sequencing, and core amplicon data analysis were performed by the Earth Microbiome Project (www.earthmicrobiome.org), and all amplicon sequence data and metadata have been made public through the data portal (qiita.microbio.me/emp). Data available from: https://doi.org/10.1038/nature24621 (Thompson

et al., 2017). We used the EMP closed-reference (Greengenes 13.8) OTU dataset classified at 97% sequence similarity to reduce computational time (instead of the open-reference dataset). The dataset contains 27,751 samples, with a median depth of 54,091 sequences per sample. We excluded archaea from the analysis because, relative to bacteria, they make up a small portion of any given community (median = 0.018% of sequences).

Habitat designations. We used the EMP Ontogeny metadata to classify the habitat and, for soil, biome of each sample based on the EMP metadata (Fig. 1.2). When the existing metadata were unclear, we used the latitude and longitude coordinates to assess the environmental context. Samples with insufficient data about their location or habitat were removed from the analysis. Host-associated samples were also removed. Further, we only retained samples that could be classified into one of the following habitats: soil, sediment, marine water, inland water (e.g., rivers and lakes), air, and biofilms/mats. These habitat types were chosen because they represent a wide range of environmental conditions and are well sampled within the EMP dataset. Within the soil habitat, we further classified samples into forest, hot desert, cold desert, grassland, shrubland, tundra, and agricultural soil. We also classified inland water and sediment samples as saline and non-saline. See supplemental materials for descriptions of sample locations (Appendix 1B). After removing samples with less than 15,000 sequences (rarefaction depth in this study), 11,680 free-living (non-host system) samples remained.

Alpha-diversity analysis. To account for differences in sequencing depth, the samples were rarefied to 15,000 sequences with 1,000 resamplings in QIIME (Caporaso et al., 2010). This rarefaction depth provided a high sequence count per sample while minimizing sample loss to

4.74% of samples. All samples with less than 15,000 sequences were removed leaving 11,680 free-living (non-host system) samples. For each resampling, we calculated 24 alpha-diversity metrics on the rarified OTU table in QIIME (Fig. 1.3A). These metrics characterized the community in five general ways: observed richness, estimated richness, evenness/dominance, phylogenetic diversity, and coverage of sampling. We used all 24 metrics throughout the alpha-diversity analysis to ensure that our final conclusions were not dependent on the type of metric. We calculated the median value of each metric across the 1,000 replicates.

To minimize the effect of unevenly spaced samples, we averaged the alpha-diversity of the samples within a single geocluster. Many of the samples are highly clumped such that some geographic regions contribute unequally to the habitat's diversity. Geoclusters (n= 172) were formed by clustering samples of the same habitat type located within 110 km of each other (distance of 1º latitude at the equator) using hclust() and cutree() from package 'stats' and rdist.earth() from package 'fields' in R (Nychka et al., 2015; R Core Team, 2020). While a smaller clustering distance would have yielded a higher geocluster sample size, this conservative distance allowed us to be more confident that our results reflected ecological processes, rather than sampling locations. We calculated the median of each diversity metric for the samples within each geocluster of the same habitat type. The averaged alpha-diversities were then cube root transformed to achieve normality and homoscedasticity. Finally, we tested for significant differences in alpha-diversity among habitats by performing a one-way ANOVA and Tukey's HSD in R.

We tested whether the alpha-diversity results depended on the diversity metric by running a correlation with every pairwise combination of diversity metrics using all 11,680 samples. Likewise, we tested whether our results depended on the resolution of OTU clustering by

comparing the 97% similarity OTU table with the single-nucleotide resolution 'sub-OTUs' dataset, Deblur, produced by the EMP. We rarefied the Deblur dataset to 15,000 sequences per samples (1,000 times), calculated Exact Sequence Variance (ESV) richness per sample, and calculated the mean richness across the 1,000 replicates. We then ran a correlation between the OTU richness and ESV richness for every sample present in both datasets (n = 11,137).

To further explore what factors might be driving alpha-diversity, we compared bacterial richness with pH and temperature at each sample site. We chose pH and temperature because these were the most widely included in the EMP dataset. For soil samples, however, the temperature data were often missing from the EMP dataset. Thus, for the analysis with just soil samples, we used temperature data from WorldClim. Metadata (pH and temperature) from the EMP were taken at the site at the time of sample collection. Data from WorldClim, a publicly-available dataset, included mean annual temperature averaged from 1970 – 2000 with spatial resolution of 10 minutes. Data is available from WorldClim Version2: http://doi.org/10.1002/joc.5086 (Fick & Hijmans, 2017). We assigned external temperature data to soil samples using latitude and longitude with extract() from package 'raster' in R (Hijmans, 2018). Temperature and pH were correlated with OTU richness using a second-degree polynomial in R. We tested whether temperature and pH differed among habitats and biomes using an ANOVA in R.

Beta-diversity analysis. To reduce computation time, we used a subset (150 rarefied tables) of the 1000 rarefied OTU tables generated during the alpha-diversity analysis to analyze beta-diversity. The OTU tables were first square root transformed to increase weight given to the rare taxa (Magurran, 1988) that make up the majority of microbial communities (Jennifer B.H. Martiny & Walters, 2018). For each of the 150 rarefied, square root transformed OTU tables, we

calculated the median abundance for each taxon across all the samples within a single geocluster. We then calculated a Bray-Curtis dissimilarity matrix for each of the 150 OTU-by-geocluster tables in QIIME. Finally, we calculated the median of the 150 dissimilarity matrices to yield one median Bray-Curtis dissimilarity matrix.

To visualize compositional differences among habitats, we used NMDS in PRIMER6 (Clarke & Gorley, 2006). We also tested community composition differences among habitats using PERMANOVA in PERMANOVA+ (M.J. Anderson et al., 2008). Because we averaged OTU abundances for all samples of the same habitat type located within the same geographic area (geocluster), beta-diversity provides an approximation of the amount of community variation from location to location within one habitat (as opposed to variation from sample to sample within one location).

To compare beta-diversity across habitats, we analyzed the variance within each habitat using the function PERMDISP in PERMANOVA+. To determine if unequal sampling among habitats biased these results, we re-calculated the Bray-Curtis values based on a selection of only 20 geoclusters for each habitat from a rarified OTU-by-geocluster table. We chose 20 geoclusters because that depth included five of the six habitats (excluding air) but avoided the biases expected with sample sizes less than ten (M.J. Anderson et al., 2008). We repeated these subsamplings 100 times and tested for differences in beta-diversity among habitats using betadisper(), the PERMDISP test implemented in the R package ‘vegan’ (Oksanen et al., 2019). We compared the relative rankings of these rarefied beta-diversity results to the unrarefied results to determine if rarefaction changed the relationships of variance among habitat groups. Specifically, we considered the rarefied results to match the unrarefied results if 95-100

subsampled tests were significant and showed the same beta-diversity rankings (based on mean distance to centroid) as the unrarefied test.

We tested whether the beta-diversity results depended on the diversity metric by running a correlation with every pairwise combination of nine diversity metrics. For each of the 150 OTU-by-geocluster tables, we calculated nine beta-diversity metrics using vegdist() from package 'vegan' in R (Oksanen et al., 2019). We then took the mean matrix (of the 150 matrices) for each diversity metric. We performed a Spearman's mantel test for every pairwise comparison of the nine averaged beta-diversity matrices using mantel() from package 'stats' (R Core Team, 2020). To compare Raup-Crick to the other beta-diversity metrics, we calculated the dissimilarity within, but not among, habitats. Raup-Crick is not an appropriate metric when communities do not share the same species pool (Chase et al., 2011). To calculate Raup-Crick, we took the mean of the matrices computed with raupcrick(…, chase = TRUE) and raupcrick(…, chase = FALSE) from package 'vegan' in R (Oksanen et al., 2019) to follow the method recommended by Chase et al. (Chase et al., 2011). We calculated one Raup-Crick matrix for each habitat for each of the 150 OTU-by-geocluster tables, took the mean for each habitat across the 150 matrices, and then calculated the mean and SE distance from centroid in PRIMER6 (Clarke & Gorley, 2006). We used the mean and SE to compare the trends in dissimilarity to those generated with the Bray-Curtis metric.

Gamma-diversity analysis. To assess gamma-diversity by habitat, we plotted an OTU accumulation curve for each habitat with specaccum() from package 'vegan' in R (Oksanen et al., 2019) using the 150 OTU-by-geocluster tables. The OTU accumulation curve displays the numbers of geoclusters sampled on the x-axis and observed OTU richness on the y-axis. This

plot allowed us to compare cumulative diversity levels across multiple samples distributed across the world. We examined whether habitats likely exhibit different gamma-diversity levels by calculating error bars equal to 1.96 times the standard deviation.

RESULTS

Alpha-Diversity

Out of the six habitats compared, soils contained the highest observed richness (i.e., number of observed taxa rarefied at 15,000 sequences) for a single sample, with a median of 1,842 taxa (97% OTUs) per sample given this depth of sequencing (one-way ANOVA: $F = 39.13$, $P < 0.001$, $r^2 = 0.541$; Fig. 1.2A). Sediments were the second most diverse habitat with an average of 1,137 taxa. Marine water, air, inland water, and biofilms/mats had a significantly lower richness (averaging 571, 500, 478, and 342 taxa, respectively) than soils and sediments ($P < 0.001$; Fig. 1.2A) but could not be distinguished from one another by richness .

Because salinity influences bacterial community composition (Lozupone & Knight, 2007), we further tested whether taxon richness varied between non-saline and saline habitats. We found that salinity had no impact on alpha-diversity for sediments (one-way ANOVA: $F = 0.433$, $P = 0.516$; Fig. S1.1) or inland water ($F = 0.093$, $P = 0.763$; Fig. S1.1).

Within the soil habitat, we further compared alpha-diversity among seven biomes (agricultural, grassland, shrubland, forest, hot desert, cold desert, and tundra soil). Within soil samples, richness differed significantly among biomes. Agricultural soils supported the highest richness in a sample, along with hot desert, grassland, and shrubland biomes. Forest soils were less diverse than agricultural soil, and tundra and cold deserts supported the lowest richness (one-way ANOVA: $F = 42.62$, $P < 0.001$, $r^2 = 0.642$; Fig. 1.2B). Notably, the cold desert biome

was only represented by two geoclusters (averaging across 117 samples); thus, more data are needed to assess that particular biome's diversity.

The above results were robust to the alpha-diversity metric used. On a sample-by-sample basis, 24 alpha-diversity indices, including observed richness, were all correlated with each other (r^2 = 0.09 – 1.00, P < 0.0001) with a mean r^2 of 0.63 (Fig. 1.3A). The metrics grouped into two main clusters. One cluster encompassed the richness/coverage metrics such as OTU richness, Faith's Phylogenetic Diversity, and Chao1 (r^2 = 0.86 – 1.00, mean r^2 = 0.97). The other cluster included the evenness/dominance metrics such as Simpson's and McIntosh dominance index (r^2 = 0.31 – 1.00, mean r^2 = 0.85). Further, each metric ranked the habitats from highest to lowest alpha-diversity in the same way, with the exception of air. Air communities were more even than other habitats, given their relative richness level (Fig. 1.3B). Excluding air samples, OTU richness (ANCOVA: F = 54.68, P < 0.0001, r^2 = 0.229), but not habitat (F = 2.14, P = 0.0775) was a predictor of evenness. When air samples were included, both OTU richness (F = 57.09, P < 0.0001, r^2 = 0.173) and habitat (F = 7.0989, P < 0.0001, r^2 = 0.107) were significant predictors of evenness.

The alpha-diversity patterns were also robust to the taxonomic binning method (Fig. 1.4). We compared the 97% OTU and Exact Sequence Variant (ESV) Deblur datasets provided by the EMP. Not only were OTU richness and ESV richness strongly correlated (r^2 = 0. 933, P < 0.0001), but, on a sample-by-sample basis, they were also nearly identical (slope = 0.953, P < 0.0001). However, the relationship between OTU richness and ESV richness varied among habitats (ANOVA: F = 786.5, P < 0.0001, r^2 = 0.331). In particular, non-saline sediments and inland water demonstrated a higher ESV:OTU richness ratio than other habitats (Fig. 1.4).

Taxon richness displayed a weak hump-shaped relationship with pH and a peak in diversity at a neutral pH (non-linear regression: $P < 0.0001$, $r^2 = 0.047$; Fig. S1.2A). In contrast, taxon richness only weakly correlated with temperature, and this relationship was driven by low-diversity biofilm/mat samples sampled from high temperatures ($P < 0.0001$, $r^2 = 0.036$; Fig. S1.2C). Overall, the bacterial alpha-diversity patterns across all habitats were not obviously related to pH or temperature. Most of the samples were, on average, at a neutral pH, with the soil samples more acidic and the biofilm/mat samples more basic (Fig. S1.2B). Despite this, temperature (ANOVA: $F = 272.3$, $P < 0.0001$, $r^2 = 0.198$) and pH ($F = 245.2$, $P < 0.0001$, $r^2 = 0.215$) differed significantly among habitats (Fig. S1.2D).

Similar to the pattern observed across all habitats, richness within just the soil samples also peaked at a neutral pH (non-linear regression: $P = 0.001$, $r^2 = 0.121$; Fig. S1.3A). In contrast, richness in soils also peaked at a temperature around 10°C ($P < 0.0001$, $r^2 = 0.288$; Fig. S1.3C). Both temperature (ANOVA: $F = 640.2$, $P < 0.0001$, $r^2 = 0.153$) and pH ($F = 199.8$, $P < 0.0001$, $r^2 = 0.594$) differed among soil samples by biome (Fig. S1.3B and Fig. S1.3D). Soils from both hot and cold deserts tended to be basic while agricultural fields, forests, and tundra were acidic. Tundra and cold deserts were the coldest biomes, and shrubland, agriculture, and hot deserts were among the hottest biomes.

Beta-Diversity

Sediment, biofilm/mat, and inland water habitats displayed the highest beta-diversity among geographic locations or geoclusters (not within a single sample), whereas soil, air, and marine water exhibited 17% lower beta-diversity (PERMDISP: $F = 10.7$, $P = 0.001$; Fig. 1.2C). To test that these patterns were not influenced by unequal sampling (number of geoclusters) of

the habitats, we subsampled the habitats (to 20 geoclusters per habitat) and retested the patterns. All 100 subsamplings produced the same beta-diversity rankings, and all models were significant, indicating that unequal sampling did not influence within-habitat beta-diversity. Within the soil habitat, beta-diversity did not differ by biome (P = 0.526; Fig. 1.2D).

Overall, bacterial community composition differed significantly by habitat (PERMANOVA: P = 0.001, Pseudo-F = 9.8601, r^2 = 0.210; Fig. 1.2E) and by biome for soils (P = 0.001, Pseudo-F = 4.221, r^2 = 0.227; Fig. 1.2F). Because salinity influenced the community composition for both sediments (P = 0.002, Pseudo-F = 2.0168, r^2 = 0.075) and inland water (P = 0.02, Pseudo-F = 1.7702, r^2 = 0.085), we tested whether salinity likewise influenced beta-diversity within these habitats. Beta-diversity did not differ between saline and non-saline samples within sediments (PERMDISP: P = 0.21, F = 2.4404) or inland water (P = 0.849, F = 0.27406; Fig. S1.4).

The above results did not depend on the beta-diversity metric used. On a sample-by-sample basis, nine beta-diversity indices, including Bray-Curtis, were correlated with each other (r = 0.435 – 1.00, P = 0.001, mean r = 0.88; Fig. S1.5A). Raup-Crick has been suggested as a more appropriate metric when comparing groups with different alpha-diversity levels (Chase et al., 2011). Because Raup-Crick assumes that all communities are part of the same regional species pool, we calculated the mean and standard error within each habitat (excluding between habitat comparisons) and compared the trend to that generated by the Bray-Curtis metric. Both Bray-Curtis and Raup-Crick metrics showed the same trend of beta-diversity among habitats (Fig. S1.5B).

Gamma-Diversity

Considering the accumulation of taxon richness across geoclusters, sediments exhibited the highest gamma-diversity of any habitat, followed by soils and inland water (Fig. 1.5). The sediment rarefaction curve showed little sign of flattening out, indicating that most taxa are yet to be sampled. In contrast, the soil curve noticeably leveled off, even at a similar level of sampling. The gamma-diversity of marine water, biofilms/mats, and air were not statistically distinguishable from one another but, as a group, exhibited lower gamma-diversity than inland water, soils, and sediments.

DISCUSSION

Here, we tested which habitat contains the most bacterial taxa within a single sample (alpha-diversity), which exhibits the most variation among samples (beta-diversity), and which contains the most taxa across all samples (gamma-diversity). We show that a single sample of soil on average contained higher bacterial alpha-diversity than any other habitat, including sediment (Fig. 1.2A). However, sediment had higher gamma-diversity, with much of its diversity yet to be sampled (Fig. 1.5). Within soils, we found that agricultural soils had among the highest richness and exhibited just as much compositional variation (beta-diversity) as other biomes.

Although both sediments and soil were previously known to be highly diverse microbial habitats, previous studies demonstrated conflicting results about their relative ranking (Fierer & Lennon, 2011; Lozupone & Knight, 2007; Thompson et al., 2017; Torsvik et al., 2002). Using 11,680 samples and minimizing geographic biases, this analysis suggests that soil contains higher alpha-diversity than sediments. In contrast, marine water, inland water, air, and biofilms/mats contain the lowest alpha-diversity (Fig. 1.2A).

Within-habitat heterogeneity is a known driver of plant and animal diversity (Johnson & Simberloff, 1974; Tews et al., 2004) and has been correlated with microbial communities as well (Cheeke et al., 2015; Curd et al., 2018). Here, we show that the alpha-diversity patterns are consistent with the idea that habitat heterogeneity may drive bacterial diversity at a single sample. The highly mixed water and air environments harbor lower diversity, consistent with previous smaller-scale studies (Asakawa & Kimura, 2008; Crump et al., 2012; Torsvik et al., 2002). While both sediments and soil are not as well mixed, sediments contain higher water content than soils. Water content increases connectivity and thus reduces environmental heterogeneity and promotes dispersal, both of which can result in lower diversity (Bickel et al., 2019; Carson et al., 2010; Freestone & Inouye, 2006; Vellend, 2010). At the same time, biofilms and mats, despite being spatially structured, also displayed low alpha-diversity (Hall-Stoodley et al., 2004). However, the biofilm/mat samples from the EMP dataset encompassed samples with the highest pH and temperature (Fig. S1.2). We therefore speculate that these abiotic extremes contribute to low alpha-diversity (Campbell & Kirchman, 2013; Miller et al., 2009; Sharp et al., 2014). In fact, diversity in many habitats is lowest at extreme temperatures (Bahram et al., 2018; Campbell & Kirchman, 2013; Fierer & Jackson, 2006; Griffiths et al., 2011; Milici et al., 2016; Miller et al., 2009; Raes et al., 2018; Sharp et al., 2014). Yet ultimately, little is known about the environmental conditions, and their heterogeneity, at the spatial scale that matters for microorganisms (Vos et al., 2013). To test the importance of within-sample heterogeneity on microbial diversity directly, finer-scale data are needed.

Our analysis is the first to quantify bacterial beta-diversity among habitats across many parts of the globe. While soils contained the highest alpha-diversity within a single sample, sediments displayed higher beta-diversity among geoclustered samples within a habitat (Fig.

1.2C). Sediment beta-diversity was also similar to that of inland water and biofilms/mats. Additional environmental data associated with the individual samples would be needed to distinguish whether these beta-diversity patterns might be driven by dispersal limitation (Carson et al., 2010; S. Evans et al., 2016) or spatial variation in environmental conditions (Asakawa & Kimura, 2008; Barberán & Casamayor, 2010).

Given their high alpha- and beta-diversity, it is not surprising that sediments are also estimated to contain the highest gamma-diversity (Fig. 1.5). While extracellular DNA (eDNA) may be particularly prevalent in ocean sediments (Dell'Anno & Danovaro, 2005), evidence thus far suggests that eDNA has minimal effect on sediment diversity estimates from sequencing surveys (Ramírez et al., 2018). The taxa accumulation curves also suggest that, while we may have observed most bacterial taxa in soil (at least from highly sampled continents), there is much more diversity to discover in sediments. Similarly, other than soil, the accumulation curves suggest that air, water habitats, and biofilms/mats remain undersampled as well. Of course, the number and localities of samples available will influence the diversity estimates. Thus, a limitation to these conclusions is that the samples are highly concentrated in North America and Europe (Fig. 1.1), and continued sampling is needed to test the robustness of these diversity patterns.

Because plant diversity and productivity are shown to impact microbial communities (Lamb et al., 2011; Zak et al., 2003), we further characterized alpha- and beta-diversity trends among biomes from which the soil samples were collected. Agricultural soils contained among the highest alpha-diversity, as previously noted in smaller scale studies (Rodrigues et al., 2013; Upchurch et al., 2008). Indeed, some agricultural practices, such as application of manure, are known to increase bacterial diversity (Ding et al., 2016; Sharma et al., 2010). Even more notable,

however, is that agricultural soils encompassed similar levels of beta-diversity to those of other biomes. While some agricultural practices have been shown to homogenize communities within a single field (Rodrigues et al., 2013), not all practices have a homogenizing effect (O'Brien et al., 2016). Further, the diversity of agricultural practices around the world (Sharma et al., 2010; Soman et al., 2017) seems to select for as much variation in bacterial composition (beta-diversity) as different types forests or deserts.

Contrary to our hypothesis, biomes with higher plant diversity or productivity, such as forest or shrubland soils, were no more diverse within a sample than other biomes (Fig. 1.2B). These results support previous findings that grasslands contain more bacterial diversity than forests (Delgado-Baquerizo & Eldridge, 2019) and, overall, plant and soil diversity are uncoupled (Prober et al., 2015). We therefore propose that abiotic factors may be more important for soil bacterial alpha-diversity than plant biomes. Biomes differed significantly in pH and temperature, and soil alpha-diversity was strongly correlated with both factors (Fig. S1.3). These results are consistent with previous studies (Bahram et al., 2018; Fierer & Jackson, 2006; Griffiths et al., 2011) that find bacterial richness in soils peaks at a neutral pH and at mid-temperatures (around 10ºC). Of course, other unmeasured environmental factors and/or ecological interactions are likely influencing soil diversity.

Finally, these diversity patterns appear to be robust to two key methodological issues. First, diversity trends did not depend on the particular alpha- or beta-diversity metrics used (Fig. 1.3A and Fig. S1.5A). Air, as the only exception, was more even than expected, given its richness (Fig. 1.3B). We speculate that the movement of air contributes to its evenness as air likely picks up a sampling of bacteria from many different habitats (Miletto & Lindow, 2015; Rintala et al., 2008). Second, the results were robust to the degree of clustering of the amplicon

sequences (Fig. 1.4). Both the 97% OTU and the ESV datasets yielded the same alpha-diversity trends, as previously noted in a smaller scale study (Glassman & Martiny, 2018). Most of those outliers in this analysis came from non-saline sediment or inland water samples, which had higher ESV richness than 97% OTU richness. While this pattern could suggest higher finer-scale diversity within these habitats, we caution that these samples originated from only four geoclusters. Overall, while ESVs can be useful for resolving finer diversity among specific taxonomic groups (Needham et al., 2017), broad-scale alpha-diversity patterns do not seem to be altered by these particular operational definitions.

With the largest dataset created with consistent methodology and a geographically widespread sampling effort, we show that soils support the highest diversity within a single sample (alpha-diversity) and that sediments are more variable in composition among locations (beta-diversity) and likely support the most bacterial taxa at a larger spatial scale (gamma-diversity). Within soils, we find biome type impacts soil alpha-diversity but not beta-diversity. Many of these results appear consistent with the idea that spatial heterogeneity and dispersal limitation promote bacterial diversity. These baseline patterns set the stage for new research on the mechanisms driving the generation and maintenance of bacterial diversity.

ACKNOWLEDGEMENTS

We would like to thank Luke Thompson, Jack Gilbert, and the Earth Microbiome Project team for their vision and hard work in assembling this dataset and making it widely available. We also thank Steve Allison and Cascade Sorte for their input on analyses and methods, and Michaeline Albright and Alex Chase for help with statistical analyses and computational

methods. We further thank Alex Chase, Cynthia Rodriguez, Sarai Finks, Claudia Weihe, Joia Capocchi, Kazuo Isobe, and Pauline Nguyen for their guidance with earlier drafts.

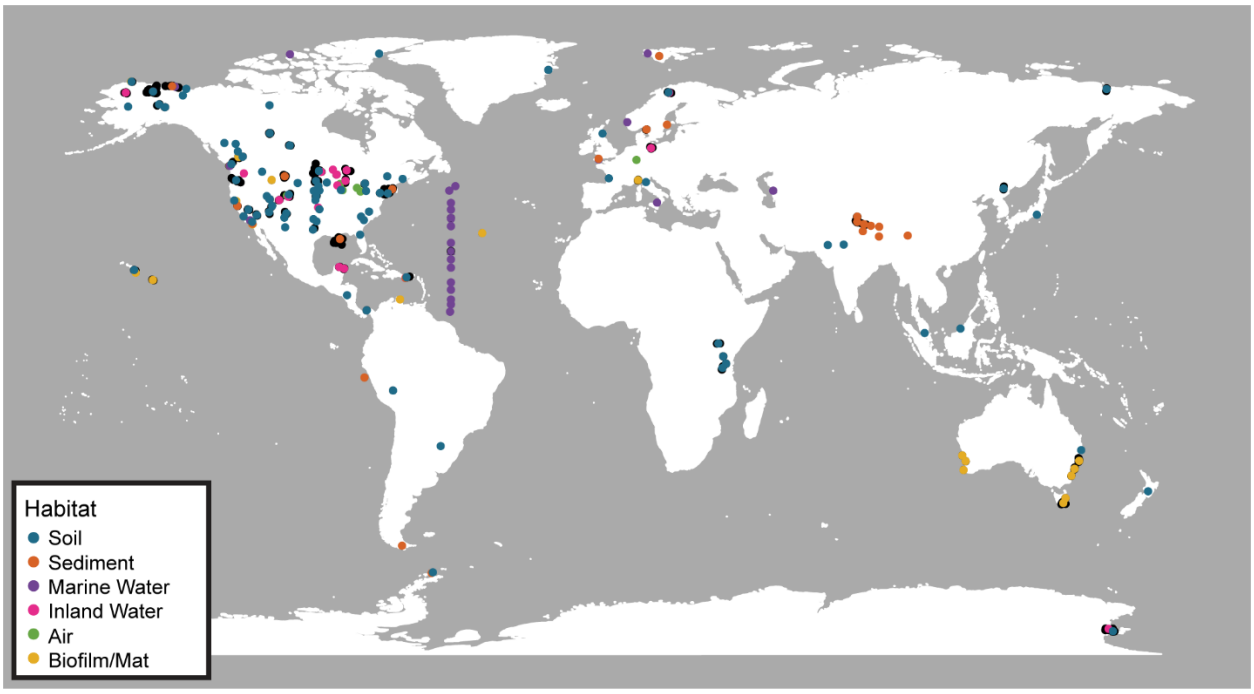

Figure 1.1 Sample and geocluster locations. Map showing locations of each of the EMP samples used in this study (black dots) and the geoclustered samples for each habitat (colored dots). Geoclusters were created from samples located within 110 km of each other.

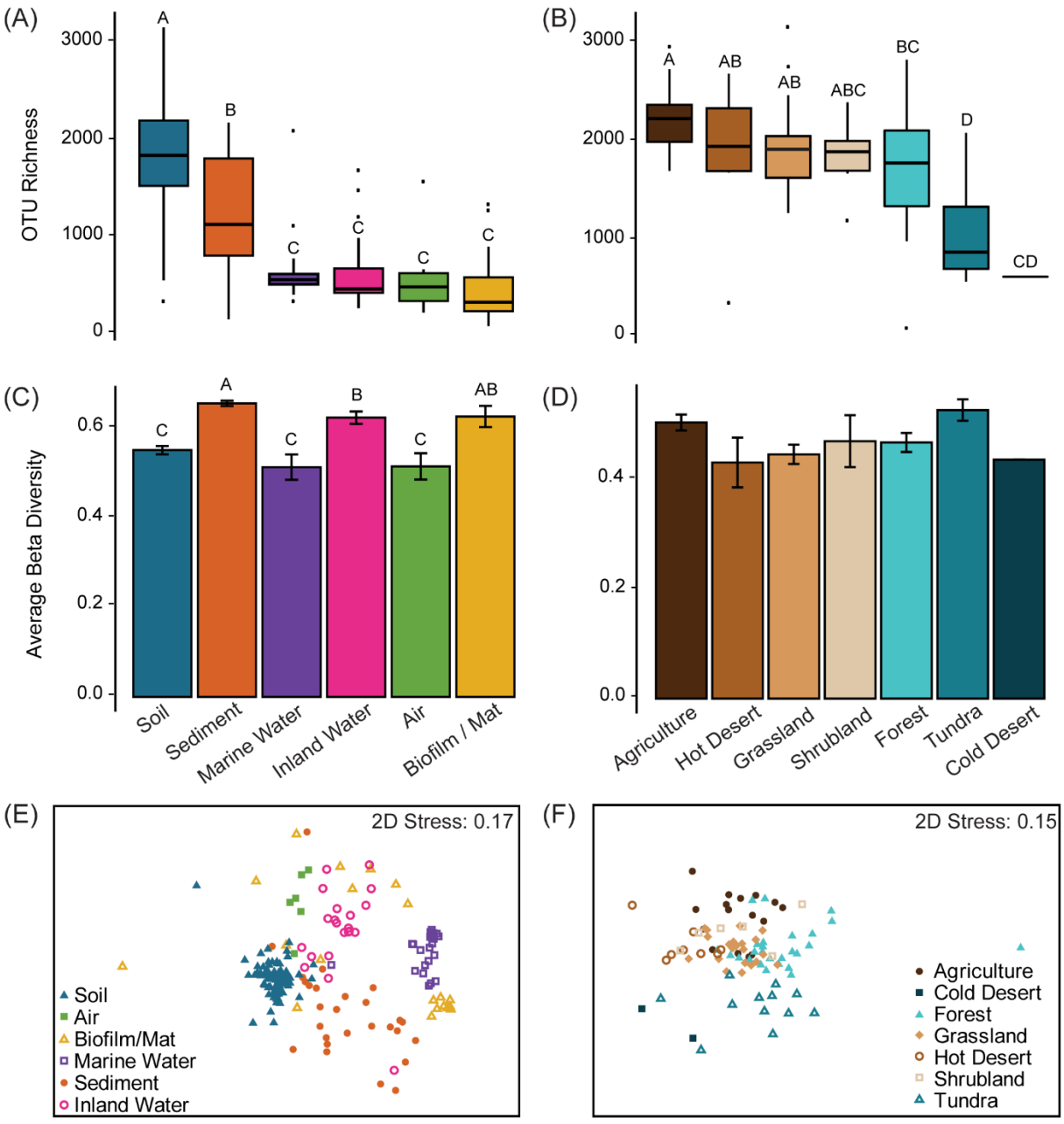

Figure 1.2 Alpha- and beta-diversity patterns. Alpha- and beta-diversity per habitat for all geoclusters used in study (A, C, and E) and per biome for soil geoclusters (B, D, and F). (A and B) Boxplot of alpha-diversity (OTU richness). (C and D) Mean beta-diversity (distance from centroid) ± standard error. For all bar and boxplots, letters above indicate significant differences among groups (Tukey test) where groups that share a letter are not significantly different from each other. (E and F) NMDS of geoclusters.

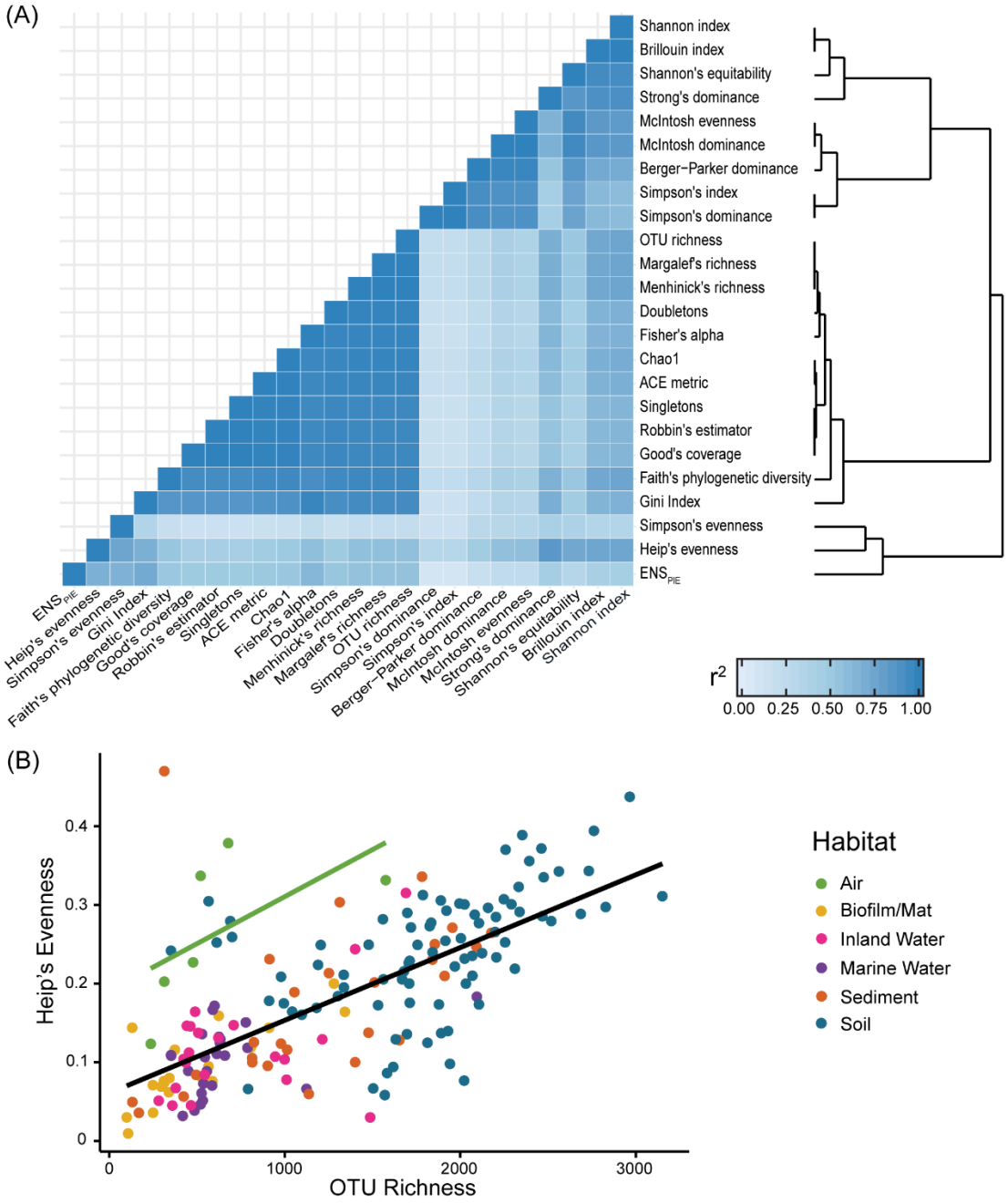

Figure 1.3 Comparison of diversity metrics. (A) Heatmap shows degree of correlation (r^2 from linear regression with all EMP samples used in analysis). Dendrogram shows relatedness of metrics based on their correlation strength. Note that the metrics are clustered into two groups: one composed of mainly evenness metrics (top cluster on dendrogram) and one composed of mainly richness metrics (bottom cluster on dendrogram). Simpson's evenness, Heip's evenness, and ENS_{pie} fall outside of those two clusters. (B) Dot plot showing relationship between Heip's evenness and OTU richness metrics for geoclusters of each habitat. The green line is a linear regression for air geoclusters, and black line is a linear regression for all geoclusters except for air.

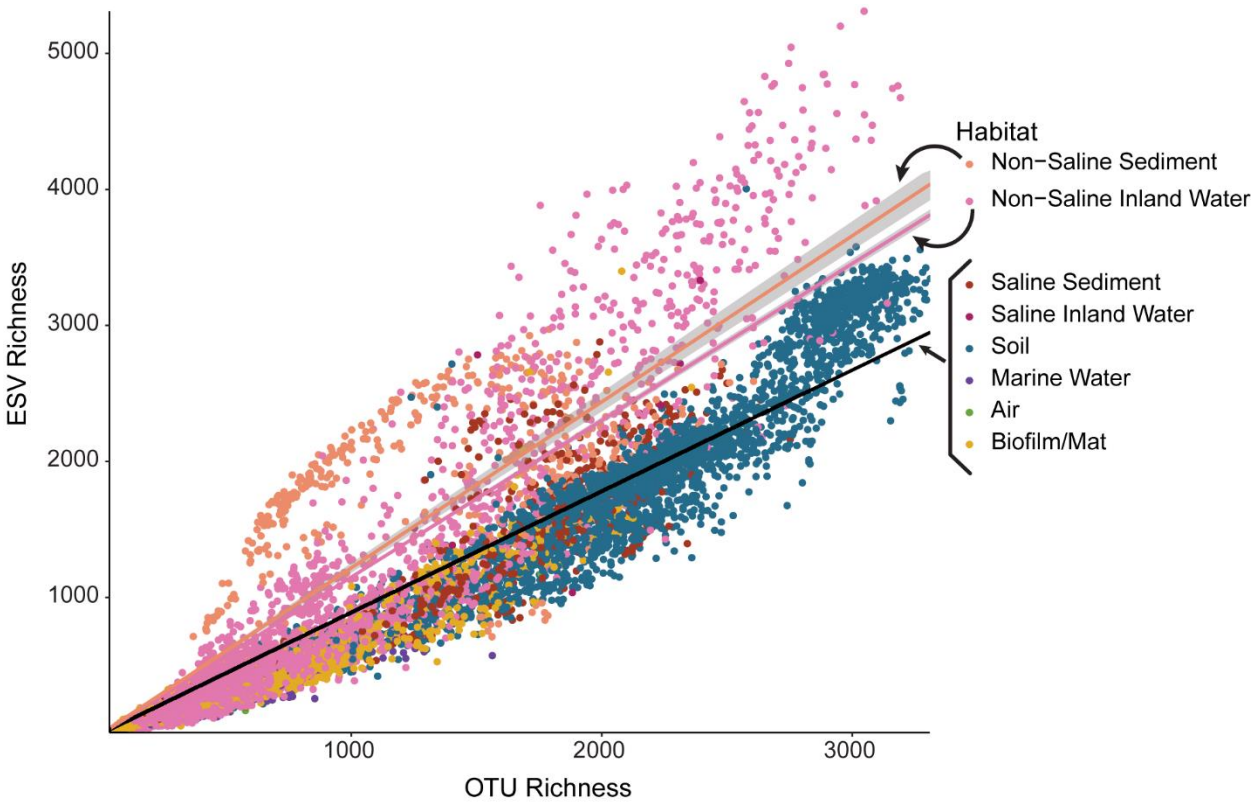

Figure 1.4 Comparison of OTU and ESV richness. On a sample-by-sample basis, OTU richness and ESV richness are highly correlated ($r^2 = 0.933$, $P < 0.0001$). On average, for any given sample, ESV richness is equal to 95.26% of OTU richness. However, not all habitats showed the exact same relationship between OTU and ESV richness. Non-saline sediment and inland water samples have significantly higher ESV richness given their relative OTU richness. All samples except for non-saline sediment and inland water have a regression line with a slope of 0.891 ($P < 0.0001$, $r^2 = 0.971$), shown as the black regression line on the graph. Non-saline sediment samples have a regression line with a slope of 1.220 ($P < 0.0001$, $r^2 = 0.895$), shown as light orange on the graph. Non-saline inland water samples have a regression line with a slope of 1.152 ($P < 0.0001$, $r^2 = 0.902$), shown as light pink on the graph.

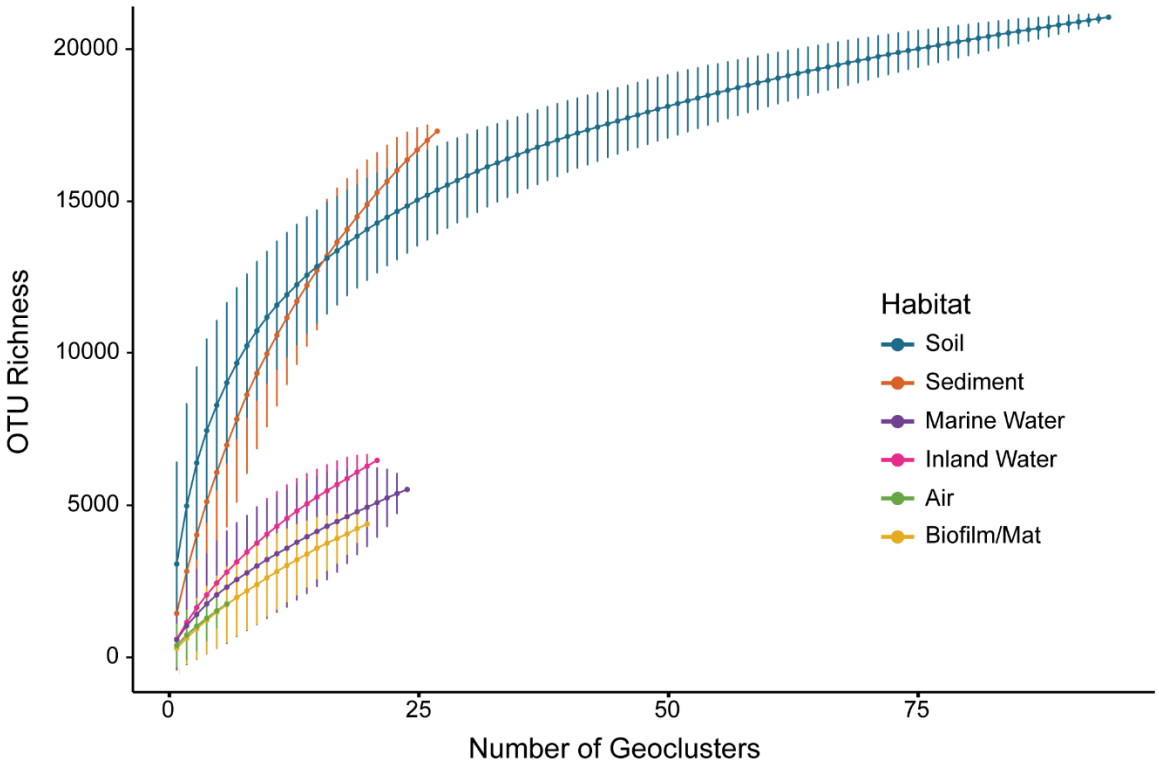

Figure 1.5 Geocluster accumulation curves. Geocluster accumulation curves (gamma-diversity) for mean OTU richness from a random sampling of geoclusters (permutations = 999) with 95% confidence intervals drawn for each habitat.

SUPPLEMENTAL MATERIALS

Appendix 1A: Supplemental Figures

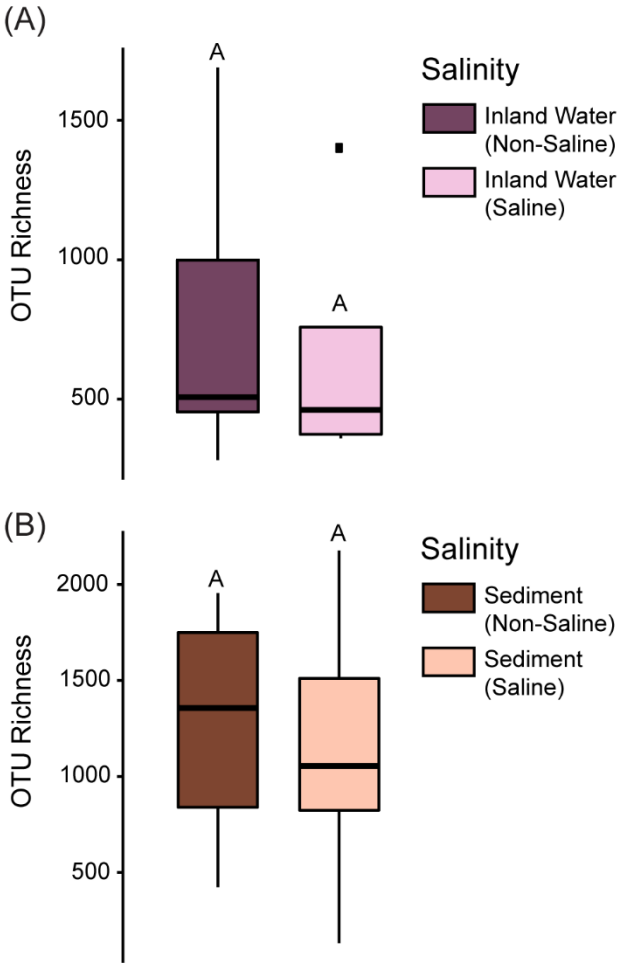

Figure S1.1 The influence of salinity on alpha-diversity. Taxon richness did not differ between saline and non-saline samples from inland water or sediment habitats.

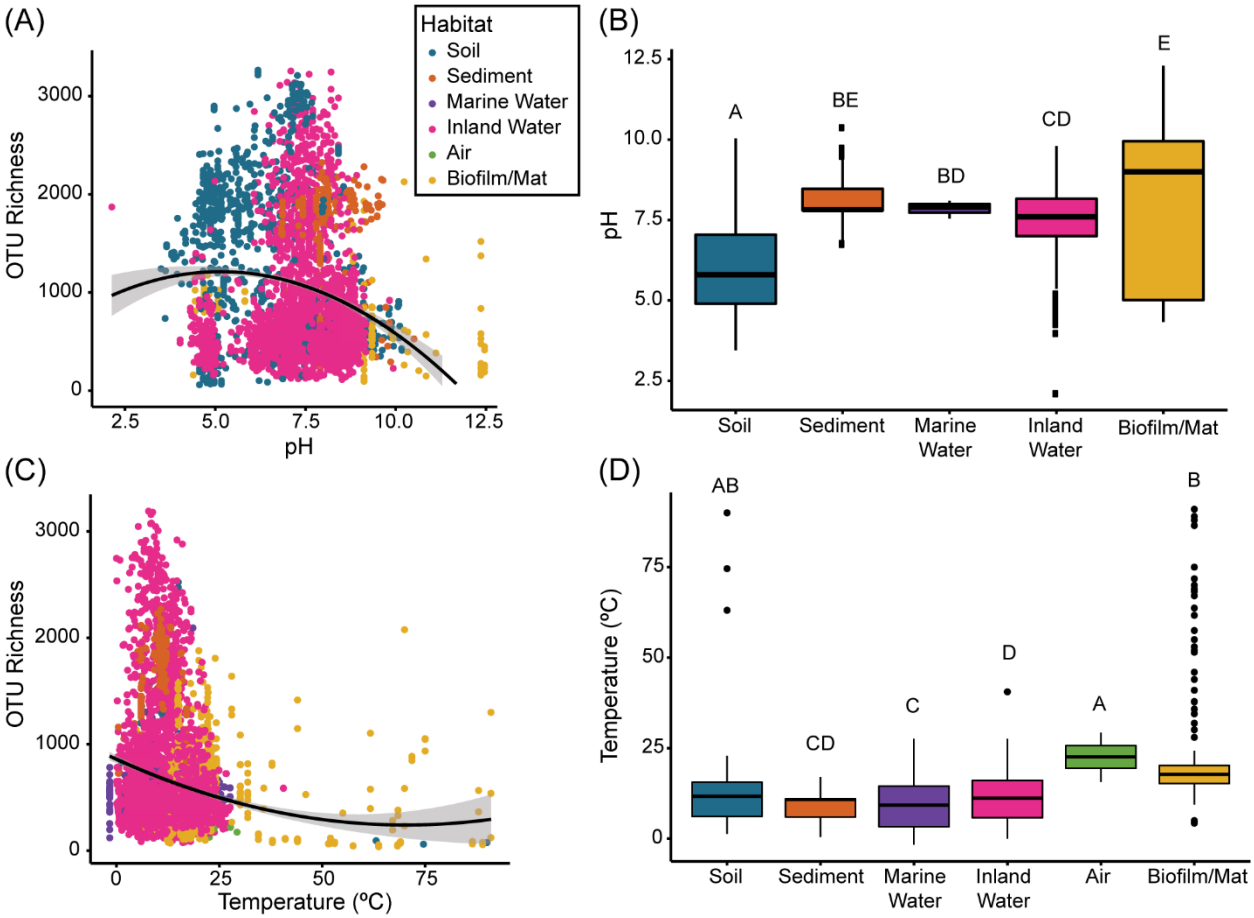

Figure S1.2 The influence of abiotic factors on taxon richness. (A) pH significantly impacts taxon richness ($P < 0.0001$). Each point represents an individual EMP sample (not a geocluster) and is colored by habitat. (B) pH differs among habitats (ANOVA, $p < 0.0001$). Out of all the EMP environmental samples used in this study, 30.7% had associated pH metadata: 0% of air samples, 60.0% of inland water samples, 19.6% of sediment samples, 6.5% of marine water samples, 5.1% of biofilm/mat samples, and 23.4% of soil samples. (C) Temperature significantly influences taxon richness ($P < 0.0001$). Each point represents an individual EMP sample (not a geocluster) and is colored by habitat. (D) Temperature recorded when samples were collected (EMP metadata) differs among habitats (ANOVA, $p < 0.0001$). Out of all the EMP environmental samples used in this study, 47.3% had associated temperature metadata: 9.0% of air samples, 77.0% of inland water samples, 17.5% of sediment samples, 48.8% of marine water samples, 81.0% of biofilm/mat samples, and 0.6% of soil samples.

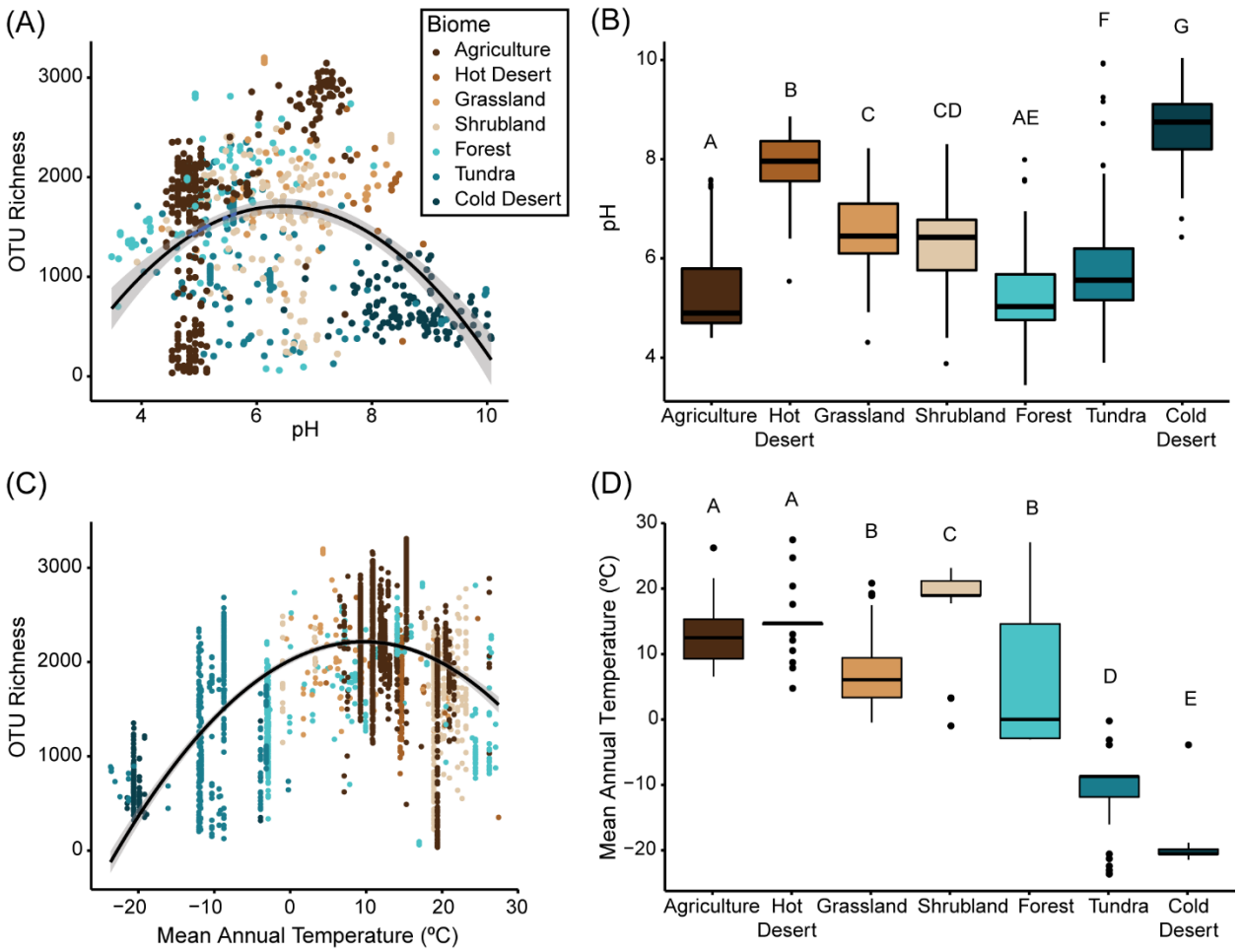

Figure S1.3 Abiotic factors influence taxon richness in soil. (A) pH significantly impacts taxon richness (P = 0.001). Each point represents an individual EMP soil sample (not a geocluster) and is colored by biome. (B) pH differs among biomes (ANOVA, p < 0.0001). (C) Mean annual temperature significantly impacts taxon richness (P < 0.0001). Each point represents an individual EMP soil sample (not a geocluster) and is colored by biome. Temperature data was retrieved from WorldClim, a publicly available data source. (D) Temperature differs among biomes (ANOVA, p < 0.0001).

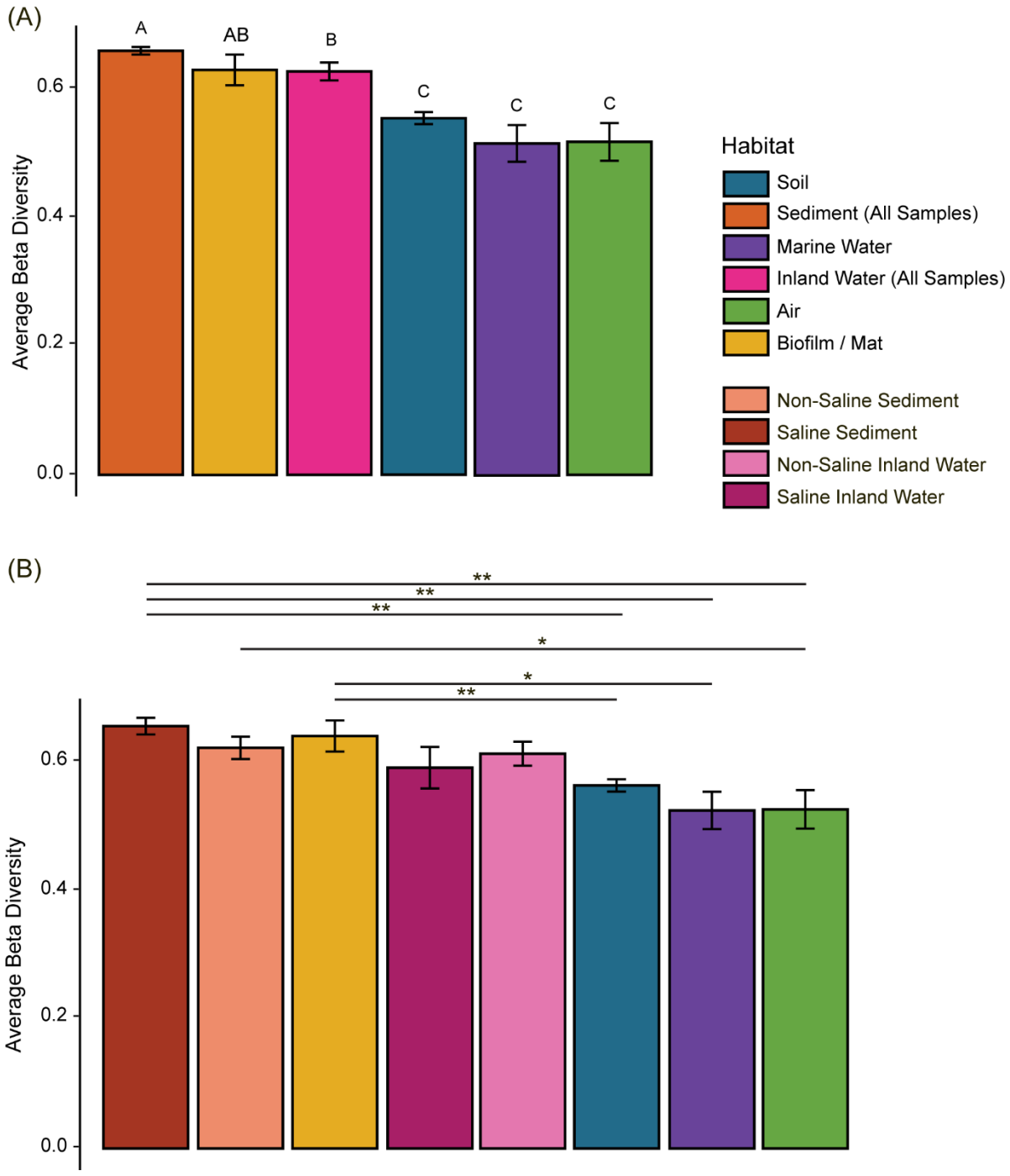

Figure S1.4 The influence of salinity on beta-diversity. The level of beta-diversity within sediments and inland water is not driven by the combination of saline and non-saline samples within a single habitat. (A) Mean beta-diversity (distance from centroid) ± standard error of the six habitats, ranking habitats from highest to lowest beta-diversity. (B) Mean beta-diversity ± standard error of the six habitats with sediment and inland water habitats split into saline and non-saline samples.

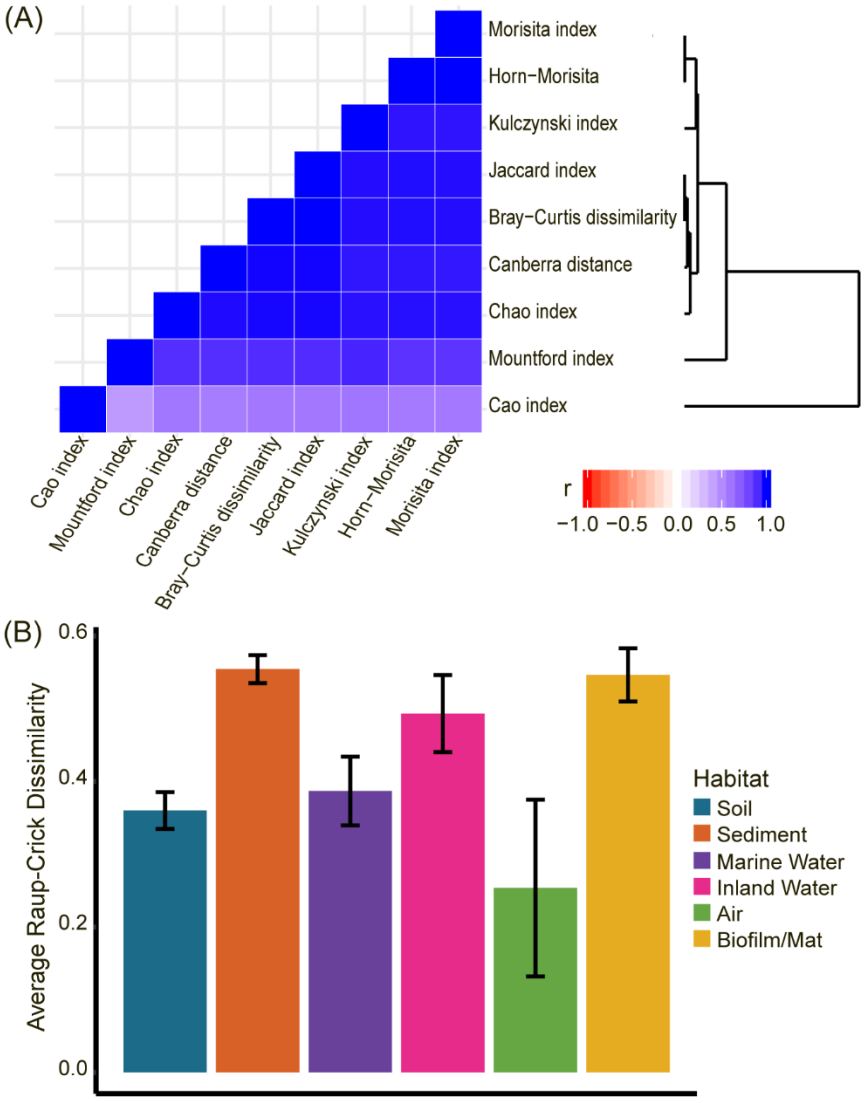

Figure S1.5 Comparison of beta-diversity metrics. Beta-diversity patterns are not dependent on the beta-diversity metric used. (A) Heatmap shows degree of correlation (r from a Spearman's mantel test with all EMP samples used in analysis). Dendrogram shows relatedness of metrics based on their correlation strength. (B) Mean Raup-Crick dissimilarity (distance from centroid) ± standard error. The patterns shown by Raup-Crick analysis match those demonstrated by Bray-Curtis metric.

SUPPLEMENTAL INFORMATION

Appendix 1B. Sample Locations

Sediment samples came from freshwater ponds/lakes and rivers/streams, saline ponds/lakes, coastal, mangrove, and oil-contaminated marine sediments. Marine samples came from northern seas to ocean waters near the equator. Inland water samples included rivers/streams, non-saline lakes/ponds, saline lakes/ponds, well water, groundwater, and sinkholes. Air samples were taken from urban regions in North America and Europe. Biofilm/mat samples were from kelp forests, hot springs, volcanic geothermal vents, coral reefs, Arctic lakes/streams, marine and travertine hydrothermal vents, tundra peat soils, and intertidal regions.

Agricultural soils included samples from vineyards, fertilized temperate grasslands, soybean fields, wheat fields, rice paddies, a legume cropland, a palm plantation, green roofs, an urban park, a coffee plantation, a FACE pasture, a potato farm, a maize field, and a fertilized urban field soil. Grassland soils are from North America, Hawai'i, and Argentina. Shrubland soils came from desert shrublands, mountain shrublands, and tropical volcanic shrublands. Forest soils are from taiga, temperature forests, tropical forests, volcanic forests from Hawaii, and unknown forests. Hot desert soils are from North America and India. Cold desert soils are from Antarctica. Tundra soils are from North America, Asia, Europe, Greenland, and Antarctica.

CHAPTER 2

The routes and rates of bacterial dispersal impact surface soil microbiomes and their functioning

ABSTRACT

Recent evidence suggests that, similar to larger organisms, dispersal is a key driver of microbiome assembly; however, our understanding of the rates and taxonomic composition of microbial dispersal in natural environments is limited. Here, we characterized the rate and composition of bacteria dispersing into the litter layer of surface soil via three dispersal routes (through the air, from nearby vegetation, and up from the bulk soil). We then quantified the impact of those routes on microbial community composition and functioning in soil. The dispersal rate onto the soil surface was low (7,900 cells/cm^2/day) relative to the abundance of the resident community. While bacteria dispersed through all three routes at the same rate, only dispersal through the air and vegetation routes brought new taxa onto leaf litter. Further, only the air and vegetation routes impacted microbiome composition, suggesting that the composition, not rate, of dispersal influenced community assembly. Dispersal also impacted microbiome functioning. When exposed to dispersal, leaf litter decomposed faster than when dispersal was excluded, although neither decomposition rate nor litter chemistry differed by route. Overall, we conclude that individual dispersal routes transport distinct bacterial communities that influence soil microbiome composition.

INTRODUCTION

Dispersal is the movement of individuals or propagules with potential consequences for gene flow (Ronce, 2007). This process has long been recognized as fundamental to the ecology and evolution of plant and animal communities (Shmida & Wilson, 1985; Slatkin, 1987; Vellend, 2010). More recently, evidence has accumulated that dispersal may also be important for microbiomes. Contrary to the long-standing assumption that microbial dispersal is so pervasive that it can be ignored (Baas-Becking, 1934), biogeographic patterns suggest that dispersal limitation influences the evolution and distribution of microbial diversity (Andam et al., 2016; Choudoir et al., 2017; Hanson et al., 2019; Jennifer B. Hughes Martiny et al., 2006; Peay et al., 2012). Likewise, recent experiments that exclude immigration or artificially introduce cells demonstrate the potential for dispersal to alter microbiome composition and functioning (Albright & Martiny, 2018; S. E. Evans et al., 2019; Svoboda et al., 2018). While this evidence demonstrates the potential impacts of dispersal, we still have not measured the rates and taxonomic composition of dispersing bacteria, or the impact of multiple dispersal routes on natural communities.

A dispersal route can be defined as the combination of the source community (e.g., soil or vegetation) and the physical vector (e.g., rain or wind) that moves individual cells. Two main attributes of dispersal routes – the rate at which individual bacteria move through them and the composition of those bacteria – are key to their influence on resident communities (Cevallos-Cevallos et al., 2012; Lindström & Langenheder, 2012; Rime et al., 2016). In laboratory microcosms, higher dispersal rates generally cause greater changes in resident microbiomes (Declerck et al., 2013; Lindström & Östman, 2011; Souffreau et al., 2014). However, the impact of dispersal also depends on taxonomic composition; a route transporting taxa that easily

establish and grow in the resident community can have an outsized impact even under low dispersal rates (Albright et al., 2020; Comte et al., 2017). Thus, it is important to characterize both the rate and composition of a dispersal route to determine its impact.

Yet tracking the movement of microorganisms in the field, let alone characterizing individual dispersal rates, is a challenge. A handful of studies have followed subsets of microbial taxa, such as the accumulation of thermophilic endospores in sediments or the dispersal kernel for a single microbial taxon (E. Bell et al., 2018; Galès et al., 2014). Separately, dispersal routes have been inferred by the similarity between a focal community and potential source communities (Leung et al., 2014; Maignien et al., 2014; Rime et al., 2016), although such inferences conflate the influence of (unknown) environmental selection and dispersal (T. Bell, 2010). Other studies have experimentally blocked all dispersal or specific dispersal routes and then characterized changes in microbial composition. For instance, Kaneko & Kaneko (2004) covered branches of beech trees to investigate the influence of dispersal on endophytic fungi, and Vannette and Fukami (2017) caged flowers to restrict pollination to test for differential effects of animal pollinators on nectar-inhabiting microbial communities. Nonetheless, we still lack direct quantification of both the rates and composition of microbial dispersal routes in natural ecosystems, let alone tests of their differential impacts on microbiome composition and functioning.

To address these gaps, we characterized three potential bacterial dispersal routes and their impact on the surface soil microbiome of a southern California grassland. We asked three questions: **(1)** At what rate and by what routes are bacteria dispersing into the surface soil? **(2)** How do these dispersal routes influence microbiome composition in the leaf litter, the topmost layer of soil? **(3)** Do the routes differentially influence microbiome functioning? We first

characterized the abundance and taxonomic composition of bacteria immigrating onto sterile "traps" – glass microscope slides that allowed little or no cell growth. Three potential routes by which bacteria might disperse into the soil surface are: (1) the regional pool of **air**, settling out or raining down during precipitation events; (2) local live and dead **vegetation**, moved by wind, rain, or gravity; and (3) **soil**, transported by wind or rain up into the surface litter layer. To manipulate the exposure of the slides to these routes, the slides were placed in different locations and enclosed in bags made of material that either allowed (2mm window screen) or blocked (0.22 µm nylon) the immigration of bacteria and larger organisms (Figure 1). In a parallel experiment, we filled a second set of bags with freshly cut grass litter from the field site to test whether the routes differentially altered the composition (bacteria and fungi) and functioning (decomposition rate and litter chemistry) of the resident surface soil microbiome.

MATERIALS AND METHODS

Field site. The experiment was conducted at the Loma Ridge Global Change Experiment in California, USA (33°44′ N, 117°42′ W, 365 m elevation) from April 14, 2018 – October 26, 2018. The site is a Mediterranean grassland (dry warm summers and cool wet winters), with 325 mm mean annual precipitation and 17°C mean temperature, and is dominated by non-native grasses *Bromus madritensis* and *Avena* sp., non-native forbs *Hirschfeldia incana* and *Erodium* sp., and the native forb *Deinandra fasciculata* (Finks et al., 2021; Kimball et al., 2014).

Dispersal slides and litterbags. We characterized dispersal through three routes: settling down from the air, horizontally from nearby surrounding vegetation (shrubs, standing dead grass, and the litter layer), and up from the bulk soil. We measured dispersal onto two substrates: (1) sterile

glass microscope slides to identify immigrating taxa; and (2) green grass leaf litter collected on April 11, 2018, from the site. In total, the experiment encompassed eight dispersal treatments (four dispersal route manipulations x two substrates) and one death rate treatment. This design was replicated in seven experimental blocks (2 m by 2 m each) in an 11 m by 5 m field site.

The dispersal bags were made up of 0.22 μm nylon (Tisch, SPEC17970) and/or 2 mm window screen (Phifer, Model # 3002212) depending on the treatment (Fig. 2.1). Glass microscope slides (2.5 cm x 7.5 cm) were sterilized in 70% ethanol, dried, sealed into bags (5 cm x 10 cm), and autoclaved. Autoclaved litterbags (10 cm x 10 cm) were filled with green grass clipped into 2 cm segments and stored at 4°C for up to three days before placement in the field. The bags were set out in the field on April 14, 2018, and either staked into the ground or stapled to the field tables to secure in place. At each timepoint (May 23rd, June 13th, July 23rd, September 12th, and October 26th, 2018), we collected one bag from each treatment from each experimental block (9 treatments x 7 experimental blocks = 63 samples per collection). Glass slides were transferred to a sterile plastic bag with 2 mL of 1% Pi-buffered GTA and 220 μL of 0.1 M tetrasodium pyrophosphate and processed for community composition and bacterial abundance. Leaf litter samples were weighed, ground to homogenize, and processed for community composition, bacterial abundance, mass loss, and litter chemistry.

Death rate slides. Bacterial death rate on the glass slides was measured so that immigration rate could be calculated using the bacterial abundance on the dispersal slides. To measure death rate, glass slides containing a known number of bacterial cells were placed into the field and sampled alongside the dispersal bags. Bacterial cells were extracted from grass litter from the field site by steeping in 1 L of 0.9% saline solution overnight. The litter was then filtered through cheese cloth, and the filtrate was aliquoted in 2 mL volumes, further concentrated and washed by

pelleting the cells and resuspending into 100 µL sterile 0.9% saline solution. Each aliquot was then spread onto an ethanol-sterilized glass microscope slide and allowed to dry before being sealed into autoclaved nylon (0.22 µm) bag that was closed to dispersal. The slides were kept at 4°C overnight until deployed in the field (on the soil surface). Timepoint 0 samples were suspended in 1% phosphate-buffered glutaraldehyde (Pi-buffered GTA) to preserve cell abundance for flow cytometry. One death rate sample was collected from each experimental block at each timepoint (described above) for a total of 35 samples (7 blocks x 5 timepoints). Additionally, a second set of death rate slides were deployed on both the soil and table surfaces on September 12, 2018, and sampled on September 19th, 26th, October 3rd, 10th, and 26th to calculate the death rate over a finer temporal scale and to capture the difference in death rate between soil and table surface. All samples were processed for bacterial abundance and community composition.

Dispersal sources. To characterize potential sources of dispersal, we also collected air, soil, and environmental leaf litter samples from the field site (N = 3) at each timepoint (for a total of 15 samples each). To collect the air samples, we used the QuickTake® 30 sampling pump with the BioStage® single-stage impactor (SKC, Inc) fitted with a sterile agar plate, collecting air from 4.5 m above ground for 30 min at 2.8 L/min flow rate. For the last timepoint, air samples were collected by directing airflow from a sterilized portable fan (O2COOL, model FD10101A) towards three vertical sterile agar plates for 30 min. All agar plates were kept at 4°C for up to a week after collection. We removed a 4 cm x 4 cm area of the top mm of agar using a sterile razor blade. To collect soil and litter, we pooled samples taken from the top layer of the bulk soil or the litter layer from three corners of each experimental block at each timepoint. Air and soil

samples were processed for community composition, and litter samples were processed for bacterial abundance and community composition.

Bacterial abundance. At the time of sample collection, an aliquot of 0.1 g of ground and homogenized leaf litter was preserved in 1% Pi-buffered GTA and 550 µL of 0.1 M tetrasodium pyrophosphate. Glass slides were transferred to a sterile plastic bag with 2 mL of 1% Pi-buffered GTA and 220 µL of 0.1 M tetrasodium pyrophosphate. All samples were stored in the dark at 4°C for up to two days before being sonicated for 30 minutes in the dark at 4°C, filtered through a 4-µm-pore-size vacuum filter to remove large particulates, and stored in the dark at 4°C for up to one day before being measured on the flow cytometer. To process samples on the flow cytometer, 2 µL of SYBR green (200x, Invitrogen Life Science Technologies, S756) was added to 400 µL of each sample, and samples were incubated in the dark at room temperature for 10 minutes. Samples were run for 30 seconds at 40 µL/min, using a SYBR-Green-H threshold value of 1,500 and SSC-H threshold value of 1,000. Gating parameters were used to count particles in the size of typical bacterial cells, optimized by Khalili *et al.* (2019).

Amplicon Sequencing. Aliquots of 0.05 g of leaf litter, 0.1 g of soil, 250 µL from the glass slide solution taken immediately before vacuum filtering, and the agar scraped from the air samples were frozen at -70°C and stored until extraction. DNA was extracted following the ZymoBIOMICS Microprep DNA Extraction Kit protocol, with the following modifications: (1) for all samples, maximum centrifuge speed was 2808 x g, instead of 3500 x g, and centrifuge time was increased at that speed from 3 min to 4 min and from 5 min to 7 min; (2) for glass slide and air samples only, bead beating was reduced to 3 minutes, instead of 5 minutes, to avoid shearing the DNA in these low biomass samples; and (3) proteinase K was added to glass slide

samples (with the exception of 33 randomly selected samples) to help uncrosslink proteins caused by the addition of the GTA.

To characterize the bacterial community, we amplified the V4 – V5 region of the 16S ribosomal DNA gene using the 515F (GTGYCAGCMGCCGCGGTAA) – 926R (CCGTCAATTCCTTTRAGTTT) primers, described in (Caporaso et al., 2012; Lane et al., 1985). For leaf litter and soil samples, we used 1 µL of template DNA for the PCR. For air and glass slide samples, we used 5 µL of template DNA. For all samples, the remaining PCR reaction contained 12.5 µL of AccustartII PCR ToughMix (Quanta BioSciences, Inc), 0.5 µL forward primer, 0.5 µL reverse primer, and H_2O to reach a final volume of 25 µL. Following an initial denaturation step at 94°C for 3 min, the PCR was cycled 30 times at 94°C for 45 s, 55°C for 30 s, and 72°C for 60 s, with a final extension at 72°C for 10 min.

To characterize the fungal community, we amplified the ITS2 region of the Internal Transcribed Spacer (ITS) using the ITS9F (GAACGCAGCRAAIIGYGA) – ITS4 (TCCTCCGCTTATTGATATGC) primer combination (Looby et al., 2016). For leaf litter and soil samples, we used 1 µL of template DNA for the PCR. For air and glass slide samples, we used 5 µL of template DNA. For all samples, the remaining PCR reaction contained 12.5 µL of AccustartII PCR ToughMix (Quanta BioSciences, Inc), 0.75 µL forward primer, 0.75 µL reverse primer, and H_2O to reach a final volume of 25 µL. After an initial denaturing step at 94°C for 3 min, the PCR was cycled 35 times at 95°C for 45 s, 50°C for 60 s, and 72°C for 90 s, with a final extension at 72°C for 10 min.

To prepare libraries for sequencing, PCR products were pooled at different volumes based on amplification brightness on gel pictures: high (1 µL), medium (2 µL), low (3 µL), very low (5 µL), and none (8 µL). Glass slide, environmental litter, air, soil, and death rate samples

were pooled together for one sequencing run. Litterbag samples and environmental litter samples were pooled together for a second sequencing run. We sequenced the environmental litter in both runs to control for differences in sequencing runs. After pooling, libraries were purified using SpeedBeads magnetic carboxylate-modified particles (GE Healthcare UK Limited). Purified libraries were sequenced in two paired-end Illumina MiSeq runs (2 x 250 bp) at the Genomics High Throughput Facility, UC Irvine, Irvine, CA, USA.

Bioinformatic processing and analysis. Sequence data were processed in QIIME2 (Bolyen et al., 2019), version 2018.11. For samples in the first sequencing run, the forward reads were trimmed to 6 – 279 base pairs of the 16S segment. For samples in the second sequencing run, the forward reads were trimmed to 6 – 300 base pairs. The reverse reads from both runs were discarded. We used DADA2 (Callahan et al., 2016) to define exact sequence variants and assigned taxonomy using classify-sklearn with, for bacteria, the reference Greengenes database (13_8) at a 99% OTU level (McDonald et al., 2012) and, for fungi, the UNITE (v. 7.2) database (Nilsson et al., 2019).

Decomposition and litter chemistry. Decomposition, or mass loss, was measured as the percent decrease in dry weight. Dry weight was calculated by multiplying the wet weight of the leaf litter (both pre- and post-experiment) by the ratio of litter dry weight/wet weight. The ratio of dry weight/wet weight was calculated by drying a 1 g subset of the wet litter (taken at time of sample collection) overnight in a 60°C oven until constant mass and dividing the dry weight by the initial wet weight. Ground and oven-dried litter collected at 2 months (on June 13, 2018) was analyzed for litter chemistry using attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy.

Temperature, precipitation, and light. Precipitation, temperature, and wind speed data for the entire field site was collected from the weather station at the site maintained by the Center for Environmental Biology at the University of California – Irvine. We deployed Onset HOBO Pendant data loggers (UA-002-64) on the tables and soil surface in the dispersal experiment. Although these data were collected outside of our field experiment timeframe (February 15th – March 5th, 2019), they allowed us to compare the relative light exposure of the soil and table surfaces.

Statistical analyses. We accounted for differences in sequencing depth by rarefying samples to 1,000 sequences (bacterial communities) or 3,500 sequences (fungal communities) with 1,000 resamplings. For each resampling, we calculated a Bray-Curtis dissimilarity matrix, taking the median similarity values of all 1,000 matrices using the 'vegan' package in R (Oksanen et al., 2019). Using this median matrix, we performed a PERMANOVA and *post-hoc* test in PRIMER+ (M.J. Anderson et al., 2008; Clarke & Gorley, 2006) to test for differences in community composition by treatment and collection month. Bacterial composition on the glass slides was further linked to the potential dispersal sources using SourceTracker, a Bayesian approach that estimates the proportions of a community that come from possible sources (Knights et al., 2011). To analyze alpha-diversity, we used the rounded rarefied dataset to calculate the Shannon index in R and an ANOVA in R to test for differences among treatments. Beta-diversity (Bray-Curtis metric) was analyzed using the betadisper() function in the 'vegan' package in R. Finally, to analyze similarity of litterbag samples to environmental (surrounding) litter over time, we calculated the centroid of the environmental litter samples within each collection month using PRIMER+ and then used an ANOVA in R on the similarity between each litterbag sample and the environmental litter centroid at each collection month.

To isolate the contribution of each dispersal route on univariate data (bacterial abundance, mass loss), we subtracted data between treatments in a nested fashion (Fig. 2.1). To analyze the contribution of dispersal route and time on bacterial cell abundance or mass loss, we performed a Type III ANOVA using the 'car' package (Fox & Weisberg, 2011) and Tukey's HSD in R (R Core Team, 2020) for each subtracted dataset (bacterial abundance on glass slides; bacterial abundance in litterbags; mass loss). We also tested whether dispersal increased subtracted bacterial abundance or mass loss by performing a one-sample t-test that tested the difference from zero for each dispersal route overall and by route at each timepoint. As a complement, we analyzed the contribution of dispersal route and time on the non-subtracted bacterial abundance and mass loss data using a Type III ANOVA and Tukey's HSD in R. To test the relationship between bacterial abundance in the litterbags and the decomposition rate, we performed a Spearman's correlation on the two variables. To analyze litter chemistry, we calculated the Euclidean distance between samples of the baseline-corrected dataset (between 900-1800 nm) and used a PERMANOVA from the 'vegan' package in R to test for differences in chemistry among treatments, using the FTIR spectral data. To compare the light levels received on the soil and table surfaces, we performed a two-sampled Welch's t-test with that alternative hypothesis that the table surface receives higher light intensity than the soil or ground surface.

RESULTS

The rate and composition of bacteria dispersing into soil depends on the route

We destructively sampled the glass slides (n = 7) from four treatments (Elevated, Overhead, Open, and Closed) over six months, comparing cell abundances and taxonomic

composition between the treatments. The four treatments allowed us to assess the contribution of the three dispersal routes to surface litter communities in a nested manner (Fig. 2.1). Closed samples were exposed to no dispersal. Elevated samples were exposed to only the air route, Overhead to the air and vegetation routes, and Open to all three routes (air, vegetation, and soil). Where we can subtract data between treatments (i.e., for univariate data such as cell abundance and decomposition), we report estimates by route. However, for data that cannot be subtracted between treatments (i.e., multivariate data including community composition and litter chemistry), we report the results by treatment – Closed, Elevated, Overhead, and Open – and infer route effects.

The average immigration rate of bacterial cells was similar through all three dispersal routes but varied over time in a route-dependent manner (Table S2.1; Fig. 2.2B). After accounting for cell death (3.34% of cells died daily; Fig. S2.1; see Supplemental Text), we calculated an average of 7,900 cells/cm^2/day immigrated onto the Open glass slides, equivalent to 0.47% of the average abundance in Open litter communities. Not all dispersal routes transferred cells onto the glass slides at every timepoint. Immigration through air was significantly greater than zero during all months (one-sample t-test: $P < 0.01$) except October ($P = 0.84$), whereas immigration through vegetation only occurred during May, June, and October ($P < 0.05$). Further, immigration through the soil route was highest at the October collection immediately after a large rainfall ($P = 0.015$; Fig. S2.2), suggesting that rain transferred cells from the soil up onto the slides. In contrast, immigration through the soil route was not significantly different from zero in July or September ($P > 0.05$) when the site received less than one mm of rain per month.

The taxonomic composition of dispersing bacteria also depended on the route (Table S2.2, Fig. 2.3), with large differences between the air and vegetation routes. Bacterial composition on the Elevated slides appeared most similar to composition detected in environmental air samples (Fig. 2.3A), although the two communities were still distinct (*post-hoc* test: P = 0.005). These communities were dominated by the genera *Methylobacterium* and *Bacillus* (Fig. 2.3B), which are known to be viable in the atmosphere and have a high tolerance to desiccation and UV (Bryan et al., 2019; Smith et al., 2018). Bacterial composition on the Overhead and Open slides differed from Elevated slides (P = 0.001) but not from one another (P = 0.35). These communities most closely resembled those in the nearby, environmental leaf litter (Fig. 2.3A). Their composition comprised a relatively even abundance of the genera *Hymenobacter*, *Sphingomonas*, *Janthinobacterium*, *Curtobacterium*, and *Methylobacterium* (Fig. 2.3B), taxa commonly observed in Loma Ridge leaf litter (Finks et al., 2021; Matulich et al., 2015). Thus, the signature of dispersal into the soil surface from the vegetation route was strong compared to the soil or air routes. In support of this result, a SourceTracker analysis (Knights et al., 2011) estimated that 40% of the Overhead and Open communities could be traced to environmental leaf litter, whereas fewer taxa were traced to environmental soil (3.9%) and air (9.8%) (Fig. S2.3).

This taxonomic composition of the glass slides also changed significantly over time and by the interaction between route and time (Table S2.2). In particular, Elevated samples showed higher variability throughout the course of the experiment (Fig. 2.3B; Fig. S2.4D), possibly reflecting the temporal variability of air communities (Bertolini et al., 2013).

Dispersal routes differentially influence the leaf litter bacterial community

Next, we investigated the influence of different dispersal routes on the surface litter microbiome. We constructed litterbags containing freshly cut grass from the field site to follow the successional shift from a phyllosphere to a decomposer community (Voříšková & Baldrian, 2013) in treatments parallel to the glass slide experiment (Closed, Elevated, Overhead, Open). The phyllosphere, which includes the leaf surface and the apoplast (Rastogi et al., 2013), contains diverse microbial communities that may contribute to the decomposer community as the leaf goes through senescence.

Bacterial abundance on the leaf litter was altered by dispersal in a route-specific manner (Table S2.3; Fig. 2.4B). In the absence of dispersal (Closed litterbags), abundance increased and then decreased over the course of succession (Fig. 2.4A). Dispersal from the air route did not increase bacterial abundance on top of this baseline pattern (one-sample t-test: $P = 0.99$; Fig. 2.4B). In contrast, dispersal from the vegetation route resulted in significantly higher cell abundance ($P < 0.0001$), where the increase was greatest during the first month of the experiment (May) and diminished throughout the experiment. As a comparison, the bacterial abundance in the surrounding litter increased throughout the experiment, although remained relatively low compared to the abundance within the litterbags (Fig. 2.4A). Moreover, additional dispersal from the soil route did not further increase bacterial abundance over that of the vegetation route ($P = 0.61$), indicating that dispersal from local vegetation was likely responsible for the effect on bacterial abundance in the litter.

As with abundance, the composition of the litter microbiome was influenced by the dispersal route (Table S2.4A; Fig. 2.5A). Both the air and vegetation routes influenced composition, whereas the soil route did not (the Overhead and Open communities did not differ

in community composition; *post hoc* comparison: P = 0.81). Bacterial community composition also changed over time, as expected during the succession of decomposing litter (Table S2.4A). However, dispersal also impacted this successional pattern, as indicated by a significant time-by-route interaction (Table S2.4A). Specifically, communities exposed to the vegetation route (Overhead and Open) resembled the composition of the surrounding environmental litter (a later stage of decomposition than the litterbags) after only a month (Fig. S2.5). In contrast, the Elevated treatment did not converge on the environmental litter composition until three months of decomposition, and the Closed litterbag communities remained the most dissimilar throughout the experiment. However, the Elevated treatment was also exposed to higher light intensities (Welch's t-test: $P < 0.0001$, $t = 26.10$) than the other treatments, and the effects from dispersal exposure cannot be separated here from the possible effects from increased radiation.

The temporal dynamics of specific abundant genera further illustrate the importance of dispersal route for microbiome succession on the leaf litter (Fig. 2.5B). *Erwinia*, a dominant grass phyllosphere member (Lindow & Leveau, 2002), was the most abundant bacterial taxa (85%) in the litter communities across all treatments after the first month. Over time, however, the communities that received dispersal through the vegetation route (Open and Overhead) showed an increase (9% to 68%) in *Janthinobacterium*, an abundant genus in surrounding leaf litter at the site (Finks et al., 2021). In contrast, the treatment exposed only to the air route (Elevated) showed an increase in *Hymenobacter* from undetectable to 64% at the end of the experiment, whereas in the absence of all dispersal (Closed), *Paenibacillus* and *Sanguibacter* became more abundant (8% and 5%, respectively) than in the treatments open to dispersal (1%). Further, dispersal through the vegetation route altered the overall alpha-diversity of the litter communities in ways consistent with later stages of litter decomposition (Purahong et al., 2016);

communities exposed to the vegetation route (Overhead and Open) exhibited higher Shannon diversity (alpha-diversity) than Closed litter ($P < 0.005$; Fig. S2.4B). Overhead and Open treatments generally showed lower compositional variation (beta-diversity) than treatments not exposed to the vegetation route (Closed and Elevated) (Fig. S2.4E; betadisper: $P = 0.001$). These results, as well as previous studies (Cadotte, 2006; Souffreau et al., 2014), support the metacommunity theory predictions that increased dispersal will increase alpha-diversity and decrease beta-diversity (Leibold et al., 2004).

Dispersal via the air and vegetation routes accelerates initial litter decomposition

To test the influence of dispersal routes on community functioning, we tracked the rate of decomposition within the grass litterbags, comparing among treatments. Whereas all treatments plateaued in mass loss during the dry summer months, Elevated leaf litter exhibited a linear increase throughout the entire experiment (Fig. 2.6A). Elevated samples were exposed to higher light intensities (Welch's t-test: $P < 0.0001$, $t = 26.10$), and by the end of the experiment, radiation-resistant genus *Hymenobacter* (Lee et al., 2014; Sedláček et al., 2019) made up 53.6% of the community (Fig. 2.5B). Thus, we hypothesize that this high rate of decomposition reflected increased photodegradation (Austin & Vivanco, 2006), rather than biotic decomposition, and we removed this treatment from further decomposition and litter chemistry analysis.

Amongst the remaining treatments, the litterbags lost on average 30.3% of their mass throughout the six-month experiment. Exposure to dispersal through the air and/or vegetation routes increased mass loss during the first month of the experiment (one-sample t-test: $P = 0.015$;

Table S2.5; Fig. 2.6B), during which Overhead and Open communities degraded leaf litter 2.5 times faster than Closed communities (Fig. 2.6A). The soil route did not further increase mass loss above and beyond the vegetation and air routes during the first month (P = 0.070; air and vegetation dispersal routes cannot be separated here because the Elevated treatment was removed from the analysis). Since bacterial abundance in the litterbags was not correlated with decomposition rate (Spearman correlation: P = 0.16), dispersal from the air and/or vegetation routes likely introduced specific taxa that were able to decompose dead plant litter more quickly. By the second month, however, all treatments degraded the leaf litter at the same rate (Table S2.5; Fig. 2.6A) and, likewise, the chemical composition of the litter did not differ by treatment (Table S2.6; Fig. S2.6).

Dispersal routes also influence fungal composition

Although our study focused on bacterial dispersal, fungi are also important decomposers in leaf litter at this research site (Glassman et al., 2018). We therefore also characterized fungal community composition in the grass-filled litterbags through ITS metabarcoding. Similar to the bacterial communities, dispersal via the air and vegetation routes, but not soil, influenced fungal community composition (PERMANOVA: P = 0.001; Table S2.4B; Fig. S2.7). Communities exposed to dispersal via air (Elevated, Overhead, and Open treatments) differed from Closed communities (*post-hoc* test: P = 0.001) and showed higher proportions of the air-associated genera *Aureobasidium* (4.0%) and *Filobasidium* (4.1%) (Elmassry et al., 2020; Punnapayak et al., 2003) than the Closed treatment (1.3% and 0.7%, respectively). Likewise, the vegetation-exposed Overhead and Open communities differed from Elevated communities (P = 0.001), but not one another (P = 0.884), and showed higher proportions of the decomposer taxon

Paraconiothyrium (15.1%) than the Elevated or Closed communities (2.0%) (Van Diepen et al., 2017). Further, like bacterial communities, increasing exposure to dispersal (Closed to Elevated to Overhead to Open) increased fungal alpha-diversity (Fig. S2.4C). The only pattern that differed from bacteria was that exposure to the vegetation route increased (rather than decreased) fungal beta-diversity relative to the Closed and Elevated communities (PERMDISP: $P = 0.001$; Fig. S2.4F). This discrepancy might be caused by the higher spatial and temporal heterogeneity of fungal versus bacteria community composition in this system (Finks et al., 2021), which could lead to heightened priority effects (Schmidt et al., 2014; Vannette & Fukami, 2017).

DISCUSSION

Here, we measured bacterial dispersal into the soil surface through three dispersal routes and characterized their effects on microbiome composition and functioning. The overall rate of dispersal through the three routes was similar but remarkably low, despite the high abundance of bacteria in the surrounding potential sources of dispersing taxa, such as leaf litter and bulk soil, which both contain roughly 10^8 cells g^{-1} dry weight (Khalili et al., 2019). Incoming bacterial cells made up only a small fraction (0.5% per day) of the open litterbag community abundances. Given that leaf litter is often a much deeper layer than that in the litterbags, this dispersal rate is likely a highly conservative percentage of the resident litter community, suggesting that dispersal does not swamp the resident surface soil communities and lead to mass effects (*sensu* (Leibold et al., 2004)). It is difficult to directly compare this continuous rate to previous mesocosm and modeling studies as those manipulate dispersal by adding pulses, or discrete additions, of cells to a resident community. However, in those studies, minimum pulses of 2.5-25% of total abundance were needed to impact community composition (S. Evans et al., 2016; Lindström &

Östman, 2011; Souffreau et al., 2014). Thus, the dispersal rate that we quantified in this natural system altered the composition and functioning of the leaf litter microbiome despite a seemingly low (albeit continuous) rate.

Two dispersal routes – air and vegetation – transported a unique composition of bacteria into the soil surface, and, in turn, both routes had unique impacts on microbiome composition. Microbial taxa dispersing through the air route most closely resembled taxa sampled directly from the air and, likewise, taxa moving through the vegetation route resembled those found in the surrounding leaf litter. In contrast, taxa dispersing through the soil route were indistinguishable from the vegetation route and did not resemble surrounding bulk soil, suggesting that most of the cells dispersing through the soil route may be derived from the soil surface interface rather than dispersing up from the bulk soil below. Furthermore, exposure to the air and vegetation routes changed the course of bacterial and fungal succession, shifting these communities towards the composition of the surrounding leaf litter, presumably reflecting a later stage of leaf litter succession. Given that the dispersal rates between these routes did not differ, we conclude that the taxonomic identities, rather than the rates, of dispersing bacteria changed the resident community by outcompeting resident phyllosphere taxa. For instance, the soil route did not affect leaf litter composition even though it increased the overall rate of immigration into the litter communities, suggesting that immigrating taxa influence the resident community through biotic interactions rather than mass effects (Albright et al., 2020).

The differential impacts of the dispersal routes on microbiome composition also had initial functional consequences. Exposure to the air and/or vegetation routes accelerated decomposition in the Overhead treatment during the first month of the experiment. Although we cannot tease apart the effect of the air route (having removed the Elevated treatment in this

analysis), we hypothesize that the increase in decomposition was due to dispersal through the vegetation route and in particular, immigration of well-adapted litter decomposers to the phyllosphere communities on the fresh green leaves. In support of this interpretation, dispersal through the vegetation route alone increased bacterial abundance in the first three months of decomposition; past studies in this system demonstrate that higher decomposition rates lead to high bacterial abundance (rather than vice versa) (Baker et al., 2018; Glassman et al., 2018). However, after two months, all communities – including those without dispersal – experienced similar levels of mass loss and similar litter chemical makeup. This attenuation shows support for previous findings that dispersal impacts are strongest during early succession (T. Bell, 2010).

This study demonstrates the possibility of directly characterizing bacterial dispersal by different routes in an ecosystem; however, we also acknowledge limitations in the extent to which the results can be generalized, even to other surface soils. Dispersal may be more important during environmental shifts or stressors because immigrating taxa adapted to the new conditions may outcompete resident taxa (Comte et al., 2017). In particular, our experiment used freshly cut litter, which may be more susceptible to dispersal than litter in later stages of decomposition as the transition from green leaf to litter represents significant changes in the leaf environment (e.g., nutrient content, water availability). Further, the environmental context may impact dispersal routes and their influence on communities (Albright et al., 2020). For instance, litter decomposition generally stops during the dry summer months at our site (Jennifer B. H. Martiny et al., 2017), and our results may have differed if the experiment was performed during the wet season instead. In fact, after an unusually heavy rainstorm for the season in October, the rate at which bacteria dispersed through the soil route increased significantly, indicating that abiotic conditions may have a strong impact on dispersal rate by route. Along these lines, the

dispersal routes in our study may also be specific to this leaf litter system. Our site comprises a thick leaf litter layer with minimal exposed bulk soil, perhaps explaining why dispersal through soil did not impact community assembly. A site with more exposed bulk soil would likely show different results, as we hypothesize that environmental context has a strong influence on dispersal. For example, a study conducted along the coast would likely observe influence from sea spray dispersal (Santander et al., 2021), yielding a different impact from dispersal through air on microbiomes.

CONCLUSION

Our study suggests that dispersal impacts the surface soil microbiome in a route-dependent way, driven by differences in the taxonomic composition of the bacteria immigrating through different routes rather than by differences in dispersal rates. By adopting the experimental approaches demonstrated here to quantify dispersal routes in other systems, we can start to ask how differences in the environment (e.g., precipitation, degree of wind) or ecosystem (e.g., plant community) may impact microbial dispersal. For example, in this study, we observed seasonal variation in dispersal through the three routes, observing that dispersal rates and taxonomic composition vary over time. We hypothesize that meteorological conditions may strongly influence microbial dispersal and, thus, impact microbiome assembly. In this litter system, microbiome structure was mainly impacted by dispersal from surrounding vegetation, which was primarily composed of older leaf litter. If new litter communities are primarily colonized, or "seeded," by older decomposer communities, then disturbances such as fire that remove the litter layer "seed bank" are likely to significantly delay or prevent the assembling of decomposers in new leaf litter, a major source of nutrient addition to soil (Hobbie, 2015).

ACKNOWLEDGMENTS

We thank the Center for Environmental Biology at UCI for maintaining the weather station and making available the weather data used in this analysis. We also thank Steve Allison and Cascade Sorte for their input on analyses and methods, and Claudia Weihe for her guidance with methodological development and laboratory techniques. We further thank Cynthia Rodriguez, Sarai Finks, Nick Scales, Claudia Weihe, Tarif Rashid, Lizandro Herandez, and Khanh Huynh for comments on earlier drafts. Funding was provided by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (DE-SC0020382). This work was supported by the Ridge to Reef National Science Foundation Research Traineeship, award DGE-1735040 and a National Science Foundation Graduate Research Fellowship Program (DGE-1321846) to KEW. We also thank the Irvine Ranch Conservancy for facilitating research at Loma Ridge. The LRGCE is located on the unceded territory of the Kizh and Acjachemen indigenous peoples.

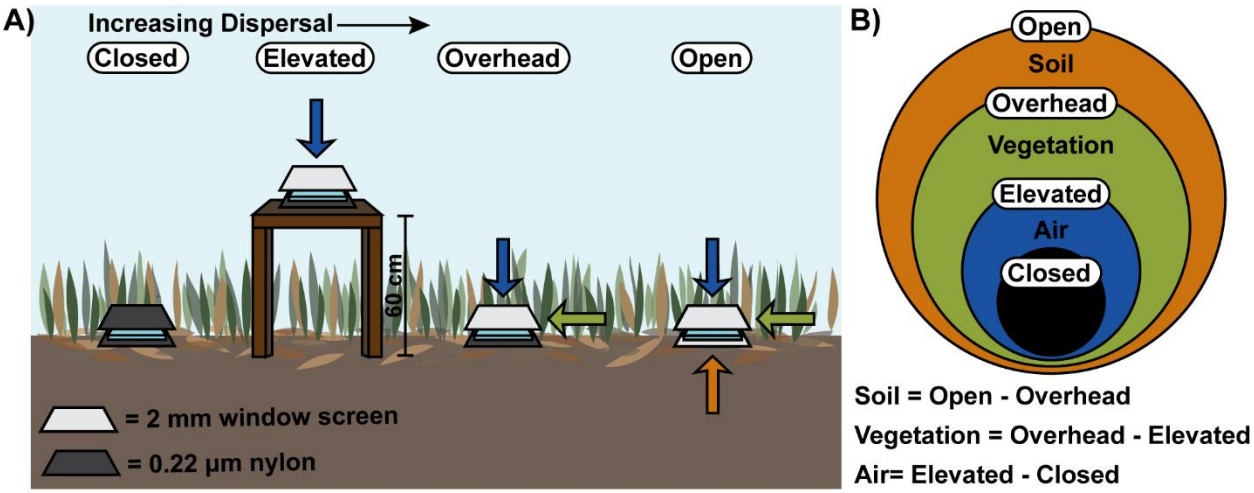

Figure 2.1 Overview of the dispersal experiment. (A) The four dispersal bag treatments contained one of two substrates, either glass slides or grass litter. The Closed treatment, a negative control, was placed on the ground and closed to all dispersal using 0.22um nylon mesh bag. The Elevated treatment was placed on 60 cm above the surrounding vegetation in bags open to dispersal from above (2 mm window screen) but not below (0.22 um nylon). The Overhead treatment was placed on the ground and open to dispersal from above (2 mm window screen) but not below (0.22 um nylon). The Open treatment was placed on the ground and open to all dispersal (2 mm window screen). (B) Dispersal routes were isolated by comparing the treatments in a nested manner: contribution by the air route was inferred by the difference between the Elevated and Closed treatments; the vegetation route by the difference between the Overhead and Elevated treatments; and the soil route by the difference between the Open and Overhead treatments.

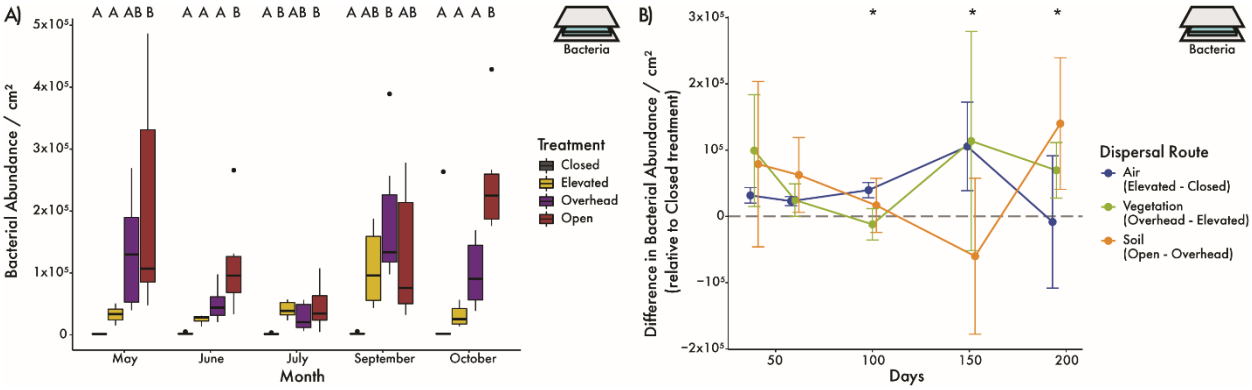

Figure 2.2 Bacterial abundance on the glass dispersal slides. (A) Treatment (ANOVA: $P < 0.0001$, $r^2 = 0.30$, $F = 26.18$), month ($P < 0.0001$, $r^2 = 0.11$, $F = 7.18$), and the interaction between treatment and month ($P < 0.001$, $r^2 = 0.15$, $F = 3.20$) are significant predictors for bacterial abundance on dispersal slides. Abundance differs by treatment in all months ($P < 0.05$)**.** Letters indicate pairwise significance within each significant month (Tukey's HSD). **(B)** Abundance by dispersal route is calculated by the difference between treatments (Table S2.1). Asterisks indicate that abundance differed by dispersal route (*post hoc* ANOVA: $P < 0.05$). Error bars represent 95% confidence intervals.

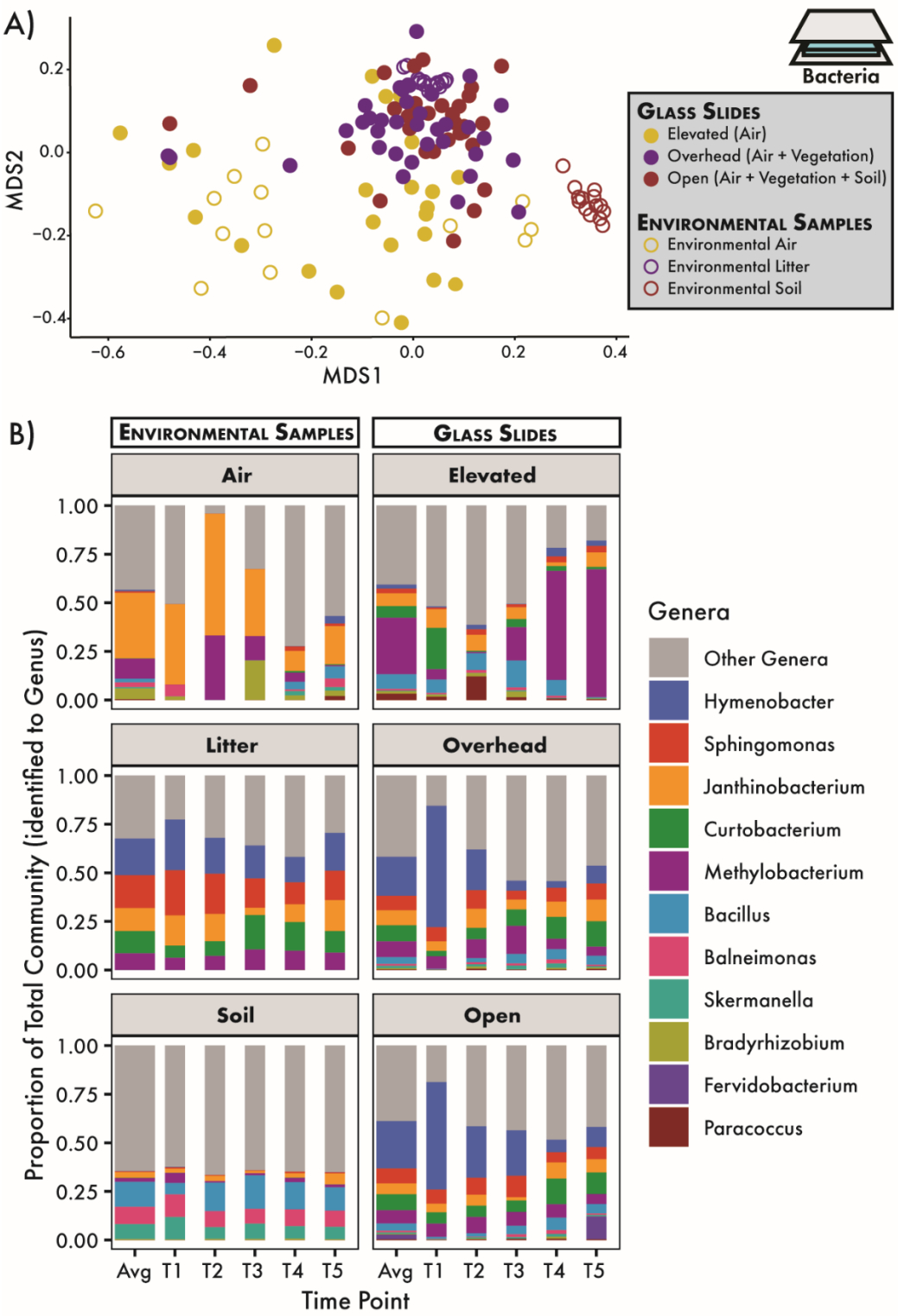

Figure 2.3 Community composition of dispersing bacteria. (A) Visualization (NMDS) of the composition of bacteria dispersing onto the glass slides and of bacteria from potential environmental sources. (B) The most abundant bacterial genera (relative abundance > 9% of community) in the environmental and glass slides samples, averaged (Avg) and by time point.

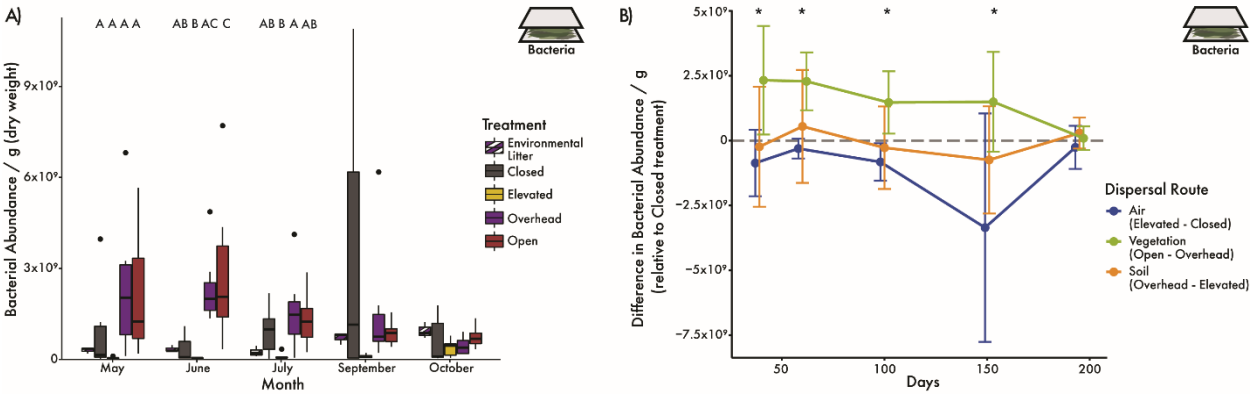

Figure 2.4 Bacterial abundance in the litterbags. (A) Both treatment (ANOVA: P = 0.00026, r^2 = 0.12, F = 6.87) and the interaction between treatment and month (P = 0.015, r^2 = 0.19, F = 2.06) are significant predictors for bacterial abundance. Abundance differs by treatment for the first three months (P < 0.05) but not September (P = 0.12) or October (P = 0.53). Letters indicate pairwise significance within each significant month (Tukey's HSD). Environmental litter is provided for context but was not included in statistical analyses. **(B)** Abundance by dispersal route is calculated by the difference between treatments (Table S2.3). Asterisks indicate that abundance differed by dispersal route for all timepoints (*post hoc* ANOVA: P < 0.05), except October (P = 0.36). Error bars represent 95% confidence intervals.

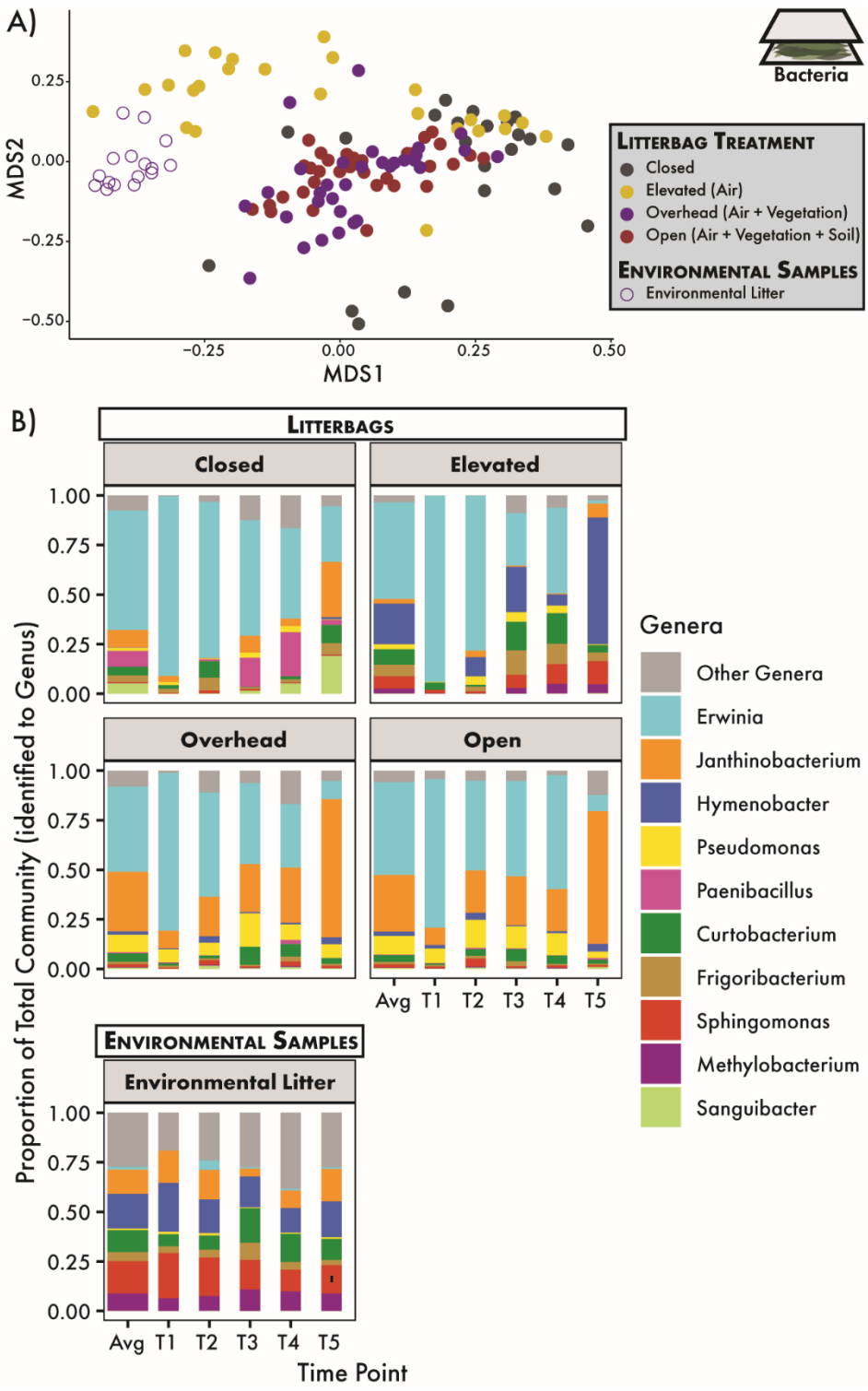

Figure 2.5 Community composition of litterbag bacteria. (A) Visualization (NMDS) of the composition of litterbag bacterial communities exposed to different dispersal routes including environmental (surrounding) litter as a comparison. (B) The most abundant bacterial genera (relative abundance > 9%) in the litterbag communities and the environmental litter, averaged (Avg) and by time point.

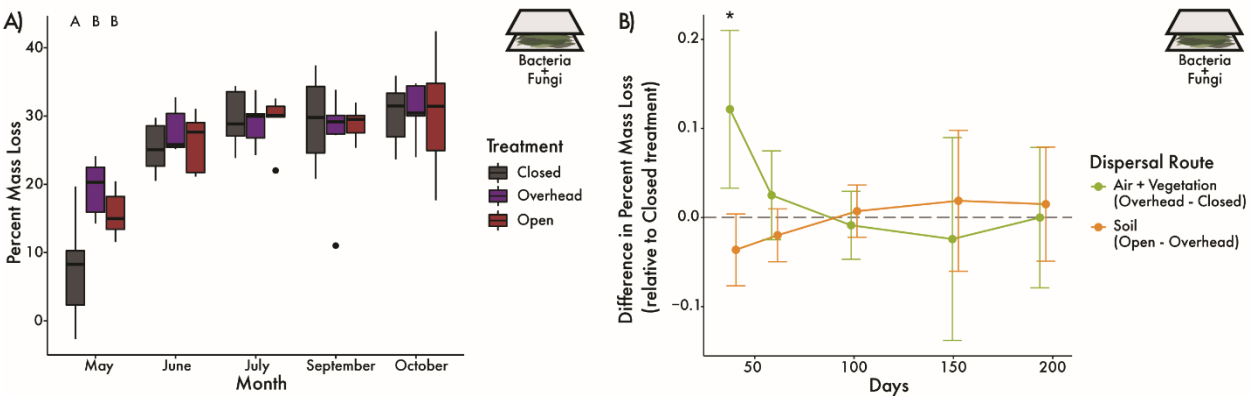

Figure 2.6 Mass loss within litterbags. (A) Treatment impacted mass loss in May (ANOVA: F = 10.02, P = 0.0012, r^2 = 0.53) but no other months. Letters indicate pairwise significance within collection month (Tukey's HSD). **(B)** Mass loss by dispersal route is calculated by the difference between treatments, grouping air and vegetation routes together. May is the only month when mass loss differed significantly by dispersal route (Table S2.5), represented by an asterisk. Error bars represent 95% confidence intervals.

SUPPLEMENTAL MATERIALS

Appendix 2A: Supplemental Figures

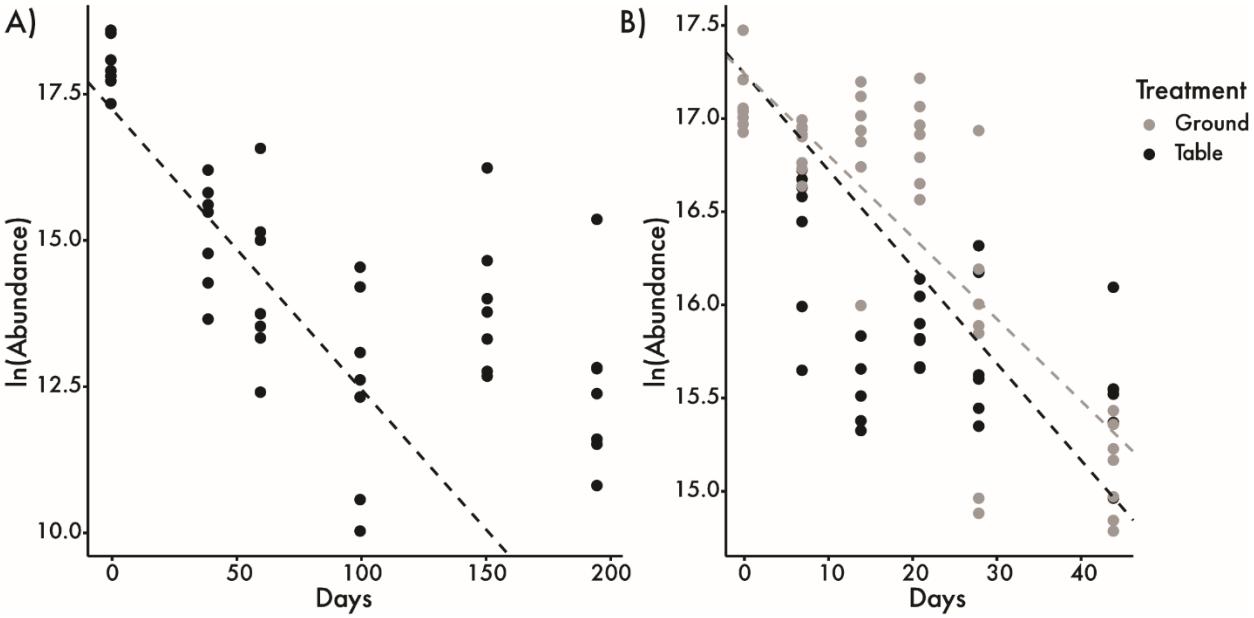

Figure S2.1 Results of the two death rate experiments. (A) The abundance of bacteria cells remaining on the glass slides on the ground for the entirety of the field experiment (April – October; linear model: $r^2 = 0.81$, $P < 0.0001$); (**B**) Comparison bacterial cells remaining on glass slides on the ground ($r^2 = 0.81$, $P < 0.0001$) and on tables ($r^2 = 0.80$, $P < 0.0001$) for the last 1.5 months of the field experiment (September – October).

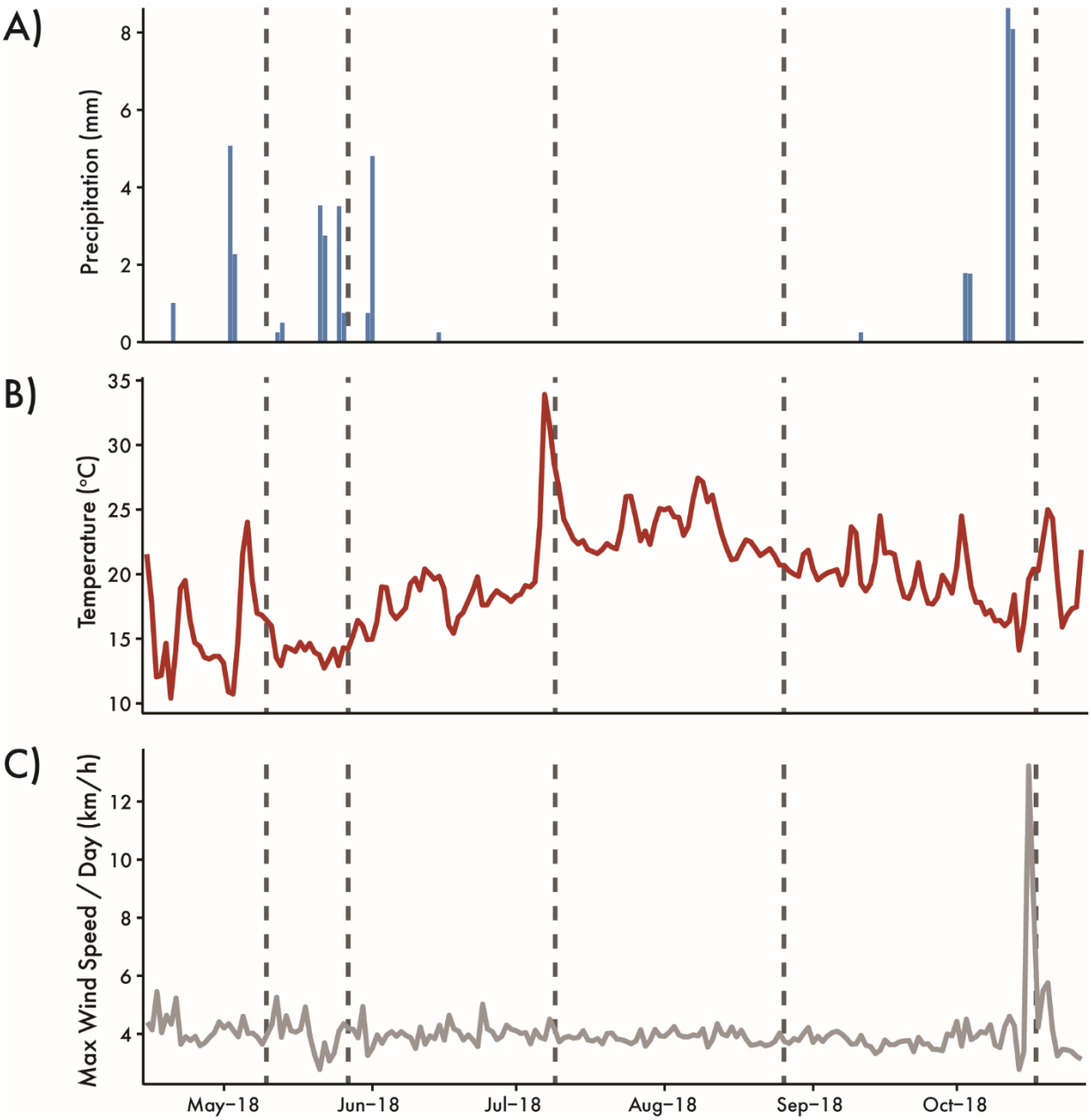

Figure S2.2 Weather data during the field experiment. (A) Cumulative precipitation at the field site per day; (**B**) Average daily temperature at the field site; (C) Maximum daily wind speed at the field site. Vertical dashed lines represent sampling dates.

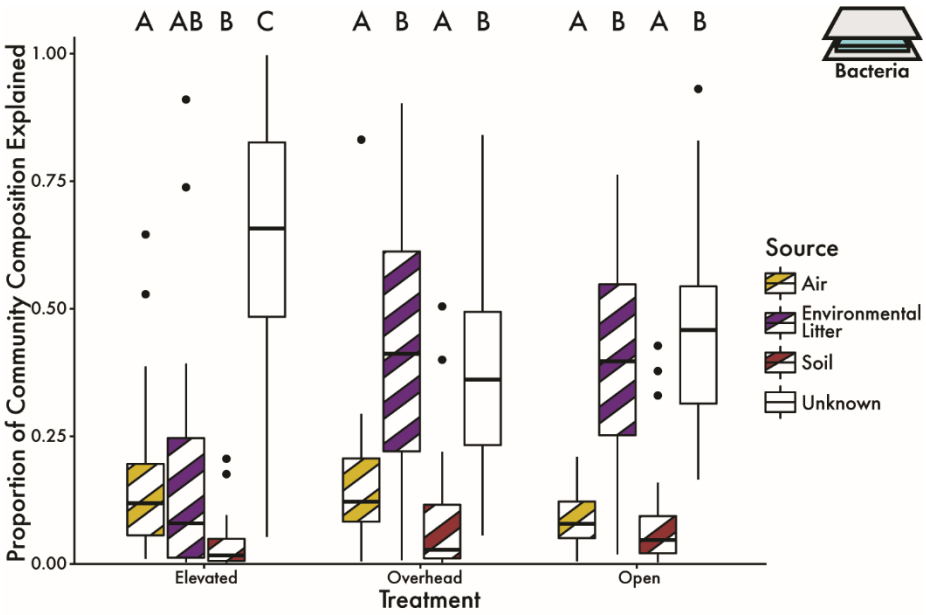

Figure S2.3 SourceTracker results. Proportion of bacterial community composition that was attributed to different source communities for Elevated (Kruskal-Wallis: $P < 0.0001$), Overhead ($P < 0.0001$), and Open ($P < 0.0001$) treatments. Letters indicate significant pairwise differences by dispersal route using Dunn's post hoc test with a Bonferroni correction.

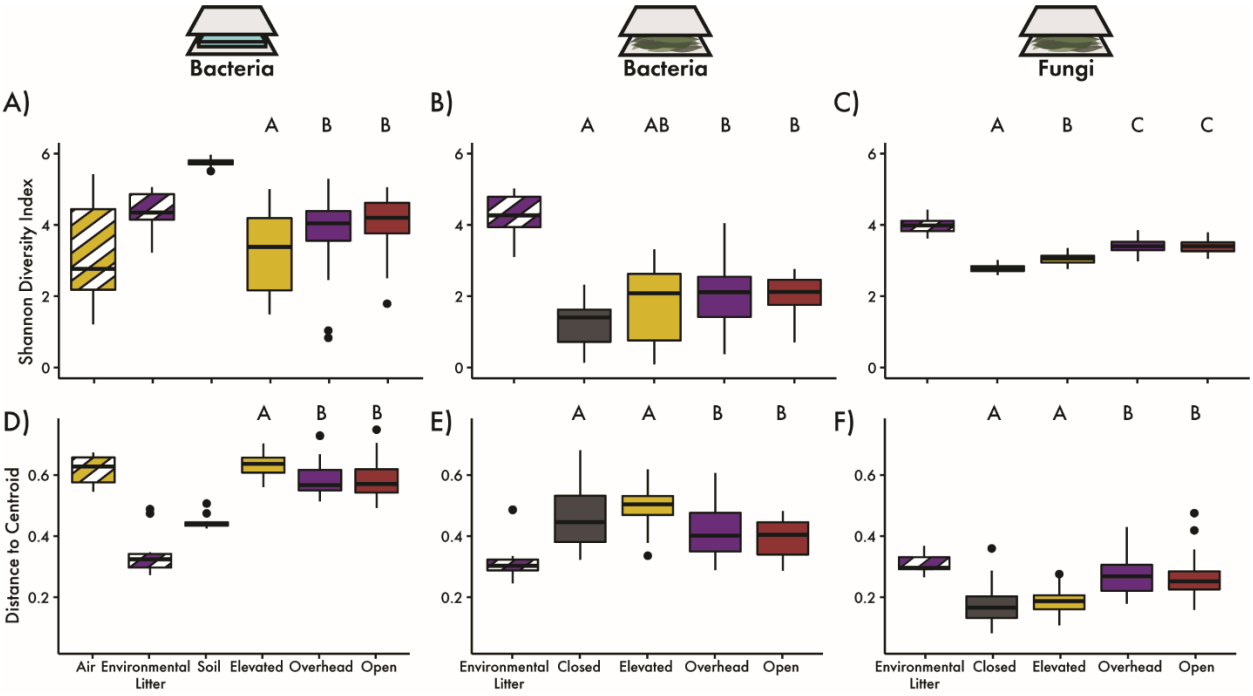

Figure S2.4 Alpha- and beta-diversity by dispersal route. Shannon diversity differs by dispersal routes for **(A)** bacterial communities on the glass slides (ANOVA: $F = 5.22$, $P = 0.0072$, $r^2 = 0.11$), **(B)** bacterial communities in the litterbags ($F = 6.25$, $P = 0.00056$, $r^2 = 0.14$), and **(C)** fungal communities in the litterbags ($F = 62.88$, $P < 0.0001$, $r^2 = 0.67$). Letters for panels A, B, and C indicate pairwise significance for each treatment (Tukey's HSD). Beta-diversity (distance to group centroid) differs by dispersal route for **(D)** bacterial communities on the glass slides (betadisper: $F = 8.2518$, $P = 0.002$, $r^2 = 0.16$), **(E)** bacterial communities in the litterbags ($F = 10.533$, $P = 0.001$, $r^2 = 0.21$), and **(F)** fungal communities in the litterbags ($F = 27.03$, $P = 0.020$, $r^2 = 0.37$). Letters indicate pairwise differences using a *post-hoc* test. Environmental samples are provided for context and are the same across all columns but rarefied independently.

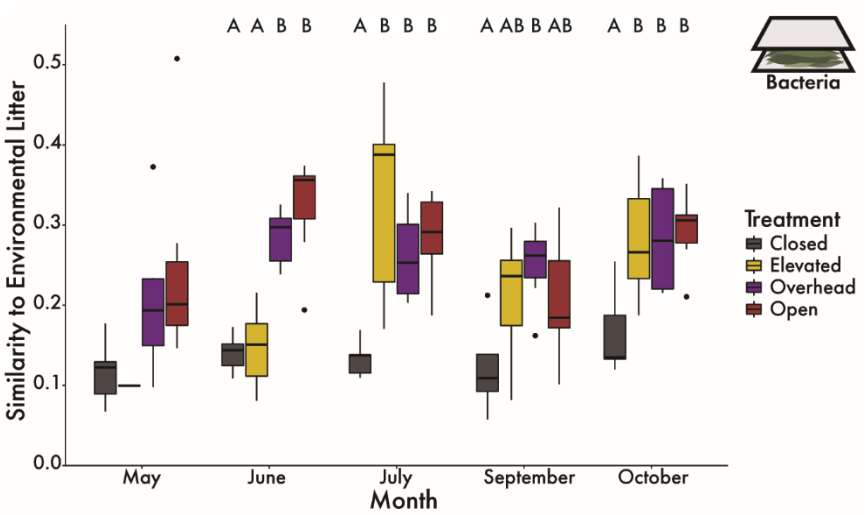

Figure S2.5 Similarity to environmental litter. Similarity to environmental litter was calculated as the similarity (1 – Bray-Curtis) between each sample and the centroid of the environmental litter samples each month. The distance to the environmental litter differs by treatment for all months (ANOVA: $P < 0.05$), except May ($P = 0.16$). Letters indicate pairwise significance within each significant month (Tukey's HSD).

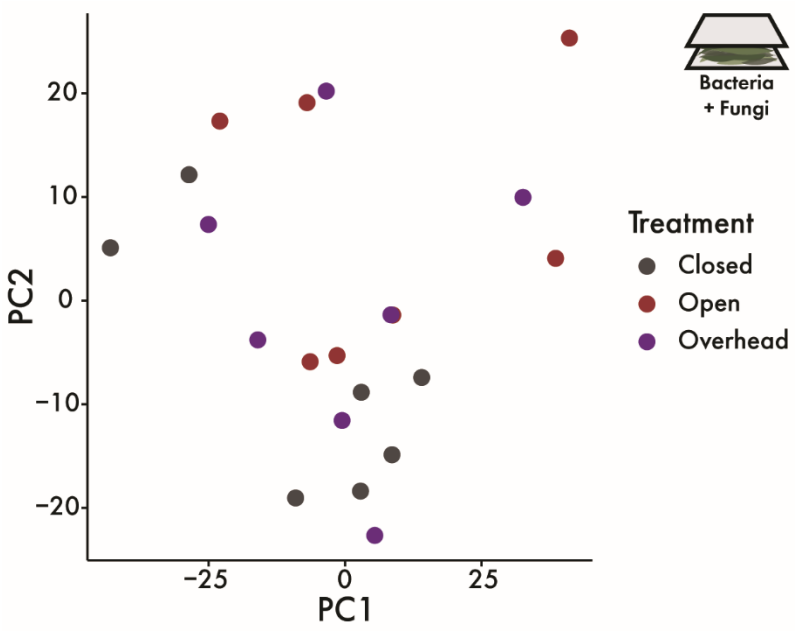

Figure S2.6 Litter chemistry composition by timepoint. Visualization (PCA) of the composition of litterbag litter chemistry exposed to different dispersal routes, including environmental (surrounding) litter and initial litter (phyllosphere) as a comparison.

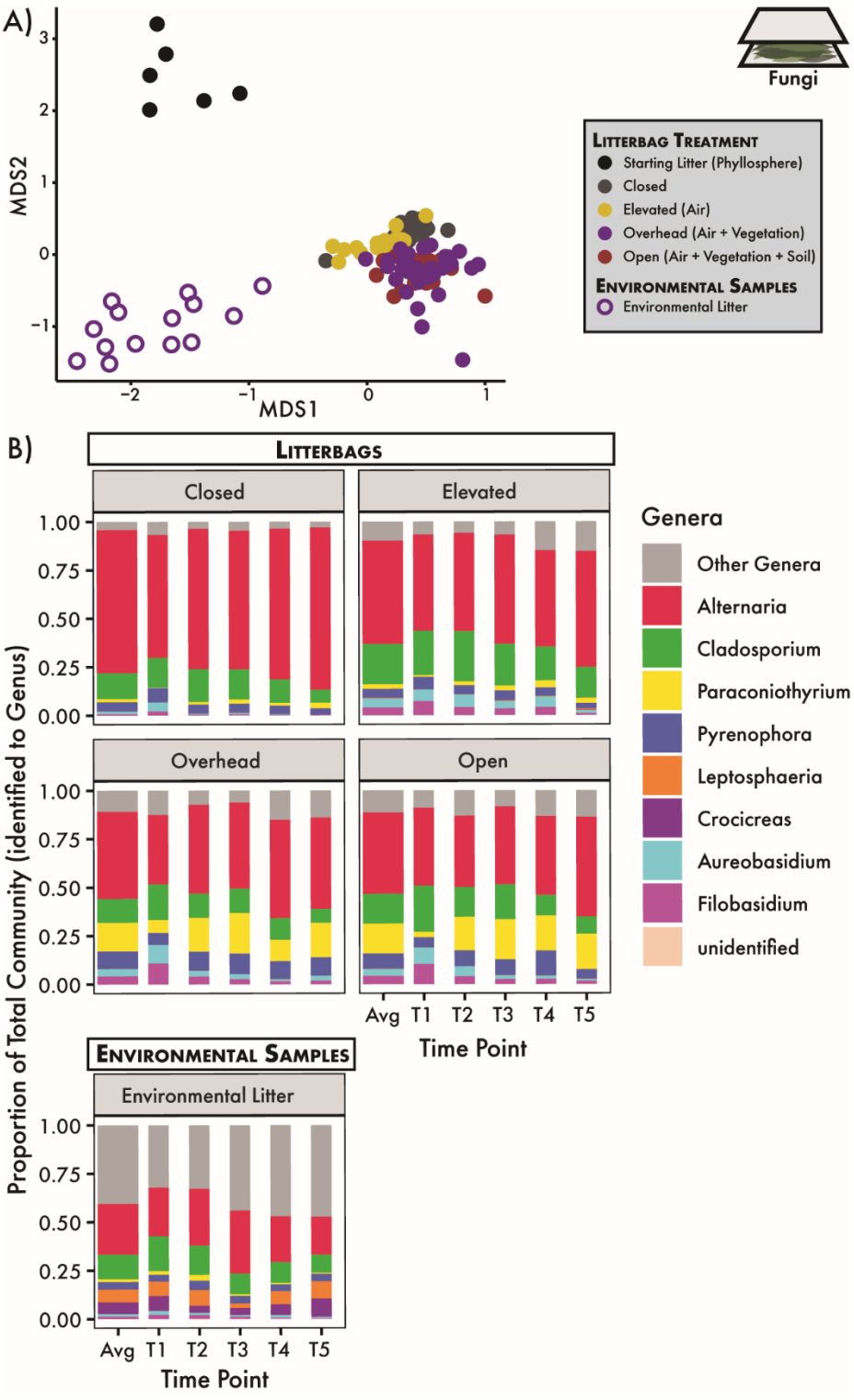

Figure S2.7 Fungal community composition of litterbag communities. (A) Visualization (NMDS) of the composition of litterbag fungal communities exposed to different dispersal routes, including environmental (surrounding) litter as a comparison. **(B)** The most abundant bacterial genera (relative abundance > 4%) in the litterbag communities and environmental litter, averaged (Avg) and by time point.

Appendix 2B: Supplemental Tables

Table S2.1. ANOVA results for the effects of dispersal route, time (sampling timepoint), and dispersal route-by-time interaction on subtracted bacterial abundance on glass slides.

	Df	SumsOfSqs	F	r^2	Pr(>F)	
Route	2	$7.15x10^{9}$	0.63		0.54	
Month	4	$4.37x10^{10}$	1.92		0.12	
Route-by-Month Interaction	8	$2.10x10^{11}$	4.60	0.30	**0.00011**	***
Residuals	85	$4.85x10^{11}$				

Table S2.2. PERMANOVA results for the effects of dispersal route, time (sampling timepoint), and dispersal route-by-time interaction on bacterial community composition on glass slides.

	Df	SumsOfSqs	MeanSqs	Pseudo-F	r^2	P(Perm)	
Route	2	1.45	0.72	2.07	0.043	**0.001**	**
Time	4	2.41	0.60	1.72	0.072	**0.001**	**
Route-by-Time Interaction	8	3.59	0.45	1.29	0.11	**0.001**	**
Residuals	75	26.18	0.35		0.78		
Total	89	33.82			1		

Table S2.3. ANOVA results for the effects of dispersal route, time (sampling timepoint), and dispersal route-by-time interaction on subtracted bacterial abundance in litterbags.

	Df	SumsOfSqs	F	r^2	Pr(>F)	
Route	2	$1.23x10^{20}$	15.32	0.26	**$1.95x10^{-06}$**	***
Time	4	$3.33x10^{19}$	2.07		0.091	
Route-by-Time Interaction	8	$4.05x10^{19}$	1.26		0.27	
Residuals	88	$3.53x10^{20}$				

Table S2.4. PERMANOVA results for the effects of dispersal route, time (sampling timepoint), and dispersal route-by-time interaction on community composition in litterbags for A) bacteria and B) fungi.

A) Bacteria

	Df	SumsOfSqs	MeanSqs	Pseudo-F	r^2	P(Perm)	
Route	3	3.71	1.24	7.49	0.14	**0.001**	**
Time	4	3.17	0.79	4.81	0.12	**0.001**	**
Route-by-Time Interaction	12	3.23	0.27	1.63	0.12	**0.001**	**
Residuals	98	16.16	0.16		0.62		
Total	117	27.61			1		

B) Fungi

	Df	SumsOfSqs	MeanSqs	Pseudo-F	r^2	P(Perm)	
Route	3	3.78	1.26	22.64	0.31	**0.001**	**
Time	4	0.84	0.21	3.75	0.069	**0.001**	**
Route-by-Time Interaction	12	0.98	8.17e-02	1.47	0.081	**0.002**	**
Residuals	117	6.51	5.57e-02		0.54		
Total	136	12.11			1		

Table S2.5. ANOVA results for the effects of dispersal route, time (sampling timepoint), and dispersal route-by-time interaction on subtracted mass loss (decomposition rate) in litterbags for A) all timepoints, B) May, C) June, D) July, E) September, F) October.

A) All timepoints

	Df	SumsOfSqs	F	r^2	Pr(>F)	
Route	1	0.011	2.25		0.14	
Time	4	0.020	1.03		0.40	
Route-by-Time Interaction	4	0.089	4.64	0.25	**0.0026**	**
Residuals	56	0.27				

B) May

	Df	SumsOfSqs	F	r^2	Pr(>F)	
Route	1	0.087	15.73	0.57	**0.0019**	**
Residuals	12	0.066				

C) June

	Df	SumsOfSqs	F	r^2	Pr(>F)
Route	1	0.0071	3.57		0.083
Residuals	12	0.024			

D) July

	Df	SumsOfSqs	F	r^2	Pr(>F)
Route	1	0.00088	0.65		0.44
Residuals	12	0.016			

E) September

	Df	SumsOfSqs	F	r^2	Pr(>F)
Route	1	0.0065	0.58		0.46
Residuals	12	0.13			

F) October

	Df	SumsOfSqs	F	r^2	Pr(>F)
Route	1	0.00057	0.17		0.69
Residuals	8	0.027			

Table S2.6. PERMANOVA results for the effects of dispersal route on leaf litter chemical makeup of the litterbags from June (2nd collection month).

	Df	SumsOfSqs	MeanSqs	Pseudo-F	r^2	P(Perm)
Treatment	2	2435.8	1217.89	1.35	0.13	0.21
Residuals	18	16264.2	903.57		0.87	
Total	20	18700			1	

Appendix 2C: Equilibrium between death and immigration rate

The cells observed on the glass slides are not just the product of immigration but the difference between immigration and death rate. Therefore, to estimate immigration rate, we first needed to estimate death rate on the glass slides. At the start of the experiment, we set out sample bags closed to dispersal that contained a glass slide with a known number of bacteria on it. The bacteria were a community derived from leaf litter to keep the taxonomic composition consistent with the rest of the experiment. We measured the decrease in abundance over time at each time point to model death rate.

Overall, 3.34% of the community dies every day, or 21.2% per week (Fig. S2.1A). To calculate immigration rate, we used the average cell abundance on the Open dispersal samples and the estimated death rate. Because abundance on dispersal samples does not increase or decrease over time, we assume that equilibrium between death and immigration was established before the first time point. At equilibrium, the number of cells immigrating is equal to the death rate times the number of cells observed (mean over the five timepoints). Using that equation, we calculated 7,900 cells/cm^2/day immigrating every day, or 55,000 cells/cm^2/week.

These calculations assume that death rate is constant for every sample, whether samples were placed on the ground or the table surface. We tested whether this was true by measuring weekly death rate on the table and the ground from September – October. Overall, the death rates did not differ between ground and table. The weekly death rate on the table (28.05%) was slightly, but non-significantly, higher than the death rate on the ground (23.92%; ANCOVA, $P = 0.090$; Fig. S2.1B). Therefore, we can conclude that our assumptions are appropriate. However, there is still a chance we are underestimating dispersal from regional rain/air. The table surface is exposed to higher light intensities (Welch's t-test: $P < 0.0001$, $t = 26.10$), which might result in a

higher death rate but would not be picked up by our nylon death rate samples as the nylon blocks UV exposure.

CHAPTER 3

Staying local: Comparing the rate and composition of microbial dispersal into soil between two ecosystems

Abstract

Dispersal is a key process influencing soil microbiomes. The taxonomic composition of microbial dispersal influences community composition and functioning, including decomposition rate. However, where dispersing taxa originate from or how far they disperse remains unknown because microbial dispersal has yet to be studied in a spatially explicit way. Here we measured the rate and composition of microbial dispersal into the soil surface along three 30 m transects that span the boundary of two ecosystems: a grassland and a shrubland. We estimated the sources and dispersal distances by comparing the species composition of immigrating microorganisms with the surrounding plant composition measured at different radii around each sampling location. Simultaneously, we characterized the microbial composition in potential sources of dispersal – air, leaf litter, and soil. The taxonomic composition of immigrating bacteria and fungi were distinct between ecosystems. Bacterial dispersal rate, however, did not differ by ecosystem. The local vegetation was correlated with the taxonomic composition of dispersing microorganisms with the strongest correlation at 1 m from the dispersal location. Vegetation explained 35% of variation in dispersing fungal composition. Bacteria, however, were more strongly associated with soil, suggesting that the dispersing bacteria and fungi emigrated from different sources. Overall, this research provides evidence that microbial dispersal is spatially heterogeneous and is influenced by the immediate surroundings. In particular, these results suggest that microorganisms exhibit smaller dispersal kernels (within 1 meter) than previously assumed. Further research on microbial dispersal kernels is needed to predict how dispersal will

influence microorganisms' ability to respond to environmental shifts such as climate or land-use change.

INTRODUCTION

Microbial communities are primary decomposers of leaf litter that play essential roles in nutrient and carbon cycling (Sayer, 2006). The composition of microbial communities – the fungal and bacterial taxa – strongly influence decomposition processes, including the rate of decomposition and which compounds are broken down (Glassman et al., 2018). Thus, understanding the ecological drivers of leaf litter microbiomes provides deeper understanding of ecosystem health. Dispersal, in particular, is a key ecological process, but its impacts on soil microbiomes is not well understood.

Recent studies have demonstrated that manipulating dispersal into leaf litter influences the microbiome composition and litter decomposition (Albright & Martiny, 2018; Walters et al., 2021), thus impacting ecosystem processes. However, we have yet to answer even basic questions about microbial dispersal. For example, at what rate do microorganisms disperse into soil? Dispersal rates can inform how strongly dispersal acts on a resident community in relation to other ecological processes, such as selection. Further, we have only a preliminary understanding of what sources contribute to microbial dispersal. Past research suggests that soil or sediment (Rime et al., 2016; Yang & van Elsas, 2018), air, and vegetation (Lindow & Andersen, 1996; Peay, Garbelotto, & Bruns, 2010; Walters et al., 2021) are important sources of dispersal into soil. However, the degree of spatial heterogeneity of these sources in unknown. In other words, do microorganisms disperse far enough that dispersal is homogenous across a landscape? Or do cells disperse short distances, yielding high heterogeneity? Because the microbiomes of potential dispersal sources vary over small spatial scales, dispersal is likely heterogeneous, although at what scale is unknown.

Different microorganisms are also likely to exhibit different dispersal patterns,

particularly bacteria and fungi. Bacteria and fungi exhibit different life history strategies that may influence how they disperse. In particular, fungi form strong symbiotic relationships with plants (Selosse et al., 2004) and thus may show greater spatial heterogeneity. Further, past research suggests that fungi may be more dispersal limited than bacteria (Chen et al., 2020; Li et al., 2020), potentially because fungi have larger cell size and patchier species distributions (Schmidt et al., 2014). For example, at our research site, fungal communities in leaf litter show stronger heterogeneity across a landscape than bacterial communities (Finks et al., 2021). Because traits like cell size and species distributions are known to impact a species dispersal ability, we expect that bacteria and fungi may exhibit different dispersal abilities.

In this study, we tested three hypotheses by characterizing the abundance and taxonomic composition of the bacteria and fungi immigrating onto sterile glass slides (dispersal "traps") in two ecosystems (a grassland and a shrubland). First, we hypothesized that both the taxonomic composition and the rate of microorganisms dispersing onto the surface soil varies between ecosystems. Microbiome composition and abundance of important dispersal sources, such as leaf litter and soil, differ between these ecosystems (Finks et al., 2021; Khalili, Weihe, Kimball, Schmidt, & Martiny, 2019; Barbour et al., *in prep*), and these differences in source communities may impact dispersal composition and rates. Second, we hypothesized that vegetation differences between the ecosystems drive compositional differences in the microorganisms dispersing onto the surface soil. Past research suggests that local vegetation is the main dispersal source that may influence heterogeneity on a small spatial scale (Walters et al., 2021). Finally, we hypothesized that dispersing fungi show stronger associations with the surrounding plant community than dispersing bacteria because fungi form specific symbiotic relationships with plants (Selosse et al., 2004).

METHODS

Field site. The experiment was conducted at the Loma Ridge Global Change Experiment in California, USA (33°44′ N, 117°42′ W, 365 m elevation) from June 18th, 2020 – August 11th, 2020. The site encompasses a Mediterranean grassland and shrubland, which experiences dry warm summers and cool wet winters, with 325 mm mean annual precipitation and 17°C mean temperature. The grassland is dominated by non-native grasses *Bromus madritensis* and *Avena* sp., non-native forbs *Hirschfeldia incana* and *Erodium* sp., and the native forb *Deinandra fasciculata*. The shrubland is dominated by native shrubs *Salvia mellifera*, *Malosma laurina* and *Artemisia californica*, native herb *Acmispon glaber*, and invasive grass *Bromus madritensis* (Finks et al., 2021; Kimball et al., 2014).

Experimental Approach. To identify the taxa dispersing across the landscape, we deployed dispersal "traps" at eight locations along each of three 30 m transects that spanned the boundary between the grassland and shrubland ecosystems (Fig. 3.1). These dispersal samples were placed on top of the soil and thus captured microorganisms immigrating onto the soil surface. To estimate bacterial dispersal rate, we measured the number of bacterial cells that accumulated on the dispersal samples over time; however, the cell abundance on these samples is affected by both immigration and death rates. Therefore, to estimate immigration rate, we simultaneously deployed samples that allowed us to calculate in-field death rates. These death rate samples contained a known number of bacterial cells and were closed to immigration, allowing us to measure the decrease in abundance over time to model death rate. Death rate was measured independently in the grassland and shrubland ecosystems (Fig. 3.1). In addition, along each

transect, we deployed negative control samples that excluded immigration and thus provided a measure of potential contamination.

Plant composition and geographic sampling locations. We characterized the plant composition and geographic sampling locations by taking aerial drone (Holy Stone, HS100) photos of the study area. Photos were stitched together using Hugin (2019.2.0) and aligned to a Google Earth satellite image of the field site to ensure accurate aspect ratio. Sampling locations were identified on the composite aerial photo by the presence of red flagging, and polygons were drawn around individual plants and areas of exposed soil. Plant polygons were identified to species in the field, including the species composition of the grassland which was applied to all areas of grass coverage.

Dispersal slides. The dispersal samples were made of 2 mm window screen (Phifer, Model # 3002212) and the negative controls were made of 0.22 µm nylon (Tisch, SPEC17970). Glass microscope slides (2.5 cm x 7.5 cm) were sterilized in 70% ethanol, dried, sealed into bags (5 cm x 10 cm), and autoclaved. The bags were set out in the field on June 18, 2020 and were secured under chicken wire and staked into the ground to prevent disturbance. At four timepoints (June 30th, July 14th, August 4th, August 11th), we collected one bag from each treatment from each location along the transects. All samples were processed for amplicon sequencing and bacterial abundance.

Death rate slides. To measure bacterial death rate, slides containing a known number of bacterial cells were placed into the field and sampled at the same time as the dispersal samples. Bacterial cells were extracted separately from both grass and shrub litter from the field site by steeping the litter in 1 L of 0.9% saline solution overnight. We used bacterial communities derived from leaf litter to keep the taxonomic composition consistent with the rest of the experiment. The litter

was then filtered through cheese cloth, and the filtrate was aliquoted in 50 mL volumes, further concentrated and washed by pelleting the cells and resuspending into 0.5 mL sterile 0.9% saline solution. Aliquots of 100 µL of either the grass- or shrub-derived cells were then spread onto ethanol-sterilized glass microscope slides and allowed to dry before being sealed into autoclaved nylon (0.22 µm) bags that were closed to dispersal. The slides were stored at room temperature overnight until being deployed in the field.

Dispersal sources. To characterize the potential sources of dispersal, we collected soil, leaf litter, and air samples from the field site at the second timepoint (July 14th). To collect soil samples, we used a sterile scoop to collect approximately 2 g of the top layer of soil at four locations across each transect (N = 12). To collect environmental litter samples, we used sterile clippers to collect approximately 3 g of leaf litter at four locations along each transect (N = 12). To collect the air samples, we deployed glass slide samples open to dispersal (enclosed in 2 mm window screen bags; N = 3) and Mini FTA cards (Whatman, WB120356; N = 3) elevated at 3 m on June 18th and then collected on July 14th. All samples were processed for amplicon sequencing.

Bacterial abundance. At the time of sample collection, glass slides were transferred to a sterile plastic bag with 2 mL of 0.9% sterile saline. After aliquots were removed (2 x 250 µL) for amplicon sequencing, 167 µL of 10% phosphate-buffered GTA was added to bring the solution to a final concentration of 1% phosphate-buffered GTA. Samples were filtered through a 4-µm-pore-size vacuum filter to remove large particulates and stored in the dark at 4°C overnight before being measured on the flow cytometer. To process samples on the flow cytometer, 2 µL of SYBR green (200x, Invitrogen Life Science Technologies, S756) was added to 400 µL of each sample, and samples were incubated in the dark at room temperature for 10 minutes. Samples were run for 30 seconds at 40 µL/min, using a SYBR-Green-H threshold value of 1,500

and SSC-H threshold value of 1,000. Gating parameters were used to count particles in the size of typical bacterial cells, optimized by Khalili *et al.* (2019).

Amplicon Sequencing. Aliquots of 0.05 g of leaf litter, 0.1 g of soil, 250 µL from the glass slide saline solution, and half of the FTA card rounds were frozen at -70°C and stored until extraction. DNA was extracted following the ZymoBIOMICS Microprep DNA Extraction Kit with tubes for glass slide samples and FTA cards, and plates for litter and soil samples. The kit protocols were followed with the following modifications. For glass slide and FTA card samples only, bead beating was reduced to 3 minutes, instead of 5 minutes, to avoid shearing the DNA in these low biomass samples. For all tube extraction samples, the Zymo-Spin™ III-F filter was not used and water for elution was heated to 60°C. For all plate extraction samples, maximum centrifuge speed was 2808 x g, instead of 3500 x g, and centrifuge time was increased at that speed from 3 min to 4 min and from 5 min to 7 min.

To characterize the bacterial community, we amplified the V4 – V5 region of the 16S ribosomal DNA gene using the 515F (GTGYCAGCMGCCGCGGTAA) – 926R (CCGTCAATTCCTTTRAGTTT) primers, described in (Caporaso et al., 2012; Lane et al., 1985). For leaf litter and soil samples, we used 1 µL of template DNA for the PCR. For glass slide and FTA card samples, we used 5 µL of template DNA. For all samples, the remaining PCR reaction contained 12.5 µL of AccustartII PCR ToughMix (Quanta BioSciences, Inc), 0.5 µL forward primer, 0.5 µL reverse primer, and H_2O to reach a final volume of 25 µL. Following an initial denaturation step at 94°C for 3 min, the PCR was cycled 30 times at 94°C for 45 s, 55°C for 30 s, and 72°C for 60 s, with a final extension at 72°C for 10 min. Because 29 glass slide and FTA card samples did not amplify with the above protocol, the PCR was repeated for those samples with 8 µL of template DNA (final volume of 25 µL) and 34 cycles.

To characterize the fungal community, we amplified the ITS2 region of the Internal Transcribed Spacer (ITS) using the ITS9F (GAACGCAGCRAAIIGYGA) – ITS4 (TCCTCCGCTTATTGATATGC) primer combination (Looby *et al.* 2016). For leaf litter and soil samples, we used 1 µL of template DNA for the PCR. For air and glass slide samples, we used 5 µL of template DNA. For all samples, the remaining PCR reaction contained 12.5 µL of AccustartII PCR ToughMix (Quanta BioSciences, Inc), 0.75 µL forward primer, 0.75 µL reverse primer, and H_2O to reach a final volume of 25 µL. After an initial denaturing step at 94°C for 3 min, the PCR was cycled 35 times at 95°C for 45 s, 50°C for 60 s, and 72°C for 90 s, with a final extension at 72°C for 10 min.

To prepare libraries for sequencing, PCR products were pooled at different volumes based on amplification brightness on gel pictures: high (1 µL), medium high (2 µL), medium (3 µL), medium weak (4 µL), weak (5 µL), and none (10 µL). All samples were pooled together for one sequencing run – both 16S and ITS amplicons. After pooling, the library was purified using SpeedBeads magnetic carboxylate-modified particles (GE Healthcare UK Limited) and sequenced in a paired-end Illumina MiSeq run (2 x 250 bp) at the Genomics High Throughput Facility, UC Irvine, Irvine, CA, USA.

Bioinformatic processing and analysis. Sequence data were processed in QIIME2 (Bolyen et al., 2019), version 2020.8. The forward reads were trimmed to 6 – 224 base pairs of the 16S segment and 5 – 235 base pairs of the ITS segment. The reverse reads were discarded. We used DADA2 (Callahan et al., 2016) to define exact sequence variants and assigned taxonomy using classify-sklearn with, for bacteria, the reference Greengenes database (13_8) at a 99% OTU level (McDonald et al., 2012) and, for fungi, the UNITE (v. 7.2) database (Nilsson et al. 2019).

Statistical analyses. We accounted for differences in sequencing depth by rarefying samples to 971 sequences (bacterial communities) or 2,400 sequences (fungal communities) with 1,000 resamplings. For each resampling, we calculated a Bray-Curtis dissimilarity matrix, taking the median similarity values of all 1,000 matrices using the 'vegan' package in R (Oksanen et al., 2019; R Core Team, 2020). Using this median matrix, we performed a PERMANOVA and *post-hoc* test in R using the 'vegan' and 'pairwiseAdonis' packages (Martinez Arbizu, 2017; Oksanen et al., 2019) to test for differences in community composition by ecosystem and timepoint. Microbial composition on the glass slides was further linked to the potential dispersal sources using SourceTracker, a Bayesian approach that estimates the proportions of a community that came from possible sources (Knights et al., 2011). To analyze alpha-diversity, we used the rounded rarefied datasets to calculate richness and an ANOVA to test for differences among treatments. Beta-diversity (Bray-Curtis metric) was analyzed with a PERMDISP test using the betadisper() function. To identify taxa that were driving changes over time, we used an indicator species analysis in the 'indicspecies' package in R (De Caceres & Legendre, 2009).

Immigration rates were calculated using two different equations that describe the bacterial abundance on the glass slides as a function of immigration and death rates. Both equations used the measured in-field death rates as a parameter. Death rate was calculated as the slope of a linear model fitted to the decrease in the log bacterial abundance on the death slides over time. Death rate was measured separately in the grassland and shrubland ecosystems. To test whether death rate differed by ecosystem, we used an ANCOVA to test the influence of time and a time-by-ecosystem interaction on the bacterial abundance on the death slides, where a significant time-by-ecosystem interaction would indicate different death rates between ecosystems. Further, to test whether microbial taxa exhibited different death rates, we performed

a PERMANOVA testing the influence of time, ecosystem, and a time-by-ecosystem interaction on the bacterial and fungal communities on the death slides, where a significant influence of time would indicate that microbial taxa die at different rates.

The first equation used to calculate immigration rate assumed a dynamic relationship between immigration and death rates (not at equilibrium). That equation modeled the abundance on the glass slides (n) as $n(t) = \frac{i}{d} \times (1 - e^{-d \times t})$, where d is the death rate measured for each ecosystem, t is time, and i is the immigration rate estimated using the nls() function in the 'stats' package in R (R Core Team, 2020). The second approach to calculate immigration rate assumed that immigration and death rates were at equilibrium such that the abundance on the glass slides was constant over time. This calculation only used the abundance at the third timepoint for which abundance was highest. At equilibrium, the number of cells immigrating is equal to the death rate times the number of cells observed. Because the abundance on the glass slides did, in fact, change over time (ANOVA: P = 0.02; indicating that immigration and death rate were not at equilibrium), we expect the first equation to provide more a more accurate estimate of immigration rate; however, we report both to provide a sense of uncertainty around these estimates.

To examine the scale to which local vegetation influences dispersing microorganisms, we tested the correlation between plant composition and geographic distance on the bacterial and fungal communities on the glass slides using a partial Mantel test. Plant composition was measured by calculating the area of each plant species within circles of 1-m radius around each sampling location. Geographic distance was measured as the pairwise distance among locations. This analysis was repeated at a series of increasing radii of the circles, from 0.1 m to 4 m. By graphing the strength of these correlations against the radii of the circles for which plant

composition was measured, we identified the distance for which plant composition had the strongest influence on immigrating taxa. In other words, we identified the average dispersal distance (dispersal kernel) of bacteria and fungi immigrating from vegetation.

RESULTS

Hypothesis 1: The composition, but not rate, of immigrating taxa differed by ecosystem.

Composition of immigrating microorganisms.

In support of our first hypothesis, the bacterial and fungal taxa dispersing onto the surface soils (captured by the glass slide "traps") were distinct between the shrubland and grassland ecosystems (Table S3.3; Fig. 3.2). Overall, immigrating fungi had higher richness in the shrubland than the grassland (ANOVA: $P < 0.0001$, $F = 27.65$), but immigrating bacteria showed no difference in richness between ecosystems ($P = 0.28$). Similarly, beta-diversity was higher in the shrubland than the grassland for immigrating fungi (PERMDISP: $P < 0.0001$; Fig. 3.2 shows greater spread in fungal community data points in the shrubland). Beta-diversity in bacterial communities, however, did not differ by ecosystem ($P = 0.45$).

Further, the composition of dispersing bacteria and fungi communities changed over time, although more strongly for fungi than bacteria (Table S3.3; Fig. S3.1). Specifically, the first timepoint differed from each of the last three timepoints (post-hoc comparisons: $P < 0.05$), which did not differ from one another ($P > 0.05$). However, this statistical difference may be driven by changes in the variance in bacterial community composition among the glass slides (the spread of the points in Fig. S3.1), where the variance decreased over time (PERMDISP: $P = 0.033$). The change in variance may contribute to the effect of time on composition (Marti J.

Anderson & Walsh, 2013). In contrast, variance in fungal community composition did not change significantly over time ($P = 0.087$).

Further, a subset of the fungal taxa appears to be driving this change over time. For both the grassland and shrubland ecosystems, the last three timepoints are characterized by an increase in *Alternaria* and *Cladosporium*, from 37% or 12%, respectively, in the first timepoint to 55% or 17%, respectively, in later timepoints (Fig. 3.3A). In the grassland, we also saw a decrease in the relative abundance of *Teichospora* from 0.5% to 0.02%. In the shrubland, the first timepoint showed a high proportion of *Caryophylloseptoria* (2.6%), which decreased to 0.03% in later timepoints, and there were also decreases in relative abundances of *Aureobasidium*, *Truncatella*, and *Seimatosporium* (8, 3, and 8% to 2, 1, and 2%, respectively). An indicator species analysis further supported these trends (Fig. 3.3B). In particular, multiple fungal OTUs in the *Alternaria* and *Cladosporium* genera were also identified as indicator species of the later timepoints in both ecosystems.

Dispersal Rate into Leaf Litter.

In contrast to our first hypothesis, bacterial immigration rates did not differ by ecosystem. Bacteria immigrated onto the top layer of soil at a rate of 1,170 cells/cm^2/day ($P < 0.0001$) in the grassland and 940 cells/cm^2/day ($P < 0.0001$) in the shrubland (Fig. 3.4). These rates represent 0.05% and 0.04% of the average bacterial abundance in the grass and shrub litterbags, respectively. However, the grassland immigration rate was not statistically higher than the shrubland rate (difference between the two rates: 230 ± 375 cells/cm^2/day). This immigration rate calculation was performed assuming a dynamic number of cells on the glass slides, where immigration and death rates were not required to be at equilibrium. To complement this

calculation, immigration rates were also calculated assuming an equilibrium between immigration and death rates. This second method yielded similar rates at 1,100 cells/cm^2/day in the grassland and 600 cells/cm^2/day in the shrubland.

To estimate bacterial death rates for the above calculations, we measured death rate on the glass slides in both ecosystems. Bacterial death rates were marginally lower in the shrubland than the grassland (Table S3.1; Fig. 3.4). On average, 3.16% of the community died per day in the grassland but only 2.40% in the shrubland. Bacterial abundance on the death rate slides was low, and thus many samples did not sequence successfully. Still, bacterial taxa on the glass slides appeared to exhibit different death rates, as indicated by a significant change in the community composition on the slides between the beginning of the experiment, the first timepoint and the three later timepoints grouped together (PERMANOVA: $P < 0.0001$; Table S3.2; Fig. S3.2). Similarly, fungal taxa also exhibited different death rates ($P = 0.02$; Table S3.2; Fig. S3.2).

<u>Hypothesis 2</u>: Plant composition was positively correlated with dispersing microorganisms.

Plant composition and immigrating microorganisms.

In support of our second hypothesis, plant composition was correlated with the composition of dispersing microorganisms, although more strongly with fungi than bacteria. Dispersing bacterial communities were positively correlated with the plant composition within a 1-m radius of each sampling location (Pearson's $r = 0.13$, $P = 0.027$) but not geographic distance ($P = 0.15$). Fungal communities, on the other hand, were positively correlated with both plant composition ($r = 0.61$, $P = 0.0020$) and geographic distance ($r = 0.20$, $P = 0.0030$). When the variation from plant composition was removed, geographic distance was not a significant

predictor of dispersing fungi (partial Mantel test: $P = 0.13$), whereas plant composition remained positively correlated with fungal community composition when variation from geographic distance was removed ($r = 0.59$, $P = 0.0020$).

To estimate the distance at which bacteria and fungi disperse from the vegetation, we calculated a series of partial Mantel tests between microbial communities (on the glass slides) and plant composition calculated at radii between 0.1 m and 4 m around each sampling location (with variation from geographic distance removed for fungi). The strength of the correlation between the microbial community and the surrounding plant composition peaked at 1 m for both bacteria and fungi (Fig. 3.5), although the strength of the correlations with plant composition were generally much higher for fungal composition at all radii. Note, however, that the measures of plant composition for the sampling locations closest to the ecosystem boundary were not independent because circles of radii larger than 1 m overlapped substantially.

Sources of dispersal differed for bacteria and fungi.

In partial support of our second hypothesis, leaf litter was a significant source of dispersing fungi but not bacteria. We characterized the microbial communities in the air, soil, and leaf litter at the field site and used SourceTracker to identify contributions from each potential source. Soil significantly contributed to the immigrating bacterial communities (those on the glass slides) in both the grassland and shrubland (Fig. 3.6). Air and grass litter were the most likely sources of fungal dispersal onto the glass slides in the grassland while all sources – air, soil, grass litter, and shrub litter – seemed to contribute equally in the shrubland.

<u>Hypothesis 3:</u> Dispersing fungi showed stronger associations with the surrounding plant community than bacteria.

Bacteria and fungi comparison.

In support of our third hypothesis, plant composition explained 35% of the variation in fungal composition on the glass slides but only 2% in bacterial composition (Fig. 3.5). Based on the SourceTracker analysis, bacteria on the glass slides may be dispersing primarily from soil (Fig. 3.6).

DISCUSSION

While studies show that microbial dispersal influences community assembly, we do not know at what rates bacteria disperse nor what sources contributed to dispersal. Here, we compared the microbial dispersal community composition and rates between two ecosystems and characterized the influence of the plant composition on microbial dispersal. In support of our first hypothesis, the taxa dispersing into the surface soil differed between the grassland and shrubland ecosystems for both bacteria and fungi. We expect that this difference may be driven by distinct sources in the two ecosystems. In particular, immigrating bacteria were primarily coming from soil and fungi were dispersing from leaf litter, soil, and air. Leaf litter and soil in the shrubland and grassland harbor distinct microbiomes (Finks et al., 2021; Barbour et al., *in prep*), which may contribute to the distinct taxa dispersing in each ecosystem. Air currents lift cells from soil and leaf surfaces (Bowers, McLetchie, et al., 2011; Bowers, Sullivan, et al., 2011; Tignat-Perrier et al., 2019), and the same mechanism may be at work here to transfer cells from soil and leaves to the glass slides. However, the importance of sources likely depends on

conditions (e.g., weather, amount of bare soil, litter, etc.). While 34% of immigrating bacteria in this experiment were attributed to bulk soil, a previous experiment at the same study area found that soil was not an important source of immigrating bacteria (Walters et al., 2021). The transects in this study covered significant areas of exposed soil (9% of the area in the grassland and 19% in the shrubland), which may explain why soil is a strong contributor to immigrating bacteria.

In contrast to our first hypothesis, however, the rate of bacterial dispersal did not differ by ecosystem. Both ecosystems showed equally low dispersal rates, relative to the average bacterial abundance in the leaf litter microbiome. At 1,000 bacterial cells/cm^2/day, bacterial dispersal represents around 0.05% of the grass litter microbiome community per day and 0.04% of the shrub litter community. This rate is about an order of magnitude lower than previous bacterial deposition rates (Jones, Newton, & McMahon, 2008; Walters et al., 2021). We expect that the dispersal rates are similar between these ecosystems because, being in close proximity, these two ecosystems experience similar abiotic conditions, particularly wind direction and speed, precipitation, temperature, and landscape topography which influence microbial dispersal (Cevallos-Cevallos et al., 2012; Galès et al., 2014; Julianne Lindemann & Upper, 1985; Tignat-Perrier et al., 2019).

To determine whether differences in vegetation were driving differences in dispersal communities between ecosystems, we correlated the dispersal communities to the surrounding plant communities while controlling for geographic distance. In support of our second hypothesis, plant composition corresponded with the dispersal community for both fungi and bacteria, although more strongly for fungi than bacteria. Leaf surfaces are an important source of airborne bacteria, whose dispersal is facilitated by wind (J. Lindemann et al., 1982; Julianne Lindemann & Upper, 1985; Šantl-Temkiv et al., 2018). This analysis, however, does not exclude

the possibility that the dispersing cells may be instead coming from leaf litter, which has microbiomes that are correlated with the plant species composing the litter (Urbanová et al., 2015; Veen et al., 2019). Further, it is likely that other sources are also driving the distinct dispersal communities for each ecosystem. This is especially true for bacteria for which only 2% of the variation can be explained by the plant community. The bulk soil and litter layer, important dispersal sources in this system, harbor distinct microbiomes by ecosystem at this field site (Barbour et al., *in prep*). Thus, these sources are also likely contributing to the distinct dispersal communities between the shrubland and grassland ecosystems.

Because plant communities were a key source of immigrating taxa, we estimated the distance at which vegetation contributes most to dispersal communities. We found that the strongest correlation between plant composition and microbial dispersal community occurred around 1 meter, indicating a small average dispersal distance for both bacteria and fungi. Thus, while some microbial taxa are able to disperse over large distances, even between continents (Smith et al., 2013; Wilkinson et al., 2012), many dispersing microorganisms in this experiment were likely highly localized, moving directly from the surrounding vegetation to the surface soil. This study adds to the growing body of literature that microorganisms are often dispersal-limited within a kilometer (Adams et al., 2013; Galante et al., 2017; Norros et al., 2012; Peay et al., 2010, 2012) but further highlights that dispersal limitation for both fungi and bacteria can occur on a much smaller scale – within a single meter. Previous research suggests that fungi are more dispersal limited than bacteria (Chen et al., 2020), potentially because of their larger cell size (Schmidt et al., 2014); however, this study found both bacteria and fungi had similar dispersal distances from vegetation. For such a short dispersal distance, habitat fragmentation, even in small patches, is likely to have a significant impact on microbial dispersal.

Our study also provided evidence that fungi and bacteria disperse in different ways. In support of our third hypothesis, immigrating fungi were more strongly associated with the surrounding plant community than bacteria. Specifically, 35% of the variation in immigrating fungal taxa could be explained by the surrounding plant composition within 1 m, compared to only 2% of the bacterial variation. In agreement, fungi show higher beta-diversity in the shrubland (heterogeneous plant composition) than the grassland (homogeneous plant community). Bacteria, however, showed equal beta-diversity between the two ecosystems. Past research has shown that fungi form more specific symbiotic relationships with plants than bacteria (Schmidt et al., 2014), likely driving this pattern. Soil, on the other hand is a more likely driver of bacterial dispersal in this system. Dispersal from soil onto leaf litter may occur through direct contact or through facilitation by wind picking up individual cells or cells attached to soil particles (Yang & van Elsas, 2018).

CONCLUSIONS

In this study, we demonstrated that microbial dispersal shows spatial heterogeneity on a small scale. Bacteria and fungi dispersing in two different ecosystems are distinct and vegetation may be driving that pattern, at least for fungi. However, we only examined two ecosystems in a relatively small geographic area. Thus, the question remains whether plant composition holds the same explanatory power in other ecosystems as well. This study also provided estimates of bacterial immigration rates. We expect that the dispersal rates did not differ by ecosystem because these ecosystems experienced similar abiotic conditions (e.g., precipitation, wind speed). However, research on controls on microbial dispersal are limited, and future research could continue to characterize influences on bacterial immigration rates. Further, more research is

needed to estimate dispersal distances of microorganisms. While we showed that fungi and bacteria disperse only 1 meter from the local vegetation, dispersal distances from different sources, during different seasons, and in different ecosystems may vary. All in all, this research suggests that local vegetation is an important dispersal source and that microorganisms exhibited smaller dispersal kernels than previously assumed.

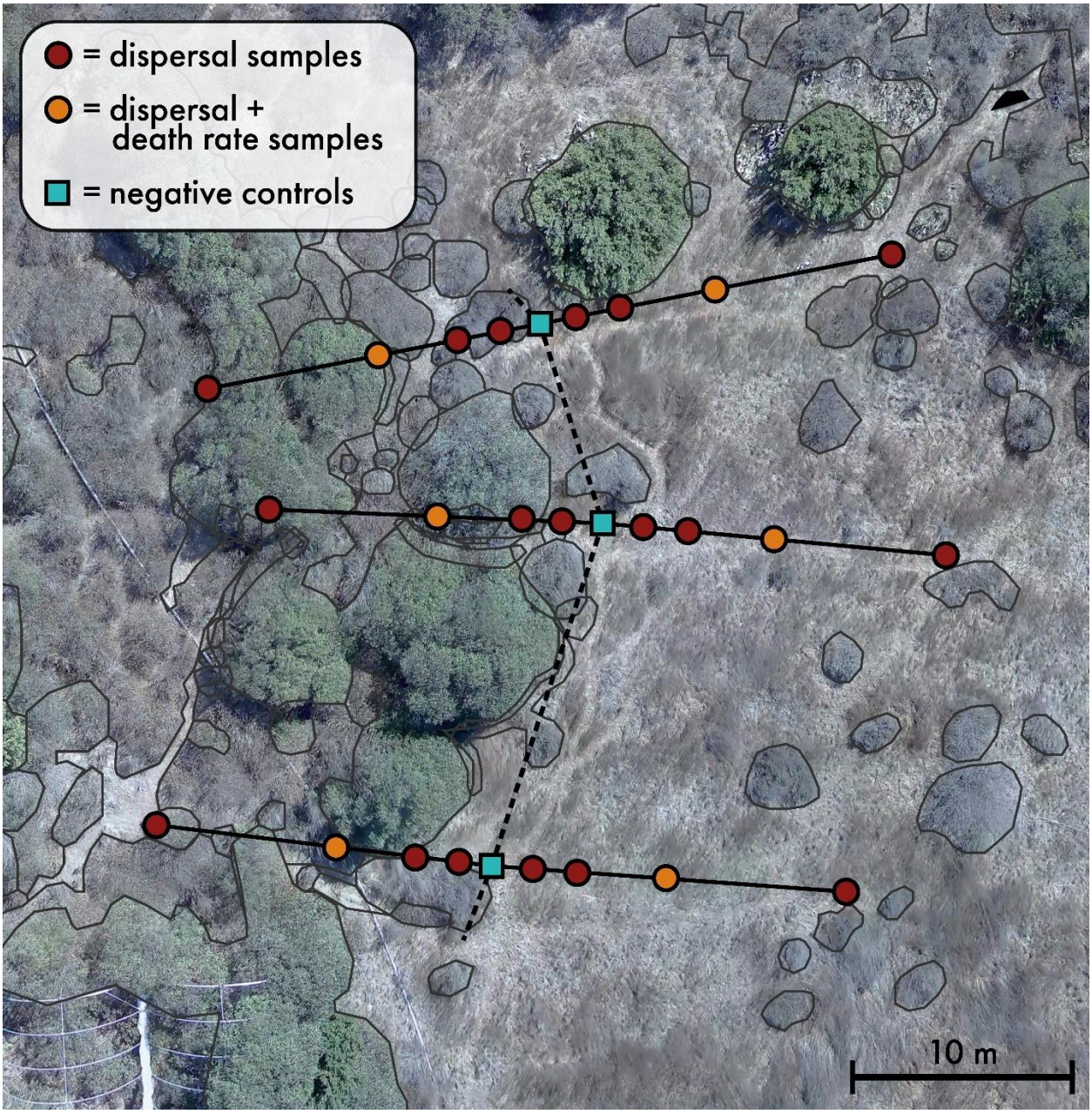

Figure 3.1 Experimental design. An aerial photo of the field site showing the three transects with eight locations per transect. The dashed line represents the boundary between the shrubland and grassland. Dispersal samples (glass slides collecting immigrating microorganisms) were placed at each location along the transects (locations marked in red or orange). Death rate samples were deployed at two locations per transect (locations marked in orange). Negative controls (glass slide samples closed to dispersal) were placed at the ecosystem boundary along each transect (marked in blue squares). Polygons around each plant were used to calculated plant composition around each location.

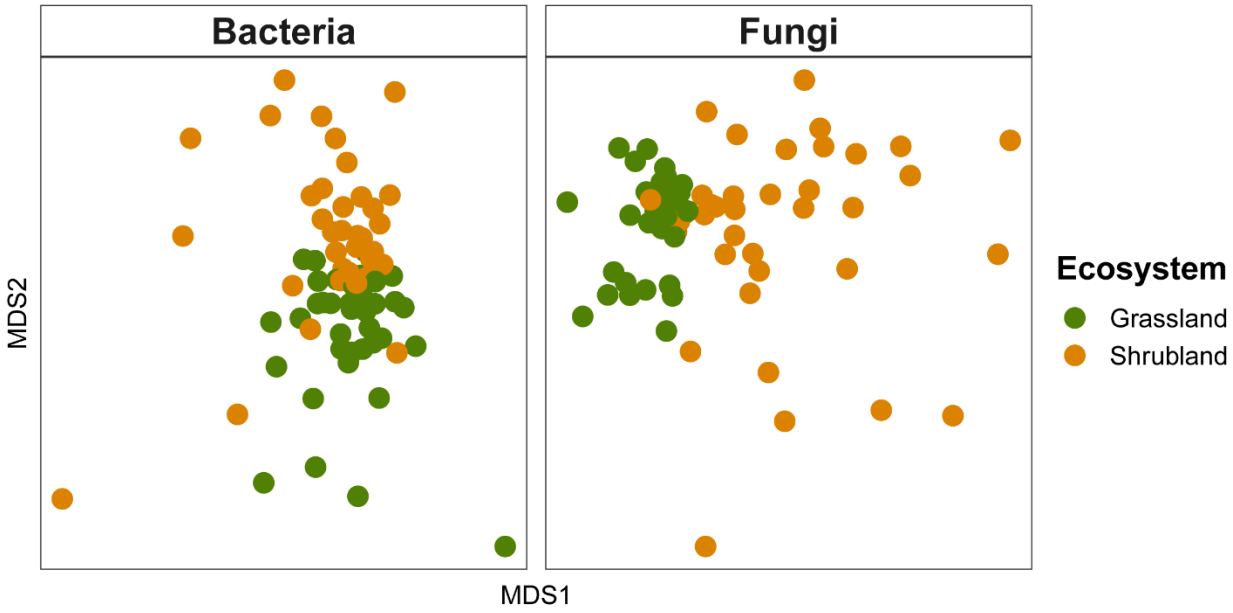

Figure 3.2 Dispersing microbial taxa by ecosystem. NMDS of the bacterial and fungal communities on the dispersal slides. Community composition differed by ecosystem for both bacteria and fungi (PERMANOVA: $P < 0.0001$; Table S3.3).

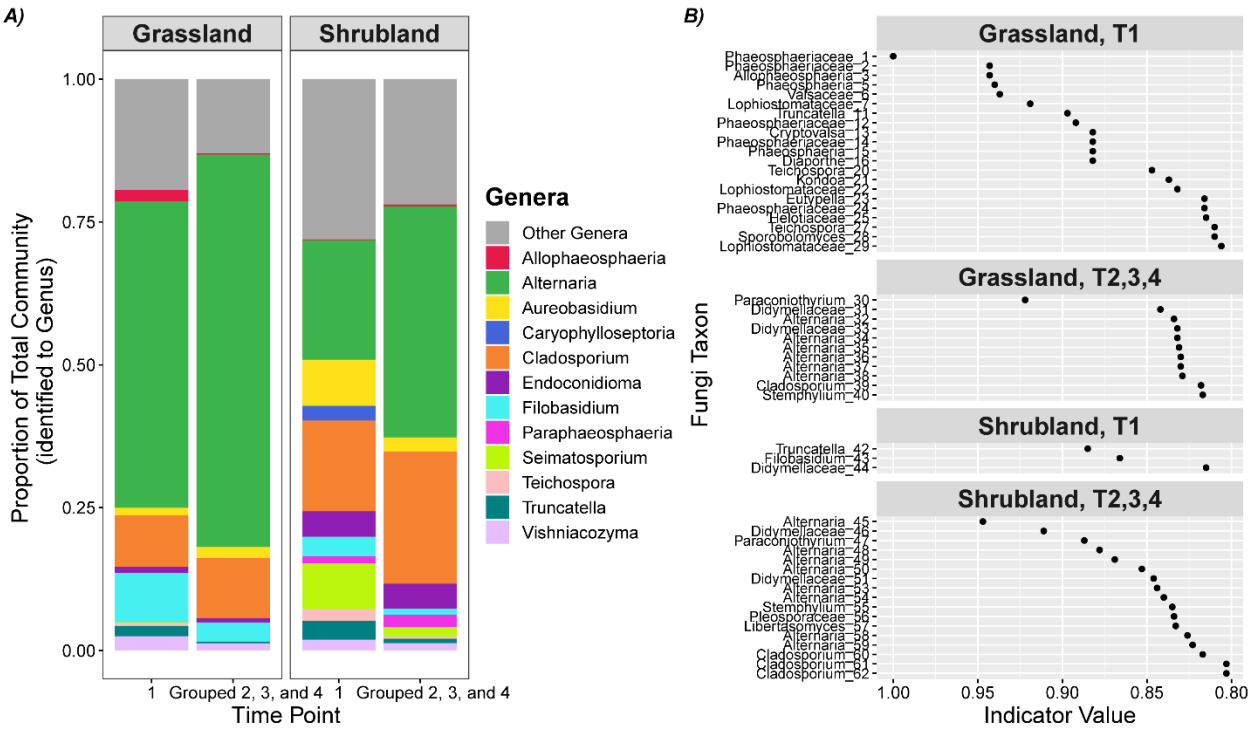

Figure 3.3 Changes in dispersing fungi over time. (A) The fungal genera within the grassland and shrubland ecosystems during the first timepoint and the last three timepoints. **(B)** Indicator species that significantly ($P < 0.05$) characterized the grassland and shrubland ecosystems during the first timepoint and the other three timepoints of the field experiment. The indicator species represented here include all OTUs that had an indicator value of at least 0.8 and were identified at least to the family level.

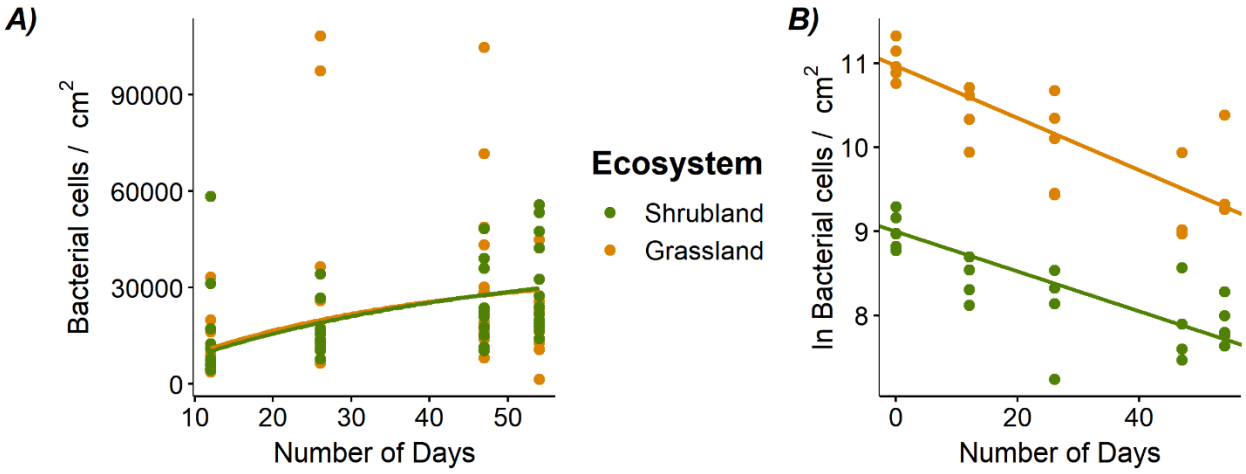

Figure 3.4 Bacterial abundance on glass and death slides. (A) Bacterial abundance on the death slides over time. The lines are linear regressions, from which the slopes are used to model immigration rates in panel B. **(B)** Bacterial abundance on the dispersal slides over time. The curved lines represent the immigration model that does not assume equilibrium between the immigration and death rates: $n(t) = i/d * (1 - e^{(-d*t)})$, where n is the abundance on the dispersal slides, i is the immigration rate, d is the death rate, and t is time.

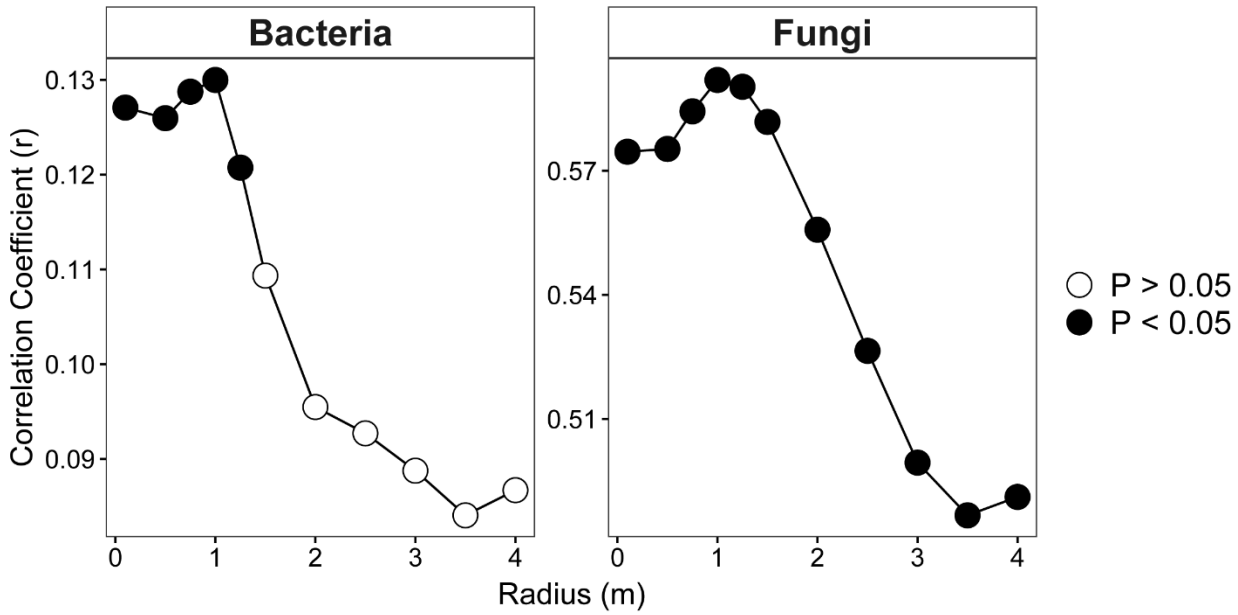

Figure 3.5 Plant composition and dispersing microorganisms. Correlation between bacterial and fungal taxon richness and the surrounding plant composition, measured within circles of increasing radii around each transect location. Each point represents one partial Mantel test. The y-axis shows the Pearson correlation coefficients.

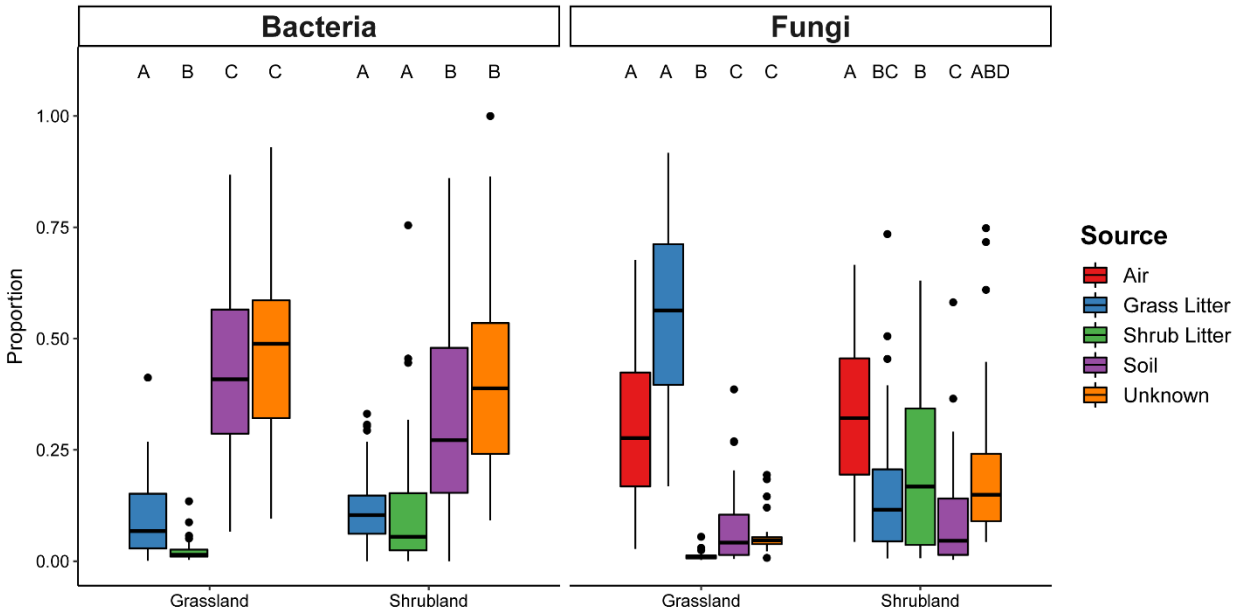

Figure 3.6 Dispersal sources. The proportions of the bacterial and fungal communities that were attributed to different source communities for the grassland (Kruskal-Wallis: $P < 0.0001$) and shrubland ($P < 0.0001$) ecosystems. Letters indicate significant pairwise differences by ecosystem using Dunn's post-hoc test with a Bonferroni correction.

SUPPLEMENTAL MATERIALS

Appendix 3A: Supplemental Figures

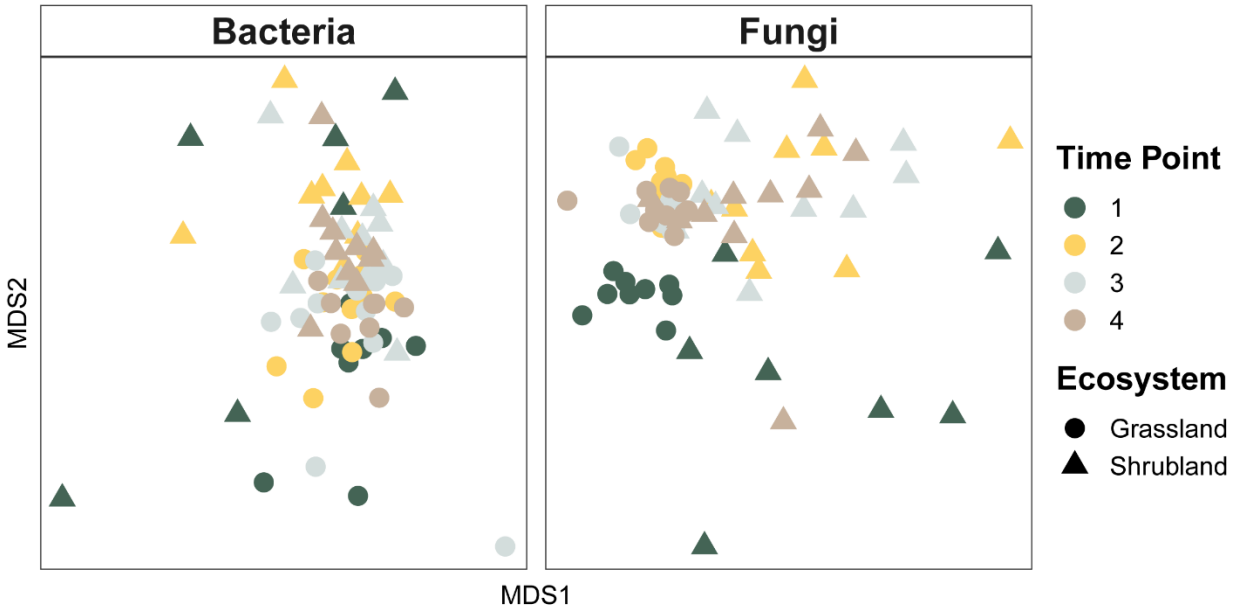

Figure S3.1 Changes in dispersing taxa over time. NMDS of the bacterial and fungal communities on the dispersal slides. Dispersal community composition differs by ecosystem and time for both bacteria and fungi, and by an ecosystem-by-time interaction for fungi (PERMANOVA: $P < 0.05$; Table S3.3).

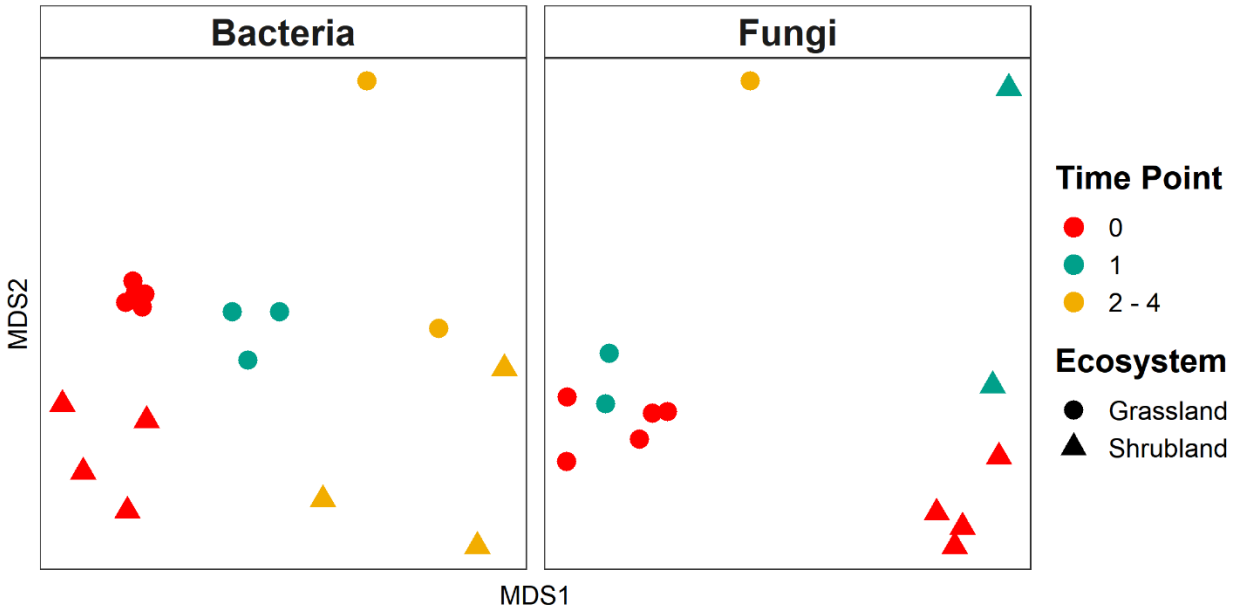

Figure S3.2 Changes in composition on death slides over time. NMDS of the bacterial and fungal communities on the death slides. Community composition differed by ecosystem and time (comparing timepoint 0, 1, and 2 – 4) for both bacteria and fungi, and by a time-by-ecosystem interaction for bacteria (PERMANOVA: $P < 0.05$; Table S3.2).

Appendix 3B: Supplemental Tables

Table S3.1. ANCOVA results for the effects of time and time-by-ecosystem interaction on bacterial abundance on the death rate glass slides.

	Df	SumsOfSqs	F	r^2	Pr(>F)	
Time	1	35.84	211.86	0.82	< **0.0001**	***
Time-by-Ecosystem Interaction	1	0.65	3.87		0.056	
Residuals	43	7.27		0.17		

Table S3.2. PERMANOVA results for the effects of ecosystem, time (grouped into Timepoint 1 and Timepoint 2+3+4), and ecosystem-by-time interaction on community composition on the death slides for A) bacteria and B) fungi.

A) Bacteria

	Df	SumsOfSqs	MeanSqs	Pseudo-F	r^2	P(Perm)	
Time (grouped)	2	1.57	0.79	2.75	0.28	< **0.0001**	***
Ecosystem	1	0.81	0.81	3.69	0.15	**0.0003**	***
Interaction	1	0.54	0.54	2.47	0.10	**0.007**	**
Residuals	12	2.65	0.22		0.47		
Total	16	5.58			1		

B) Fungi

	Df	SumsOfSqs	MeanSqs	Pseudo-F	r^2	P(Perm)	
Time (grouped)	2	0.54	0.29	2.19	0.17	**0.02**	*
Ecosystem	1	1.27	1.27	10.35	0.41	< **0.0001**	***
Interaction	1	0.21	0.21	1.70		0.11	
Residuals	9	1.11	0.12		0.35		
Total	13	3.12			1		

Table S3.3. PERMANOVA results for the effects of ecosystem, time, and ecosystem-by-time interaction on community composition on the dispersal slides for A) bacteria and B) fungi.

A) Bacteria

	Df	SumsOfSqs	MeanSqs	Pseudo-F	r^2	P(Perm)	
Ecosystem	1	0.78	0.78	2.11	0.028	< **0.0001**	***
Time	1	0.54	0.54	1.46	0.019	**0.0008**	***
Interaction	1	0.41	0.41	1.10		0.13	
Residuals	72	26.62	0.37		0.94		
Total	74	27.94			1		

B) Fungi

	Df	SumsOfSqs	MeanSqs	Pseudo-F	r^2	P(Perm)	
Ecosystem	1	1.50	1.50	10.31	0.12	< **0.0001**	***
Time	1	0.63	0.63	4.36	0.050	**0.0002**	***
Interaction	1	0.27	0.27	1.86	0.02	**0.016**	*
Residuals	71	10.30	0.15		0.81		
Total	74	12.69			1		

REFERENCES

Adams, R. I., Miletto, M., Taylor, J. W., & Bruns, T. D. (2013). Dispersal in microbes: Fungi in indoor air are dominated by outdoor air and show dispersal limitation at short distances. *ISME Journal*, *7*(7), 1262–1273. https://doi.org/10.1038/ismej.2013.28

Albright, M. B. N., & Martiny, J. B. H. (2018). Dispersal alters bacterial diversity and composition in a natural community. *The ISME Journal*, *12*(1), 296–299. https://doi.org/10.1038/ismej.2017.161

Albright, M. B. N., Sevanto, S., Gallegos-Graves, L. V., & Dunbar, J. (2020). Biotic interactions are more important than propagule pressure in microbial community invasions. *MBio*, *11*(5), 1–16. https://doi.org/10.1128/mBio.02089-20

Andam, C. P., Doroghazi, J. R., Campbell, A. N., Kelly, P. J., Choudoir, M. J., & Buckley, D. H. (2016). A latitudinal diversity gradient in terrestrial bacteria of the genus Streptomyces. *MBio*, *7*(2), e02200-15. https://doi.org/10.1128/mBio.02200-15

Anderson, M.J., Gorley, R. N., & Clarke, K. R. (2008). *PERMANOVA+ for PRIMER: Guide to Software and Statistical Methods*.

Anderson, Marti J., Ellingsen, K. E., & McArdle, B. H. (2006). Multivariate dispersion as a measure of beta diversity. *Ecology Letters*, *9*(6), 683–693. https://doi.org/10.1111/j.1461-0248.2006.00926.x

Anderson, Marti J., & Walsh, D. C. I. (2013). PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing? *Ecological Monographs*, *83*(4), 557–574. https://doi.org/10.1890/12-2010.1

Apprill, A., McNally, S., Parsons, R., & Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology*, *75*(2), 129–137. https://doi.org/10.3354/ame01753

Asakawa, S., & Kimura, M. (2008). Comparison of bacterial community structures at main habitats in paddy field ecosystem based on DGGE analysis. *Soil Biology and Biochemistry*, *40*(6), 1322–1329. https://doi.org/10.1016/J.SOILBIO.2007.09.024

Auguet, J.-C., Barberan, A., & Casamayor, E. O. (2010). Global ecological patterns in uncultured Archaea. *The ISME Journal*, *4*(2), 182–190. https://doi.org/10.1038/ismej.2009.109

Austin, A. T., & Vivanco, L. (2006). Plant litter decomposition in a semi-arid ecosystem controlled by photodegradation. *Nature*, *442*(7102), 555–558. https://doi.org/10.1038/nature05038

Baas-Becking, L. G. M. (1934). Geobiology or Introduction to Environmental Science (Translated from Dutch). *WP Van Stockum & Zoon: The Hague, Netherlands*.

Bahram, M., Hildebrand, F., Forslund, S. K., Anderson, J. L., Soudzilovskaia, N. A., Bodegom, P. M., Bengtsson-Palme, J., Anslan, S., Coelho, L. P., Harend, H., Huerta-Cepas, J., Medema, M. H., Maltz, M. R., Mundra, S., Olsson, P. A., Pent, M., Põlme, S., Sunagawa, S., Ryberg, M., … Bork, P. (2018). Structure and function of the global topsoil microbiome. In *Nature* (Vol. 560, Issue 7717, pp. 233–237). Nature Publishing Group.

https://doi.org/10.1038/s41586-018-0386-6

Baker, N. R., Khalili, B., Martiny, J. B. H., & Allison, S. D. (2018). Microbial decomposers not constrained by climate history along a Mediterranean climate gradient in southern California. *Ecology*, *99*(6), 1441–1452. https://doi.org/10.1002/ecy.2345

Barberán, A., & Casamayor, E. (2010). Global phylogenetic community structure and β-diversity patterns in surface bacterioplankton metacommunities. *Aquatic Microbial Ecology*, *59*(1), 1–10. https://doi.org/10.3354/ame01389

Barbour, K. M., Weihe, C., & Martiny, J. B. H. (2021). Bacterial community composition in leaf litter and soil in grassland and shrubland ecosystems. Unpublished raw data.

Bates, S. T., Berg-Lyons, D., Caporaso, J. G., Walters, W. A., Knight, R., & Fierer, N. (2011). Examining the global distribution of dominant archaeal populations in soil. *The ISME Journal*, *5*, 908–917. https://doi.org/10.1038/ismej.2010.171

Bell, E., Blake, L. I., Sherry, A., Head, I. M., & Hubert, C. R. J. (2018). Distribution of thermophilic endospores in a temperate estuary indicate that dispersal history structures sediment microbial communities. *Environmental Microbiology*, *20*(3), 1134–1147. https://doi.org/10.1111/1462-2920.14056

Bell, T. (2010). Experimental tests of the bacterial distance-decay relationship. *The ISME Journal*, *4*(11), 1357–1365. https://doi.org/10.1038/ismej.2010.77

Bell, T., Newman, J. A., Silverman, B. W., Turner, S. L., & Lilley, A. K. (2005). The contribution of species richness and composition to bacterial services. *Nature*, *436*(7054), 1157–1160. https://doi.org/10.1038/nature03891

Bertolini, V., Gandolfi, I., Ambrosini, R., Bestetti, G., Innocente, E., Rampazzo, G., & Franzetti, A. (2013). Temporal variability and effect of environmental variables on airborne bacterial communities in an urban area of Northern Italy. *Applied Microbiology and Biotechnology*, *97*(14), 6561–6570. https://doi.org/10.1007/s00253-012-4450-0

Bickel, S., Chen, X., Papritz, A., & Or, D. (2019). A hierarchy of environmental covariates control the global biogeography of soil bacterial richness. *Scientific Reports*, *9*(1), 1–10. https://doi.org/10.1038/s41598-019-48571-w

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. In *Nature Biotechnology* (Vol. 37, Issue 8, pp. 852–857). Nature Publishing Group. https://doi.org/10.1038/s41587-019-0209-9

Bowers, R. M., McLetchie, S., Knight, R., & Fierer, N. (2011). Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments. *ISME Journal*, *5*(4), 601–612. https://doi.org/10.1038/ismej.2010.167

Bowers, R. M., Sullivan, A. P., Costello, E. K., Collett, J. L., Knight, R., & Fierer, N. (2011). Sources of bacteria in outdoor air across cities in the midwestern United States. *Applied and Environmental Microbiology*, *77*(18), 6350–6356. https://doi.org/10.1128/AEM.05498-11

Bryan, N. C., Christner, B. C., Guzik, T. G., Granger, D. J., & Stewart, M. F. (2019). Abundance and survival of microbial aerosols in the troposphere and stratosphere. *ISME Journal*, *13*(11), 2789–2799. https://doi.org/10.1038/s41396-019-0474-0

Bryant, J. A., Lamanna, C., Morlon, H., Kerkhoff, A. J., Enquist, B. J., & Green, J. L. (2008). Microbes on mountainsides: Contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences*, *105*(Supplement 1), 11505–11511. https://doi.org/10.1073/pnas.0801920105

Cadotte, M. W. (2006). Dispersal and Species Diversity: A Meta-Analysis. *The American Naturalist*, *167*(6), 913–924.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Campbell, B. J., & Kirchman, D. L. (2013). Bacterial diversity, community structure and potential growth rates along an estuarine salinity gradient. *The ISME Journal*, *7*(1), 210–220. https://doi.org/10.1038/ismej.2012.93

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., … Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, *7*(5), 335–336. https://doi.org/10.1038/nmeth.f.303

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G., & Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME Journal*, *6*(8), 1621–1624. https://doi.org/10.1038/ismej.2012.8

Carson, J. K., Gonzalez-Quiñones, V., Murphy, D. V, Hinz, C., Shaw, J. A., & Gleeson, D. B. (2010). Low pore connectivity increases bacterial diversity in soil. *Applied and Environmental Microbiology*, *76*(12), 3936–3942. https://doi.org/10.1128/AEM.03085-09

Cevallos-Cevallos, J. M., Danyluk, M. D., Gu, G., Vallad, G. E., & van Bruggen, A. H. C. (2012). Dispersal of Salmonella Typhimurium by Rain Splash onto Tomato Plants. *Journal of Food Protection*, *75*(3), 472–479. https://doi.org/10.4315/0362-028X.JFP-11-399

Chase, J. M., Kraft, N. J. B., Smith, K. G., Vellend, M., & Inouye, B. D. (2011). Using null models to disentangle variation in community dissimilarity from variation in α-diversity. *Ecosphere*, *2*(2), art24. https://doi.org/10.1890/ES10-00117.1

Cheeke, T. E., Schütte, U. M., Hemmerich, C. M., Cruzan, M. B., Rosenstiel, T. N., & Bever, J. D. (2015). Spatial soil heterogeneity has a greater effect on symbiotic arbuscular mycorrhizal fungal communities and plant growth than genetic modification with Bacillus thuringiensis toxin genes. *Molecular Ecology*. https://doi.org/10.1111/mec.13178

Chen, J., Wang, P., Wang, C., Wang, X., Miao, L., Liu, S., Yuan, Q., & Sun, S. (2020). Fungal community demonstrates stronger dispersal limitation and less network connectivity than bacterial community in sediments along a large river. *Environmental Microbiology*, *22*(3), 832–849. https://doi.org/10.1111/1462-2920.14795

Choudoir, M. J., Barberán, A., Menninger, H. L., Dunn, R. R., & Fierer, N. (2017). Variation in range size and dispersal capabilities of microbial taxa. *Ecology*, *0*(0), 1–13. https://doi.org/10.1002/ecy.2094

Clarke, K. R., & Gorley, R. N. (2006). *Primer v6: User Manual/Tutorial*.

Comte, J., Langenheder, S., Berga, M., & Lindström, E. S. (2017). Contribution of different dispersal sources to the metabolic response of lake bacterioplankton following a salinity change. *Environmental Microbiology*, *19*(1), 251–260. https://doi.org/10.1111/1462-2920.13593

Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannan, B. J. M., & Relman, D. A. (2012). The application of ecological theory toward an understanding of the human microbiome. In *Science* (Vol. 336, Issue 6086, pp. 1255–1262). American Association for the Advancement of Science. https://doi.org/10.1126/science.1224203

Crump, B. C., Amaral-Zettler, L. A., & Kling, G. W. (2012). Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *The ISME Journal*, *6*(9), 1629–1639. https://doi.org/10.1038/ismej.2012.9

Curd, E. E., Martiny, J. B. H., Li, H., & Smith, T. B. (2018). Bacterial diversity is positively correlated with soil heterogeneity. *Ecosphere*. https://doi.org/10.1002/ecs2.2079

De Caceres, M., & Legendre, P. (2009). Associations between species and groups of sites: indices and statistical inference. *Ecology*, *90*(12), 3566–3574. http://sites.google.com/site/miqueldecaceres/

Declerck, S. A. J., Winter, C., Shurin, J. B., Suttle, C. A., & Matthews, B. (2013). Effects of patch connectivity and heterogeneity on metacommunity structure of planktonic bacteria and viruses. *ISME Journal*, *7*(3), 533–542. https://doi.org/10.1038/ismej.2012.138

Delgado-Baquerizo, M., & Eldridge, D. J. (2019). Cross-Biome Drivers of Soil Bacterial Alpha Diversity on a Worldwide Scale. *Ecosystems*, *22*(6), 1220–1231. https://doi.org/10.1007/s10021-018-0333-2

Dell'Anno, A., & Danovaro, R. (2005). Ecology: Extracellular DNA plays a key role in deep-sea ecosystem functioning. *Science*, *309*(5744), 2179. https://doi.org/10.1126/science.1117475

Ding, J., Jiang, X., Ma, M., Zhou, B., Guan, D., Zhao, B., Zhou, J., Cao, F., Li, L., & Li, J. (2016). Effect of 35 years inorganic fertilizer and manure amendment on structure of bacterial and archaeal communities in black soil of northeast China. *Applied Soil Ecology*, *105*, 187–195. https://doi.org/10.1016/j.apsoil.2016.04.010

Elmassry, M. M., Ray, N., Sorge, S., Webster, J., Merry, K., Caserio, A., Vecellio, D. J., Kruczek, C., Dowd, S., Ardon-Dryer, K., Vanos, J., & San Francisco, M. J. (2020). Investigating the culturable atmospheric fungal and bacterial microbiome in West Texas: Implication of dust storms and origins of the air parcels. *FEMS Microbes*, *1*(1). https://doi.org/10.1093/femsmc/xtaa009

Evans, S. E., Bell-Dereske, L. P., Dougherty, K. M., & Kittredge, H. A. (2019). Dispersal alters soil microbial community response to drought. *Environmental Microbiology*. https://doi.org/10.1111/1462-2920.14707

Evans, S., Martiny, J. B. H., & Allison, S. D. (2016). Effects of dispersal and selection on

stochastic assembly in microbial communities. *The ISME Journal*, *11*(1), 1–10. https://doi.org/10.1038/ismej.2016.96

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(3), 626–631. https://doi.org/10.1073/pnas.0507535103

Fierer, N., Leff, J. W., Adams, B. J., Nielsen, U. N., Bates, S. T., Lauber, C. L., Owens, S., Gilbert, J. A., Wall, D. H., & Caporaso, J. G. (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences*, *109*(52), 21390–21395. https://doi.org/10.1073/pnas.1215210110

Fierer, N., & Lennon, J. T. (2011). The generation and maintenance of diversity in microbial communities. *American Journal of Botany*, *98*(3), 439–448. https://doi.org/10.3732/ajb.1000498

Finks, S. S., Weihe, C., Kimball, S., Allison, S. D., Martiny, A. C., Treseder, K. K., & Martiny, J. B. H. (2021). Microbial community response to a decade of simulated global changes depends on the plant community. *Elementa: Science of the Anthropocene*, *9*(1). https://doi.org/10.1525/elementa.2021.00124

Finlay, B. J. (2002). Global dispersal of free-living microbial eukaryote species. In *Science* (Vol. 296, Issue 5570, pp. 1061–1063). American Association for the Advancement of Science. https://doi.org/10.1126/science.1070710

Foissner, W. (2006). Biogeography and dispersal of micro-organisms: A review emphasizing protists. In *Acta Protozoologica* (Vol. 45, Issue 2, pp. 111–136). https://doi.org/10.1.1.465.6672

Fox, J., & Weisberg, S. (2011). *An {R} Companion to Applied Regression* (Second). Sage. http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Freestone, A. L., & Inouye, B. D. (2006). Dispersal Limitation and Environmental Heterogeneity Shape Scale-Dependent Diversity Patterns in Plant Communities. *Ecology*, *87*(10), 2425–2432. https://doi.org/10.1890/0012-9658(2006)87[2425:DLAEHS]2.0.CO;2

Galante, T. E., Horton, T. R., & Swaney, D. P. (2017). 95% of basidiospores fall within 1 m of the cap: a field-and modeling-based study. *Https://Doi.Org/10.3852/10-388*, *103*(6), 1175–1183. https://doi.org/10.3852/10-388

Galès, A., Latrille, E., Wéry, N., Steyer, J. P., & Godon, J. J. (2014). Needles of pinus halepensis as biomonitors of bioaerosol emissions. *PLoS ONE*, *9*(11), e112182. https://doi.org/10.1371/journal.pone.0112182

Gimmler, A., Korn, R., de Vargas, C., Audic, S., & Stoeck, T. (2016). The Tara Oceans voyage reveals global diversity and distribution patterns of marine planktonic ciliates. *Scientific Reports*, *6*(1), 33555. https://doi.org/10.1038/srep33555

Glassman, S. I., & Martiny, J. B. H. (2018). *Broadscale Ecological Patterns Are Robust to Use of Exact Sequence Variants versus Operational Taxonomic Units*.

https://doi.org/10.1128/mSphere

Glassman, S. I., Weihe, C., Li, J., Albright, M. B. N., Looby, C. I., Martiny, A. C., Treseder, K. K., Allison, S. D., & Martiny, J. B. H. (2018). Decomposition responses to climate depend on microbial community composition. *Proceedings of the National Academy of Sciences*, *115*(47), 11994–11999. https://doi.org/10.1073/PNAS.1811269115

Griffiths, R. I., Thomson, B. C., James, P., Bell, T., Bailey, M., & Whiteley, A. S. (2011). The bacterial biogeography of British soils. *Environmental Microbiology*, *13*(6), 1642–1654. https://doi.org/10.1111/j.1462-2920.2011.02480.x

Hall-Stoodley, L., Costerton, J. W., & Stoodley, P. (2004). Bacterial biofilms: From the natural environment to infectious diseases. In *Nature Reviews Microbiology* (Vol. 2, Issue 2, pp. 95–108). Nature Publishing Group. https://doi.org/10.1038/nrmicro821

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., & Martiny, J. B. H. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature Reviews Microbiology*, *10*(7), 1–10. https://doi.org/10.1038/nrmicro2795

Hanson, C. A., Müller, A. L., Loy, A., Dona, C., Appel, R., Jørgensen, B. B., & Hubert, C. R. J. (2019). Historical factors associated with past environments influence the biogeography of thermophilic endospores in Arctic marine sediments. *Frontiers in Microbiology*, *10*(245). https://doi.org/10.3389/fmicb.2019.00245

Hijmans, R. J. (2018). *raster: Geographic Data Analysis and Modeling* (R package version 2.8-4.). https://cran.r-project.org/package=raster

Hobbie, S. E. (2015). Plant species effects on nutrient cycling: revisiting litter feedbacks. In *Trends in Ecology and Evolution* (Vol. 30, Issue 6, pp. 357–363). Elsevier Ltd. https://doi.org/10.1016/j.tree.2015.03.015

Holdridge, L. R. (1967). Life zone ecology. In *Life zone ecology.* (Issue (rev. ed.)). Tropical Science Center.

Ibarbalz, F. M., Henry, N., Brandão, M. C., Martini, S., Busseni, G., Byrne, H., Coelho, L. P., Endo, H., Gasol, J. M., Gregory, A. C., Mahé, F., Rigonato, J., Royo-Llonch, M., Salazar, G., Sanz-Sáez, I., Scalco, E., Soviadan, D., Zayed, A. A., Zingone, A., … Zinger, L. (2019). Global Trends in Marine Plankton Diversity across Kingdoms of Life. *Cell*, *179*(5), 1084–1097. https://doi.org/10.1016/j.cell.2019.10.008

Jangid, K., Williams, M. A., Franzluebbers, A. J., Sanderlin, J. S., Reeves, J. H., Jenkins, M. B., Endale, D. M., Coleman, D. C., & Whitman, W. B. (2008). Relative impacts of land-use, management intensity and fertilization upon soil microbial community structure in agricultural systems. *Soil Biology and Biochemistry*, *40*(11), 2843–2853. https://doi.org/10.1016/j.soilbio.2008.07.030

Johnson, M. P., & Simberloff, D. S. (1974). Environmental determinants of island species numbers in the british Isles. *Journal of Biogeography*, *1*(3), 149–154. https://doi.org/10.2307/3037964

Jones, S. E., Newton, R. J., & McMahon, K. D. (2008). Potential for atmospheric deposition of bacteria to influence bacterioplankton communities. *FEMS Microbiology Ecology*, *64*(3), 388–394. https://doi.org/10.1111/j.1574-6941.2008.00476.x

Kaneko, R., & Kaneko, S. (2004). The effect of bagging branches on levels of endophytic fungal infection in Japanese beech leaves. *Forest Pathology*, *34*(2), 65–78. https://doi.org/10.1111/j.1439-0329.2004.00350.x

Khalili, B., Weihe, C., Kimball, S., Schmidt, K. T., & Martiny, J. B. H. (2019). Optimization of a Method To Quantify Soil Bacterial Abundance by Flow Cytometry. *MSphere*, *4*(5). https://doi.org/10.1128/msphere.00435-19

Kier, G., Mutke, J., Dinerstein, E., Ricketts, T. H., Küper, W., Kreft, H., & Barthlott, W. (2005). Global patterns of plant diversity and floristic knowledge. *Journal of Biogeography*, *32*(7), 1107–1116. https://doi.org/10.1111/j.1365-2699.2005.01272.x

Kimball, S., Goulden, M. L., Suding, K. N., & Parker, S. (2014). Altered water and nitrogen input shifts succession in a southern California coastal sage community. *Ecological Applications*, *24*(6), 1390–1404. https://doi.org/10.1890/13-1313.1

Knights, D., Kuczynski, J., Charlson, E. S., Zaneveld, J., Mozer, M. C., Collman, R. G., Bushman, F. D., Knight, R., & Kelley, S. T. (2011). Bayesian community-wide culture-independent microbial source tracking. *Nature Methods*, *8*(9), 761–763. https://doi.org/10.1038/nmeth.1650

Lamb, E. G., Kennedy, N., & Siciliano, S. D. (2011). Effects of plant species richness and evenness on soil microbial community diversity and function. *Plant and Soil*, *338*(1), 483–495. https://doi.org/10.1007/s11104-010-0560-6

Lane, D. J., Pace, B., Olsen, G. J., Stahl, D. A., Sogin, M. L., & Pace, N. R. (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences of the United States of America*, *82*(20), 6955–6959. https://doi.org/10.1073/pnas.82.20.6955

Lee, J. J., Srinivasan, S., Lim, S., Joe, M., Lee, S. H., Kwon, S. A., Kwon, Y. J., Lee, J., Choi, J. J., Lee, H. M., Auh, Y. K., & Kim, M. K. (2014). Hymenobacter swuensis sp. nov., a gamma-radiation-resistant bacteria isolated from mountain soil. *Current Microbiology*, *68*(3), 305–310. https://doi.org/10.1007/s00284-013-0478-3

Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., Holt, R. D., Shurin, J. B., Law, R., Tilman, D., Loreau, M., & Gonzalez, A. (2004). The metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters*, *7*(7), 601–613. https://doi.org/10.1111/j.1461-0248.2004.00608.x

Leung, M. H. Y., Wilkins, D., Li, E. K. T., Kong, F. K. F., & Lee, P. K. H. (2014). Indoor-air microbiome in an urban subway network: Diversity and dynamics. *Applied and Environmental Microbiology*, *80*(21), 6760–6770. https://doi.org/10.1128/AEM.02244-14

Li, S., Wang, P., Chen, Y., Wilson, M. C., Yang, X., Ma, C., Lu, J., Chen, X., Wu, J., Shu, W., & Jiang, L. (2020). Island biogeography of soil bacteria and fungi: similar patterns, but different mechanisms. *The ISME Journal 2020 14:7*, *14*(7), 1886–1896. https://doi.org/10.1038/s41396-020-0657-8

Lindemann, J., Constantinidou, H. A., Barchet, W. R., & Upper, C. D. (1982). Plants as sources of airborne bacteria, including ice nucleation-active bacteria. *Applied and Environmental Microbiology*, *44*(5), 1059–1063. https://doi.org/10.1128/aem.44.5.1059-1063.1982

Lindemann, Julianne, & Upper, C. D. (1985). Aerial Dispersal of Epiphytic Bacteria over Bean Plants. *Applied and Environmental Microbiology*, *50*(5), 1229–1232. https://doi.org/10.1128/aem.50.5.1229-1232.1985

Lindow, S. E., & Andersen, G. L. (1996). Influence of immigration on epiphytic bacterial populations on navel orange leaves. *Applied and Environmental Microbiology*, *62*(8), 2978–2987. https://doi.org/10.1128/aem.62.8.2978-2987.1996

Lindow, S. E., & Leveau, J. H. J. (2002). Phyllosphere microbiology. In *Current Opinion in Biotechnology* (Vol. 13, Issue 3, pp. 238–243). Elsevier Ltd. https://doi.org/10.1016/S0958-1669(02)00313-0

Lindström, E. S., & Langenheder, S. (2012). Local and regional factors influencing bacterial community assembly. In *Environmental Microbiology Reports* (Vol. 4, Issue 1, pp. 1–9). Blackwell Publishing Ltd. https://doi.org/10.1111/j.1758-2229.2011.00257.x

Lindström, E. S., & Östman, Ö. (2011). The importance of dispersal for bacterial community composition and functioning. *PLoS ONE*, *6*(10), e25883. https://doi.org/10.1371/journal.pone.0025883

Logares, R., Deutschmann, I. M., Junger, P. C., Giner, C. R., Krabberød, A. K., Schmidt, T. S. B., Rubinat-Ripoll, L., Mestre, M., Salazar, G., Ruiz-González, C., Sebastián, M., De Vargas, C., Acinas, S. G., Duarte, C. M., Gasol, J. M., & Massana, R. (2020). Disentangling the mechanisms shaping the surface ocean microbiota. *Microbiome*, *8*(1), 1–17. https://doi.org/10.1186/s40168-020-00827-8

Looby, C. I., Maltz, M. R., & Treseder, K. K. (2016). Belowground responses to elevation in a changing cloud forest. *Ecology and Evolution*, *6*(7), 1996–2009. https://doi.org/10.1002/ece3.2025

Lozupone, C. A., & Knight, R. (2007). Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences*, *104*(27), 11436–11440. https://doi.org/10.1073/pnas.0611525104

Magurran, A. (1988). *Ecological diversity and its measurement*. Princeton university press.

Maignien, L., DeForce, E. A., Chafee, M. E., Murat Eren, A., & Simmons, S. L. (2014). Ecological succession and stochastic variation in the assembly of Arabidopsis thaliana phyllosphere communities. *MBio*, *5*(1), e00682-13. https://doi.org/10.1128/mBio.00682-13

Martinez Arbizu, P. (2017). *pairwiseAdonis: Pairwise multilevel comparison using adonis*.

Martiny, Jennifer B. H., Martiny, A. C., Weihe, C., Lu, Y., Berlemont, R., Brodie, E. L., Goulden, M. L., Treseder, K. K., & Allison, S. D. (2017). Microbial legacies alter decomposition in response to simulated global change. *The ISME Journal*, *11*(2), 1–10. https://doi.org/10.1038/ismej.2016.122

Martiny, Jennifer B. Hughes, Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M., Krumins, J. A., Kuske, C. R., Morin, P. J., Naeem, S., Øvreås, L., Reysenbach, A.-L., Smith, V. H., & Staley, J. T. (2006). Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, *4*(2), 102–112. https://doi.org/10.1038/nrmicro1341

Martiny, Jennifer B.H., & Walters, K. E. (2018). Towards a Natural History of Soil Bacterial

Communities. In *Trends in Microbiology* (Vol. 26, Issue 4, pp. 250–252). Elsevier Ltd. https://doi.org/10.1016/j.tim.2018.02.010

Matulich, K. L., Weihe, C., Allison, S. D., Amend, A. S., Berlemont, R., Goulden, M. L., Kimball, S., Martiny, A. C., & Martiny, J. B. (2015). Temporal variation overshadows the response of leaf litter microbial communities to simulated global change. *The ISME Journal*, *9*(10), 2477–2489. https://doi.org/10.1038/ismej.2015.58

McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., Desantis, T. Z., Probst, A., Andersen, G. L., Knight, R., & Hugenholtz, P. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME Journal*, *6*(3), 610–618. https://doi.org/10.1038/ismej.2011.139

Miletto, M., & Lindow, S. E. (2015). Relative and contextual contribution of different sources to the composition and abundance of indoor air bacteria in residences. *Microbiome*, *3*(1), 61. https://doi.org/10.1186/s40168-015-0128-z

Milici, M., Tomasch, J., Wos-Oxley, M. L., Wang, H., Jáuregui, R., Camarinha-Silva, A., Deng, Z. L., Plumeier, I., Giebel, H. A., Wurst, M., Pieper, D. H., Simon, M., & Wagner-Döbler, I. (2016). Low diversity of planktonic bacteria in the tropical ocean. *Scientific Reports*, *6*(1), 1–9. https://doi.org/10.1038/srep19054

Miller, S. R., Strong, A. L., Jones, K. L., & Ungerer, M. C. (2009). Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. *Applied and Environmental Microbiology*, *75*(13), 4565–4572. https://doi.org/10.1128/AEM.02792-08

Mittelbach, G. G., Schemske, D. W., Cornell, H. V., Allen, A. P., Brown, J. M., Bush, M. B., Harrison, S. P., Hurlbert, A. H., Knowlton, N., Lessios, H. A., McCain, C. M., McCune, A. R., McDade, L. A., McPeek, M. A., Near, T. J., Price, T. D., Ricklefs, R. E., Roy, K., Sax, D. F., … Turelli, M. (2007). Evolution and the latitudinal diversity gradient: Speciation, extinction and biogeography. In *Ecology Letters* (Vol. 10, Issue 4, pp. 315–331). John Wiley & Sons, Ltd (10.1111). https://doi.org/10.1111/j.1461-0248.2007.01020.x

Needham, D. M., Sachdeva, R., & Fuhrman, J. A. (2017). Ecological dynamics and co-occurrence among marine phytoplankton, bacteria and myoviruses shows microdiversity matters. *The ISME Journal*, *11*(7), 1614–1629. https://doi.org/10.1038/ismej.2017.29

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., Knelman, J. E., Darcy, J. L., Lynch, R. C., Wickey, P., & Ferrenberg, S. (2013). Patterns and Processes of Microbial Community Assembly. *Microbiology and Molecular Biology Reviews*, *77*(3), 342–356. https://doi.org/10.1128/MMBR.00051-12

Neufeld, J. D., & Mohn, W. W. (2005). Unexpectedly High Bacterial Diversity in Arctic Tundra Relative to Boreal Forest Soils , Revealed by Serial Analysis of Ribosomal Sequence Tags Unexpectedly High Bacterial Diversity in Arctic Tundra Relative to Boreal Forest Soils , Revealed by Serial Ana. *Applied and Environmental Microbiology*, *71*(10), 5710–5718. https://doi.org/10.1128/AEM.71.10.5710

Nilsson, R. H., Larsson, K. H., Taylor, A. F. S., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., Kennedy, P., Picard, K., Glöckner, F. O., Tedersoo, L., Saar, I., Kõljalg, U., & Abarenkov, K. (2019). The UNITE database for molecular identification of fungi: Handling

dark taxa and parallel taxonomic classifications. *Nucleic Acids Research*, *47*(D1), D259–D264. https://doi.org/10.1093/nar/gky1022

Norros, V., Penttilä, R., Suominen, M., & Ovaskainen, O. (2012). Dispersal may limit the occurrence of specialist wood decay fungi already at small spatial scales. *Oikos*, *121*(6), 961–974. https://doi.org/10.1111/j.1600-0706.2012.20052.x

Nychka, D., Furrer, R., Paige, J., & Sain, S. (2015). *fields: Tools for spatial data.* https://doi.org/10.5065/D6W957CT

O'Brien, S. L., Gibbons, S. M., Owens, S. M., Hampton-Marcell, J., Johnston, E. R., Jastrow, J. D., Gilbert, J. A., Meyer, F., & Antonopoulos, D. A. (2016). Spatial scale drives patterns in soil bacterial diversity. *Environmental Microbiology*, *18*(6), 2039–2051. https://doi.org/10.1111/1462-2920.13231

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2019). *vegan: Community Ecology Package* (R package version 2.5-6). https://cran.r-project.org/package=vegan

Peay, K. G., Garbelotto, M., & Bruns, T. D. (2010). Evidence of dispersal limitation in soil microorganisms: Isolation reduces species richness on mycorrhizal tree islands. *Ecology*, *91*(12), 3631–3640. https://doi.org/10.1890/09-2237.1

Peay, K. G., Schubert, M. G., Nguyen, N. H., & Bruns, T. D. (2012). Measuring ectomycorrhizal fungal dispersal: Macroecological patterns driven by microscopic propagules. *Molecular Ecology*, *21*(16), 4122–4136. https://doi.org/10.1111/j.1365-294X.2012.05666.x

Philippot, L., Spor, A., Hénault, C., Bru, D., Bizouard, F., Jones, C. M., Sarr, A., & Maron, P.-A. (2013). Loss in microbial diversity affects nitrogen cycling in soil. *The ISME Journal*, *7*(8), 1609–1619. https://doi.org/10.1038/ismej.2013.34

Prober, S. M., Leff, J. W., Bates, S. T., Borer, E. T., Firn, J., Harpole, W. S., Lind, E. M., Seabloom, E. W., Adler, P. B., Bakker, J. D., Cleland, E. E., Decrappeo, N. M., Delorenze, E., Hagenah, N., Hautier, Y., Hofmockel, K. S., Kirkman, K. P., Knops, J. M. H., La Pierre, K. J., … Fierer, N. (2015). Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide. *Ecology Letters*, *18*(1), 85–95. https://doi.org/10.1111/ele.12381

Punnapayak, H., Sudhadham, M., Prasongsuk, S., & Pichayangkura, S. (2003). Characterization of Aureobasidium pullulans isolated from airborne spores in Thailand. *Journal of Industrial Microbiology & Biotechnology*, *30*(2), 89–94. https://doi.org/10.1007/s10295-002-0016-y

Purahong, W., Wubet, T., Lentendu, G., Schloter, M., Pecyna, M. J., Kapturska, D., Hofrichter, M., Krüger, D., & Buscot, F. (2016). Life in leaf litter: novel insights into community dynamics of bacteria and fungi during litter decomposition. *Molecular Ecology*, *25*(16), 4059–4074. https://doi.org/10.1111/mec.13739

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

Raes, E. J., Bodrossy, L., van de Kamp, J., Bissett, A., & Waite, A. M. (2018). Marine bacterial richness increases towards higher latitudes in the eastern Indian Ocean. *Limnology and*

Oceanography Letters, *3*(1), 10–19. https://doi.org/10.1002/lol2.10058

Ramette, A., & Tiedje, J. M. (2007). Biogeography: An emerging cornerstone for understanding prokaryotic diversity, ecology, and evolution. *Microbial Ecology*, *53*(2), 197–207. https://doi.org/10.1007/s00248-005-5010-2

Ramírez, G. A., Jørgensen, S. L., Zhao, R., & D'Hondt, S. (2018). Minimal Influence of Extracellular DNA on Molecular Surveys of Marine Sedimentary Communities. *Frontiers in Microbiology*, *9*, 2969. https://doi.org/10.3389/fmicb.2018.02969

Rastogi, G., Coaker, G. L., & Leveau, J. H. J. (2013). New insights into the structure and function of phyllosphere microbiota through high-throughput molecular approaches. *FEMS Microbiology Letters*, *348*(1), 1–10. https://doi.org/10.1111/1574-6968.12225

Rime, T., Hartmann, M., & Frey, B. (2016). Potential sources of microbial colonizers in an initial soil ecosystem after retreat of an alpine glacier. *The ISME Journal*, *10*(7), 1625–1641. https://doi.org/10.1038/ismej.2015.238

Rintala, H., Pitkaranta, M., Toivola, M., Paulin, L., & Nevalainen, A. (2008). Diversity and seasonal dynamics of bacterial community in indoor environment. *BMC Microbiology*, *8*(1), 56. https://doi.org/10.1186/1471-2180-8-56

Rodrigues, J. L. M., Pellizari, V. H., Mueller, R., Baek, K., Jesus, E. d. C., Paula, F. S., Mirza, B., Hamaoui, G. S., Tsai, S. M., Feigl, B., Tiedje, J. M., Bohannan, B. J. M., & Nusslein, K. (2013). Conversion of the Amazon rainforest to agriculture results in biotic homogenization of soil bacterial communities. *Proceedings of the National Academy of Sciences*, *110*(3), 988–993. https://doi.org/10.1073/pnas.1220608110

Ronce, O. (2007). How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annual Review of Ecology, Evolution, and Systematics*, *38*, 231–253. https://doi.org/10.1146/annurev.ecolsys.38.091206.095611

Santander, M. V., Mitts, B. A., Pendergraft, M. A., Dinasquet, J., Lee, C., Moore, A. N., Cancelada, L. B., Kimble, K. A., Malfatti, F., & Prather, K. A. (2021). Tandem Fluorescence Measurements of Organic Matter and Bacteria Released in Sea Spray Aerosols. *Environmental Science & Technology*, acs.est.0c05493. https://doi.org/10.1021/acs.est.0c05493

Šantl-Temkiv, T., Gosewinkel, U., Starnawski, P., Lever, M., & Finster, K. (2018). Aeolian dispersal of bacteria in southwest Greenland: Their sources, abundance, diversity and physiological states. *FEMS Microbiology Ecology*, *94*(4). https://doi.org/10.1093/femsec/fiy031

Sayer, E. J. (2006). Using experimental manipulation to assess the roles of leaf litter in the functioning of forest ecosystems. *Biological Reviews of the Cambridge Philosophical Society*, *81*(1), 1–31. https://doi.org/10.1017/S1464793105006846

Schmidt, S. K., Nemergut, D. R., Darcy, J. L., & Lynch, R. (2014). Do bacterial and fungal communities assemble differently during primary succession? In *Molecular Ecology* (Vol. 23, Issue 2, pp. 254–258). John Wiley & Sons, Ltd. https://doi.org/10.1111/mec.12589

Sedláček, I., Pantůček, R., Králová, S., Mašlaňová, I., Holochová, P., Staňková, E., Vrbovská, V., Švec, P., & Busse, H. J. (2019). Hymenobacter amundsenii sp. nov. resistant to

ultraviolet radiation, isolated from regoliths in Antarctica. *Systematic and Applied Microbiology*, *42*(3), 284–290. https://doi.org/10.1016/j.syapm.2018.12.004

Selosse, M. A., Baudoin, E., & Vandenkoornhuyse, P. (2004). Symbiotic microorganisms, a key for ecological success and protection of plants. *Comptes Rendus Biologies*, *327*(7), 639–648. https://doi.org/10.1016/J.CRVI.2003.12.008

Sharma, S. K., Ramesh, A., Sharma, M. P., Joshi, O. P., Govaerts, B., Steenwerth, K. L., & Karlen, D. L. (2010). Microbial Community Structure and Diversity as Indicators for Evaluating Soil Quality. In *Biodiversity, Biofuels, Agroforestry and Conservation Agriculture* (pp. 317–358). Springer, Dordrecht. https://doi.org/10.1007/978-90-481-9513-8_11

Sharp, C. E., Brady, A. L., Sharp, G. H., Grasby, S. E., Stott, M. B., & Dunfield, P. F. (2014). Humboldt's spa: microbial diversity is controlled by temperature in geothermal environments. *The ISME Journal*, *8*(6), 1166–1174. https://doi.org/10.1038/ismej.2013.237

Shmida, A., & Wilson, M. V. (1985). Biological Determinants of Species Diversity. In *Source: Journal of Biogeography* (Vol. 12, Issue 1). https://about.jstor.org/terms

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, *236*(4803), 787–792. https://doi.org/10.1126/science.3576198

Smith, D. J., Ravichandar, J. D., Jain, S., Griffin, D. W., Yu, H., Tan, Q., Thissen, J., Lusby, T., Nicoll, P., Shedler, S., Martinez, P., Osorio, A., Lechniak, J., Choi, S., Sabino, K., Iverson, K., Chan, L., Jaing, C., & McGrath, J. (2018). Airborne Bacteria in Earth's Lower Stratosphere Resemble Taxa Detected in the Troposphere: Results From a New NASA Aircraft Bioaerosol Collector (ABC). *Frontiers in Microbiology*, *9*(AUG), 1752. https://doi.org/10.3389/fmicb.2018.01752

Smith, D. J., Timonen, H. J., Jaffe, D. A., Griffin, D. W., Birmele, M. N., Perry, K. D., Ward, P. D., & Roberts, M. S. (2013). Intercontinental dispersal of bacteria and archaea by transpacific winds. *Applied and Environmental Microbiology*, *79*(4), 1134–1139. https://doi.org/10.1128/AEM.03029-12

Soman, C., Li, D., Wander, M. M., & Kent, A. D. (2017). Long-term fertilizer and crop-rotation treatments differentially affect soil bacterial community structure. *Plant and Soil*, *413*(1–2), 145–159. https://doi.org/10.1007/s11104-016-3083-y

Souffreau, C., Pecceu, B., Denis, C., Rummens, K., & De Meester, L. (2014). An experimental analysis of species sorting and mass effects in freshwater bacterioplankton. *Freshwater Biology*, *59*(10), 2081–2095. https://doi.org/10.1111/fwb.12408

Strickland, M. S., Lauber, C., Fierer, N., & Bradford, M. A. (2009). Testing the functional significance of microbial community composition. *Ecology*, *90*(2), 441–451. https://doi.org/10.1890/08-0296.1

Svoboda, P., Lindström, E. S., Ahmed Osman, O., & Langenheder, S. (2018). Dispersal timing determines the importance of priority effects in bacterial communities. *ISME Journal*, *12*(2), 644–646. https://doi.org/10.1038/ismej.2017.180

Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M. C., Schwager, M., & Jeltsch, F. (2004). Animal species diversity driven by habitat heterogeneity/diversity: The importance

of keystone structures. In *Journal of Biogeography* (Vol. 31, Issue 1, pp. 79–92). Wiley/Blackwell (10.1111). https://doi.org/10.1046/j.0305-0270.2003.00994.x

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., Prill, R. J., Tripathi, A., Gibbons, S. M., Ackermann, G., Navas-Molina, J. A., Janssen, S., Kopylova, E., Vázquez-Baeza, Y., González, A., Morton, J. T., Mirarab, S., Xu, Z. Z., Jiang, L., … Zhao, H. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*, *551*(7681), 457–463. https://doi.org/10.1038/nature24621

Tignat-Perrier, R., Dommergue, A., Thollot, A., Keuschnig, C., Magand, O., Vogel, T. M., & Larose, C. (2019). Global airborne microbial communities controlled by surrounding landscapes and wind conditions. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-51073-4

Torsvik, V., Øvreås, L., & Thingstad, T. F. (2002). Prokaryotic diversity--magnitude, dynamics, and controlling factors. *Science (New York, N.Y.)*, *296*(5570), 1064–1066. https://doi.org/10.1126/science.1071698

Upchurch, R., Chi, C. Y., Everett, K., Dyszynski, G., Coleman, D. C., & Whitman, W. B. (2008). Differences in the composition and diversity of bacterial communities from agricultural and forest soils. *Soil Biology & Biochemistry*, *40*(6), 1294–1305. https://doi.org/10.1016/j.soilbio.2007.06.027

Urbanová, M., Šnajdr, J., & Baldrian, P. (2015). Composition of fungal and bacterial communities in forest litter and soil is largely determined by dominant trees. *Soil Biology and Biochemistry*, *84*, 53–64. https://doi.org/10.1016/J.SOILBIO.2015.02.011

Van Diepen, L. T. A., Frey, S. D., Landis, E. A., Morrison, E. W., & Pringle, A. (2017). Fungi exposed to chronic nitrogen enrichment are less able to decay leaf litter. *Ecology*, *98*(1), 5–11. https://doi.org/10.1002/ecy.1635

Vannette, R. L., & Fukami, T. (2017). Dispersal enhances beta diversity in nectar microbes. *Ecology Letters*, *20*(7), 901–910. https://doi.org/10.1111/ele.12787

Veen, G. F. (Ciska), Snoek, B. L., Bakx-Schotman, T., Wardle, D. A., & Putten, W. H. van der. (2019). Relationships between fungal community composition in decomposing leaf litter and home-field advantage effects. *Functional Ecology*, *33*(8), 1524–1535. https://doi.org/10.1111/1365-2435.13351

Vellend, M. (2010). Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology*, *85*(2), 183–206. https://doi.org/10.1086/652373

Voříšková, J., & Baldrian, P. (2013). Fungal community on decomposing leaf litter undergoes rapid successional changes. *The ISME Journal*, *7*(3), 477–486. https://doi.org/10.1038/ismej.2012.116

Vos, M., Wolf, A. B., Jennings, S. J., & Kowalchuk, G. A. (2013). Micro-scale determinants of bacterial diversity in soil. In *FEMS Microbiology Reviews* (Vol. 37, Issue 6, pp. 936–954). https://doi.org/10.1111/1574-6976.12023

Wagg, C., Bender, S. F., Widmer, F., & van der Heijden, M. G. A. (2014). Soil biodiversity and soil community composition determine ecosystem multifunctionality. *Proceedings of the National Academy of Sciences*, *111*(14), 5266–5270.

https://doi.org/10.1073/pnas.1320054111

Walters, K. E., & Martiny, J. B. H. (2020). Alpha-, beta-, and gamma-diversity of bacteria varies across habitats. *PLOS ONE*, *15*(9), e0233872. https://doi.org/10.1371/JOURNAL.PONE.0233872

Walters, K. E., Capocchi, J. K., Albright, M. B. N., Hao, Z., Brodie, E. L., Martiny, J. B. H. (2021). The routes and rates of bacterial dispersal impact surface soil microbiomes and their functioning [Unpublished manuscript]. Department of Ecology and Evolutionary Biology, University of California, Irvine.

Whitman, W. B., Coleman, D. C., & Wiebe, W. J. (1998). Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences*, *95*(12), 6578–6583. https://doi.org/10.1073/pnas.95.12.6578

Wilkinson, D. M., Koumoutsaris, S., Mitchell, E. A. D., & Bey, I. (2012). Modelling the effect of size on the aerial dispersal of microorganisms. *Journal of Biogeography*, *39*(1), 89–97. https://doi.org/10.1111/j.1365-2699.2011.02569.x

Yang, P., & van Elsas, J. D. (2018). Mechanisms and ecological implications of the movement of bacteria in soil. *Applied Soil Ecology*, *129*(1), 112–120. http://www.rug.nl/research/portal.

Zak, D. R., Holmes, W. E., White, D. C., Peacock, A. D., & Tilman, D. (2003). Plant diversity, soil microbial communities, and ecosystem function: Are there any links? *Ecology*, *84*(8), 2042–2050. https://doi.org/10.1890/02-0433